

금융이해력, 거주지와 개인연금 및 보장성 보험 가입의 관계*

The Relationship between Financial Literacy, Residing Area, and the Demand for Private Pension and Insurance

손 준 호**·박 소 정***

Junho Son·Sojung Carol Park

이 연구에서는 제13차 재정패널조사 결과를 활용하여, 금융소비자들의 이자 개념에 관한 금융이해력(Financial Literacy)을 평가하고, 금융이해력, 거주지와 보장성 보험 및 개인연금에 가입이 가지는 관계에 대해 분석하였다.

31%의 응답자들이 상대적으로 간단한 금융 개념인 이자에 대한 이해도가 낮은 것으로 나타났고, 금융이해력은 교육수준이나 성별, 직업, 소득수준, 순자산 등과 같은 인구특성에 따라 다르게 나타났다. 금융이해력이 금융상품을 가입하는 금융행동(Financial Behavior)에 미치는 영향을 겉보기 무관 회귀(Seemingly Unrelated Regression; SUR)모형으로 분석한 결과, 금융이해력은 전체 설문 대상자의 12%만이 가입한 개인연금 가입과는 유의적인 관계가 있으나, 86%가 가입한 보장성 보험의 가입과는 관계가 없는 것으로 나타났다.

그리고, 연령, 교육수준, 금융이해력 등을 통제한 후에도 수도권 거주자가 비수도권 거주자와 비교하면 개인연금 가입 비율이 2% 이상 높은 것으로 나타났는데, 이는 개인의 금융이해력뿐 아니라 거주지의 사회·문화적 인프라가 금융역량(Financial Capacity)에 영향을 미친 결과로 보인다.

국문 색인어: 금융이해력, 금융역량, 개인연금, 보장성보험

한국연구재단 분류 연구분야 코드: B030700, B051602, C030805, C120304

* 본 연구는 손준호(지도교수: 박소정)의 공동연구로 서울대학교 경영대학 석사학위논문으로 사용될 논문을 기반으로 작성되었음. 본 논문은 서울대학교 경영연구소 연구비 지원으로 수행되었음.

** 서울대학교 경영학과 석사과정(junhoson@snu.ac.kr), 제1저자

*** 서울대학교 경영학과 교수(sojungpark@snu.ac.kr), 교신저자

논문 투고일: 2022. 2. 3, 논문 최종 수정일: 2022. 6. 13, 논문 게재 확정일: 2022. 8. 18

I. 서론

한국경제연구원(2022)에 따르면, 대한민국의 노인빈곤율¹⁾은 '20년 기준 40.4%로, 조사 대상 OECD 37개국 중 최고 수준이라고 밝혔다. 이는 G5 국가²⁾의 평균인 14.4%의 대략 3배에 달하는 높은 수치이다. 정부 차원에서는 2015년 「노후준비 지원법」을 시행하며 노후준비에 대한 인식 개선을 위한 기반을 마련하고자 하였으며, 학계에서도 개인연금 가입에 영향을 미치는 요인들을 분석하고 연금 활성화 방안을 논의하며 노후 대비와 관련한 연구를 계속하여 진행하고 있다. 연금 개혁을 통해 새로운 정책이 효과를 발휘하기 위해서는 개개인이 노후 대비의 필요성에 대해 먼저 인지할 필요가 있는데, 이와 같은 문제의식을 바라보는 것은 기초적인 금융 지식에 대하여 올바른 이해를 하는 것으로부터 시작된다. 확정급여형(Defined Benefit)이 주를 이루던 미국의 기업연금 시장이 확정기여형(Defined Contribution)으로 급격히 전환되기 시작하면서, 개인들이 직접 은퇴소득과 관련된 중요한 의사결정을 해야 하는 시대로 접어들었고, 개인이 과연 그러한 의사결정을 잘 할 수 있는지는 중요한 질문을 시작하게 되면서, 금융이해력(Financial Literacy)이라는 개념이 중요하게 등장하게 되었다. 2000년대에는 전 세계적으로 다양한 형태로 금융이해력을 측정하고, 금융이해력과 노후, 은퇴 설계, 재산 형성과의 관계를 보고자 하는 실증연구들이 진행됐다(Lusardi and Mitchell 2007; Bucher-Koenen and Lusardi 2011; Lusardi and Mitchell 2011a, 2011b; Beckmann 2013). 금융이해력은 은퇴 계획에만 국한된 이야기는 아니며, 대출 의사결정(Lusardi and Tufano 2009; Moore 2003; Campbell 2006), 주식시장 참여(Van Rooij et al. 2011a), 생명보험 가입(Lin et al. 2017), 생명보험 해지(Nolte and Schneider 2017)와 관련된 연구에서도 의사결정과 관련하여 유의미한 관계가 있음을 보여주고 있다.

기존의 연구가 주로 이자, 인플레이션, 분산 투자 등의 금융 지식에 대해 개인이 직접 이해하고 있는 정도를 측정하고 분석하였다면, Lachance(2014)는 주변의 환경이 개인의 금융 이해도에 미치는 영향과, 바람직한 금융행동(Financial Behavior)으로 이끄는 의사

1) 66세 이상 인구 중 가처분소득이 전체 인구 중위소득의 50% 이하인 자의 비율.

2) 미국, 일본, 독일, 영국, 프랑스 5개국. '18년 기준: 일본, 독일. '19년 기준: 그 외.

결정을 내릴 수 있는 역량인 금융역량(Financial Capacity)에 미치는 영향에 주목하였다. Lachance(2014)는 교육수준이 높은 이웃이 개인의 금융이해력 향상에도 영향을 미칠 수 있으며, 주변 환경 자체가 금융이해력을 넘어서 더욱더 바람직한 금융 행동으로 이끄는 데 영향을 미칠 수 있다는 결과를 보여주었다.

본 연구에서는 제13차 「재정패널조사」의 설문 결과를 활용하여, 이자 개념과 관련된 금융이해력을 측정하고, 금융이해력과 개인연금, 보장성 보험 가입 간의 관계를 보았으며, 수도권 거주 여부와 금융이해력, 금융상품 가입과의 관계를 분석하였다. 분석결과, 31%의 응답자들은 이자에 대해 정확히 이해하지 못하고 있었으며, 선행연구들과 같이 성별, 연령, 교육수준, 소득수준 등의 인구통계학적 특성에 따라 금융이해력이 다른 것으로 나타났다. 단, 거주지의 경우 기존 선행연구들과는 달리 수도권 거주자들의 금융이해력이 비수도권 거주자들보다 더 높은 것은 아니었다. 금융이해력과 개인연금, 보장성 보험 가입 간의 관계는 상반되게 나타났다. 개인연금은 제13차 「재정패널조사」 설문대상자의 12%, 상대적으로 일부의 금융소비자들이 활용하고 있는 금융상품이고, 실손, 질병, 사망 보장 등의 보장성 보험은 80% 이상의 국민이 적어도 하나는 가입하고 있는 훨씬 대중적인 금융상품이다. 개인연금 가입에서는 금융이해력과 거주지가 유의한 상관관계를 보였지만, 보장성 보험의 경우 금융이해력이나 거주지 모두 관련이 없거나 오히려 음의 상관관계를 가지는 특성을 보여주었다. 즉, 본 연구 결과는 누구나 알고, 모두가 가입하는 금융상품에 비해, 상대적으로 복잡하고 잘 알려지지 않은 상품의 경우, 금융이해력과 주변 환경이 더 큰 영향을 미친다는 것을 보여준다.

이 연구는 다음과 같이 기여하고 있다. 첫째, 금융이해력 관련 연구들이 그간 보여왔던 것처럼, 꽤 많은 사람이 간단한 금융 개념조차도 제대로 이해하지 못하고 있다는 것을 금융이해력에 대한 독립적인 다른 측정치를 사용하여 보여준다.

둘째, 금융이해력이 인구특성별로 다르게 나타난다는 연구들에 또 하나의 실증 근거를 제시하고 있다. 기존 연구와 같이 30, 40대의 금융이해력이 20대나 50대 이상보다 높게 나타나는 형태를 보였고, 교육수준도 금융이해력에 영향을 미치는 것으로 나타남에 따라, 세부집단별로 차별화된 금융교육이 필요하다는 것을 보여준다.

셋째, 금융이해력이 금융상품의 가입 의사결정, 즉 금융 행동에 영향을 미친다는 것을

보여주었다. 내생성 문제를 엄밀하게 다루지 못했다는 문제가 남아있긴 하나, 기존 해외 연구들과 동일하게 개인연금의 가입과 금융이해력이 양의 상관관계를 가진다는 것을 보여 주고 있다. 이찬희·정홍주(2013)는 둘 간의 관계를 제시하였지만, 다소 제한된 표본 크기와 사무관리직 종사자가 많은 표본 특성으로 전체를 대표하지는 못하였다. 본 연구는 충분한 표본 크기를 확보하고, 최근에 조사된 설문의 결과에서 금융이해력 변수를 활용한 분석을 통하여 추가적인 근거를 제시하였다.

다섯째, 기존의 국내의 문헌에서는 금융이해력이 금융상품별로 가입 의사결정이 어떻게 달리 나타나는지를 보여준 연구는 희박한데, 본 연구에서는 잘 알려진 상품과 그렇지 않은 상품에 대한 가입 의사결정에 있어서, 상대적으로 잘 알려지지 않은 상품의 경우 금융이해력과 주변 환경이 더 큰 영향을 미치고 있다는 것을 처음으로 보여주었다.

마지막으로 거주지가 금융이해력 자체와는 큰 관련이 없으나, 금융이해력과 소득, 순자산, 성별, 연령 등 다른 인구통계학적 특성 등을 통제한 뒤에도 거주지가 개인연금 가입에 양의 상관관계를 보인다는 것은 수도권과 비수도권 사이의 다양한 환경의 차이가 금융역량에 유의미한 영향을 미치고 있다는 것을 보여주었다.

이 연구는 다음과 같은 정책적 함의를 가진다. 첫째, 많은 사람이 기본적인 금융 개념에 대한 이해력이 부족한 것으로 나타났다. 둘째, 금융이해력은 세부집단에 따라 다르게 나타나므로, 금융이해력 향상을 위한 노력을 할 때 세분화된 프로그램을 운영하는 것이 필요하다는 것을 알려준다. 셋째, 금융이해력은 실질적인 금융 행동과 관련이 있으며, 이는 자산의 형성과 금융행복지수(Financial Wellbeing)에까지 영향을 미칠 수 있으므로 금융이해력을 높이는 것은 중요하다. 넷째, 보장성보험과 개인연금 가입에 금융이해력이 미치는 차이를 보면 새로운 금융상품을 만들어 내는 것도 중요하지만, 그러한 상품이 광범위하게 다양한 집단에서 활용될 수 있게 하기 위해서는 금융이해력 향상에 대한 노력도 함께 이루어져야 한다는 것을 알려준다. 마지막으로 거주지와 개인연금 가입 간의 관계를 보았을 때, 금융역량은 금융이해력 향상으로만 이루어지는 것은 아니며, 주변의 여러 요인도 큰 영향을 미칠 수 있다는 것을 이 연구는 보여준다. 본 연구에서는 정확한 원인에 대한 분석은 없으나, 비수도권보다 수도권 거주자들이 금융상품 활용을 더 잘하고 있음을 보였다. 이는, 국민의 균형 있는 금융역량 향상을 위해서 다양한 연구와 정책 개선이 필요함을 시사한다.

이 연구는 다음과 같이 구성된다. 2장에서는 국내외의 관련 기존 문헌을 소개하고, 3장에서는 본 논문에서 사용된 데이터와 본 논문에서 사용한 금융이해력 변수를 설명한다. 4장에서는 금융이해력, 거주지와 개인연금 및 보장성보험 간의 관계에 관한 연구 결과를 보여준다. 5장에서는 강건성 검사를 하고, 마지막 6장에서는 연구의 결론 및 미래에 필요한 연구에 대해 논한다.

II. 선행연구

1. 금융이해력(Financial Literacy)

개인들이 금융시장에 적극적으로 참가하고, 최근 코로나 19와 금융의 디지털화로 흐름은 가속화되고 있다. 개인이 직접 저축, 투자, 대출, 부동산 매매, 보험상품 가입 등 다양한 금융 의사결정을 해나가야 하며, 그러한 의사결정의 결과가 개인의 자산 형성에 지대한 영향을 미치게 된다. 기존 금융이해력에 대한 일련의 연구들에서는 개인이 금융 지식과 금융에 대한 이해력 자체가 부족하다는 것을 보여주고 있으며, 행동경제학이나 행동재무학은 개인의 재무적 의사결정이 전통 경제학 틀 안에서의 의사결정자와는 달리 이성적이지 않을 수 있다는 결과를 제시하고 있다.

초기의 금융이해력 측정은 복리 이자, 인플레이션 등 금융에 대한 기초적인 개념을 물어보는 형태로 진행되었으며, Lusardi and Mitchell(2007)에서 처음 제안되고 Lusardi and Mitchell(2008, 2011a, 2011b)에서 발전된 형태의 질문지는 전세계적으로 사용되고 있는 금융이해력 측정 기준으로 자리를 잡고 있다. Lusardi and Mitchell의 방법론은 1)복리 이자, 2)인플레이션, 3)위험 분산의 세 영역에 대한 이해를 측정한다.³⁾ OECD/INFE⁴⁾는 금융이해력을 금융지식, 금융태도, 금융행동의 3가지 범주로 정의하고

3) 세 영역을 다룬 대표적인 금융이해력 측정 문항을 “Big 3”이라고 하며, “Big 3”을 포함하고 대출과 채권 가격에 대한 개념을 물어보는 문항이 추가된 설문지를 “Big 5”라고 한다. 실제 문항은 Appendix 2를 참고하기를 바란다.

4) International Network on Financial Education: 금융교육에 대한 국제적인 가이드라인 제공을 목적으로 2008년에 설립된 OECD 산하 기구이다.

금융이해력 지수를 산출하도록 하고 있으며, 우리나라에서는 한국은행-금융감독원 「전국민 금융이해력 조사」를 2012년부터 2년 주기로 시행하며 금융이해력 지수 조사에 참여하고 있다.

금융이해력 수준에 관한 연구들은 공통되게 금융이해력이 모두에게 주어지는 것이 아니라 상당수의 개인이 금융에 대한 기본 개념들에 대한 이해가 부족하고, 이자 계산능력 등이 떨어진다는 사실을 보고한다. Lusardi and Mitchell(2011a)는 약 67% 정도만이 이자에 대해 이해하고, 75%가 인플레이션에 대해 이해를 하고 있으며 다소 복잡한 개념인 위험 분산은 52%만이 이해하고 있는 것을 보였다. Lusardi(2011)은 조금 더 복잡한 개념인 이자와 채권 가격 간 관계를 물어본 문항에서는 21%만이 알고 있다는 것을 보여주고 있다. 국가마다 다소 차이는 있으나, 다른 국가들에서 진행된 연구에서도 유사한 수준의 이해도를 보여주고 있다.

금융이해력은 연령, 성별, 교육수준, 소득수준 등에 따라 다르게 나타나는 것으로 나타난다. 대체로 여성이 남성보다 이해도가 낮으며 (Lusardi and Mitchell 2008), 교육수준이 높을수록 이해도가 높으며 (Lusardi and Mitchell 2007; Behrman et al. 2012), 소득수준이 높을수록 (Lusardi and Tufano 2009), 자산이 많을수록 (Monticone 2010), 직장인으로 고용되어 있을수록 (Lusardi and Mitchell 2011a, 2011b) 더 이해도가 높다고 보고하고 있다. 그리고, 금융이해력은 실질적인 금융 행동에 영향을 미칠 수 있음을 보였다. Lusardi and Mitchell(2007)는 금융이해력이 높을수록 은퇴에 대한 계획을 수립하며, 실제로 자산의 형성과 유의미한 관계가 있는 것을 보여주었다. 금융이해력이 낮은 사람들이 주식시장에 덜 참가하고 (Van Rooij et al. 2011a), 뮤추얼 펀드 선택에서도 수수료가 높은 펀드를 선택하며 (Hastings and Tejada-Ashton 2008), 모기지 및 각종 대출의 비용도 높게 지출하고 있는 것으로 나타났다(Moore 2003; Lusardi and Tufano 2009). 또한, 금리가 하락하는 시기에 리파이낸싱을 덜 하는 경향을 보였으며 (Campbell 2006), 금융위기 시기에도 더 큰 어려움을 겪은 것으로 나타났다(Klapper et al. 2013).

본 연구와 직접 관련이 있는 연금상품 가입, 은퇴소득과 관련된 금융 의사결정과 금융이해력과의 관계에 관한 실증연구도 다양한 나라에서 진행되어 왔다. Fornero and Monticone(2011)는 이탈리아 국민의 금융이해력은 낮지만, 금융이해력이 높은 집단에서

는 연금 수요가 높음을 보였으며, Klapper and Panos(2011)는 러시아 국민의 개인연금 가입을 장려하기 위해서는 금융이해력을 높이기 위한 노력이 필요하다고 주장하였다. Bucher-Koenen and Lusardi(2011)은 독일 국민의 금융이해력을 측정하였으며, 동유럽 대비 서유럽 거주자에게서 금융이해력이 높게 나타났고, 금융이해력이 높은 집단에서는 은퇴 설계를 하는 것으로 나타났다. Sekita(2011)는 일본 국민의 전반적인 금융이해력 수준은 높지 않다고 하였으며, 금융이해력을 높이는 것은 은퇴를 위한 저축 계획을 수립할 확률을 높인다고 주장하였다. Van rooij et al.(2011b)은 네덜란드의 많은 가정에서 은퇴에 대해 준비하지 않고 있으나, 금융이해력이 높은 집단에서는 노후 대비를 하고 있음을 보였다. 그리고, Song(2020)은 복리 이자 개념에 대해 알려진 집단과 그렇지 않은 집단 간에 연금 납입액의 차이가 있으며, 학습을 통해 연금 납입액이 40%가량 증가함을 보였다. 하지만, 이러한 연구들은 금융상품 가입의 경험과 관심 자체가 금융이해력을 높일 수도 있기에 금융이해력이 올바른 금융 행동으로 이어진다고 결론을 내리기에는 역인과성(Reverse Causality)과 내생성(Endogeneity) 문제가 남아 있다. 이러한 문제를 해결하기 위해, 다양한 연구들은 도구변수(Instrument Variable)를 활용하여 내생성 문제에 대한 강건성 분석을 하였는데(Behrman et al. 2012; Van Rooij et al. 2011b; Bucher-Koene and Lusardi 2011), 도구변수를 활용할 경우 오히려 그 설명력이 더 높게 나타나는 경향이 있어서 역인과성을 완전히 배제하기는 어려울지라도, 개인의 금융이해력이 금융 의사결정에 영향을 미칠 것이라고 결론을 내리고 있다.

2. 금융역량(Financial Capability)과 거주지

금융이해력과 연관된 개념으로 금융역량(Financial Capability)이라는 개념이 있는데, Financial Industry Regulatory Authority(FINRA)는 2009년부터 3년마다 미국인의 금융역량을 측정하는 설문을 진행하고 있다. 금융이해력 질문은 개인의 금융 지식에 대한 이해도를 측정하는 것이라면, 금융역량은 조금 더 포괄적인 개념으로, 바람직한 금융 의사결정을 내릴 수 있는 능력을 의미한다. 개인의 금융 지식은 금융 행동으로 완벽히 같은 형태로 이어지는 않다 보니, 금융이해력보다 금융역량이 개인의 재무적 결과 및 금융 행복 지수에 더욱더 연관되는 개념일 수 있다. 그러나, 금융이해력과 비교하면 선행연구는 많지

않은 편이다. Collins(2012)는 금융이해력과 금융역량이 낮은 사람은 금융 조언을 덜 받아들이려는 경향이 나타났으며, 반대로 높은 경우에는 투자, 보험, 세금 등의 조언을 잘 받아들이는 결과를 보였다. Lin et al.(2017)은 금융이해력이 높고, 금융 컨설팅을 받은 경험이 있는 경우에는 생명보험 가입률이 높고 은퇴에 대해 준비를 하고 있다는 결과를 보였다. 금융이해력과 금융역량은 스스로 만들어 가는 부분도 있으나, 사회적 환경에 의해 영향을 받을 수 있다. Yan et al.(2021)은 도시의 출생 및 거주자가 금융이해력이 더 높음을 밝혔고, 이들은 비도시와 도시 이주자들에게 금융 주제에 더 많은 관심을 기울이고 금융 활동에 더 참여할 수 있도록 장려하면서, 사회 연결망을 통해 금융이해력 파급효과(Spillover effect)를 가져온다고 주장하였다.

주변 환경이 금융이해력 및 금융역량에 영향을 미친다는 것을 보여준 대표적인 연구로 Lachance(2014)가 있다. Lachance(2014)는 우편번호(zip code) 수준에서 이웃들의 교육수준이 개인의 금융이해력과 금융역량에 어떠한 영향을 미치는지를 분석하였다. 다양한 지식의 접근이 쉽고 정교한 사회 연결망의 정도를 측정하기 위해 지역별 평균 학력 수준을 대리변수로 사용하여 개인의 금융이해력과의 관계를 측정한 결과, 주변 환경은 개인의 금융이해력에 긍정적인 영향을 미칠 수 있음을 보였다. 또한, Lachance(2014)는 단순한 지식보다 주변의 도움이 금융 행동에 더 큰 영향을 미칠 수도 있다고 주장하는데, 예를 들면, 증권계좌를 개설할 때, 증권계좌에 관한 개인의 이해수준보다 지인이 여는 방법을 쉽게 알려준다거나, 믿을 수 있는 재무 설계사가 더 큰 영향을 미칠 수 있다는 것이다. 저자는 사회적 학습(social learning)이 실질적으로 다양한 금융행동 - 은행계좌, 증권계좌, 모기지 대출, 건강보험, 개인연금가입, 수수료 높은 대출 사용, 신용카드 연체 수수료 지불 등 - 과 같은 행동과 통계적으로 유의한 상관관계를 가지는 것을 보이는 것을 통해 사회적 학습이 중요할 수 있다는 것을 보였다. 단, 이 연구에서도 내생성이나 역인과성 문제는 완전히 해소되지 않아 추후 더 연구가 필요하다고 언급하였다.

Lachance(2014)와 같이 우편번호 수준에서의 세밀한 분석이 이루어지지는 않았지만, 거주지와 금융이해력, 금융 행동의 관계에 관한 연구들은 다양하게 존재한다. 러시아(Klapper and Panos 2011), 폴란드(Beckmann 2013), 미국(Lusardi and Mitchell 2014) 그리고 전 세계적으로 금융이해력을 측정한 연구에서도(Klapper and Lusardi

2020), 비도시 지역보다는 도시지역, 특히 수도권 거주자의 금융이해력이 더 높다는 결과를 보여주고 있다. 우리나라는 수도권과 비수도권의 사회 및 문화적 인프라 차이가 크게 나타나며, 금융교육에서도 지역 간 격차가 존재하는 것으로 여겨진다. 한경동(2016)은 경제 개설했던 학교 수와 경제 수업을 희망하는 학생 수가 수도권에서 훨씬 많음을 보였고, 민세진·홍진주(2021)는 수도권에 위치하면서 학교의 규모가 크고 지역의 소득수준이 높으면 경제 과목이 개설될 가능성이 크다는 결과를 보였다. 그러나, 거주지와 개인연금 가입 간의 관계를 살펴본 선행연구들의 결과는 통계적으로 유의수준이 10% 내외로 낮거나 유의하지 않았으며, 표본 크기가 충분히 확보되지 못하여 안정적인 연구 결과라고 보기에는 어려웠다. 차경욱(2015)은 거주지 간 개인연금 가입 여부는 통계적으로는 차이가 없으나, 가입자의 납입액은 수도권 거주자가 더 적음을 보였다. 여운경·이남희(2012)는 비수도권에서 개인연금 가입이 더 높게 나타남을, 개인연금보험으로 분석대상을 한정한 백은영(2012)은 중소 및 대도시 지역의 거주자가 더 높은 가입률이 나타남을 보였다.

III. 사용자료 및 변수설명

1. 사용자료

본 연구는 이경희·전병욱(2020) 등 최근 개인연금 가입에 대한 의사결정 관련 선행연구에서 사용한 것과 동일한 자료인 한국조세재정연구원 「재정패널조사」(National Survey of Tax and Benefit)를 활용하였다. 「재정패널조사」는 2008년을 시작으로 매년 시행되고 있으며, 설문지를 활용하여 대면조사를 통해 가계의 재무적, 비재무적 자료를 수집한다. 국민연금, 퇴직연금, 개인형 퇴직연금(IRP), 연금저축, 연금보험 등의 가입 여부, 그리고 가입자의 경우 연금 혹은 일시금 수령과 같은 항목 또한 다루고 있다. 13차에서는 개인의 시간선호를 측정하는 문항과 출생지 문항이 신규로 추가되었는데, 본 연구에서는 개인의 시간선호를 측정하는 문항을 활용하여 금융이해력 변수를 생성하기 위해 13차 데이터만을 사용하였다.

본 패널조사는 1차년도와 2차년도에서 조사한 가구를 원표본 가구라고 정의한다. 시계열 유지를 위해 같은 가구를 계속해서 추적하는 특성이 있으며, 13차 기준 원표본 유지율은 71.1%에 달한다. 13차 데이터는 기존 설문을 진행한 4,762개의 가구와 4,030개의 신규 가구가 추가되어, 총 8,792개의 가구와 14,759명의 가구원으로 구성되었다. 신규 가구는 이번 차시에서만 이례적으로 추가되었고 이후에는 추적하지 않는다. 비록 13차 데이터만을 활용하였으나, 상당한 수준의 표본이 포함되었고, 기존의 그 어떤 금융이해력 관련 연구보다도 큰 표본을 가진 연구라고 할 수 있다.

설문지는 가구 조사와 가구원 조사로 구성되어 있다. 가구 조사는 가구원 중 한 명이 대표로 가구원의 나이, 학력 등과 같은 가구의 기본 현황과 가구 전체의 경제상태를 확인할 수 있는 자산, 부채, 지출 등의 값을 써넣는다. 가구원 조사는 개인별로 직접 문항에 응답하며, 개인의 경제활동 상태, 소득 및 연금/보험 관련 지출 현황뿐만 아니라, 우리나라의 복지 수준 정도, 개인의 위험회피 정도, 건강 상태 및 예상 은퇴 연령 등의 다양한 문항에 대해 답을 기재하게 된다. 본 연구에서는 가구 조사와 가구원 조사를 결합하여 데이터를 사용하였다. 본 연구에서는 경제활동인구의 금융상품 가입 의사결정을 중점으로 다루고자 연령을 20세 이상, 60세 미만으로 범위를 한정하였다. 그중에서 연구에 필요한 변수를 빠짐없이 가지고 있는 관측치는 총 7,579명의 가구원이었다.

2. 금융이해력 변수 정의

금융이해력은 「재정패널조사」 13차에서 다음의 개인의 시간에 대한 선호를 측정하기 위해 추가된 문항에서 착안하여 변수를 생성하였다.

〈Table 1〉 Questionnaire for measuring time preference

* Please read and respond to the text below.

Suppose that you have 20 tokens. If you exchange one token for cash tomorrow, you'll get 100,000 won. On the other hand, if you exchange it in one month from now (that is, one month + 1 day), the value of token is assessed as follows.

Q. How many tokens are expected to be exchanged tomorrow and how many in a month?

Category	Exchange value for one token (tomorrow)	Exchange value for one token (in a month)	① number of token exchanged tomorrow	② number of token exchanged in a month	Total (①+②)
1	100,000 won	101,000 won			20
2	100,000 won	101,250 won			20
3	100,000 won	101,500 won			20
4	100,000 won	101,750 won			20
5	100,000 won	102,000 won			20

Source: Question h1 on page 28 in the 13th questionnaire

문항에 따라, 한 달 뒤 교환 시 토큰 1개의 가치가 101,000원에서 102,000까지 증가하는데, 이는 이자율로 변환하였을 때, 월 1%, 1.25%, 1.5%, 1.75%, 2% 상황과 같다. 응답자는 5개의 주어진 상황에서 각각의 문항별로 내일, 한 달 뒤 수령 할 토큰의 개수를 각각 기재하게 된다.

문항별로 한 달 뒤 받고자 하는 토큰의 증감을 살펴보았으며, 이를 토대로 시간선호와 이자에 대한 이해도를 측정하였다. 1번에서 5번으로 갈수록 한 달 뒤 토큰의 가치가 증가하므로 이자가 높아진다고 볼 수 있으며, 이 개념에 대해 이해하고 있다면, 1번에서 5번으로 갈수록 한 달 뒤 교환 개수를 증가시킬 것이다. 하지만, 이자에 대한 이해가 있더라도, 가장 높은 5번의 이자도 충분하지 않다고 생각하는 경우에는 모든 질문에 대해 내일 교환

할 토큰의 개수를 20이라고 응답할 수 있을 것이다. 이러한 논리를 바탕으로, <Table 2>와 같이 다섯 가지의 그룹으로 분류하였다.

<Table 2> Grouping financial literacy

Group	Definition	N	Total
1	Exchange all tokens tomorrow	1,215	16%
2	Exchange the same number of tokens each time	482	6%
3	Exchange all tokens a month later	3,035	40%
4	Exchange more tokens to higher fixed yield	2,181	29%
5	Exchange random or less number of tokens despite higher yield	666	9%
Total		7,579	100%

Group 1~3은 모든 질문에 대해 같은 답을 한 경우이다. 그 중, Group 1은 현재 시점에서, Group 3은 미래 시점에서 20개의 토큰을 모두 교환, Group 2는 현재와 미래 모두 받는 경우를 제외하면서 항상 같은 수량의 토큰을 교환한다. 이자율이 증가할수록 미래 시점에 받는 토큰의 개수가 더 증가하거나, 중간에 일정한 구간은 있더라도 처음보다는 미래 토큰을 받는 개수가 많아지는 경우는 Group 4로 분류하였다. 이와 반대로 미래에 토큰을 받는 개수가 더 적어지거나, 일관되지 않게 일정한 패턴을 보이지 않으면 이자에 대해 이해하지 못하는 집단이라 분류하고 Group 5라고 분류하였다. 다섯 가지의 그룹은 각각 이자에 대한 이해도가 다른 그룹으로 볼 수 있으며, 이는 <Table 3>과 같이 정리될 수 있다.

<Table 3> Definition of financial literacy

Group	Financial Literacy Level	Literacy	Literacy 2
1	lack of information ⁵⁾	0	1
2	low	0	0
3	high	1	1
4	high	1	1
5	low	0	0

5) Group 1에 포함된 응답자들은 설문 문항을 귀찮게 생각하고 대충 응답하였는지, 혹은 현재 시점에 높은 가치를 두고 충실히 응답하였는지가 불분명하므로, 응답 결과에 대한 신뢰성과 금융이해력 변수의 타당성 검증이 필요하다. Appendix 3에 따르면, Group 1 응답

Group 1은 1번 문항에 대해서도 이자가 충분하지 않다고 판단하는 매우 높은 시간 선호도를 보인 경우로, 잘못된 이자 개념을 가지고 있다고 판단하기에는 정보가 부족하다고 할 수 있으나, 문제에서 제공하고 있는 이자가 월 1% 이상이라는 점을 고려할 때, 금융이해력이 높지 않은 그룹으로 분류하는 것이 타당하다고 판단하였다. 하지만, Group 2나 Group 5와 같이 완벽하게 틀린 경우와는 다르다고 볼 수도 있으므로, Group 1도 금융이해력이 높은 그룹으로 분류하는 Literacy2를 추가로 정의하여 결과에 대한 강건성 분석을 하였다.

금융이해력이 낮다고 분류할 수 있는 Group 2와 Group 5 중에서, Group 5의 경우에는 질문 1에서 5로 가면서 증가하지 않고, 오히려 감소하거나 패턴이 없는 경우로 이자에 대한 이해가 부족한 경우라고 명백히 판단할 수 있다. 반면, Group 2는 이자와 무관하게 일정한 수를 수렴 하겠다고 응답한 경우로, 이자에 대한 이해가 낮은 경우라고 볼 수 있으나, 데이터를 살펴보면 현재와 미래에 각 10개씩, 반반으로 수렴을 선택 경우가 85% 이상으로 대부분을 차지하여, 설문을 성실히 응답하지 않은 집단이라고 볼 수 있을 것으로 판단하였다. 이에 대해, 전체 표본의 약 6%에 해당하는 Group 2를 제외한 강건성 분석도 추가로 수행하였다.

위의 정의를 활용하여 <Table 2> 결과에 적용하면, Group 1, Group 2, Group 5를 제외한 69%는 이자에 대해 이해를 잘하고 있는 집단으로 분류된다. 이 비율은 2020년 한국은행과 금융감독원에서 2,400명을 대상으로 한 「전국민 금융이해력 조사」에서 OECD/INFE 방법으로 조사한 결과, 금융 지식 부문에서 최소 점수를 상회한 비율인 68%와 매우 유사하다. 그리고, 금융이해력 점수는 30대와 40대가 높았고 20대가 가장 낮았으며, 남성보다 여성이 더 높다는 결과를 보였는데, 이는 두 연구 모두 연령대별, 성별로 분류하였을 때에 같은 결과가 나왔음을 확인할 수 있었다.

본 연구의 분석결과를 통해, 대략적인 금융 지식수준을 파악할 수 있었지만, 단일 문항을 사용하여 금융이해력을 측정하였다는 한계점은 남아 있다. 그러나, 한국은행-금융감독원의 조사 대상자보다 세 배 이상의 표본을 확보한 점, 가구의 다양한 비재무 데이터와 금융이해력 간 관계에 대해 추가 분석이 가능하다는 장점이 있다.

자들의 교육수준, 소득, 순자산은 상대적으로 낮게 나타났다. 낮은 교육수준, 현재 여유있지 않은 재무 상황을 미루어 보았을 때, 현재의 가치를 중요시할 것으로 예상해 볼 수 있으며, 이는 개개인의 상황이 적절히 반영된 응답 결과로 여겨진다.

IV. 분석결과

1. 기술 통계량

〈Table 5〉는 금융이해력과 여러 인구통계학적 특성 간의 관계를 보기 위하여 기초통계량을 정리하여 제시한다. 각 변수에 대한 구체적인 정의는 Appendix 1에 있다. 본 연구에서는 이자 개념을 이해하는지와 관련된 질문을 사용하여 금융이해력을 측정하고 있으나, 유사한 이자 개념을 비롯한 위험 분산, 인플레이션 등에 대해서도 묻는 기존 연구들과 다소 다른 금융이해력 정의를 사용하고 있다. 만약 본 연구에서 사용하는 금융이해력이 기존 다른 연구에서 사용하는 것과 다른 것을 측정한다면, 금융이해력과 인구통계학적 특성들과의 관계 또한 다르게 나타날 것이다. 따라서 이 장에서는 본 연구에서 사용하는 금융이해력과 기존 연구들에서 분석했던 인구통계학적 특성과의 관계를 분석하고, 추가로 거주지와 금융이해력과의 관계를 분석해본다.

분석결과 교육수준에 따라 금융이해력은 단조증가 하는 모습을 보인다. 연령은 다른 국내외 선행연구들과 마찬가지로 볼록한 혹은 모양의 형태를 보여, 30대와 40대에서 가장 높은 이해도를, 20대와 50대에서 다소 낮은 이해도를 보였다. 결혼 여부에 따라서는 기혼자의 이해력이 높았고, 소득수준과 순자산과 금융이해력도 금융이해력과 양의 상관관계를 보이는 것으로 나타나 모두 기존 연구와 같은 결과를 보여준다. 거주지의 경우는 수도권과 비수도권에서 금융이해력 점수 차가 없는 것으로 나타났다. 성별은 일반적으로 남성이 금융이해력이 더 높다는 결과를 보인 기존 선행연구와 달리, 본 연구의 결과에서는 여성이 더 높은 것으로 나타났다.

〈Table 4〉는 〈Table 2〉에서의 분류대로 성별의 분포를 제시한다. 이자가 늘어날수록 미래 수령을 늘리는 Group 4는 남녀가 거의 비슷한 비율을 보이나, 오늘 토큰을 모두 수령 하는 Group 1에는 남성이, 미래에 토큰을 모두 수령 하는 Group 3에는 여성이 더 많은 것으로 나타난다. 본 연구에서 사용한 질문은 이자 개념의 이해라는 측면에서는 금융지식을 측정하고 있으나, 현재의 소비 및 미래를 위한 저축과 관련한 금융행위 또는 금융태도 측면도 포함하고 있다. 「2020 전국민 금융이해력 조사」의 결과에 따르면, 저축보다 소비 선호, 미래보다 현재 선호 등의 문항에서는 남성보다 여성의 점수가 더 높았다.

〈Table 4〉 Distribution of financial literacy group by gender

Group	Male	Female	χ^2
1	17.36	14.50	22.32*** (0.000)
2	6.18	6.57	
3	38.24	42.13	
4	28.71	28.85	
5	9.51	7.96	

Note: *p<0.10, **p<0.05, ***p<0.01, P-value in parentheses, Male: 4061, Female: 3518

거주지와 금융이해력 관계의 경우, 수도권과 비수도권 거주자의 소득, 교육수준, 고용상태 등의 인구통계학적인 요소들을 통제하여야 Lachance(2014)나 Yan et al.(2021)에서 보여주는 파급효과(spillover effect) 또는 사회적 학습효과(social learning)에 대한 분석이 가능하다.

이에 대한 분석을 위해, 〈Table 6〉에서는 여러 재무, 비재무 변수와 금융이해력의 관계를 고려하여 진행한 로짓 회귀분석(Logit Regression) 결과를 제시하고 있다. 회귀분석에서는 중학교 졸업 이하인 집단(Edu_low)과 무직 집단(Non-employed) 비율이 매우 낮아서, 교육은 전문대학 졸업 이상과 그 외로, 직업은 상용 근로직과 그 외로 변수를 재정의하여 분석을 진행하였다. 분석결과, 교육수준, 소득수준과 금융이해력은 통계적으로 유의미한 관계가 있었으나, 연령의 경우에는 통계적 유의성이 없었다. 그리고, 다른 요소들을 통제하였을 경우 수도권 거주자의 금융이해력이 높은 것이 아니라 오히려 다소 낮게 나타나는 결과를 보여주었다. 〈Table 5〉와 〈Table 6〉을 종합해보면, 거주자의 교육수준이 전반적으로 높은 수도권에 거주하는 것 자체가 개인의 금융이해력을 높이는 것은 아니라고 결론을 내릴 수 있다.

비록 성별의 관계가 다소 다르게 나타나긴 했으나, 금융이해력이 높은 집단의 비율도 한국은행-금융감독원의 「전국민 금융이해력 조사」 결과와 매우 유사하고, 금융이해력이 높은 집단의 인구통계학적 특성도 매우 유사하게 나타나는 것으로 보아, 이 연구에서 사용하고 있는 금융이해력 변수가 간단하기는 하지만 소비자의 금융이해력을 적절히 측정하고 있다고 볼 수 있다. 단, 본 연구에서 사용한 변수는 단순히 이자에 대한 이해를 넘어서 저축에 대한 선호도를 보이는 긍정적 금융 행동 또한 어느 정도 반영이 되었다는 점에서 약간의 차이가 있다고 볼 수 있다.

〈Table 5〉 Financial literacy by socio-demographic characteristics

Variables	Total		Literacy = 0	Literacy = 1	χ^2
	N	Mean Literacy	Percentage	Percentage	
Education					
Edu_high	4802	0.703	60.26	64.76	15.65*** (0.000)
Edu_mid	2589	0.665	36.73	32.99	
Edu_low	188	0.622	3.00	2.24	
Age					
20~29	729	0.661	10.45	9.24	6.42* (0.093)
30~39	1397	0.709	17.22	18.98	
40~49	2304	0.694	29.79	30.67	
50~59	3149	0.681	42.53	41.10	
Residing Area					
Capital	3466	0.680	46.97	45.17	2.14 (0.144)
Non-capital	4113	0.695	53.03	54.83	
Gender					
Male	4061	0.670	56.79	52.13	14.22*** (0.000)
Female	3518	0.710	43.21	47.87	
Marital Status					
Married	5375	0.696	69.15	71.72	5.22** (0.022)
Single	2204	0.669	30.85	28.28	
Occupational Status					
Employed	5746	0.689	75.58	75.92	3.28 (0.194)
Self-employed	1650	0.679	22.43	21.47	
Non-employed	183	0.743	1.99	2.61	
Other Controls					
Income	7579	0.688	31.18	68.82	6.95*** (0.000)
Net_asset					2.53*** (0.011)
Age_diff					-2.84*** (0.005)
Smoke					-1.46 (0.143)

Note: *p<0.10, **p<0.05, ***p<0.01, P-value in parentheses, Other Controls: t statistics

〈Table 6〉 Logit Regression results: socio-demographic characteristics and financial literacy

Variables	Coef.	(S.E.)
Residing_Area	-0.134***	(0.051)
Income	0.282***	(0.050)
Net_asset	0.009	(0.008)
Age_20	-0.122	(0.105)
Age_30	0.128	(0.078)
Age_40	0.057	(0.061)
Age_diff	-0.010	(0.004)
Smoke	0.035	(0.074)
Married	0.042	(0.068)
Employed	-0.005	(0.060)
Male	-0.186***	(0.059)
Edu_high	0.148***	(0.057)
Constant	-1.001***	(0.290)

Note: *p<0.10, **p<0.05, ***p<0.01, N: 7,579

2. 금융이해력, 거주지와 개인연금 및 보장성보험 가입과의 관계

금융이해력은 주식투자, 대출, 보험 가입 등 다양한 금융 행동과 유의미한 상관관계를 보인다. 또한, 은퇴에 대한 계획 수립, 자발적 연금가입 등과의 인과관계 또한 전 세계의 다양한 실증연구에서 제시하고 있는데, 이러한 연구들은 올바른 금융 의사결정과 은퇴 소득확보를 위해 금융이해력 향상이 매우 중요하다는 결론을 내리고 있다.

금융상품 중에서도 개인연금 상품은 은퇴에 대한 계획, 장기적인 관점에서의 세금 지출 관리 등에 대한 고려가 있는 경우에 활용할 가능성이 크다. 최근, 국민연금의 재정 위기에 따른 연금 개혁 방안을 모색함에 따라, 사적연금의 중요도가 더욱 높아질 것으로 여겨진다. 개인연금 상품들의 세금 혜택은 〈Table 7〉에서 보여주는 것과 같이 소득이 일정 수준 이상일 경우, 연간 약 40만 원에서 150만 원 정도로 적극적으로 활용한다면 혜택이 상당한 수준이라고 볼 수 있다. 하지만, 본 연구의 설문대상자 12% 정도만이 가입하고 있으며, 아직은 활용도가 높지 않은 금융상품이다.

〈Table 7〉 Tax incentives for IRP and Qualified Contracts

Year	Global income (Total wage and Salary)	Age	Ceiling of payment eligible for tax credit (including retirement pension)	Tax Credit Rate	Tax Credit Amount
Before 19'	Up to 40 mil. (55 mil.)	under 50	4 mil. (7 mil.)	16.5%	660,000 (1,155,000)
		50 or older			
	Up to 100 mil. (120 mil.)	under 50		13.2%	528,000 (924,000)
		50 or older			
	Over 100 mil. (120 mil.)	under 50	3 mil. (7 mil.)	13.2%	396,000 (924,000)
		50 or older			
20'~22' (Temp)	Up to 40 mil. (55 mil.)	under 50	4 mil. (7 mil.)	16.5%	660,000 (1,155,000)
		50 or older	6 mil. (9 mil.)		990,000 (1,485,000)
	Up to 100 mil. (120 mil.)	under 50	4 mil. (7 mil.)	13.2%	528,000 (924,000)
		50 or older	6 mil. (9 mil.)		792,000 (1,188,000)
	Over 100 mil. (120 mil.)	under 50	3 mil. (7 mil.)	13.2%	396,000 (924,000)
		50 or older			

Source: Financial Supervisory Service, National Tax Service (unit: won)

이 장에서는 900개 정도의 관측치로 금융이해력과 개인연금 간의 관계를 분석한 이찬희·정홍주(2013)의 연구를 7,000명 이상의 관측치를 활용하여 동일한 관계가 발견되는지 살펴보고, 금융이해력 이외에 수도권 거주로 인한 효과도 추가로 분석해본다. 또한, 미래를 대비하고, 자발적 가입이 필요한 금융상품이라는 측면에서 유사한 속성을 가지고 있으나, 훨씬 보편화 되어 가입률이 매우 높은 보장성 보험 가입요인과 개인연금 가입요인을 비교하여, 상품의 특성에 따라 금융이해력이나 거주지 요인이 어떠한 다른 영향을 미칠 수 있는지도 분석한다. 〈Table 8〉과 〈Table 9〉는 각각 개인연금 가입자와 보장성보험 가입자의 특성에 대한 기초통계량을 제시한다.

사용된 변수는 다음과 같다. 종속변수로 사용한 개인연금은 개인형 퇴직연금(IRP), 연금저축, 연금보험을 모두 포함하였다. 각각의 금융상품은 세부적으로 차이가 있으나, 모두 은퇴 준비를 위한 장기 저축을 유도하고 세제 혜택을 누릴 수 있는 상품이라는 점에서 유

사하다고 보았다. 보장성보험은 질병, 상해, 사망보험을 포함한다. 보장성보험은 자동차보험과 같이 의무보험은 아니지만 비교적 최근에 소개된 개인연금에 비해 훨씬 보편화된 상품으로 더 많은 사람이 상품의 존재에 대해 알고 있는 전통적인 상품이다. 따라서, 개인연금과 보장성보험은 모두 개인의 재무 위험관리에 유용하게 사용될 수 있는 상품이나, 금융이해력과 거주지 효과가 보장성보험보다 상품에 대한 학습이 필요한 개인연금에서 더 강하게 나타날 것으로 예상된다.

개인연금 가입은 은퇴와 관련된 장수위험 관리의 수단이므로, 개인의 건강 상태 및 기대여명이 의사결정에 영향을 미칠 수 있다는 선행연구 결과들에 따라, 앞서 사용한 인구특성 변수들 이외에도 두 가지 변수를 추가하였다. 주관적 기대여명(Age_diff)은 본인이 예상하는 기대여명에서 국민생명표의 연령별 기대여명을 차감한 값으로, 이창선(2020), 이경희·전병욱(2020)이 제시한 변수를 사용하였다. 그리고, 흡연 여부(Smoke) 항목도 추가하였다. 개인연금 수요와의 관계는 주관적 기대여명은 양의 상관관계를, 흡연 여부와는 음의 상관관계를 예상하였다.

〈Table 8〉의 결과를 보면, 금융이해력이 높을수록, 수도권에 거주하는 경우에는 개인연금 가입률이 높게 나타나는 것을 확인할 수 있다. 반면, 〈Table 9〉는 금융이해력과 보장성보험 가입은 관계가 없는 것으로 보이며, 거주지에서도 오히려 비수도권 거주자들의 수도권 거주자들보다 보장성 보험 가입률이 높은 것으로 나타났다. 따라서, 개인연금과 보장성 보험은 유사한 속성을 가진 상품이나, 가입 결정요인이 매우 다르게 나타나는 것을 알 수 있었다.

〈Table 8〉 Individual Retirement Account participation by subgroups

Variables	Total		Pension = 0	Pension = 1	χ^2
	N	Mean Pension	Percentage	Percentage	
Literacy					
Literacy = 1	5216	0.128	68.21	73.30	9.73*** (0.002)
Literacy = 0	2363	0.103	31.79	26.70	
Education					
Edu_high	4802	0.142	61.82	74.62	57.48*** (0.000)
Edu_mid	2589	0.085	35.54	24.07	
Edu_low	188	0.064	2.64	1.31	
Age					
20~29	729	0.030	10.61	2.41	73.77*** (0.000)
30~39	1397	0.104	18.78	15.86	
40~49	2304	0.138	29.80	34.79	
50~59	3149	0.136	40.81	46.94	
Residing Area					
Capital	3466	0.141	44.67	53.50	25.28*** (0.000)
Non-capital	4113	0.103	55.33	46.50	
Gender					
Male	4061	0.136	52.66	60.28	18.77*** (0.000)
Female	3518	0.103	47.34	39.72	
Marital Status					
Married	5375	0.143	69.09	84.25	89.49*** (0.000)
Single	2204	0.065	30.91	15.75	
Occupational Status					
Employed	5746	0.130	75.02	81.62	22.56*** (0.000)
Self-employed	1650	0.096	22.37	17.40	
Non-employed	183	0.049	2.61	0.98	
Other Controls					
Income	7579	0.121	87.94	12.06	16.30*** (0.000)
Net_asset					9.86*** (0.000)
Age_diff					2.37** (0.018)
Smoke					0.03 (0.977)

Note: *p<0.10, **p<0.05, ***p<0.01, P-value in parentheses, Other Controls: t statistics

〈Table 9〉 Insurance participation by subgroups

Variables	Total		Insurance = 0	Insurance = 1	χ^2
	N	Mean Insurance	Percentage	Percentage	
Literacy					
Literacy = 1	5216	0.854	68.02	68.95	0.37 (0.540)
Literacy = 0	2363	0.859	31.98	31.05	
Education					
Edu_high	4802	0.855	64.51	63.17	15.83*** (0.000)
Edu_mid	2589	0.869	31.42	34.62	
Edu_low	188	0.766	4.07	2.22	
Age					
20~29	729	0.605	26.62	6.79	441.54*** (0.000)
30~39	1397	0.854	18.85	18.36	
40~49	2304	0.908	19.50	32.21	
50~59	3149	0.880	35.03	42.64	
Residing Area					
Capital	3466	0.849	48.34	45.30	3.45* (0.063)
Non-capital	4113	0.860	51.66	54.70	
Gender					
Male	4061	0.841	59.70	52.56	19.02*** (0.000)
Female	3518	0.876	40.30	47.44	
Marital Status					
Married	5375	0.900	50.09	74.39	265.47*** (0.000)
Single	2204	0.755	49.91	25.61	
Occupational Status					
Employed	5746	0.849	80.31	75.07	13.96*** (0.001)
Self-employed	1650	0.884	17.65	22.46	
Non-employed	183	0.880	2.03	2.48	
Other Controls					
Income	7579	0.857	14.28	85.72	4.74*** (0.000)
Net_asset					4.31*** (0.000)
Age_diff					0.84 (0.399)
Smoke					-2.27** (0.024)

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, P-value in parentheses, Other Controls: t statistics

〈Table 11〉에서는 개인연금과 보장성보험 가입 여부를 종속변수로 가지는 회귀분석 결과를 제시한다. 개인연금과 보장성보험 가입에 관한 의사결정은 완전히 독립적이지 않다고 보는 것이 타당하다. 따라서 동시에 일어나는 의사결정 모형을 효율적으로 분석하기 위해, 본 연구에서는 아래의 두 산식 추정을 Zellner(1962)에서 제안된 겹보기 무관 회귀(Seemingly Unrelated Regression: SUR)모형을 사용하였다.

$$Pension_{i,t} = \beta_1 Literacy_{i,t} + \beta_2 Residing_Area_{i,t} + \beta_3 X_{i,t} + \epsilon_{i,t} \quad (1)$$

$$Insurance_{i,t} = \beta_1 Literacy_{i,t} + \beta_2 Residing_Area_{i,t} + \beta_3 X_{i,t} + \epsilon_{i,t} \quad (2)$$

금융이해력 변수와 거주지 변수의 포함 여부에 따라 세 가지 모형으로 나눈 다음, 종속 변수로는 각각 개인연금과 보장성보험을 사용하였다. 〈Table 11〉의 모형 (1), (3), (5)의 종속변수는 개인연금 가입 여부, 나머지는 보장성보험 가입 여부이다. 기초통계량 분석결과와 거의 유사하게 금융이해력은 개인연금 가입과 양의 상관관계를 보였으나 보험 가입과는 무관한 것으로 나타났다. 금융이해력을 측정한 질문이 이자에 관한 내용임을 고려하였을 때, 보장성보험은 이자와 관계가 없지만, 개인연금은 이자율과 관련이 있는 금융상품 이기에 이와 같은 결과가 나왔을 수도 있을 것으로 생각해 볼 수 있다. 그리고, 거주지 변수를 포함한 경우와 포함하지 않은 경우, 금융이해력 및 다른 변수들과 두 상품 가입에 큰 차이는 없는 것으로 나타났다.

기대여명이 높을수록 개인연금을 더 많이 가입하는 것은 아니었으나, 보장성보험은 더 많이 가입하는 경향을 보였다. 장수할 것을 예상하는 사람들이 연금상품에 가입하고, 사망이나 건강 위험이 큰 사람들이 건강 관련 보장성 보험에 가입할 것이라는 역선택 이론에서 예측하는 것과는 다소 반대의 결과를 보였다. 이는 건강이 좋지 않은 사람들은 보장성보험 가입 자체가 어렵거나, 위험 회피도가 큰 사람들이 보장성보험 가입과 동시에 기대여명을 더 길게 예상하기 때문일 수도 있으나, 그 원인을 본 연구의 결과만으로 추측하기는 어렵다. 근로소득자와 남성의 경우 개인연금을 더 가입하고, 보장성보험은 덜 가입하였으며, 교육수준이 높을수록 개인연금 가입률이 높은 것과는 달리, 보장성보험 간에는 유의한 상관관계가 없는 것으로 나타났다. 두 가지 상품 모두 흡연 여부와는 큰 관련이 없었는데, 개인연금 가입 수요를 흡연량 변수를 사용하여 분석한 이경희·전병욱(2020)의 연구와 같은 결과를 보였다.

거주지의 경우, 서울, 인천, 경기도를 수도권으로 정의하고 1을, 그 외에는 0인 거주지 변수를 생성하였다. 여러 요인을 통제한 뒤에도 수도권 거주자들이 개인연금에 더 많이 가입하고 있다는 결과가 나타났는데, 그 효과가 금융이해력보다도 더 크게 나타났다. 앞선 분석에서 수도권 거주자들이 금융이해력 자체가 더 높지는 않은 것으로 나타났고, 거주지를 제외하고 분석한 모델 (1)과 모델 (5)에서 금융이해력 변수의 계수가 거의 차이가 나지 않고, 모델 (3)과 모델 (5)에서의 거주지 변수의 계수도 거의 차이가 나지 않는다는 것을 보면, 거주지가 금융이해력의 대용변수가 아니라 다른 요인을 설명하고 있다는 것을 알 수 있다. 즉, 수도권 거주자들이 같은 수준의 금융이해력을 가진 비수도권 거주자보다 개인연금 가입에 더 용이한 환경이거나, 파급효과 또는 사회적 학습 등의 사회문화적 요인으로 인하여 금융역량에서 차이가 난다고 해석할 수 있을 것이다. 단, 이러한 효과는 보장성보험 가입에는 나타나지 않았는데, 개인연금과 달리 절대다수의 사람들이 가입하고 있는 실손, 사망보험 등의 상품에 있어서 수도권/비수도권의 환경적 격차가 거의 없는 것으로 보인다. 오히려 보장성보험의 경우 통계적으로 유의하지 않으나, 금융이해력과 약간의 음의 관계에 있고, 교육수준과도 무관하다는 사실 등을 비추어 볼 때, 단순한 가입 여부가 아니라 구체적으로 가입한 상품 내용의 적정성 등을 평가하여 필요성 대비 과도하거나 과소한 가입이 있지는 않은지 분석해 볼 필요가 있다.

다음으로, 개인연금의 가입여부뿐만 아니라 개인연금 가입액도 금융이해력과 양의 상관관계를 갖는지 확인하기 위해, 종속변수만 개인연금 가입액으로 바꾼 다음에 토빗 회귀분석(Tobit Regression)을 진행하고 그 결과를 <Table 10>에 제시하였다. 결과를 살펴보면, 금융이해력과 거주지 모두 유의한 양의 상관관계를 가졌으며, 다른 통제변수들도 기존과 모두 동일한 결과를 보였다.

<Table 10> Tobit regression results: Pension amount

Variables	Model (1)		Model (2)		Model (3)	
	Coef.	(S.E.)	Coef.	(S.E.)	Coef.	(S.E.)
Literacy	0.084*	(0.045)			0.090**	(0.045)
Residing_Area			0.159***	(0.042)	0.162***	(0.042)

Note: *p<0.10, **p<0.05, ***p<0.01, N: 7,579

Control variables: Yes. Coefficients for other control variables is not shown due to space. Full result is available upon request.

〈Table 11〉 SUR regression results: IRA and insurance participation

Variables	Model (1)	Model (2)	Model (3)	Model (4)	Model (5)	Model (6)
	Pension	Insurance	Pension	Insurance	Pension	Insurance
Literacy	0.014* (0.008)	-0.005 (0.008)			0.015* (0.008)	-0.006 (0.008)
Residing_ Area			0.024*** (0.007)	-0.013* (0.008)	0.025*** (0.007)	-0.013* (0.008)
Income	0.091*** (0.007)	0.025*** (0.008)	0.089*** (0.007)	0.026*** (0.008)	0.089*** (0.007)	0.026*** (0.008)
Net_asset	0.004*** (0.001)	0.003*** (0.001)	0.004*** (0.001)	0.004*** (0.001)	0.004*** (0.001)	0.004*** (0.001)
Age_20	-0.126*** (0.016)	-0.243*** (0.017)	-0.127*** (0.016)	-0.243*** (0.017)	-0.127*** (0.016)	-0.243*** (0.017)
Age_30	-0.055*** (0.011)	-0.0105 (0.012)	-0.056*** (0.011)	-0.010 (0.012)	-0.056*** (0.011)	-0.010 (0.012)
Age_40	-0.011 (0.009)	0.029*** (0.009)	-0.011 (0.009)	0.029*** (0.009)	-0.011 (0.009)	0.029*** (0.009)
Age_diff	0.000 (0.001)	0.002*** (0.001)	0.000 (0.001)	0.002*** (0.001)	0.000 (0.001)	0.002*** (0.001)
Smoke	-0.003 (0.011)	-0.012 (0.012)	-0.004 (0.011)	-0.012 (0.012)	-0.004 (0.011)	-0.012 (0.012)
Married	0.011 (0.010)	0.065*** (0.011)	0.012 (0.010)	0.064*** (0.011)	0.012 (0.010)	0.064*** (0.011)
Employed	0.049*** (0.009)	-0.007 (0.009)	0.049*** (0.009)	-0.007 (0.009)	0.049*** (0.009)	-0.007 (0.009)
Male	0.020** (0.009)	-0.071*** (0.009)	0.019** (0.009)	-0.070*** (0.009)	0.019** (0.009)	-0.070*** (0.009)
Edu_high	0.043*** (0.008)	0.012 (0.009)	0.043*** (0.008)	0.012 (0.009)	0.043*** (0.008)	0.012 (0.009)
Constant	-0.550*** (0.043)	0.677*** (0.045)	-0.538*** (0.043)	0.672*** (0.045)	-0.542*** (0.043)	0.673*** (0.045)
N	7,579	7,579	7,579	7,579	7,579	7,579
R-squared	0.056	0.078	0.057	0.078	0.058	0.079

Note: *p<0.10, **p<0.05, ***p<0.01, Standard errors in parentheses

이경희·전병욱(2020)은 개인연금 가입 의사결정과 같은 은퇴소득에 관한 계획은 주로 어느 정도 나이가 든 이후에 하게 된다는 점을 고려하여, 여러 변수가 개인연금 가입 수요에 미치는 영향이 45세 전후로 다를 수 있다는 가설을 세우고, 두 그룹으로 나누어 금융상품의 수요를 분석하였다. 이경희·전병욱(2020)의 분석결과에 따르면, 주관적 기대여명 변수는 45세 이상 그룹에서만 유의미한 영향을 미쳤으며, 두 그룹의 의사결정이 상당히 다른 것으로 나타났다. 따라서, 본 연구에서도 같은 방식으로, 표본을 45세를 기준으로 나누어 분석을 진행해 보았다. <Table 12>의 모델 (1), (2)는 20세 이상 45세 미만, 모델 (3), (4)는 45세 이상 60세 미만으로 나누어 각각 분석한 결과이다. 매우 흥미롭게도 개인연금 가입에 있어서 거주지 효과는 45세 미만과 45세 이상의 그룹에서 큰 차이가 없었으나, 금융이해력은 45세 이상의 집단에서만 유의미한 관계를 보였다. 이는 은퇴와 장기적 자산관리에 대해 본격적으로 고민하기 시작하는 중장년층 그룹에서 은퇴소득을 관리하는 도구로 개인연금을 활용하는 데에 금융이해력이 큰 영향을 미치는 것으로 해석할 수 있다. 13차 재정패널조사를 활용하고, 금융이해력 및 지역의 효과를 추가하여 분석한 본 연구의 개인연금 수요 부분의 결과는 이전 차시의 재정패널조사를 활용하여 개인연금수요를 분석했던 이경희·전병욱(2020)의 분석결과와 거의 유사한 것으로 나타났다.

<Table 12> SUR Regression results: Holding of private pension

Variables	Model (1)	Model (2)	Model (3)	Model (4)
	Pension	Insurance	Pension	Insurance
Literacy	0.006 (0.011)	-0.010 (0.014)	0.022** (0.011)	-0.003 (0.010)
Residing_Area	0.029*** (0.010)	-0.000 (0.013)	0.024** (0.010)	-0.022** (0.010)
Age_20	-0.065*** (0.017)	-0.218*** (0.0215)		
Age_30	-0.022* (0.013)	-0.022 (0.016)		
Age_40			-0.005 (0.011)	0.030*** (0.010)
Age_diff	0.001 (0.001)	0.003*** (0.001)	0.000 (0.001)	0.002** (0.001)

Note: *p<0.10, **p<0.05, ***p<0.01, Standard errors in parentheses

N: Model(1)~(2): 3,147, Model(3)~(4): 4,432, Control variables: Yes

V. 강건성 검사

이 장에서는 앞서 제시한 결과들에 대한 몇 가지 강건성 검사 분석 결과를 제시한다. 먼저 이자율 변화와 관계없이 모든 문항의 답을 50%는 현재, 50%는 미래로 답한 사람들을 분류한 Group 2의 경우에는 금융이해력이 낮다기보다 단순히 설문에 성실히 응답하지 않았을 가능성이 있다고 판단하여, 이 그룹을 제외하고 분석한 결과를 <Table 13>의 모델 (1)~(3)에 제시하였다.

다음으로, 금융이해력 변수 정의에 있어서 Group 1을 이해력이 높은 집단으로 재정의 하였다. Group 1의 경우에는 이자율과 관계없이 현재 시점에서 모든 토큰을 수령하는 경우를 말하는데, 이는 시간 선호도는 높지만, 이자에 대한 이해도와 관련해서는 정보가 부족하다고 볼 수도 있다. Group 1을 어떻게 분류하는지에 따라 결과에 미치는 영향을 확인하고자 하였으며, 재정의한 변수 Literacy2 분석 결과는 <Table 13>의 모델 (4)~(6)에 제시하였다.

분석 결과, 일부 관측치를 제외하거나, 조금 다른 금융이해력의 정의를 사용하더라도, 금융이해력이 개인연금의 가입과는 양의 상관관계를, 보장성 보험과는 관계가 없거나 오히려 음의 상관관계를 보이는 결과에 변동은 없었다. 거주지와 관련해서도, 수도권에 거주한다는 사실이 금융이해력을 높이지는 않았지만, 개인연금 가입과는 양의 상관관계를 가진다는 결론은 그대로 유지되었다.

<Table 13> Robustness test results

Variables	Model (1)	Model (2)	Model (3)	Model (4)	Model (5)	Model (6)
	Literacy	Pension	Insurance	Literacy2	Pension	Insurance
Literacy		0.019** (0.009)	-0.008 (0.009)			
Literacy2					0.013 (0.010)	-0.052*** (0.011)
Residing_ Area	-0.171*** (0.055)	0.021*** (0.008)	-0.010 (0.008)	-0.113* (0.066)	0.024*** (0.007)	-0.014* (0.008)

Note: *p<0.10, **p<0.05, ***p<0.01, Standard errors in parentheses

N: Model(1)~(3): 7,097, Model(4)~(6): 7,579, Control variables: Yes

VI. 결론 및 토의

본 연구는 개인의 금융이해력과 거주지, 두 가지 변수와 개인연금 및 보장성 보험 수요 간의 관계를 살펴보고자 하였다. 개인연금이나 보장성보험 가입 수요에 영향을 미칠 수 있는 다양한 요소들에 대한 통제를 한 뒤에도 금융이해력이 낮은 집단에서는 개인연금 수요가 낮게 나타났으며, 수도권 거주자의 경우에는 개인연금 수요가 더 높게 나타났다. 반면 이러한 효과가 보장성보험 가입에는 나타나지 않았다.

거주지와 개인연금 가입 간 결과에 따르면, 수도권 거주자들이 금융이해력이 더 높은 것이 아님에도 개인연금 가입이 더 많다는 것은 금융역량에 있어서 개인의 지식이나 이해도 이외에 접근성이나 사회적 네트워크 효과 등 다양한 요인들이 영향을 미칠 수 있다는 것에 대한 기초적인 근거가 될 수 있다. 기존의 국내 연구에서는 재무 설계사와의 상담이 금융 행동에 미치는 영향 등과 같은 요인들만 주로 분석이 되었는데, 개인의 금융이해력 향상만큼, 또는 더 중요할 수도 있는 환경 요인에 있어서 지역별 격차가 있는지, 그 원인은 무엇인지 분석하여 금융역량을 높일 수 있는 근거를 찾기 위한 추가 연구가 필요할 것이다. 즉, 거주지 이외에 파급효과(spillover effect)나 사회적 인프라 효과 등 금융역량 향상과 관련이 있는 변수들의 영향과 인과관계를 분석할 수 있는 정교한 연구가 진행될 필요가 있다. 추가로, 최근 금융산업의 디지털화는 금융 접근성에서 지역적인 격차를 감소시킬 수 있는 요인이 되는데, 금융역량 강화에 디지털화가 어떤 영향을 미치는지와 관련된 연구도 의미가 있는 연구 방향이 될 것이다.

본 연구에서는 13차 재정패널조사에서 처음 사용된 이자와 관련된 질문 문항을 이용하여 금융이해력을 측정하였다. 재정패널조사는 비교적 큰 표본과 금융소비자를 이해할 수 있는 다수의 변수를 정의할 수 있는 설문 조사라는 장점이 있으나, 금융이해력 관련 문항은 한 가지밖에 없어서 금융이해력을 조금 더 정교하게 측정하지 못했다는 한계가 있다. 특히 여성의 비율이 절대적으로 높은 Group 3의 경우, 저축에 대한 선호를 강하게 보이는 그룹인데, 그로 인하여 보장성 보험보다는 장기 저축의 성격이 강한 개인연금 가입과 본 연구의 금융이해력 변수가 높은 상관관계를 보인 것일 가능성도 배제할 수 없다. 따라서 금융이해력과 개인연금 수요와의 관계를 더 잘 이해하기 위해서 보다 정교한 금융이해력

관련 문항을 활용한 연구가 이루어지면 좋을 것이다.

정책적인 함의는 다음과 같다. 개인의 금융이해력 정도와 거주지에 따라 연금 수요가 다르게 나타나는데 이는 노후 대비를 위해 주목해야 할 문제다. 간단한 수준의 금융 및 경제 교육이더라도 보다 많은 사람에게 꾸준히 제공된다면, 정부의 연금 정책 및 연금 개혁의 경제적인 가치는 더욱 효과적으로 나타날 것이다. 단, 교육에 있어서 세부집단별로 다른 이해도가 있음을 고려한 정책이 필요할 것이다. 금융 및 경제 지식의 교육 불평등 해소를 위한 움직임은 우리나라의 노인빈곤율 문제를 해결할 수 있는 첫걸음이 될 것이며, 이는 국민의 여유로운 노후준비를 위해 이바지하는 데 도움이 될 수 있을 것이다.

참고문헌

- 민세진·홍진주 (2021), “고등학교 특성별 경제 교육의 차이에 관한 연구”, **경제학연구**, 제69권 제1호, pp. 157-192.
- (Translated in English) Min, S., and J., Hong (2021). “A Study of Difference in the Economic Education in the High Schools”, *The Korean Journal of Economic Studies*, 69(1):157-192.
- 백은영 (2012), “개인연금 가입의향에 영향을 미치는 요인 분석”, **사회보장연구**, 제28권 제3호, pp. 63-86.
- (Translated in English) Baek, E. (2012). “An Analysis of Individual Annuity Purchasing Intension”, *Korean Social Security Studies*, 28(3):63-86.
- 여운경·이남희 (2012), “개인연금자산의 수요와 적정성에 관한 분석”, **보험금융연구**, 제68권, pp. 63-93.
- (Translated in English) Yuh, Y., and N., Lee (2012). “Demand and Adequacy of Private Pension”, *Journal of Insurance and Finance*, 68:63-93.
- 이경희·전병욱 (2020), “기대여명에 대한 주관적인식과 사적연금 수요 분석”, **리스크관리연구**, 제31권 제2호, pp. 91-131.
- (Translated in English) Lee, K., and B., Jun (2020). “Analysis of Subjective Life Expectancy and Private Pension Demand”, *The Journal of Risk Management*, 31(2):91-131.
- 이찬희·정홍주 (2013), “개인연금 가입과 선호의 결정요인분석 : 인지요인, 정서요인, 금융이해력, 신뢰, 위험감수성향을 중심으로”, **금융연구**, 제27권 제4호, pp. 25-51.
- (Translated in English) Lee, C., and H., Jung (2013). “An Analysis of Determinants of Purchase and Preference of the Individual Pension : Focusing on Cognitive and Emotional Factors Financial Literacy,

- Trust and Risk Tolerance”, *Journal of Money & Finance*, 27(4):22-51.
- 이창선 (2020), “주관적 기대여명에 따른 성별 개인연금 수요 분석”, **보험금융연구**, 제 31권 제2호, pp. 95-123.
- (Translated in English) Lee, C. (2020). “An Analysis of the Association between Subjective Life Expectancy and Personal Pension Purchase by Sex”, *Journal of Insurance and Finance*, 31(2):95-123.
- 차경욱 (2015), “개인연금 가입결정과 납입액 규모 및 가입의향에 영향을 미치는 요인”, **한국FP학회지**, 제8권 제2호, pp. 89-122.
- (Translated in English) Cha, K. (2015). “Purchase Decision, Amount, and Purchase Intention of Individual Annuity”, *Financial Planning Review*, 8(2):89-122.
- 한경동 (2016), “고등학교 경제교육의 지역적 편차”, **경제교육연구**, 제23권 제3호, pp. 1-21.
- (Translated in English) Hahn, K. (2016). “Economic Education in High School and Geographical Variation”, *Journal of Korean Economic Development*, 23(3):1-21.
- Beckmann, E. (2013). “Financial Literacy and Household Savings in Romania”, *Numeracy: Advancing Education in Quantitative Literacy*, 6(2):Article 9.
- Behrman, J., O., Mitchell, C., Soo, and D., Bravo (2012). “How Financial Literacy Affects Household Wealth Accumulation”, *The American economic review*, 102(3):300-304.
- Bucher-Koenen, T., and A., Lusardi (2011). “Financial literacy and retirement planning in Germany”, *Journal of Pension Economics and Finance*, 10(4):565-584.
- Campbell, J. (2006). “Household Finance”, *Journal of Finance*,

- 61(4):1553-1604.
- Collins, M. (2012). "Financial Advice: A Substitute for Financial Literacy?", *Financial Services Review*, 21(4):307-322.
- Fornero, E., and C., Monticone (2011). "Financial literacy and pension plan participation in Italy", *Journal of Pension Economics and Finance*, 10(4):547-564.
- Hastings, J., and L., Tejeda-Ashton (2008). "Financial literacy, information, and demand elasticity: Survey and experimental evidence from Mexico", *NBER Working Paper*, No. w14538
- Klapper, L., and A., Lusardi (2020). "Financial literacy and financial resilience: Evidence from around the world", *Financial Management*, 49(3):589-614.
- Klapper, L., A., Lusardi, and G., Panos (2013). "Financial literacy and its consequences: Evidence from Russia during the financial crisis", *Journal of Banking & Finance*, 37(10):3904-3923.
- Klapper, L., and G., Panos (2011). "Financial Literacy and Retirement Planning: The Russian Case", *Journal of Pension Economics and Finance*, 10(4):599-618.
- Lachance, M. (2014). "Financial Literacy and Neighborhood Effects", *The Journal of Consumer Affairs*, 48(2):251-273.
- Lin, C., Y., Hsiao and C., Yeh (2017). "Financial literacy, financial advisors, and information sources on demand for life insurance", *Pacific-Basin Finance Journal*, 43:218-237.
- Lusardi, A. (2011). "Americans' Financial Capability", *NBER Working Paper*, No. w17103.
- Lusardi, A., and O., Mitchell (2007). "Financial Literacy and Retirement Preparedness: Evidence and Implications for Financial Education".

- Business Economics*, 42(1):35-44.
- _____ (2008). "Planning and Financial Literacy: How Do Women Fare?", *American Economic Review*, 98(2):413-417.
- _____ (2011a). "Financial literacy and retirement planning in the United States", *Journal of Pension Economics and Finance*, 10(4):509-525.
- _____ (2011b). "Financial literacy around the world: An overview", *Journal of Pension Economics and Finance*, 10(4):497-508.
- _____ (2014). "The Economic Importance of Financial Literacy: Theory and Evidence", *Journal of Economic Literature*, 52(1):5-44.
- Lusardi, A., and P., Tufano (2009). "Debt literacy, financial experiences, and overindebtedness", *NBER Working Paper*, No. w14808
- Monticone, C. (2010). "How Much Does Wealth Matter in the Acquisition of Financial Literacy?", *The Journal of Consumer Affairs*, 44(2):403-422.
- Moore, D. (2003). "Survey of financial literacy in Washington State: Knowledge, behavior, attitudes, and experiences", *Washington State Department of Financial Institutions*, Technical report 03-39.
- Nolte, S., and J., Schneider (2017). "Don't lapse into temptation: a behavioral explanation for policy surrender", *Journal of Banking & Finance*, 79:12-27.
- Sekita, S. (2011). "Financial literacy and retirement planning in Japan", *Journal of Pension Economics and Finance*, 10(4):637-656.
- Song, C. (2020). "Financial Illiteracy and Pension Contributions: A Field Experiment on Compound Interest in China", *The Review of*

Financial Studies, 33(2), 916-949.

Van Rooij, M., A., Lusardi and R., Alessie (2011a). "Financial literacy and stock market participation", *Journal of Financial Economics*, 101(2):449-472.

(2011b). "Financial literacy and retirement planning in the Netherlands", *Journal of Economic Psychology*, 32(4):593-608.

Yan, Z., A., Wang, W., Shao and Y., Zeng (2021). "Urbanization and financial literacy: the knowledge spillover effect", *Applied Economics Letters*.

Zellner, A. (1962). "An Efficient Method of Estimating Seemingly Unrelated Regression Equations and Tests of Aggregation Bias", *Journal of the American Statistical Association*, 57:348-368.

Appendix 1: List of Variables

Variables		Definition
Dependent Variable	Pension	1: contribution of private pension, 0: otherwise
	Pensionamt	ln(contribution amount of private pension)
	Insurance	1: contribution of insurance, 0: otherwise
Explanatory Variable	Literacy	1: Group3,4, 0: otherwise
	Literacy2	1: Group1,3,4, 0: otherwise
	Residing_Area	1: capital area, 0: non-capital area
Control Variable	Income	ln(household income)
	Net_asset	ln(household asset - household debt)
	Age_20	1: 20~29 years, 0: otherwise
	Age_30	1: 30~39 years, 0: otherwise
	Age_40	1: 40~49 years, 0: otherwise
	Age_50	1: 50~59 years, 0: otherwise
	Age_diff	Subjective life expectancy - Objective life expectancy estimated from Statistics Korea
	Smoke	1: smoker, 0: non-smoker
	Married	1: married and live with spouse, 0: otherwise
	Single	1: never married, married and live without spouse, 0: otherwise
	Employed	1: employed as wage and salary worker, 0: otherwise
	Self_employed	1: self-employed, 0: otherwise
	Non_employed	1: non_employed, 0: otherwise
	Male	1: male, 0: female
	Female	1: female, 0: male
	Edu_high	1: college and post graduate 0: otherwise
	Edu_mid	1: high school, 0: otherwise
	Edu_low	1: less than high school, 0: otherwise

Appendix 2: The Big 3 and Big 5 Questions⁶⁾

1. Compounding. Suppose you had \$100 in a savings account and the interest rate was 2% per year. After 5 years, how much do you think you would have in the account if you left the money to grow?
A) More than \$102, B) Exactly \$102, C) Less than \$102,
D) Don't know, E) Prefer not to say
2. Inflation. Imagine that the interest rate on your savings account was 1% per year and inflation was 2% per year. After 1 year, how much would you be able to buy with the money in this account?
A) More than today, B) Exactly the same, **C) Less than today**,
D) Don't know, E) Prefer not to say
3. Diversification. If interest rates rise, what will typically happen to bond prices?
A) They will rise, **B) They will fall**, C) They will stay the same,
D) There is no relationship between bond prices and the interest rate,
E) Don't know, F) Prefer not to say
4. Mortgage. A 15-year mortgage typically requires higher monthly payments than a 30-year mortgage, but the total interest paid over the life of the loan will be less.
A) True, B) False, C) Don't know, D) Prefer not to say
5. Bond Pricing. Buying a single company's stock usually provides a safer return than a stock mutual fund.
A) True, B) False, C) Don't know, D) Prefer not to say

6) Source: <https://gflec.org/education/questions-that-indicate-financial-literacy/>

Appendix 3: Characteristics of Group 1 and Group 3

Variables	Total		Group 1	Group 3	χ^2
	N	Mean Literacy	Percentage	Percentage	
Education					
Edu_high	2686	0.735	58.52	65.07	18.32*** (0.000)
Edu_mid	1435	0.683	37.45	32.29	
Edu_low	129	0.620	4.03	2.64	
Age					
20~29	419	0.704	10.21	9.72	0.25 (0.969)
30~39	813	0.717	18.93	19.21	
40~49	1281	0.715	30.04	30.18	
50~59	1737	0.714	40.82	40.89	
Residing Area					
Capital	2055	0.725	46.58	49.06	2.13 (0.144)
Non-capital	2195	0.704	53.42	50.94	
Gender					
Male	2258	0.688	58.02	51.17	16.37*** (0.000)
Female	1992	0.744	41.98	48.83	
Marital Status					
Married	2988	0.728	66.83	71.70	9.84*** (0.002)
Single	1262	0.681	33.17	28.30	
Occupational Status					
Employed	3215	0.723	73.17	76.64	5.68* (0.058)
Self-employed	938	0.686	24.28	21.19	
Non-employed	97	0.680	2.55	2.17	
Other Controls					
Income	4250	0.714	28.59	71.41	8.83*** (0.000)
Net_asset					5.38*** (0.000)
Age_diff					-3.65*** (0.000)
Smoke					-2.44*** (0.015)

Note: *p<0.10, **p<0.05, ***p<0.01, P-value in parentheses, Other Controls: t statistics
N: Group 1: 1,215, Group 3: 3,035

Abstract

We examined whether financial literacy and residing area affect private pension and insurance holding behaviors using the 13th wave of the National Survey of Tax and Benefits. Financial literacy was measured by whether the person understands the basic concept of interest rate. 31% of respondents had a low level of financial literacy, and financial literacy depended on demographical characteristics: education, gender, occupation, income, and net asset.

The estimating using a seemingly unrelated regression (SUR) showed that financial literacy affected holding private pensions, which only 12% of respondents possessed, but had no significant impact on having insurance, which 86% of respondents had. In addition, even after controlling age, education, and financial literacy, residents of the capital area showed a higher private pension holding rate than non-capital area residents. It suggested that the socio-cultural infrastructure of residing areas and an individual's financial literacy affect financial capacity.

※ Key words: Financial Literacy, Financial Capability, Private Pension, Protection-type Insurance

공유된 임의효과를 고려한 다중 위험 보험의 효율에 관한 연구

A Study on the Ratemaking of Multi-peril Insurance via Shared Random Effects

전 옥 현*·정 힘 찬**·최 양 호***·심 현 우****

Okhyeon Jeon·Himchan Jeong·Yang Ho Choi·Hyunoo Shim

하나의 보험 계약이 여러 개의 담보로 구성되어 서로 다른 위험에 의한 사고를 보상하는 보험을 다중 위험 보험이라 한다. 다중 위험 보험은 각 위험에 의한 청구 건수 사이에 의존성이 존재할 수 있다. 의존성이 확인된 경우 보험요율 산정에 있어 위험 간의 상관관계를 반영하는 것이 합리적이다. 본 연구에서는 다중 위험 간의 의존성 존재여부를 확인하고 이를 사후요율결정(posterior ratemaking)에 반영하기 위하여 공유된 임의효과 모델을 적용하였다. 이 모델은 신뢰도 요율을 닫힌 공식(closed formula) 형태로 제공할 수 있어 실무적 적용과 해석이 용이하다는 장점이 있다.

자동차 보험은 다양한 담보로 구성되어 다중 위험 보험의 대표적인 예로 볼 수 있다. 본 연구에서는 제안 모델을 요율산출에 적용하기 위하여 자동차 보험의 대인배상과 대물배상 담보의 실제 청구 데이터를 사용하였다. 해당 데이터를 이용하여 요율을 산정하였고 그 성능을 비교하기 위하여 표본외 검증을 수행한 결과 포아송 모델 등 비교 모델보다 성능이 우수한 것으로 나타났다.

국문 색인어: 다중 위험 보험, 공유된 임의효과, 과대산포, 청구빈도분석, 할인할증제도

한국연구재단 분류 연구분야 코드: B051605, B051608, C030805

* 한양대학교 금융보험학과 석사과정, 메리츠화재(j.okhy@meritz.co.kr), 제1저자

** Simon Fraser University 통계·계리학과 조교수(himchanj@sfu.ca), 공동저자

*** 한양대학교 보험계리학과 부교수(ychoi@hanyang.ac.kr), 공동저자

**** 한양대학교 보험계리학과 부교수(hyunooshim@hanyang.ac.kr), 교신저자

논문 투고일: 2022. 1. 12, 논문 최종 수정일: 2022. 5. 25, 논문 게재 확정일: 2022. 8. 18

I. 서론

우리나라는 물론 전 세계적으로 보아도 자동차 수는 지속적으로 증가하고 있다. 그에 따라 자동차 보험은 자연스럽게 대표적이고 보편적인 보험이 되었다. 또한 대부분의 국가에서 자동차 보험은 의무적으로 가입해야 하는 보험 상품이다. 그러므로 자동차 보험을 취급하는 보험회사 입장에서는 자동차 사고에 영향을 미치는 요인을 분석하는 것이 매우 중요하다. 자동차 보험의 계약자가 처음 보험을 가입할 시점에 적용되는 사전적인 보험요율을 결정하는 변수에는 일반적으로 연령, 성별, 직업, 거주 지역, 자동차의 유형과 용도 등이 포함된다. 그 이후에는 운전자의 과거 보험사고 이력 정보를 사후적인 보험요율 변수로 하여 경험요율을 결정한다.

할인할증제도(bonus-malus system)에서 보험료 할증이란 보험사고에 과실 책임이 있는 계약자에게 보험료를 더 높게 받는 것을 말한다. 또한 할인이란 과실이 있는 보험사고 이력이 없는 계약자에게 보험료를 더 적게 받는 혜택을 제공하는 것을 말한다. 예를 들어, 심현우·임형기·최양호(2021)는 실손 의료보험에서 과거 청구 건수에 따른 할인할증제도를 제안하였고, 실제로 2021년 한국의 4세대 실손 의료보험에서 할인할증제도를 제도적으로 도입하였다. 한국을 제외한 모든 국가의 할인할증제도에서는 보고된 보험사고의 횟수 즉, 사고 빈도에 따라 보험료를 할인 또는 할증하지만 보험사고의 크기 즉, 사고 심도는 고려하지 않고 있다. 반면 한국에서는 1989년 할인할증제도를 도입한 이후 사고 빈도뿐만 아니라 사고 심도까지 고려하여 경험요율을 결정하였다.

그러나 2018년 이후부터는 국제적인 할인할증제도를 따라 요율결정에 사고 심도 정보는 제외하고 사고 빈도만을 반영하고 있다. 따라서 자동차 보험 요율을 결정하는데 있어 사고 빈도의 영향이 매우 크고 중요하기 때문에 빈도에 대한 면밀한 분석이 뒷받침 되어야 한다.

자동차 보험의 중요한 특성 중 하나는 다중 위험(multi-peril) 보험이라는 것이다. 우리나라 자동차 보험이 보장하는 담보는 대인배상, 대물배상, 자기신체사고, 자기차량손해로 구분된다. 이러한 여러 담보에서 보장하는 손해가 하나의 보험사고에서 동시에 발생할 수 있다는 특징이 있다. 따라서 각 담보별 손해액은 독립적이라고 보기 어려우므로 담보 간

상관관계를 고려할 필요가 있다.

다양한 담보들 간에 공변량(covariate)이 존재할 때, 담보 유형별로 공변량의 영향이 다르다면 각 청구 유형에 대해 서로 다른 회귀계수를 적용해야 한다. 운전 습관과 같이 관찰되지 않은 이질성은 다양한 청구 유형들과 각각 서로 다른 상관관계를 가질 것이다. 예를 들어, 과실 배상책임 담보는 일반적으로 운전자가 어느 정도 통제할 수 있으므로 해당 청구의 발생은 운전 습관과 밀접한 관련이 있다. 그러나 폭우로 인한 침수 피해와 같이 자연재해에 의한 자기차량 손해의 경우 운전자가 거의 통제할 수 없으므로 운전 습관과의 연관성이 매우 낮다. 결국 보험회사는 공변량 정보와 각 보험 계약자의 과거 청구 내역을 활용하여 계약자의 관찰되지 않은 이질성과 보험료를 산출하는 모델을 개발해야 한다.

이에 본 연구에서는 종단면 데이터를 이용하여 관찰되지 않은 이질성을 반영한 공유된 임의효과(shared random effects) 모델을 제안한다. 이 모델은 다중 위험 보험의 특성을 효율에 보다 잘 반영하기 위한 목적으로 각 계약자의 청구이력을 통해 적절한 할인할증요인을 산출하는 모델이다. 이 모델은 기존의 할인할증제도 관점에서 해석이 용이하고 최종적인 효율 산출식도 닫힌 공식(closed formula)으로 도출되어 실무적으로 적용하기에도 편리할 것이다.

우리나라의 자동차 보험은 의무보험으로서 소비자물가지수 품목에 포함될 정도로 보편적이고 중요한 보험 상품이기 때문에 적정 보험요율 산정에 관한 여러 연구가 있었다. 예를 들어 할인할증요율 산출에 대하여 김진한·민재형(1994)은 다중상태모형을, 정중영·김태완(2004)은 기대가치보험료원칙을 이용할 것을 제안하였다. 조재린·이기형·강중철(2012)은 신뢰도 이론을 적용한 자동차 보험요율 산출 결과를 분석하였다. 선행연구에서는 대부분 자동차 보험의 담보간의 청구 건수를 독립으로 가정한 반면 본 연구에서는 담보간 의존성을 반영하였다는 특징이 있고 보험료 산출식이 기본보험료와 할인할증 요인으로 명확히 구분되어 실무적 적용과 해석이 용이하다는 장점이 있다.

본 연구의 구성은 다음과 같다. 제2장에서는 제안 모델과 관련된 신뢰도이론과 다중 위험 간의 의존성에 대한 선행연구를 살펴본다. 제3장에서는 다중 위험 간의 의존성과 과대산포를 고려한 빈도모형을 적용하여 최종적인 효율산출 모델을 도출한다. 제4장에서는 이 모델을 실제 청구데이터에 적용하여 모델에 적용되는 계수를 결정하여 보험료를 산출한

다. 그리고 그 결과를 다른 모델과 비교하여 모델의 성능을 확인한다. 마지막으로 제5장에 서는 본 연구의 결과를 요약한다.

II. 선행 연구

1. 신뢰도 이론

신뢰도 이론은 미래에 발생할 수 있는 보험사고나 그 사고로 인한 손해액 추정 및 보험료 결정 시 과거 청구 이력을 통해 관찰되지 않은 이질성을 도출하는 이론이다. 즉, 과거와 당해 연도의 데이터에 근거하여 그 다음 해의 보험료 산출 방식을 수정 보완하는 접근방법으로, 현재까지 수집된 데이터를 얼마나 믿을 수 있는가에 대한 신뢰도(credibility) 개념이 보험료 책정에 중요한 요소로 사용되고 있다. 신뢰도의 개념은 CAS(미국 손해보험협회)에서 보험료에 현재까지 수집된 데이터를 어떻게 적용할 것인가 하는 문제의 해결책으로 제시되어 보험회사의 가격정책에 적용되기 시작하였다.

Whitney(1918)는 새로운 보험료를 얻기 위하여 현재 사용되고 있는 요율과 과거 데이터를 어떻게 조합하여 사용할 것인가를 연구하여 다음과 같은 기본 개념을 정립하였다.

$$\text{사후적 보험료} = Z \times \text{청구 이력} + (1 - Z) \times \text{사전적 보험료.} \quad (1)$$

여기서 Z 를 신뢰도 요인(credibility factor, $0 \leq Z \leq 1$)이라 한다. 위 식은 청구이력으로부터 얻은 정보 즉, 평균 손해액과 과거 추정치인 사전평균(prior mean)의 가중평균 형태로 볼 수 있다. 이러한 관점에서 베이지안 패러다임과 동일하다고 할 수 있다.

이후 신뢰도 이론은 보험 계리 분야에서 주로 연구되었다. Bailey(1950)는 모델분포가 이항분포이고 사전분포가 베타분포인 경우에, 사후적 요율이 이러한 신뢰도의 가중평균이 되는 것을 보였다. Bühlmann(1967)은 서로 독립인 두 추정량의 제곱오차(squared error)를 사용하여 부분신뢰도를 구하는 방법을 제안하였다. Mayerson(1964)과 Jewell(1974)은 베이지안 관점에서 신뢰도 이론을 분석하여 관찰되지 않는 이질성을 반영

하는 방법을 제시하였다. Norberg(1986)는 관찰된 이질성과 관찰되지 않은 이질성은 모두 요율 산출에 고려되어야 한다고 제안하였다. Frees et al.(1999)은 이를 발전시켜 선형 혼합 모델(linear mixed model)을 기반으로 신뢰도 이론과 회귀분석을 통합한 일반적인 체계를 제시하였다. 선형 혼합 모델에서 반응 변수는 관찰된 공변량과 이와 관련된 회귀계수(고정효과)는 물론, 관찰되지 않은 특성(임의효과)의 영향을 받는다. 고정효과와 임의효과를 모두 사용하기 때문에 신뢰도 보험료가 할인할증제도와 밀접한 관련이 있음을 알 수 있다. Boucher and Denuit(2008) 및 Boucher et al.(2009)은 패널 데이터에 zero-inflated Poisson 분포를 사용하여 신뢰도 보험료를 도출하는 방법을 모색하였다. Jeong and Valdez(2020b)는 과거 빈도의 결합분포가 다변량음이항분포(multivariate negative binomial; MVNB)일 때, 미래 건수에 대한 신뢰도 보험료(예측분포의 기댓값)는 다음 공식으로 나타낼 수 있음을 보였다.

$$\text{사후적 보험료} = \frac{r + \text{실제 청구 경험}}{r + \text{기대 청구 경험}} \times \text{사전적 보험료} \quad (2)$$

여기서 r 은 할인할증제도의 평할 계수를 나타내고 r 값이 클수록 과거 청구 경험에 대한 가중치는 작아진다.

2. 다중 위험 보험의 의존성

그러나 실무적으로 요율결정에 적용하기 위해서는 여러 담보의 청구 간에 존재할 수 있는 의존성을 고려해야 한다. Frees et al.(2010)은 개별 위험 모델(individual risk model)하에서 다중 위험 보험에 대한 종속성 모델을 제안하였는데, 이는 한 청구 건에 대하여 복합 손실을 청구 유무와 총 청구금액으로 분해하는 모델이다. 여기서 다중 위험 청구 간의 의존성을 모델링하기 위해 Gaussian Copula를 이용한 다변량 이진회귀분석(multivariate binary regression)을 사용하였다. 반면에 Frees et al.(2016)은 총위험 모델(collective risk model)하에서 복합 손실을 빈도와 심도 요소로 분해하는 다중 위험 보험의 종속성 모델을 제안하였다. 여기서 다변량 빈도와 심도 각각의 의존성을 반영하기

위해 Copula를 이용하였다. e Silva and de Lourdes Centeno(2017)는 다중 위험의 의존성을 반영하기 위하여 여러 가지 일반화된 이변량 빈도분포를 사용할 것을 제안하였으며, 이를 통해 관찰된 청구 건수에서 과대 산포를 확인하였다. Quan and Valdez(2018)는 다중 위험 청구 금액에 대한 회귀분석에서 다변량 의사결정 트리(multivariate decision tree)를 사용하는 방법을 제안하였다.

그러나 다중 위험 간의 의존성과 종단면적 특성을 모두 포함한 연구는 많지 않다. 예를 들어, Frees(2003)는 다중 위험 계약에 대한 복합 손실의 신뢰도 보험료를 산출하는 방법을 제안하였으나 청구 빈도와 심도의 평균 및 공분산만을 고려한 비모수적 접근방식이었다. 반면 Bermúdez and Karlis(2017)는 두 가지 유형의 청구에 대한 사후적 요율결정을 위해 이변량 Poisson 분포를 적용하였다. 두 연구는 공변량을 고려하지 않은 방식이기 때문에 계약자의 관찰된 이질성과 관찰되지 않은 이질성을 모두 포함하기는 어렵다.

Pechon et al.(2018), Pechon et al.(2019) 및 Pechon et al.(2020)은 상관 로그노말(correlated lognormal) 임의효과를 이용하여 여러 유형의 청구에 대한 신뢰도 보험료 공식을 제시하였다. 이는 종단면하에서 다중 위험 청구 사이의 의존성을 고려하는 포괄적인 체계를 제공하였다. 그러나 모든 계약자 각각에 대한 신뢰도 요인과 빈도의 주변분포를 도출하기 위해서는 다중(6~7중)적분이 필요하다. 일반적으로 보험 계약자 수가 매우 많기 때문에 실무적으로 적용하기에는 부담이 있을 수 있다.

본 연구에서 제안하는 방법론은 공유된 임의효과를 통해 다중 위험 보험에 대하여 계약자의 관찰된 이질성과 관찰되지 않은 이질성을 모두를 고려하고 종단면 분석을 가능하게 한다. 또한, 다중 위험 신뢰도 보험료에 대한 닫힌 공식을 최종적으로 제시한다.

III. 연구 방법

1. 다중 위험 보험

일반적으로 다중 위험(multi-peril) 보험이란 필요한 여러 보장을 함께 묶는 보험이다. 다중 위험 보험에 가입하면 피보험자가 여러 가지 유형의 사고에 대해 보장 받을 수 있을

뿐만 아니라 함께 발생하는 손실에 대한 광범위한 보장으로 인해 전체 보험료 비용이 일반적으로 낮아진다. 다중 위험 보험에는 다양한 사례가 있으나 본 연구에서는 자동차 보험을 대상으로 분석하고자 한다. 자동차 보험은 자동차 사고에서 발생할 수 있는 다양한 유형의 경제적 손실을 보상한다. 하나의 자동차 사고로 인해 운전자, 동승자 및 보행자 등 인적 피해는 물론 자동차와 도로 위의 각종 구조물 등에 대한 물적 피해도 발생할 수 있으므로, 이러한 다양한 피해에 대한 보상을 제공하는 자동차 보험은 다중 위험 보험의 대표적 사례이다.

다중 위험 보험의 청구 건수에 대한 일반적인 특징을 표현하면 다음과 같다. 다중 위험 보험에서는 여러 유형의 사고가 발생할 수 있기 때문에 보험회사는 특정 계약자의 각 유형별 청구 건수를 시간의 흐름에 따라 관찰할 수 있다. i 번째($i = 1, \dots, I$) 계약자에 대해 t 년도($t = 1, \dots, T$)에 관찰 가능한 특성을 $\mathbf{x}_{it}$ 라 하고, 이 때 발생한 j 유형($j = 1, \dots, J$) 사고의 청구 건수를 $N_{it}^{(j)}$ 라고 나타내면 청구빈도에 대한 일반적인 데이터구조는 아래와 같이 나타낼 수 있다.

$$D = \left\{ \left(N_{it}^{(1)}, \dots, N_{it}^{(j)}, \dots, N_{it}^{(J)}, \mathbf{x}_{it} \right) \middle| i = 1, \dots, I, t = 1, \dots, T \right\}. \quad (3)$$

2. 빈도모형

일반적으로 빈도모형으로는 포아송(Poisson) 분포를 사용하는 경우가 많다. 포아송 분포는 단순하여 사용이 용이한 장점이 있으나 자료가 과대산포(overdispersion)인 경우에는 문제가 발생할 수 있다. 즉, 평균과 분산이 같다고 가정하는 포아송 분포와 달리 경험데이터는 분산이 평균보다 큰 경우에는 포아송 분포가 적합하지 않을 수 있다는 것이다.

이러한 문제를 해결하기 위해 Wedderburn(1974)은 quasi-Poisson 분포를 제안하였다. 빈도 확률변수 N 이 quasi-Poisson 분포 $QP(\nu, w)$ 를 따른다는 것은 확률변수 wN 이 포아송 분포 $P(w\nu)$ 를 따른다는 것과 동치이다.

$$N \sim QP(\nu, w) \Leftrightarrow wN \sim P(w\nu). \quad (4)$$

위 정의에 따르면 quasi-Poisson의 확률분포함수는 다음과 같이 구할 수 있다.

$$\Pr(N=n)=\Pr(wN=wn)=e^{-w\nu}\frac{(w\nu)^{wn}}{(wn)!}. \quad (5)$$

따라서 $E[wN]=wE[N]=w\nu$, $Var[wN]=w^2 Var[N]=w\nu$ 이므로 $E[N]=\nu$, $Var[N]=\nu/w$ 가 된다. 즉, 분산이 모수 w 에 반비례하기 때문에 w 에 따라 평균과 분산의 대소 관계가 바뀌어 과대산포 문제를 해결할 수 있다.

Cameron and Trivedi(1990) 등의 연구에서는 아래와 같은 가설검정을 통하여 빈도 모델의 과대산포 여부를 검정하는 방법을 제안하였다.

$$H_0 : w = 1 \text{ vs. } H_1 : w \neq 1. \quad (6)$$

위 귀무가설은 평균과 분산이 같다($w=1$)는 것이고 이를 검정하기 위하여 회귀분석을 활용하였다. 본 연구에서는 위 연구에서 제안한 방법으로 분석 데이터의 과대산포 여부를 확인한 후 quasi-Poisson 분포를 빈도모형으로 사용하였다.

3. 자동차 보험 할인할증제도

자동차 보험의 할인할증제도는 1910년대 영국인에 의해 처음 도입되었는데 처음에는 NCD(no claim discount, 무청구할인)라는 용어를 사용하였다. 문자 그대로 “우리(보험회사)를 괴롭히지 않는다면 할인을 받는다. 어떤 이유로든 우리에게 접수를 한다면 할인을 받지 못한다”라는 의미였다. 따라서 회사에 보고된 모든 청구는 과실 여부에 관계없이 계약자에게 불이익을 부과한다. 그러나 곧 보험회사들은 이 제도를 폐지하게 되었고 1960년대에 유럽의 보험회사가 할인할증제도를 도입했을 때부터는 과실 사고에만 불이익을 부과하게 되었다.

할인할증제도의 목적은 두 가지로 볼 수 있다. 첫 번째는 보험 계약자와 보험회사 간의 정보 비대칭을 줄여서 보험회사가 각 운전자의 위험 등급에 대해 더 적절하고 공정한 요금을 부과하는 것이다. 예를 들어, 주행거리가 긴 운전자는 사고 위험에 더 많이 노출되기 때문에 더 많은 사고가 발생할 것이다. 그런데 주행거리를 신고하지 않는다면 보험회사는 주행거리를 사전적으로 보험료에 반영할 수 없다. 대신, 긴 주행거리를 사후적으로 요율에

반영하여 위험을 보전한다. 이를 뒷받침하는 가정은 더 많이 운전하거나 위험한 운전 습관을 가진 사람들은 과거에 더 많은 사고를 겪었고 미래에도 많은 사고를 발생시킬 것으로 기대된다는 것이다.

현재 시행 중인 모든 자동차 보험 할인할증제도가 청구 빈도만을 사용한다는 사실은 미래 청구 예측에 있어서도 대부분 청구 빈도에만 적용된다는 것을 의미한다. 청구 심도는 사고위험을 제대로 반영하지 못한다. 그럼에도 우리나라에서 심도를 할인할증제도에 반영했던 이유는 과거 사망사고 등 인적사고가 빈발하여 이를 억제하기 위한 정책적인 목적이 있었다. 그러나 현재는 물적사고의 비중이 증가하여 이를 유지할 필요성이 없어졌기 때문에 빈도만을 반영한 할인할증제도로 전환하여 사고위험을 보다 정확하게 반영한 요율을 적용할 수 있게 하였다.

자동차 보험 할인할증제도의 두 번째 목적은 도덕적 해이를 통제하는 것이다. 자동차 보험 할인할증제도와 같은 경험요율 제도를 경험한 보험 계약자는 보험금 청구 이력에 따라 보험료가 인상될 수도 있다는 것을 알고 있고 보험금 청구빈도를 줄이기 위한 유인을 가지게 된다. 이러한 과정을 통해 도덕적 해이 문제를 통제할 수 있다.

4. 다중 위험 보험의 임의효과

식 (3)에서 나타나는 청구빈도 데이터구조의 중요한 특징은 종단면성(longitudinality)이다. 즉, 같은 계약자에 대해 시간의 흐름에 따라 청구이력을 반복적으로 관찰한다는 것이다. 따라서 각 계약자의 위험요인에 내재된 관찰되지 않은 이질성을 설명하기 위해서는 빈도모형에 임의효과를 반영하는 것이 필요하다.

관찰 가능한 특성에는 다양한 변수가 포함될 수 있다. 예컨대 연령, 성별, 차량연식, 지역, 용도(개인용, 업무용, 영업용 등) 등이 대표적이다. 이러한 다양한 특성이 각 계약자의 청구빈도를 모두 설명할 수는 없다. 즉, 관찰 가능한 특성이 모두 동일한 계약자들 사이에도 청구이력은 확연히 다를 수 있다는 것이다. 따라서 관찰된 특성이 계약자들이 가진 위험요인의 이질성을 구분하는데 충분하지 않다고 의심할 수 있다. 그런데 관찰하는 특성의 숫자를 늘린다고 해서 모든 이질성을 구분할 수는 없을 것이다. 그 특성이라는 것이 너무 많아 효율성이 떨어질 뿐 아니라 애초에 운전습관과 같이 관찰이 불가능한 특성도 존재하

기 때문이다. 그러므로 관찰되지 않은 이질성을 포함하기 위해서는 임의효과를 빈도모형에 적용할 수 있다. 임의효과 모델은 신뢰도 이론 및 할인할증제도와도 밀접하게 관련되어 있다. Gómez-Déniz and Vázquez-Polo(2005)와 Jeong and Valdez(2020b)를 비롯하여 요율산정 관점에서 임의효과를 적용한 여러 연구가 있었다.

요약하면, 청구빈도는 관찰된 공변량과 회귀계수 모두에 영향을 받고 관찰되지 않은 이질성요인 θ_i 에도 영향을 받는 것으로 가정한다. 계약자 i 에 있어서 θ_i 정보는 모든 유형의 보장과 연도별 청구 건에 공유된다. 이를 반영하여 앞서 제안한 quasi-Poisson 분포를 표현하면 다음과 같다.

$$N_{it}^{(j)} | \mathbf{x}_{it}, e_{it}, \theta_i \stackrel{\text{indep}}{\sim} QP(\theta_i \nu_{it}^{(j)}, w^{(j)}) \Leftrightarrow w^{(j)} N_{it}^{(j)} | \mathbf{x}_{it}, e_{it}, \theta_i \stackrel{\text{indep}}{\sim} P(\theta_i w^{(j)} \nu_{it}^{(j)}). \quad (7)$$

여기서 자료는 풀링하지 않고 패널모형을 기본적으로 사용하며, 모형은 $\mathbf{x}_{it}$ 가 변함에 따라 매년 달라질 수 있는 고정효과(fixed effect)와 한 보험계약자에 대해 동일하게 적용되는 θ_i , 즉 정적확률효과(static random effect)를 동시에 사용한다. 그리고 $\nu_{it}^{(j)} = e_{it} e^{\mathbf{x}_{it} \alpha^{(j)}}$ 인데 e_{it} 는 각 계약자의 t 년도의 익스포저(exposure)로 $0 < e_{it} \leq 1$ 이다. $\alpha^{(j)}$ 는 j 번째 보장 유형에 대한 회귀계수, $w^{(j)}$ 는 j 번째 보장 유형에 대한 가중치이다. $\nu_{it}^{(j)}$ 는 관찰된 공변량과 그 회귀계수로 결정되므로 각 계약자(i)의 유형(j)별로 각 기간(t)동안 위험요인이 내포하고 있는 관찰된 이질성을 설명한다. 반면 임의효과 θ_i 는 계약자(i)의 위험요인에 내포된 관찰되지 않은 이질성을 설명한다. 임의효과의 기댓값은 $E[\theta_i] = 1$ 으로 가정하는데 이는 θ_i 가 할인할증요인 역할을 하는 승법(multiplicative) 임의효과이기 때문이다. 가법(additive) 임의효과의 경우에는 $E[\theta_i] = 0$ 으로 가정하는 것이 일반적이므로 승법 임의효과에 대해서는 $E[\theta_i] = 1$ 로 설정하는 것이 자연스럽다.

위 분포에 의해 청구 건수 $N_{it}^{(j)}$ 의 평균을 구해보면 다음과 같다. 먼저 quasi-Poisson 분포의 평균을 이용하여 $E[N_{it}^{(j)} | \theta_i] = \nu_{it}^{(j)} \theta_i$ 임을 알 수 있고 $E[\theta_i] = 1$ 이므로 $E[N_{it}^{(j)}] = E[E[N_{it}^{(j)} | \theta_i]] = E[\nu_{it}^{(j)} \theta_i] = \nu_{it}^{(j)}$ 이 된다. 그러므로 $\nu_{it}^{(j)}$ 는 관찰되지 않은 이질성 θ_i 에 대한 정보가 없을 때 피보험자 i 의 시간 t 에서 청구 건수의 사전평균(prior mean)으로 해석할 수 있다.

종단면 분석에 있어서 다중 위험 청구 간에 있을 수 있는 종속성을 포함시키는 방법으로 공유된 임의효과가 유일한 것은 아니다. 대표적인 방법은 Copula를 활용하여 여러 유형 간의 의존성을 반영하는 것으로 Yang and Shi(2019)에서 연구되었다.

관찰되지 않은 이질성이 청구 빈도에 미치는 영향은 사전적으로는 알 수 없다. 따라서 특정한 사전분포(prior distribution)를 가지고 표현해야 한다. 베이지안 분석에서는 임의의 모수(parameter)에 대한 충분한 정보가 있지 않다면 일반적으로 θ 에 대한 비정보 사전분포(noninformative prior)를 사용한다. 이러한 분포 중 가장 널리 사용되는 분포가 Jeffreys의 사전분포이다. Jeffreys(1946)에 따르면 Jeffreys의 사전분포는 Fisher 정보행렬(information matrix)의 행렬식(determinant)에 대한 제곱근으로 정의된다. 위 모델에서 θ 에 대한 Jeffreys의 사전분포는 $\pi(\theta) = \frac{1}{\sqrt{\theta}}$ 이 되고 이에 대응되는 사후분포는 다음과 같이 구할 수 있다.

사후분포의 정의는 $\pi(\theta|n) = \frac{\Pr(n|\theta)\pi(\theta)}{\int_0^\infty \Pr(n|\theta)\pi(\theta)d\theta}$ 이므로 분자부분으로 결정된다. 식

(5)에서 구한 quasi-Poisson 확률분포함수를 대입하여 분자부분을 구하고 그에 비례하는 가장 단순한 식을 구해보면 다음과 같다.

$$\Pr(n|\theta)\pi(\theta) = e^{-w\nu\theta} \frac{(w\nu\theta)^{wn}}{(wn)!} \frac{1}{\sqrt{\theta}} \propto \theta^{wn - \frac{1}{2}} e^{-w\nu\theta}. \quad (8)$$

이는 감마분포의 확률분포함수에 비례한다. 즉, $\theta|n \sim \Gamma\left(wn + \frac{1}{2}, \frac{1}{w\nu}\right)$ 이 된다는 것을 알 수 있다. 그러나 위 사전분포는 식별가능성(identifiability) 문제가 있다. 즉, 분포를 유일하게 결정하는 모수가 존재하지 않는다는 것이다. 따라서 임의효과의 기댓값에 대한 조건인 $E[\theta_i] = 1$ 을 적용하여 이 문제를 해결할 수 있다. 또한 모델분포의 켄레 사전분포(conjugate prior)를 이용하는 것이 유용하므로 이를 만족하는 분포를 찾으면 다음과 같다.

모델분포가 선형지수함수 계통(linear exponential family)에 속할 때 사전분포와 사후분포가 동일한 형태인 켄레 관계가 된다는 성질을 이용한다. 즉, 모델분포를

$\Pr(N=n|\theta) = \frac{p(n)e^{r(\theta)n}}{q(\theta)}$ 로 표현할 때, 사전분포가 $\pi(\theta) = \frac{[q(\theta)]^{-k} e^{\mu k r(\theta)} r'(\theta)}{c(\mu, k)}$ 라

면 이에 대응되는 사후분포 역시 동일한 형태의 분포가 된다. quasi-Poisson 분포는 선형 지수함수 계통에 해당되므로 위의 모델분포에 대입하면 다음과 같다.

$p(n) = \frac{1}{(wn)!}$, $q(\theta) = e^{w\nu\theta}$, $r(\theta) = w \ln(w\nu\theta)$, $r'(\theta) = \frac{w}{\theta}$ 이므로 이를 대입하여 사전분포를 구하면 $\pi(\theta) \propto [e^{w\nu\theta}]^{-k} e^{\mu kw \ln(w\nu\theta)} \frac{w}{\theta} \propto e^{-kw\nu\theta} \theta^{\mu kw - 1}$ 이 된다. 이는 감마 분포이므로 $\theta \sim \Gamma\left(\mu kw, \frac{1}{kw\nu}\right)$ 임을 알 수 있다. 여기에 평균에 대한 조건을 적용하면 $E[\theta] = \frac{\mu}{\nu} = 1$ 이므로 $\mu = \nu$ 이고 따라서 감마분포의 두 모수는 서로 역수관계임을 알 수 있다. 최종적으로 $r = \mu kw = kw\nu$ 라 하고 결과를 정리하면 다음과 같다.

$$\pi(\theta) \propto \theta^{r-1} e^{-\theta r}, \text{ 즉 } \theta \sim \Gamma\left(r, \frac{1}{r}\right) \text{이므로 } E[\theta] = 1, \text{ Var}[\theta] = \frac{1}{r}. \quad (9)$$

이 결과를 통해 $t, t' = 1, \dots, T$ 와 $j, k = 1, \dots, J$ 에 대해 다음 통계량을 산출할 수 있다.

$$\begin{aligned} E[N_t^{(j)}] &= E[E[N_t^{(j)}|\theta]] \\ &= E[\nu_t^{(j)}\theta] \\ &= \nu_t^{(j)}, \end{aligned} \quad (10)$$

$$\begin{aligned} \text{Var}[N_t^{(j)}] &= E[\text{Var}[N_t^{(j)}|\theta]] + \text{Var}[E[N_t^{(j)}|\theta]] \\ &= E\left[\frac{\theta \nu_t^{(j)}}{w^{(j)}}\right] + \text{Var}[\theta \nu_t^{(j)}] \\ &= \frac{\nu_t^{(j)}}{w^{(j)}} + \frac{(\nu_t^{(j)})^2}{r}, \end{aligned} \quad (11)$$

$$\begin{aligned} \text{Cov}(N_t^{(j)}, N_{t'}^{(k)}) &= E[N_t^{(j)} N_{t'}^{(k)}] - E[N_t^{(j)}] E[N_{t'}^{(k)}] \\ &= E[E[N_t^{(j)} N_{t'}^{(k)}|\theta]] - E[N_t^{(j)}] E[N_{t'}^{(k)}] \\ &= E[\nu_t^{(j)} \nu_{t'}^{(k)} \theta^2] - \nu_t^{(j)} \nu_{t'}^{(k)} \\ &= \nu_t^{(j)} \nu_{t'}^{(k)} (E[\theta^2] - 1) \\ &= \nu_t^{(j)} \nu_{t'}^{(k)} (\text{Var}[\theta] + E[\theta]^2 - 1) \\ &= \frac{\nu_t^{(j)} \nu_{t'}^{(k)}}{r}, \end{aligned} \quad (12)$$

$$\begin{aligned}
 Corr(N_t^{(j)}, N_t^{(k)}) &= \frac{Cov(N_t^{(j)}, N_t^{(k)})}{\sqrt{Var[N_t^{(j)}]} \sqrt{Var[N_t^{(k)}]}} \\
 &= \frac{r}{\sqrt{\frac{\nu_t^{(j)}}{w^{(j)}} + \frac{(\nu_t^{(j)})^2}{r}} \sqrt{\frac{\nu_t^{(k)}}{w^{(k)}} + \frac{(\nu_t^{(k)})^2}{r}}} \\
 &= \frac{1}{\sqrt{1 + \frac{r}{w^{(j)}\nu_t^{(j)}}} \sqrt{1 + \frac{r}{w^{(k)}\nu_t^{(k)}}}}.
 \end{aligned} \tag{13}$$

따라서 제안 모델은 시간이 지남에 따라 다중 위험 빈도 사이의 의존성을 포함하고 있다. 예를 들어, $r \rightarrow \infty$ 이면 제안 모델은 독립 quasi-Poisson 모델로 축소된다. 또한, $r \rightarrow \infty$ 이고 모든 $j = 1, \dots, J$ 에 대해 $w^{(j)} = 1$ 이면 제안 모델은 독립적인 Poisson 모델로 축소된다.

빈도분석을 통한 효율산출은 빈도변수의 조건부 평균인 $E[N_t^{(j)}|\theta] = \theta\nu_t^{(j)}$ 로 볼 수 있다. 즉, 보험효율은 $\nu_t^{(j)}$ 와 임의효과 θ 의 곱으로 표현되기 때문에 θ 가 보험료 결정을 위한 할인할증요인의 역할임을 알 수 있다.

하이퍼파라미터(hyperparameter) r 의 값은 사전적 지식이나 적률법(method of moments) 등을 이용하여 결정할 수 있다. 예를 들어, Lemaire(1998)에 따르면, 빈도 효율에 대한 할인할증요인의 범위는 일반적으로 54%에서 200%이다. 따라서 제안된 사전분포의 하이퍼파라미터 r 을 선택함에 있어 이러한 사전적 지식을 적용시켜야 한다. 즉, θ 의 95% 최고사후밀도구간(highest posterior density interval, 이하, 'HPDI'라 함)이 (0.54, 2.00)을 포함할 수 있도록 해야 한다. 또한 일치추정량으로 r 을 추정할 수도 있는데 Sutradhar and Jowaheer(2003)에 의하면 $i = 1, \dots, I$ 에 대해

$\sum_{t \neq t', j \neq k} (N_{it}^{(j)} - \nu_{it}^{(j)})(N_{it'}^{(k)} - \nu_{it'}^{(k)})$ 의 기댓값은 $\frac{1}{r} \sum_{t \neq t', j \neq k} \nu_{it}^{(j)} \nu_{it'}^{(k)}$ 이고 따라서 r 에 대한 추정량은 다음과 같다.

$$\hat{r} = \frac{\sum_{i=1}^I \sum_{t \neq t', j \neq k} \hat{\nu}_{it}^{(j)} \hat{\nu}_{it'}^{(k)}}{\sum_{i=1}^I \sum_{t \neq t', j \neq k} (N_{it}^{(j)} - \hat{\nu}_{it}^{(j)}) (N_{it'}^{(k)} - \hat{\nu}_{it'}^{(k)})}. \quad (14)$$

여기서 $\hat{\nu}$ 는 미리 정해진 평균 구조를 통해 일반화추정방정식(generalized estimating equation; GEE)을 풀어 추정된 일치추정량이다(Liang and Zeger 1986, pp. 15-19). 이 가정에 따르면 다음과 같이 θ 에 대한 사후분포를 유도할 수 있다.

$$\text{사후분포의 정의는 } \pi(\theta|n) = \frac{\Pr(n|\theta)\pi(\theta)}{\int_0^\infty \Pr(n|\theta)\pi(\theta)d\theta} \text{ 이므로 모델분포와 사전분포를 대}$$

입하면 다음과 같다.

$$\begin{aligned} \pi(\theta|n_T^{(1)}, \dots, n_T^{(J)}) &\propto \pi(\theta) \prod_{j=1}^J \prod_{t=1}^T \Pr(w^{(j)} n_t^{(j)} | \theta) \\ &\propto \theta^{r-1} e^{-\theta r} \prod_{j=1}^J \prod_{t=1}^T \theta^{w^{(j)} n_t^{(j)}} e^{-w^{(j)} \nu_t^{(j)} \theta} \\ &= \theta^{\sum_{j=1}^J \sum_{t=1}^T w^{(j)} n_t^{(j)} + r - 1} e^{-\theta \left(\sum_{j=1}^J \sum_{t=1}^T w^{(j)} \nu_t^{(j)} + r \right)}. \end{aligned} \quad (15)$$

따라서 사후분포는 다음의 감마분포임을 알 수 있다. 즉,

$$\theta|n_T^{(1)}, \dots, n_T^{(J)} \sim \Gamma\left(\sum_{j=1}^J \sum_{t=1}^T w^{(j)} n_t^{(j)} + r, \frac{1}{\sum_{j=1}^J \sum_{t=1}^T w^{(j)} \nu_t^{(j)} + r}\right). \quad (16)$$

또한 $n_T^{(1)}, \dots, n_T^{(J)}$ 이 주어진 경우 $N_{T+1}^{(j)}$ 의 예측 분포를 유도할 수 있다. 이를 위해 $(w^{(1)} n_T^{(1)}, \dots, w^{(J)} n_T^{(J)})$ 와 $w^{(j)} N_{T+1}^{(j)}$ 가 제안 모델을 따른다고 가정한다. 그리고 표기 상 편의를 위해 $r_T = r + \sum_{j=1}^J \sum_{t=1}^T w^{(j)} n_t^{(j)}$ 이고 $\hat{r}_T = r + \sum_{j=1}^J \sum_{t=1}^T w^{(j)} \nu_t^{(j)}$ 라고 하면 식 (15)에서 구한 사후분포는 다음과 같이 표현된다.

$$\pi(\theta|\mathbf{n}_T^{(1)}, \dots, \mathbf{n}_T^{(J)}) \propto \theta^{\sum_{j=1}^J \sum_{t=1}^T w^{(j)} n_t^{(j)} + r - 1} e^{-\theta \left(\sum_{j=1}^J \sum_{t=1}^T w^{(j)} \nu_t^{(j)} + r \right)} = \theta^{r_T - 1} e^{-\hat{r}_T \theta}. \quad (17)$$

따라서 사후분포 함수는 다음과 같다.

$$\pi(\theta|\mathbf{n}_T^{(1)}, \dots, \mathbf{n}_T^{(J)}) = \frac{(\hat{r}_T \theta)^{r_T} e^{-\hat{r}_T \theta}}{\theta \Gamma(r_T)}. \quad (18)$$

마지막으로 과거 데이터 n 이 주어졌을 때 x 에 대한 예측분포는 $\Pr(x|n) = \int_0^\infty \Pr(x|\theta) \pi(\theta|n) d\theta$ 으로 산출되므로, 여기에 식 (5), (7)과 (18)을 대입하여 전개하면 다음과 같다.

$$\begin{aligned} & \Pr(w^{(j)} N_{T+1}^{(j)} = w^{(j)} n_{T+1}^{(j)} | \mathbf{n}_T^{(1)}, \dots, \mathbf{n}_T^{(J)}) \\ &= \int_0^\infty \Pr(w^{(j)} N_{T+1}^{(j)} = w^{(j)} n_{T+1}^{(j)} | \theta) \pi(\theta | \mathbf{n}_T^{(1)}, \dots, \mathbf{n}_T^{(J)}) d\theta \\ &= \int_0^\infty e^{-w^{(j)} \nu_{T+1}^{(j)} \theta} \frac{(w^{(j)} \nu_{T+1}^{(j)} \theta)^{w^{(j)} n_{T+1}^{(j)}}}{(w^{(j)} n_{T+1}^{(j)})!} \frac{(\hat{r}_T \theta)^{r_T} e^{-\hat{r}_T \theta}}{\theta \Gamma(r_T)} d\theta \\ &= (w^{(j)} \nu_{T+1}^{(j)})^{w^{(j)} n_{T+1}^{(j)} \hat{r}_T} \\ &= \int_0^\infty \left\{ (w^{(j)} \nu_{T+1}^{(j)} + \hat{r}_T) \theta \right\}^{w^{(j)} n_{T+1}^{(j)} + r_T - 1} e^{-(w^{(j)} \nu_{T+1}^{(j)} + \hat{r}_T) \theta} (w^{(j)} \nu_{T+1}^{(j)} + \hat{r}_T) d\theta \\ &\quad \times \frac{(w^{(j)} \nu_{T+1}^{(j)})^{w^{(j)} n_{T+1}^{(j)} \hat{r}_T}}{(w^{(j)} n_{T+1}^{(j)})! \Gamma(r_T) (w^{(j)} \nu_{T+1}^{(j)} + \hat{r}_T)^{w^{(j)} n_{T+1}^{(j)} + r_T}} \\ &= \frac{(w^{(j)} \nu_{T+1}^{(j)})^{w^{(j)} n_{T+1}^{(j)} \hat{r}_T} \Gamma(w^{(j)} n_{T+1}^{(j)} + r_T)}{(w^{(j)} n_{T+1}^{(j)})!} \end{aligned} \quad (19)$$

그러므로 $w^{(j)} N_{T+1}^{(j)} | \mathbf{n}_T^{(1)}, \dots, \mathbf{n}_T^{(J)} \sim NB\left(r_T, \frac{w^{(j)} \nu_{T+1}^{(j)}}{\hat{r}_T + w^{(j)} \nu_{T+1}^{(j)}}\right)$ 이 된다. 예측분포

가 음이항분포임을 이용하면 기댓값은 다음과 같다.

$$E[w^{(j)}N_{T+1}^{(j)}|\mathbf{n}_T^{(1)}, \dots, \mathbf{n}_T^{(J)}] = \frac{w^{(j)}\nu_{T+1}^{(j)}r_T}{\hat{r}_T + w^{(j)}\nu_{T+1}^{(j)}} \frac{\hat{r}_T + w^{(j)}\nu_{T+1}^{(j)}}{\hat{r}_T} = \frac{w^{(j)}r_T}{\hat{r}_T} \nu_{T+1}^{(j)}. \quad (20)$$

따라서 최종적으로 예측분포의 기댓값은 다음과 같다. 이 결과는 식 (2)와 동일함을 알 수 있다.

$$\begin{aligned} E[N_{T+1}^{(j)}|\mathbf{n}_T^{(1)}, \dots, \mathbf{n}_T^{(J)}] &= \frac{r_T}{\hat{r}_T} \nu_{T+1}^{(j)} \\ &= \frac{r + \sum_{j=1}^J \sum_{t=1}^T w^{(j)}n_t^{(j)}}{r + \sum_{j=1}^J \sum_{t=1}^T w^{(j)}\nu_t^{(j)}} \nu_{T+1}^{(j)}. \end{aligned} \quad (21)$$

(16)에 의해 사후분포가 감마분포이므로 사후분포의 기댓값은 다음과 같다.

$$E[\theta|\mathbf{n}_T^{(1)}, \dots, \mathbf{n}_T^{(J)}] = \frac{r + \sum_{j=1}^J \sum_{t=1}^T w^{(j)}n_t^{(j)}}{r + \sum_{j=1}^J \sum_{t=1}^T w^{(j)}\nu_t^{(j)}}. \quad (22)$$

이를 식 (21)에서 구한 예측분포의 기댓값에 적용하면 다음과 같이 표현할 수 있다.

$$E[N_{T+1}^{(j)}|\mathbf{n}_T^{(1)}, \dots, \mathbf{n}_T^{(J)}] = E[N_{T+1}^{(j)}] E[\theta|\mathbf{n}_T^{(1)}, \dots, \mathbf{n}_T^{(J)}]. \quad (23)$$

따라서 과거 T 년 동안의 청구 정보가 주어질 때, $N_{T+1}^{(j)}$ 의 예측 요율은 사전적 보험료 $\nu_{T+1}^{(j)}$ 와 신뢰도 요인의 합동추정값(pooled estimate)의 곱으로 산출됨을 알 수 있다. 여기서 사전적 요율은 관측 가능한 공변량과 회귀계수에 의해 결정되며 신뢰도 요인의 추정치는 관찰되지 않은 공유된 이질성을 설명한다.

Frangos and Vrontos(2001)는 $J=1$ 이고 $w^{(1)}=1$ 인 경우를 연구하였는데 식 (21)은 이를 확장한 형태이다. 예를 들어 $J=2$ 이고 $N^{(1)}$ 과 $N^{(2)}$ 가 각각 대인배상 및 대물배상 담보로 인한 청구 빈도를 나타낸다고 할 때, 관찰되지 않은 이질성의 합동추정값은 아래와 같다.

$$\begin{aligned}
 E[\theta | \mathbf{n}_T^{(1)}, \mathbf{n}_T^{(2)}] &= \frac{r + w^{(1)} \sum_{t=1}^T n_t^{(1)} + w^{(2)} \sum_{t=1}^T n_t^{(2)}}{r + w^{(1)} \sum_{t=1}^T \nu_t^{(1)} + w^{(2)} \sum_{t=1}^T \nu_t^{(2)}} \\
 &= \frac{\text{평활인자} + \text{실제 빈도의 가중합}}{\text{평활인자} + \text{예측 빈도의 가중합}}.
 \end{aligned} \tag{24}$$

따라서 제안 모델을 사용하면 다중 위험에 대해 다양한 가중치 $w^{(j)}$ 를 적용함으로써 사후적 요율 요인 $E[\theta | \mathbf{n}_T^{(1)}, \dots, \mathbf{n}_T^{(J)}]$ 에 대한 다양한 수준의 기여도를 고려할 수 있다.

식 (24)에서, θ 의 사전분포의 하이퍼파라미터 r 은 평활 인자의 역할을 한다. 예를 들어, $r \rightarrow \infty$ 이면 $\pi(\theta | \mathbf{n})$ 은 1에서 디랙 델타 함수(Dirac delta function)로 수렴하며, 이는 $\Pr(\theta = 1) = 1$ 을 의미하고 매우 정보적인(very informative) 점-질량(point-mass) 사전분포가 된다. 반면에 더 작은 r 을 선택하면 덜 정보적인(less informative) 사전분포를 사용하게 된다.

경험적 베이즈 방법(empirical Bayes method)을 적용하면 $j = 1, 2, \dots, J$ 에 대한 회귀계수 $\alpha^{(j)}$ 의 추정치는 다음과 같은 결합로그우도(joint loglikelihood)를 최대화하여 구할 수 있다.

$$l(\alpha^{(1)}, \dots, \alpha^{(J)} | \mathbf{n}_T^{(1)}, \dots, \mathbf{n}_T^{(J)}) = \sum_{i=1}^I \ln \Pr(w^{(1)} \mathbf{n}_{iT}^{(1)}, \dots, w^{(J)} \mathbf{n}_{iT}^{(J)}). \tag{25}$$

이 때,

$$\begin{aligned}
 \Pr(w^{(1)} \mathbf{n}_{iT}^{(1)}, \dots, w^{(J)} \mathbf{n}_{iT}^{(J)}) &= \int \prod_{j=1}^J \prod_{t=1}^T \Pr(w^{(j)} n_{it}^{(j)} | \theta_i) \pi(\theta_i) d\theta_i \\
 &\propto \prod_{j=1}^J \prod_{t=1}^T \left(\frac{w^{(j)} \nu_{it}^{(j)}}{\sum_{j=1}^J \sum_{t=1}^T w^{(j)} \nu_{it}^{(j)} + r} \right)^{w^{(j)} n_{it}^{(j)}} \left(\frac{r}{\sum_{j=1}^J \sum_{t=1}^T w^{(j)} \nu_{it}^{(j)} + r} \right)^r
 \end{aligned} \tag{26}$$

이고 이는 다변량음이항분포(MVNB)를 따른다.

IV. 실증분석

한국의 자동차 보험에서 가장 중요한 보장은 대인보상(bodily injury, 이하, 'BI'라 함)과 대물보상(property damage, 이하, 'PD'라 함) 담보이다. 자동차 보험은 한 보험계약자가 하나의 보험사고로 인해 여러 담보에서 보험금 청구가 가능하기 때문에 다중 위험(multi-peril) 보험으로 볼 수 있다. 이러한 자동차 보험의 특성을 반영하여 적절한 보험요율 산정 방법을 연구하고자 한다.

본 연구에서는 국내 한 손해보험회사에서 2014년부터 2019년까지 개인용 자동차 보험을 가입한 전체 계약의 데이터를 사용하였다. 1년 미만인 계약을 포함하며, 업무용 및 영업용 차량은 제외한다. 보험개발원 보험통계조회서비스 자료에 따르면 2014년부터 2019년까지 해당 보험회사의 자동차 책임보험의 원수보험료 기준 시장점유율은 5%가 넘는다. 본 연구에서는 1년 평균 55만 건이 넘는 관측치의 자료를 사용하였는데, 기승도·김대환(2009)은 약 50만 건의 1년 자료를 사용하였고 김영화·이현수(2010)는 1년 평균 약 30만 건의 3년 간 자료를 사용한 것과 비교하여 충분한 크기의 표본일 것으로 판단된다. 1년 미만인 데이터에는 개별 계약자에 대한 성별, 보험가입경력, 할인할증등급, 차량크기, 연령, 차량연식, 대인보상 및 대물보상 보험금 청구 건수 정보가 포함되어 있다. 보험개발원 보험통계조회서비스의 전체 모집단의 성별, 연령대별 자동차보험 건수 자료의 분포와 비교하여서도 큰 차이가 없으므로, 분석대상 데이터의 대표성에도 문제가 없을 것으로 판단된다.

일반적으로 자동차 보험료에는 차량의 종류, 배기량, 피보험자의 성별, 연령뿐만 아니라 운전자 연령한정, 범위한정 특약 등의 특약 가입여부, 보험가입경력, 과거 사고발생 실적을 반영한 할인할증요율 등 다양한 요인들이 영향을 미칠 수 있다. 기승도·김대환(2009)은 성별, 연령, 지역, 가입경력, 교통법규위반경력, 할인할증등급을 변수로 고려하였고, 김영화·김미정·김명준(2011)은 성별, 연령, 운전경력, 할인할증등급, 운전자한정 특약가입 여부를 변수로 고려하였다. 본 연구는 이를 참고하여 자료에서 얻을 수 있는 변수들을 선택하고 공변량으로 사용하였다. 본 연구의 목적이 다중 위험 보험의 빈도분석에 공유된 임의효과 모델을 적용한 효과를 확인하는 것이므로, 임의효과를 사용함으로써 유의성이 높다고 판단된 변수들을 제외한 나머지 변수와 관찰되지 않는 특성들도 간접적으로 반영될

수 있게 하였다.

2014년부터 2018년까지의 데이터를 훈련자료(training data)로, 2019년 데이터를 시험자료(test data)로 선택하였고 관측 건수는 각각 2,933,233건, 490,940건이다. 시간이 지남에 따라 보험 계약자의 청구를 추적할 수 있는 고유한 식별자가 자료에 포함되어 있기 때문에 본 자료는 패널데이터이며, 계약을 지속한 경우와 지속하지 않은 경우를 포함한 모든 계약이 자료에 포함되기 때문에 특히 불균형 패널데이터이다. 손해보험, 그 중에서도 특히 자동차보험의 자료는 본질적으로 불균형 패널데이터인 경우가 많은데, 본 연구의 방법론은 패널의 관측길이에 관계없이 적용 가능하다는 장점이 있다. 또한 데이터에는 각 계약자의 연도별 계약정보와 BI 및 PD의 청구 정보가 포함되어 있기 때문에 다변수 패널데이터 빈도 모델링을 시도할 수 있다.

〈Table 1〉, 〈Table 2〉는 훈련자료에서 관찰 가능한 계약 특성을 정리한 표이며, 각각 범주형 변수와 양적 변수의 기술통계량에 대해 요약한 결과이다. 범주형 변수로는 성별, 보험가입경력, 할인할증등급, 차량크기를, 양적 변수로는 연령과 차량연식을 공변량으로 선택하였다. 성별(Gender)은 남성(0), 여성(1)으로 구분하였고 보험가입경력(YDE)은 ‘3년 미만’(0), ‘3년 이상 5년 미만’(1), ‘5년 이상 7년 미만’(2), ‘7년 이상’(3)의 4가지 범주로 구분하였다. 할인할증등급(BM)은 자료에 세분화된 등급이 있으나 분석 시 크게 할증(0)과 할인(신규 포함)(1)으로 구분하였다. 마지막으로 차량크기(AutoSize)는 배기량으로 구분하였는데 1,000cc 미만을 경형(0), 1,600cc 미만을 소형(1), 2,000cc 미만을 중형(2), 2,000cc 이상을 대형(3)으로 구분하였다. 양적 변수인 연령(Age)의 평균은 48.9세이며 차량연식(AutoAge)의 평균은 8.2년이다.

〈Table 1〉 Observable policy characteristics (categorical variables)

Categorical variable	Value	Number of policies	Proportion to total
Gender	(0) Male (base)	2,133,111	72.72%
	(1) Female	800,122	27.28%
	Total	2,933,233	100.00%
YDE (years of driving experience)	(0) Under 3 (base)	320,887	10.94%
	(1) 3 to Under 5	202,168	6.89%
	(2) 5 to Under 7	189,949	6.48%
	(3) 7 and over	2,220,229	75.69%
	Total	2,933,233	100.00%
BM (bonus-malus)	(0) Malus (base)	187,400	6.39%
	(1) Bonus	2,745,833	93.61%
	Total	2,933,233	100.00%
AutoSize	(0) Light (base)	658,726	22.46%
	(1) Compact	1,162,854	39.64%
	(2) Medium	793,902	27.07%
	(3) Large	317,751	10.83%
	Total	2,933,233	100.00%

〈Table 2〉 Observable policy characteristics (quantitative variables)

Quantitative variable	Mean	Minimum	Maximum	S. D.
Age	48.92	18.00	100.00	11.17
AutoAge	8.22	0.00	40.00	5.19
BI claim freq.	0.05	0.00	9.00	0.23
PD claim freq.	0.16	0.00	14.00	0.44

BI와 PD가 독립 포아송 분포를 따른다고 할 때, 청구 건수에 대한 포아송 회귀분석의 결과는 〈Table 3〉과 같다. BI 청구 건수 결과를 보면 남성에 비하여 여성의 빈도가 더 높은 것으로 나타났다. 가입경력은 3년 미만에 비하여 ‘3년 이상 5년 미만’, ‘5년 이상 7년 미만’, ‘7년 이상’으로 경력이 길어질수록 청구 건수가 적게 나타났다. 할인할증 등급은 할증에 비하여 할인등급이 청구 건수가 적었다. 연령은 1차항의 계수가 음수이고 2차항의 계수가 양수이므로 연령이 증가함에 따라 청구 건수가 감소하다가 고연령대에서 다시 증

가하는 형태임을 알 수 있다. 이상의 변수들에 대해서는 PD 청구 건수도 이와 동일한 형태를 보였다. 이러한 결과는 유사한 변수로 사고빈도를 분석한 선행연구들의 결과와도 그 경향이 일치한다. 예를 들어, 김동훈·이계연(1998)은 유의미한 변수 중 가입경력이 짧을수록, 할증률이 클수록, 배기량이 클수록, 차량연식이 짧을수록 청구 건수가 증가하는 것으로 분석하였다. 기승도·김대환(2009)은 남자에 비해 여자가, 가입경력이 짧을수록, 할증률이 클수록 청구 건수가 증가하고, 연령이 증가할수록 그 건수가 감소하였다가 장년층에서 다시 증가하는 양상을 보이며, 차종에 따라서 청구 건수가 달라짐을 보여주었다.

반면 차량크기는 BI 청구 건수에서는 중형차량이 청구 건수가 가장 적었으나 PD 청구 건수는 대형차량이 가장 적은 것으로 나타나는 등 차이를 보였다. 또한 차량연식도 BI는 연식이 증가함에 따라 청구 건수가 증가하다가 감소하나 PD는 지속적으로 감소하는 추세를 보였다.

〈Table 3〉 Coefficient estimates of Poisson model

Variables(<i>l</i>)	BI (<i>j</i> =1)		PD (<i>j</i> =2)	
	Estimate ($\alpha_l^{(1)}$)	Standard error	Estimate ($\alpha_l^{(2)}$)	Standard error
(Intercept)	-1.8210***	0.0398	-0.8575***	0.0221
Gender(1)	0.1049***	0.0060	0.1260***	0.0033
YDE(1)	-0.2866***	0.0120	-0.3119***	0.0068
YDE(2)	-0.3619***	0.0126	-0.3745***	0.0071
YDE(3)	-0.4602***	0.0091	-0.4390***	0.0051
BM(1)	-0.5574***	0.0092	-0.3873***	0.0054
AutoSize(1)	-0.0188**	0.0068	0.0287***	0.0038
AutoSize(2)	-0.0915***	0.0076	0.0531***	0.0041
AutoSize(3)	-0.0424***	0.0098	-0.1913***	0.0057
Age	-0.0138***	0.0016	-0.0086***	0.0009
Age ²	0.0002***	0.0000	0.0002***	0.0000
AutoAge	0.0277***	0.0018	-0.0053***	0.0010
AutoAge ²	-0.0018***	0.0001	-0.0011***	0.0001

Note: Significance '***' 0.001, '**' 0.01, '*' 0.05

Base: Gender - Male(0)

YDE(years of driving experience) - Under 3(0)

BM(bonus-malus) - Malus(0)

AutoSize - Light(0)

〈Table 4〉 Contingency table for BI and PD claims

Frequency of PD claims	Frequency of BI claims					
	0	1	2	3	4	5+
0	2,493,835	21,719	235	7	0	0
1	271,039	89,816	1,380	21	2	0
2	28,412	15,814	3,164	94	5	0
3	2,996	2,286	815	140	3	2
4	389	451	179	46	9	4
5+	84	132	90	36	14	14

〈Table 5〉 Measures of association between frequencies of BI and PD claims

Item	Pearson correlation	Kendall's τ	Spearman's ρ
Estimate	0.4626	0.4421	0.4457
p-value	$< 10^{-6}$	$< 10^{-6}$	$< 10^{-6}$

식 (7)과 (9)에 명시된 제안 모델은 과대 산포를 갖는 다중 위험 보험의 청구 건수 사이에 연관성이 존재한다고 가정하고 있다. 따라서 단순한 독립 포아송 모델보다 더 복잡한 모델을 적용하는 것이 적절한지 검증하기 위해서는 실제로 그 의존성이 존재하는지 확인해야 한다. 〈Table 4〉는 BI와 PD의 청구에 대한 빈도 테이블을 나타낸다. 이 표를 통해 두 청구 건수 간의 의존성을 추측해 볼 수 있다. 예를 들어 BI 건수가 1건일 때 PD 건수가 1건인 경우의 빈도가 가장 높고 BI 건수가 2건, 3건의 경우도 유사한 형태를 보인다. 〈Table 5〉는 이러한 두 담보 간의 연관성을 계량적으로 확인하기 위해 일반적인 측정 방법인 Pearson correlation, Kendall's τ , Spearman's ρ 를 산출한 결과이다. 이를 통해 적어도 공변량의 영향을 제어하지 않을 때에는 BI 및 PD 청구 건수 간에 유의한 양의 상관관계가 있음을 확인할 수 있다.

quasi-Poisson 분포에서 산포 모수를 추정하는 방법을 제시하는 몇 가지 연구가 있다.

예를 들어 McCullagh and Nelder(1989)는 $\hat{\phi} = \frac{1}{m-p} \sum_{k=1}^m \frac{(n_k - \hat{\nu}_k)^2}{\hat{\nu}_k}$ 을 제안하였고

Cameron and Trivedi(1990)는 $\tilde{\phi} = \frac{1}{m} \sum_{k=1}^m \tilde{R}_k$ 를 제안하였다. 여기서

$\tilde{R}_k = \frac{(n_k - \hat{\nu}_k)^2}{\hat{\nu}_k} - n_k$ 이고, m 은 총 관측 수, p 는 추정된 모수의 개수, n_k 는 k 번째 관

측의 청구 건수, $\hat{\nu}_k$ 는 모델에서 추정된 모수를 가지고 산출한 k 번째 관측의 청구 건수 기댓값이다. 위 추정치가 1보다 크다면 과대산포가 존재한다고 볼 수 있다.

〈Table 6〉은 이를 산출한 결과인데 BI, PD 청구 건수의 $\hat{\phi}$, $\tilde{\phi}$ 모두 산포 모수가 1 이상으로 추정되는 것을 확인할 수 있다. 따라서 분석 대상인 청구 데이터에는 과대산포가 존재한다고 볼 수 있다.

〈Table 6〉 Estimates of overdispersion parameters $\hat{\phi}$ and $\tilde{\phi}$ for BI and PD claims

Overdispersion parameter	BI	PD
$\hat{\phi}$	1.0422	1.0807
$\tilde{\phi}$	1.0402	1.1069

그리고 Cameron and Trivedi(1990)는 m 이 충분히 크고 $k = 1, \dots, m$ 에 대해 n_k 가 독립일 때, 다음이 성립함을 보였다.

$$H_0 : \phi = 1 \text{ 하에서 } \hat{Z} = \frac{\tilde{\phi}}{\sqrt{\frac{1}{m} \sum_{k=1}^m (\tilde{R}_k - \tilde{\phi})^2}} \simeq N(0, 1). \quad (27)$$

주어진 데이터에서 BI 및 PD 청구에 대한 $\hat{Z}$ 값은 각각 21.0456과 66.4850이다. 따라서 두 담보의 청구에 과대 산포가 없다는 귀무가설을 유의하게 기각할 수 있으므로 과대 산포가 존재한다고 볼 수 있다.

위와 같이 데이터에서 과대 산포의 유의성과 위험 간의 의존성을 확인하였으므로 이를 포함하는 제안 모델을 적용하여 요율을 산출한다.

먼저 $w^{(j)}$ 의 값은 다중 위험의 특성에 대한 사전 지식이나 관측된 데이터의 통계량으로 결정해야 한다. 식 (24)에서 확인할 수 있듯이 신뢰도 요인의 합동 추정치는 예측 빈도의 가중합(weighted sum)에 대한 실제 빈도의 가중합의 비율로 도출된다. 따라서 $w^{(j)}$ 는 특

정 위험에 대한 계수로 해석될 수 있다. 이 계수는 위험과 관찰되지 않은 이질성 사이의 연관성에 대한 정보를 제공한다.

예를 들면 자동차 보험에서 대물보상 담보와 자기차량손해 담보가 있다. 대물보상 담보는 보험 계약자의 과실이 인정되는 자동차 사고로 인한 손해를 보상하는 반면 자기차량손해 담보는 홍수와 같은 천재지변에 의해 침수가 되는 경우에 대해서도 보상한다. 그러므로 이 두 담보를 분석한다면 자기차량손해 담보에 비해 대물보상 위험이 관찰되지 않은 이질성과 비교적 큰 양의 상관관계가 있다고 볼 수 있다. 이와 관련하여, 신뢰도 계수의 합동 추정치를 산출할 때 자기차량손해 담보의 청구이력보다 대물보상 담보의 과거 청구이력에 더 많은 가중치를 두는 것이 합리적이다. 즉, 대물보상 담보의 $w^{(j)}$ 를 자기차량손해 담보의 $w^{(j)}$ 보다 높게 설정한다는 것이다. 따라서 제안된 방법은 실무자가 모델링 과정에서 다중 위험의 특성에 대한 사전 지식을 유연하게 반영할 수 있도록 한다.

사전 지식으로 $w^{(j)}$ 의 값을 미리 지정할 수 있는 강력한 근거가 부족한 경우에는 과대 산포 모수의 추정치를 활용하고 적률법을 사용할 수 있다. 제안 모델에서

$$\sum_{t=1}^{T_i} (N_{it}^{(j)} - \nu_{it}^{(j)})^2 \text{의 기댓값은 } j = 1, \dots, J \text{와 } i = 1, \dots, I \text{에 대해 } \sum_{t=1}^{T_i} \frac{\nu_{it}}{w^{(j)}} + \frac{(\nu_{it}^{(j)})^2}{r}$$

이므로 식 (14)과 유사한 방식으로 $w^{(j)}$ 의 일치추정량을 구하면 다음과 같다.

$$\hat{w}^{(j)} = \frac{\sum_{i=1}^I \sum_{t=1}^{T_i} \hat{\nu}_{it}^{(j)}}{\sum_{i=1}^I \sum_{t=1}^{T_i} \left[(N_{it}^{(j)} - \hat{\nu}_{it}^{(j)})^2 - \frac{(\hat{\nu}_{it}^{(j)})^2}{\hat{r}} \right]}. \quad (28)$$

$j = 1, 2$ 에 대한 과대 산포 모수 $w^{(j)}$ 가 사전 지식이나 적률법을 통해 결정되면 식 (25)와 (26)을 이용하여 공유된 임의효과를 포함한 결합 우도(joint likelihood)로부터 $\alpha^{(1)}$ 와 $\alpha^{(2)}$ 를 추정할 수 있다. 본 연구의 훈련자료에 적률법을 사용한 결과 r , $w^{(1)}$, $w^{(2)}$ 가 각각 6.1330, 0.9663, 0.9150으로 추정되었다. 따라서 $w^{(1)} > w^{(2)}$ 이기 때문에 제안된 사후적 효율결정 체계에서 상대적 기여도는 BI 빈도가 PD 빈도보다 높은 것으로 나타났다.

제안 모델의 회귀계수인 $\alpha^{(j)}$ 를 추정한 결과는 <Table 7>과 같다. 이 결과는 포아송 회

귀분석 결과인 <Table 3>과 유사하게 나타났다.

<Table 7> Coefficient estimates of proposed model

Variables(<i>l</i>)	BI (<i>j</i> =1)		PD (<i>j</i> =2)	
	Estimate($\alpha_l^{(1)}$)	Standard error	Estimate($\alpha_l^{(2)}$)	Standard error
(Intercept)	-1.8242***	0.0398	-0.8548***	0.0221
Gender(1)	0.1079***	0.0060	0.1252***	0.0033
YDE(1)	-0.2758***	0.0120	-0.3046***	0.0068
YDE(2)	-0.3525***	0.0126	-0.3713***	0.0071
YDE(3)	-0.4565***	0.0091	-0.4349***	0.0051
BM(1)	-0.5476***	0.0092	-0.3776***	0.0054
AutoSize(1)	-0.0182	0.0068	0.0296***	0.0038
AutoSize(2)	-0.0891***	0.0076	0.0525***	0.0041
AutoSize(3)	-0.0348*	0.0098	-0.1887***	0.0057
Age	-0.0139***	0.0016	-0.0086***	0.0009
Age ²	0.0002***	0.0000	0.0002***	0.0000
AutoAge	0.0271***	0.0018	-0.0051***	0.0010
AutoAge ²	-0.0018***	0.0001	-0.0011***	0.0001

Note: Significance ‘***’ 0.001, ‘**’ 0.01, ‘*’ 0.05

Base: Gender - Male(0)

YDE(years of driving experience) - Under 3(0)

BM(bonus-malus) - Malus(0)

AutoSize - Light(0)

모델의 모든 계수가 추정되었으므로 이를 청구 빈도의 예측에 적용해볼 수 있다. 연구방법에서 사후분포의 기댓값을 식 (22)와 같이 도출하였으므로 이를 이용한다. $N^{(j)}_{T+1}$ 에 대한 예측치는 사후적 보험료를 산출한다는 의미이므로 예측분포의 기댓값 즉, $\hat{N}^{(j)}_{T+1} := E[N^{(j)}_{T+1} | \mathbf{n}_T^{(1)}, \mathbf{n}_T^{(2)}]$ 로 추정하는 것이 타당하다. 이렇게 산출한 결과는 표본 외(out-of-sample) 검증을 통하여 2019년 시험자료의 실제 청구 건수와 비교하여 예측 성능을 확인해볼 수 있다. 성능을 비교하기 위하여 독립 포아송 모델(naive Poisson)과 독립 할인할증(univariate credibility)을 적용한 보험료도 함께 산출하였다.

독립 포아송 모델은 과거 사고이력을 반영하지 않은 방식으로 제안 모델에서 $w^{(1)} = w^{(2)} = 1$ 이고 $r = \infty$ 인 특별한 경우이다. 독립 할인할증 보험료는 BI, PD의 사

고 이력을 각각 반영한 보험료이다. 제안 모델(multivariate credibility)과 비교하면 $j = 1$ 일 때에는 $w^{(1)} = 1$, $w^{(2)} = 0$ 이고 $j = 2$ 일 때에는 $w^{(1)} = 0$, $w^{(2)} = 1$ 인 경우이다. 이를 정리하여 나타내면 <Table 8>과 같다.

<Table 8> Predictive means of three models for out-of-sample validation

Model	Predictive mean
Naive Poisson	$E[N_{T+1}^{(j)} \mathbf{n}_T^{(1)}, \mathbf{n}_T^{(2)}] = E[N_{T+1}^{(j)}] = e^{\mathbf{x}_{T+1}\alpha^{(j)}}$
Univariate credibility	$E[N_{T+1}^{(j)} \mathbf{n}_T^{(1)}, \mathbf{n}_T^{(2)}] = \frac{r + \sum_{t=1}^T n_t^{(j)}}{r + \sum_{t=1}^T \nu_t^{(j)}} \cdot e^{\mathbf{x}_{T+1}\alpha^{(j)}}$
Multivariate credibility	$E[N_{T+1}^{(j)} \mathbf{n}_T^{(1)}, \mathbf{n}_T^{(2)}] = \frac{r + w^{(1)} \sum_{t=1}^T n_t^{(1)} + w^{(2)} \sum_{t=1}^T n_t^{(2)}}{r + w^{(1)} \sum_{t=1}^T \nu_t^{(1)} + w^{(2)} \sum_{t=1}^T \nu_t^{(2)}} \cdot e^{\mathbf{x}_{T+1}\alpha^{(j)}}$

미래 청구 빈도에 대한 예측치로서 산출된 각 보험료의 예측력을 평가하기 위해 다음과 같이 정의된 제곱근평균제곱오차(root mean square error, 이하, 'RMSE'라 함)와 평균 절대오차(mean absolute error, 이하, 'MAE'라 함)를 사용하였다.

$$\text{RMSE} = \sqrt{\frac{1}{2I} \sum_{j=1}^2 \sum_{i=1}^I \left(N_{i, T_i+1}^{(j)} - \hat{N}_{i, T_i+1}^{(j)} \right)^2}, \quad (29)$$

$$\text{MAE} = \frac{1}{2I} \sum_{j=1}^2 \sum_{i=1}^I \left| N_{i, T_i+1}^{(j)} - \hat{N}_{i, T_i+1}^{(j)} \right|. \quad (30)$$

여기서 I 는 검증 세트에서 관찰된 보험 계약자 수를 의미한다. 세 가지 모델에 대해 산출한 위 통계량을 <Table 9>에 나타내었는데, 세 모형의 RMSE와 MAE가 모두 커서 빈도를 정확히 예측하는 데에는 한계를 가짐을 알 수 있다. RMSE와 MAE는 오차를 나타내는 지표이므로 일반적으로 값이 더 작은 모델이 선호되는데, 제안 모델의 RMSE와 MAE가 모두 가장 작은 값을 보여 세 모델 중에서는 가장 좋은 예측력을 보였다. 좀 더 정확히 모형

들 간 비교를 하기 위하여 BIC(Bayesian information criterion)를 산출한 후 <Table 9>에 나타내었는데, 제안모형의 BIC가 가장 낮아서 다른 모형들에 비해 근소하게 더 나은 것으로 보인다.

<Table 9> Out-of-sample validation for frequency prediction

Statistic	Naive Poisson	Univariate credibility	Multivariate credibility
RMSE	0.31496	0.31417	0.31373
MAE	0.16955	0.16814	0.16807
BIC	3,797,314	4,138,017	3,714,699

V. 결론

한 보험계약에서 서로 다른 위험에 의한 사고를 보상하는 다중 위험 보험은 매우 흔한 보험 상품이다. 한 사고로 인하여 복수의 위험으로부터 손해가 발생할 수 있으므로 각 위험으로 인한 청구 건수가 서로 의존성을 가질 수 있다. 또한 보험회사가 여러 해 동안 동일한 보험 계약자를 반복적으로 관찰하는 것은 매우 당연한 일이다. 따라서 보험회사는 계약자의 과거 청구 이력에 대한 종단면 데이터를 자연스럽게 보유하게 되므로 이를 보험요율에 적절히 반영할 수 있어야 할 것이다. 특히 자동차보험과 일반손해보험은 보험기간이 짧고 보험기간이 종료되면 갱신 또는 재가입 형태로 새로운 요율을 적용하는 보험 상품이므로 계약자의 과거 청구 이력을 보험요율에 반영하는 할인할증제도와 신뢰도이론을 활용할 수 있을 것이다.

본 연구에서는 이러한 특성을 고려하여 다중 위험 보험의 각 위험에 내재된 특성을 의미하는 고정효과와 시간의 흐름에 걸쳐 공유된 임의효과를 통합하면서 다중 위험의 청구 건간에 자연스러운 의존 구조를 반영한 요율산출 모델을 제안하였다. 실제 청구 데이터를 확인한 결과 과대산포가 존재하는 것을 확인하여 이를 반영하기 위해 빈도 모델분포로 quasi-Poisson 분포를 채택하였다. 임의효과의 사전분포는 켈레 사전분포의 성질을 이용

하기 위하여 감마분포를 가정하였고, 이로부터 자연스럽게 감마분포인 사후분포를 도출하였다. 최종적으로 예측분포는 음이항 분포임을 유도하였다. 다중 위험 보험의 위험 간의 존성, 관찰된 이질성 및 관찰되지 않은 이질성, 그리고 과대산포 건수 데이터의 분포에 대해서는 다수의 선행연구가 존재하지만 이를 모두 포함하는 연구는 거의 이루어지지 않았다. 본 연구는 이를 모두 반영한 효율산출 모델을 제시하였다는 점에 의의가 있다. 또한, 예측분포가 음이항 분포이므로 제안 모델은 닫힌 형태(closed form)의 다중 위험 보험료 공식을 제공할 수 있고 실무적으로 적용과 해석이 용이하다는 장점이 있다.

본 연구에서는 대인보상 및 대물보상 담보의 청구 건수 데이터를 훈련자료로 사용하여 제안 모델의 계수를 산출하였고 이를 검증하기 위하여 시험자료로 표본의 검증을 수행하였다. 그 결과 청구 빈도 예측에 있어 기존 방법보다 우수한 결과를 보이는 것으로 나타났다. 이 결과는 제안 모델이 보다 적절한 보험요율을 산출할 수 있다는 것을 의미한다. 그러나 이 모델을 실제 우리나라 자동차보험의 보험료 산출에 사용하기에는 한계가 있다. 우리나라의 자동차보험은 의무보험으로서 감독당국에 의한 규제를 많이 받기 때문에 할인할증 제도 역시 정해진 규정에 따라 적용해야 한다. 따라서 제안 모델을 그대로 실무적으로 적용하는 데에는 무리가 있을 수 있으나, 효율의 적정성을 검증하거나 회사 내부적인 분석 목적으로는 유용하게 활용할 수 있을 것이다.

제안 모델에서는 빈도 데이터의 과대산포를 반영하기 위해 quasi-Poisson 분포를 모델 분포로 적용하였다. 이 분포는 과대산포를 적용하면서도 켈레 사전분포를 이용할 수 있고 그 결과 예측 분포가 음이항 분포를 따른다는 해석적인 장점을 가지고 있어 모델 분포로 채택하였다. 최종후·고인미·전수영(2011)에 따르면 과대산포를 고려한 빈도분포에는 음이항(negative binomial), 제로팽창 포아송(zero-inflated Poisson), 제로팽창 음이항(zero-inflated negative binomial) 분포 등 여러 가지가 있고 이에 대한 선행연구도 다수 존재한다. 따라서 빈도 분포로써 quasi-Poisson 분포가 아닌 다른 분포를 본 연구의 제안 모델에 적용하는 것을 시도해볼 수 있다. 또한 본 연구의 분석에서는 계약자 모두에게 동일한 관찰기간 동안 발생한 과거 청구 이력에 대하여 모두 동일한 가중치를 모델에 반영하였는데, 할인할증 요율 산출에 가장 적합한 관찰기간을 계약자집단별로 다르게 선택한 후 그에 따라 서로 다른 기간별 가중치를 모델에 반영하는 방향으로 분석을 발전시킬

수도 있을 것이다.

한편, 본 연구에서는 개인용 자동차보험의 주요 담보인 대인보상과 대물보상, 두 가지 담보만을 분석 대상으로 하였는데, 자동차보험에는 다른 용도와 여러 담보가 존재하므로 후속연구에서는 자동차상해나 자기신체사고와 같은 담보를 추가해서 분석대상 담보를 확장하여 본 연구모형의 적용가능성을 확인해 보는 것을 기대한다. 그리고 일반보험이나 인보험에도 다중 위험 보험이 다양하게 존재하기 때문에, 다른 보험상품으로까지 연구를 확장해보고 본 연구와 비교 및 대조해 보는 것 또한 기대한다.

참고문헌

- 기승도·김대환 (2009), “일반화선형모형(GLM)을 이용한 자동차보험 요율상대도 산출방법 연구”, **보험연구원 연구보고서**, 2009-05.
- (Translated in English) Ki, S., and D., Kim (2009). “Automobile Insurance Ratemaking Using Generalized Linear Model(GLM)”, *Korea Insurance Research Institute*, 2009-05.
- 김동훈·이제연 (1998), “로짓모형을 이용한 자동차보험 사고확률 추정”, **리스크관리연구**, 제10권 제1호, pp. 75-112.
- (Translated in English) Kim, D., and K., Lee (1998). “An Estimation of Automobile Insurance Accident Probability by Logit Model”, *The Journal of Risk Management*, 10(1):75-112.
- 김영화·김미정·김명준 (2011), “여러가지 신뢰도에 근거한 자동차 보험료 예측”, **응용통계연구 논문집**, 제24권 제2호, pp. 279-292.
- (Translated in English) Kim, Y., M., Kim and M., Kim (2011) “Estimating the Automobile Insurance Premium Based on Credibilities”, *The Korean Journal of Applied Statistics*, 24(2):279-292.
- 김영화·이현수 (2010), “신뢰도에 근거한 자동차 보험 가격 산출 비교”, **한국통계학회 논문집**, 제17권 제5호, pp. 713-724.
- (Translated in English) Kim, Y., and H., Lee (2010). “A Comparison Study for the Pricing of Automobile Insurance Premium Based on Credibility”, *Communications for Statistical Applications and Methods*, 17(5):713-724.
- 김진한·민재형 (1994), “마코프체인을 이용한 자동차보험료의 할인·할증시스템에 관한 연구”, **한국경영과학회 학술대회논문집**, 1994년 제1권, pp. 79-95.
- (Translated in English) Kim, J., and J., Min (1994). “An Exploratory Study on the Bonus-Malus System of Automobile Insurance Premium

through the Method of Markov Chain”, *Proceedings of the Korean Operations and Management Science Society Conference*, 1994, 1:79-95.

보험개발원 (2014), “자동차보험 할인할증제도 개선 정책토론회 개최”, 2014. 2. 24.

(Translated in English) Korea Insurance Development Institute (2014).

“Policy Discussion Session to Improve the Automobile Insurance Bonus-Malus Systems”, 2014. 2. 24.

_____ (2022), 보험통계조회서비스, 2022. 3. 30. (Retrieved from <https://incos.kidi.or.kr:5443>).

(Translated in English) Korea Insurance Development Institute (2022).

Insurance Statistics Consumer Service, 2022. 3. 30. (Retrieved from <https://incos.kidi.or.kr:5443>).

심현우·임형기·최양호 (2021), “부분적 청구경력 고지 간소화에 따른 건강보험 요율 차등화에 관한 연구”, **보험금융연구**, 제32권 제2호, pp. 43-75.

(Translated in English) Shim, H., H., Im and Y., Choi (2021). “A Study on Premium Rate Differentiation of Health Insurance by Simplified Partial Disclosure of Claim History”, *Journal of Insurance and Finance*, 32(2):43-75.

정중영·김태완 (2004), “자동차보험의 효율적인 할인·할증제도 분석 연구”, **한국자료분석학회지**, 제6권 제6호, pp. 1665-1675.

(Translated in English) Jeong, J., and T., Kim (2004). “An Analysis of Adequate and Efficient Bonus-Malus System in the Korean Automobile Insurance”, *Journal of The Korean Data Analysis Society* 6(6):1665-1675.

조재린·이기형·강중철 (2012), “신뢰도를 적용한 자동차보험요율산출 연구”, **한국자료분석학회지**, 제14권 제4호, pp. 2267-2276.

(Translated in English) Cho, J., K., Lee and J., Kang (2012). “A Study on the

- Pricing of Car Insurance with Credibility”, *Journal of The Korean Data Analysis Society*, 14(4):2267-2276.
- 최종후·고인미·전수영 (2011), “제로팽창 모형을 이용한 보험데이터 분석”, *응용통계연구*, 제24권 제3호, pp. 485-494.
- (Translated in English) Choi, J., I., Ko and S., Cheon (2011). “A Zero-Inflated Model for Insurance Data”, *The Korean Journal of Applied Statistics*, 24(3):485-494.
- Bailey, A. (1950). “Credibility procedures: Laplace’s generalization of Bayes’ rule and the combination of collateral knowledge with observed data”, *Proceedings of the Casualty Actuarial Society*, 37:7-23.
- Bermúdez, L., and D., Karlis (2017). “A posteriori ratemaking using bivariate Poisson models”, *Scandinavian Actuarial Journal*, 2017(2):148-158.
- Boucher, J.-P., and M., Denuit (2008). “Credibility premiums for the zero-inflated Poisson model and new hunger for bonus interpretation”, *Insurance: Mathematics and Economics*, 42(2):727-735.
- Boucher, J.-P., M., Denuit and M., Guillen (2008). “Models of insurance claim counts with time dependence based on generalization of Poisson and negative binomial distributions”, *Variance*, 2(1):135-162.
- _____ (2009). “Number of accidents or number of claims? An approach with zero-inflated Poisson models for panel data”, *Journal of Risk and Insurance*, 76(4):821-846.
- Bühlmann, H. (1967). “Experience rating and credibility”, *ASTIN Bulletin: The Journal of the IAA*, 4(3):199-207.
- Cameron, A., and P., Trivedi (1990). “Regression-based tests for

- overdispersion in the Poisson model”, *Journal of Econometrics*, 46(3):347-364.
- e Silva, J., and M., de Lourdes Centeno (2017). “Ratemaking of dependent risks”, *ASTIN Bulletin: The Journal of the IAA*, 47(3):875-894.
- Frangos, N., and S., Vrontos (2001). “Design of optimal bonus-malus systems with a frequency and a severity component on an individual basis in automobile insurance”, *ASTIN Bulletin: The Journal of the IAA*, 31(1):1-22.
- Frees, E. (2003). “Multivariate credibility for aggregate loss models”, *North American Actuarial Journal*, 7(1):13-37.
- Frees, E., G., Lee and L., Yang (2016). “Multivariate frequency-severity regression models in insurance”, *Risks*, 4(1):4.
- Frees, E., G., Meyers and A., Cummings (2010). “Dependent multi-peril ratemaking models”, *ASTIN Bulletin: The Journal of the IAA*, 40(2):699-726.
- Frees, E., V., Young and Y., Luo (1999). “A longitudinal data analysis interpretation of credibility models”, *Insurance: Mathematics and Economics*, 24(3):229-247.
- Gómez-Déniz, E., and F., Vázquez-Polo (2005). “Modelling uncertainty in insurance bonus-malus premium principles by using a bayesian robustness approach”, *Journal of Applied Statistics*, 32(7):771-784.
- Jeffreys, H. (1946). “An invariant form for the prior probability in estimation problems”, *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 186(1007):453-461.
- Jeong, H. (2020). “Testing for random effects in compound risk models via Bregman divergence”, *ASTIN Bulletin: The Journal of the IAA*, 50(3): 777-798.

- Jeong, H., and E., Valdez (2020a). "Bayesian credibility premium with GB2 copulas", *Dependence Modeling*, 8(1):157-171.
- _____ (2020b). "Predictive compound risk models with dependence", *Insurance: Mathematics and Economics*, 94:182-195.
- Jewell, W. (1974). "Credible means are exact bayesian for exponential families", *ASTIN Bulletin: The Journal of the IAA*, 8(1):77-90.
- Klugman, S., H., Panjer and G., Willmot (2012). *Loss Models: From Data to Decisions*, 4th Edition, Wiley.
- Lemaire, J. (1998). "Bonus-malus systems: the european and asian approach to merit-rating", *North American Actuarial Journal*, 2(1):26-38.
- Liang, K.-Y., and S., Zeger (1986). "Longitudinal data analysis using generalized linear models", *Biometrika*, 73(1):13-22.
- Mayerson, A. (1964). "A bayesian view of credibility", *Proceedings of the Casualty Actuarial Society*, 51:7-23.
- McCullagh, P., and J., Nelder (1989). *Generalized linear models*, 2nd Edition, Chapman and Hall.
- Norberg, R. (1986). "Hierarchical credibility: analysis of a random effect linear model with nested classification", *Scandinavian Actuarial Journal*, 1986(3-4):204-222.
- Pechon, F., M., Denuit and J., Trufin (2019). "Multivariate modelling of multiple guarantees in motor insurance of a household", *European Actuarial Journal*, 9(2):575-602.
- _____ (2020). "Home and motor insurance joined at a household level using multivariate credibility", *Annals of Actuarial Science*, 2020:1-33.
- Pechon, F., J., Trufin and M., Denuit (2018). "Multivariate modelling of

- household claim frequencies in motor third-party liability insurance”, *ASTIN Bulletin: The Journal of the IAA*, 48(3):969-993.
- Quan, Z., and E., Valdez (2018). “Predictive analytics of insurance claims using multivariate decision trees”, *Dependence Modeling*, 6(1):377-407.
- Sutradhar, B., and V., Jowaheer (2003). “On familial longitudinal poisson mixed models with gamma random effects”, *Journal of multivariate analysis*, 87(2):398-412.
- Wedderburn, R. (1974). “Quasi-likelihood functions, generalized linear models, and the gauss-newton method”, *Biometrika*, 61(3):439-447.
- Whitney, A. (1918). “The theory of experience rating”, *Proceedings of the Casualty Actuarial Society*, 4:274-292.
- Yang, L., and P., Shi (2019). “Multiperil ratemaking for property insurance using longitudinal data”, *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 182(2):647-668.

Abstract

Multi-peril insurance consists of multiple coverages which compensate for accidents caused by various perils. In multi-peril insurance, there may be dependence among the claims from different perils. When dependence is confirmed, it is reasonable to consider the correlation among perils in ratemaking. In this study, we confirmed such dependence and proposed a shared random effects model to capture and reflect this in posterior ratemaking. The model can provide closed forms of credibility premium, which makes it easy to implement in actuarial practice.

We calibrated the model using the actual claim data of bodily injury and property damage of automobile insurance as an example of multi-peril insurance. Then we used the model for posterior ratemaking and evaluated its performance by out-of-sample validation. The results showed that the performance of the proposed model is superior to that of comparative models such as the naive Poisson model.

※ Key words: Multi-peril Insurance, Shared Random Effects, Overdispersion, Claim Frequency Analysis, Bonus-malus System

손해보험회사의 리스크 특성과 회계정보의 가치관련성

Risk Characteristics of Non-life Insurers and Value Relevance of Accounting Information

조 석 희*

Seokhee Cho

본 연구는 국내 손해보험회사의 리스크 특성을 현행의 RBC체계 내에서 분석하고 리스크 특성의 차이에 따라 회계정보의 가치관련성이 차별화되는지에 대하여 분석하였다. 분석결과는 다음과 같다.

첫째, 회계정보의 전반적인 가치관련성은 해당 보험회사의 보험 및 운영리스크 그리고 금리 및 신용리스크 특성에 따라 서로 다른 방향으로 차별화되었다. 보험리스크와 운영리스크의 비중이 높을수록 회계정보의 전반적 가치관련성은 낮아지는 반면, 금리리스크나 신용리스크의 비중이 높을수록 회계정보의 전반적 가치관련성은 높아지는 것으로 나타났다.

둘째, 회계이익의 상대적 가치관련성과 관련하여, 보험회사의 보험 및 운영리스크의 비중이 높을수록 순이익의 상대적 가치관련성은 높아지는 반면에, 신용리스크의 비중이 높을수록 순이익의 상대적 가치관련성은 낮아지는 것으로 나타났다.

보험회계에 관한 다수의 선행연구에서는 보험회계정보의 낮은 이익품질에 대해 보고한바 있으나 아직까지 그 원인을 명쾌하게 제시한 연구는 없다. 이러한 상황에서 본 연구는 보험회사 회계정보의 전반적 가치관련성과 회계이익의 상대적 가치관련성이 낮거나 높아질 수 있는 조건을 탐색함으로써 보험회계정보의 품질에 대한 연구에 참고할 만한 정보를 제공하였다는 측면에서 의의를 가진다.

국문 색인어: 리스크 특성, RBC체계, 요구자본, 가치관련성

한국연구재단 분류 연구분야 코드: B051600, B051605

* 안동대학교 경영회계학부 부교수(csh@anu.ac.kr), 제1저자

논문 투고일: 2022. 1. 12, 논문 최종 수정일: 2022. 5. 11, 논문 게재 확정일: 2022. 8. 18

I. 서론

본 논문에서는 손해보험회사의 리스크 특성을 파악한 후 파악된 특성에 따라 해당 보험 회사 회계정보의 가치관련성이 차별적으로 나타나는지 여부에 대하여 분석하고자 한다.

현행의 보험감독체계에서는 보험회사의 리스크를 보험리스크, 금리리스크, 신용리스크, 시장리스크 그리고 운영리스크로 구분하여 내재된 리스크량을 산출하도록 하고 있다. 이렇게 각각 산출된 리스크량은 개별 리스크 간의 상관관계 유무 및 그 정도를 반영하여 합산하는 방식으로 회사전체의 요구자본으로 결정된다.

보험회사의 회계정보를 연구한 많은 선행연구에서는 보험회사의 순이익(분기순손익 또는 당기순손익) 정보의 기업가치 설명력이 유의하지 않거나 낮은 것으로 보고되고 있다.¹⁾ 이러한 현상은 제조업을 중심으로 한 비금융회사²⁾를 대상으로 한 연구결과와 일관되지 않는다. 비금융업을 대상으로 한 대부분의 관련분야 선행연구에 따르면, 순이익 정보의 기업가치 설명력은 상당한 정도로 유의하기 때문이다.

본 논문에서의 분석은 회계이익의 가치관련성 수준에 영향을 미칠 수 있는 것으로서 보험회사의 리스크 특성을 고려하였다. 그리고 보험회사의 리스크 특성은 RBC체계상 산출되는 요구자본의 구성상 특징으로 규정하고 보험리스크, 금리리스크, 신용리스크, 시장리스크 그리고 운영리스크가 해당 보험회사의 전체 리스크량에서 얼마나 중요한지를 계량화하여 분석하고자 한다.

본 논문에서는 여러 가지 개별 리스크 특성의 고저(高低)에 따라 보험회사 회계정보의 전반적인 가치관련성 및 회계이익의 상대적 가치관련성이 차별화되는지를 실증적으로 분석하기로 한다. 그리고 이러한 분석의 결과를 통하여 회계이익의 가치관련성이 높거나 낮

-
- 1) 손해보험업에서 순이익의 상대적 가치관련성이 낮을 수 있음을 이론적으로 주장한 연구로 김호중 외(1997)가 있으며, 이후 손해보험업의 회계정보에 대한 가치관련성을 다룬 오창수·변재웅(2015), 조석환(2016), 이원아·심현우(2018) 등의 실증분석 결과에서는 모두 순자산의 가치관련성은 높은데 반하여 순이익의 가치관련성은 유의하지 않게 보고되고 있다.
 - 2) 기업가치 평가과정에서 회계이익의 중요성이 낮다는 점은 보험업을 포함한 모든 금융회사들에게 공통적으로 적용되는 현상으로 알려져 있다. 그러나 본 논문에서는 회계정보의 가치관련성과 연관된 보험회사의 고유특성을 심도 있게 논의하고자 하는 취지에서 분석범위를 전 금융업으로 확장하지 않고, 보험업만을 대상으로 하기로 한다. 따라서 은행 및 금융투자업에 대한 논의는 하지 않는 것으로 한다.

아지는 조건을 탐색하고, 보험회사의 가치 평가 및 리스크 관리 관점에서 시사하는 바에 관하여 논의하고자 한다.

논문의 나머지 부분은 다음과 같이 구성된다. 제2장에서는 RBC제도상의 요구자본 및 이와 관련된 이론적 배경을 정리하고, 요구자본의 세부적 구성과 기업가치 평가와의 관련성에 대해 논의한다. 그리고 이를 토대로 실증분석을 통한 검정의 대상이 되는 연구가설을 설정한다. 그리고 제3장에서는 실증분석을 위한 연구모형을 설계한 후, 제4장에서 그 결과의 보고와 함께 보고된 결과와 관련된 시사점에 대해 논의한다. 마지막으로 제5장에서는 연구결과를 요약하고, 연구의 의의 등에 대해 논의하며 결론을 맺는다.

II. 이론적 배경 및 연구가설 설정

1. RBC제도와 지급여력비율

금융기관의 경우 예상치 못한 손실로 인한 지급불능사태가 발생하는 경우 사회 전반에 미치는 파급효과가 작지 않으므로 자기자본에 대한 엄격한 규제가 필요하다. 금융기관에 대한 자기자본규제의 핵심은 금융계약 상대방에 대한 지급능력(solvency)의 유지에 있다. 이러한 취지로 국내 각각의 금융권역별로 고유한 자기자본규제제도를 가지고 있다. 구체적으로 은행의 경우 BIS 자기자본 규제제도를, 금융투자업에 대해서는 영업용 순자본비율 규제 그리고 보험업에 대해서는 지급여력제도를 운영하고 있다(박경원·장지인 2012).

보험회사의 지급능력은 보험회사가 보험계약자에게 미리 약정한 보험금을 지급할 수 있는 재무적 능력을 말한다(정요섭 1999). 한편 지급여력제도는 예상할 수 없는 금융회사의 손실에 대비하고자 하는 것이므로 미리 예상할 수 있는 급부금에 대한 지급능력만을 유지하는 것으로 보험회사에 대한 자기자본규제의 목적을 달성할 수는 없는 것이다. 따라서 지급능력 유지에 필요한 자기자본 이상의 자본을 추가적으로 보유할 필요가 있는데, 이러한 경우 지급여력(solvency margin)이 생길 수 있다. 자기자본 규제와 관련된 국내 보험감독체계의 핵심은 금융회사로 하여금 지급여력비율을 산출하여 보고토록 하고 이 비율을 이용하여 선제적인 감독을 하는 것이다.

지급여력비율은 지급여력금액을 지급여력기준금액으로 나눈 비율이다. 지급여력금액은 보험회사의 가용한 자본량을 의미하므로 통상 가용자본(available capital)이라고 부른다. 그리고 지급여력기준금액은 회사의 규모와 리스크를 고려하여 측정한 표준적 금액으로서 안정적인 지급능력 유지에 필요한 자본의 성격을 갖는다. 지급여력금액은 통상 요구자본(required capital)이라고 부르는데, 결국 지급여력비율은 요구자본에 대비한 가용자본의 비율을 의미한다.

과거 수십 년 동안 지급여력비율의 산출방식은 몇 차례 큰 폭의 변화가 있었다. 1990년대 후반 외환위기를 겪으면서 국내 보험업권에서는 지급여력제도를 서둘러 도입할 필요가 있었고, 이에 따라 유럽식 지급여력제도를 도입하여 시행하였다. 그러나 동 제도의 여러 가지 문제점이 드러났고 이를 개선할 필요가 커짐에 따라 2009년에는 새로운 지급여력제도인 RBC제도를 도입하였으며, 2년간의 시범적 시행을 거쳐 현재까지 사용하고 있다. 본 논문은 기본적으로 RBC체계에 따라 산출되는 요구자본의 세부 리스크 항목의 특성을 중심으로 연구를 진행하기로 한다.³⁾

2. 요구자본의 산출구조와 그 경제적 특성

현행 RBC제도하에서의 지급여력 요구자본은 보험회사에 내재된 각종 리스크를 계량화하여 이에 해당하는 자본량을 산출하고, 개별 리스크 항목별 상관관계와 분산효과를 고려하여 합산한 수치를 회사 전체의 요구자본으로 확정한다. 이 때 구분하여 개별적으로 산출되는 리스크 항목의 정의와 결정요인은 <Table 1>과 같다.

보험리스크는 예상하지 못한 손해율 등의 급격한 악화로 인한 손실위험을 의미하는 것으로, 손해율의 변화와 지급준비금의 적립수준 등에 의해 영향을 받는다. 금리리스크는 이자율의 변화에 따라 순자산가치가 하락함으로써 발생하는 예상치 못한 손실위험을 의미하

3) 국내 보험감독당국에서는 새로운 지급여력제도(K-ICS)를 2023년부터 시행할 예정인바, 새로운 제도의 시행으로 현행의 지급여력요구자본에 대한 산출체계는 일정부분 변화가 있을 예정이다. 본 논문은 새로운 제도의 시행전에 사용하고 있는 지급여력요구자본의 산출체계를 기초로 분석한다는 점에서 일정한 한계가 있을 수 있으나, 제도의 변화로 인한 리스크 특성의 근간은 변화되지 않는 것이므로 본 논문은 리스크 특성과 회계정보 속성의 일반적 관계에 대한 충분한 시사점을 제공할 것으로 기대한다.

는 것으로서, 보험회사의 자산과 부채가 가지는 금리민감도와 금리연동형 상품 등의 보유 비중 등에 따라 영향을 받을 수 있다.

신용리스크는 주로 금융자산에 내재된 위험으로서 채무자의 부도 등으로 인해 예상하지 못한 손실을 반영한다. 대출자산의 규모나 건전성, 담보유무 등에 따라 리스크의 크기가 결정될 수 있다. 시장리스크는 주식가격, 이자율수준 그리고 환율 등 시장변수의 불리한 변동으로 인하여 순자산가치가 예상치 못하게 하락할 위험을 의미하는 것으로서, 주로 투자포트폴리오의 구성이 어떠한지에 따라 리스크의 크기가 결정될 수 있다. 시장리스크의 경우에는 보험권역에서 차지하는 비중이 적고, 그 중요성도 크지 않은 리스크로 여겨지고 있다. 운영리스크는 이론적으로 적절하지 못한 내부통제제도나 인력운영 등을 포함한 내부프로세스 전반에 걸친 비재무적 위험뿐만 아니라 불리한 외부적 사건에 따라 입을 수 있는 손실위험을 말하는데, 동 리스크 항목은 계량화하여 측정하기 어렵기 때문에 보험감독 목적으로는 수입보험료의 일정비율 등 비교적 간편한 방법으로 산정하고 있다.

〈Table 1〉 Individual Risk Definition and Determinants

Category	Definition	Determinants
Insurance risk	Risk of loss due to an unexpected increase in the loss ratio	• Loss ratio • Level of claim reserve accumulation
Interest rate risk	Risk that will negatively affect financial status due to the fluctuation of net asset value due to interest rate fluctuations.	• Interest rate sensitivity of assets/liabilities • The proportion of interest linked products
Credit risk	Risk of loss due to the debtor's default or the counter party's default, etc.	• Credit rating, default rate • Credit reinforcement details such as collateral and guarantee
Market risk	Risk of loss due to changes in asset value caused by fluctuations in market prices (stock prices, interest rates, exchange rates, etc.).	• Appropriateness of distributed investment • The proportion of variable insurance
Operating risk	Risk of loss due to inappropriate internal procedures, personnel, systems, external events, etc.	• Appropriateness of internal control • Appropriateness of accident prevention measures

Source: FSS(2020), (Revised) Guide for RBC system for the insurers.

이상에서 살펴본 다섯 가지 리스크 항목에 대해 종합하면 다음과 같다. 보험회사의 활동은 크게 보험영업과 투자영업으로 구분하는 것이 일반적이다. 이는 감독회계에서도 그대로 받아들여져 보험회사의 손익계산구조를 설계하였는데, 영업이익을 보험영업과 투자영업으로 엄격하게 구분하는 것이 그 대표적인 예이다. 이러한 구분방식에 따라 개별 리스크 항목을 구분한다면, 보험리스크와 운영리스크는 주로 보험영업과 관련이 깊은 항목으로 볼 수 있고, 신용리스크와 시장리스크는 주로 투자자산과 연동하여 산출되므로 투자영업과 관련이 깊은 항목으로 볼 수 있다. 그리고 금리리스크의 경우에는 금융자산과 책임준비금의 대응구조에서 발생할 수 있으므로 보험영업 및 투자영업 모두와 깊게 관련되어 있다고 볼 수 있다.

3. 연결기준 지급여력제도의 시행

2016년 4분기부터는 지급여력비율 산출 시 연결대상 종속회사까지 반영하여 연결실체 전체의 지급여력비율을 산출하고 있다. 본 연구의 연구대상 기간이 2011년부터 2021년까지 이므로 연구를 수행함에 있어 2016년 3분기까지는 별도재무제표(separate F/S) 또는 개별재무제표(individual F/S)를 기준으로 분석하는 것이 바람직하고 2016년 4분기부터는 연결재무제표(consolidate F/S)를 기준으로 분석하는 것이 바람직하다. 그러나 국내 손해보험회사들의 경우 연결실체 내에서 비지배지분(non-controlling interest)의 비중이 극히 낮기 때문에 연결재무제표 전체와 연결재무제표 중 지배지분해당액 또는 별도재무제표와의 차이는 중요하지 않을 것으로 판단된다.⁴⁾ 예를 들어 2016년 4분기말 기준 국내 상장손해보험회사의 연결기준(연결대상이 없는 기업은 개별기준, 이하 동일) 자본총액에서 별도기준(연결대상이 없는 기업은 개별기준, 이하 동일) 자본총액과 연결기준 자본총액 중 지배기업소유주지분이 차지하는 비율은 각각 99.55%와 99.07%로 세 가지 자본개념 간의 차이는 중요하지 않은 것으로 나타났다. 또한 2016년 4분기 기준 연결기준 분기

4) 여기서 연결재무제표 전체라 함은 연결재무상태표상의 자기자본이나 연결포괄손익계산서상의 연결당기순이익을 의미하는 것이고, 연결재무제표 중 지배지분해당액은 연결재무상태표상의 자기자본 중 지배기업소유주지분이나 연결포괄손익계산서상의 연결당기순이익 중 지배기업소유주지분귀속분을 의미한다. 같은 취지로 별도재무제표 전체는 별도재무상태표상의 자기자본이나 별도포괄손익계산서상의 당기순이익을 의미하는 것이다.

순이익에서 별도기준 분기순이익과 연결기준 분기순이익 중 지배기업소유주귀속분이 차지하는 비율은 각각 92.49%와 99.74%로 세 가지 손익개념 간의 차이 역시 크게 중요하지 않아 보인다.

한편, 지급여력 요구자본을 연결기준으로 산출하게 됨에 따른 요구자본의 변화효과를 살펴보면 다음과 같다. 별도재무제표 기준으로 산출된 2016년 3분기말의 요구자본과 연결재무제표 기준으로 산출된 2016년 4분기말의 요구자본을 비교하면, 평균적으로 11.79%의 증가가 있었는데, 대부분 연결의 효과로 보여진다. 2016년 3분기말 전체 요구자본 대비 개별 리스크 항목별 리스크량 증가액의 비율은 보험, 금리, 신용, 시장 그리고 운영리스크 항목 각각에서 2.07%, 2.70%, 8.18%, 0.84% 그리고 0.47%수준을 보였다. 그리고 연결기준 지급여력 요구자본에 새롭게 편입된 해외 관계사 지분관련 요구자본과 국내 비보험회사 지분관련 요구자본 증가액이 2016년 3분기말의 전체 요구자본에서 차지하는 비율은 각각 0.42%와 0.53%였다. 이에 비하여 지급여력금액(가용자본)은 상대적으로 적게 증가하였고, 결과적으로 국내 손해보험회사의 지급여력비율은 하락⁵⁾하게 되었다.

이와 같은 상황을 반영하여 본 연구에서는 개별 리스크 특성의 고저(高低) 판단 시 같은 시점을 기준으로 업계의 평균적인 수준과 비교하는 방식을 사용함으로써 감독기준의 변경에 따른 요구자본 산출방식의 변화를 반영할 수 있도록 한다. 그리고 회계변수에 대해서는 2016년 3분기까지는 개별 및 별도기준 정보를 이용하고, 2016년 4분기부터는 개별 및 연결기준 정보를 이용하여, 요구자본의 산출체계와 적절히 대응되도록 하고자 한다.

4. 선행연구 검토

보험회사의 지급여력비율 및 지급여력 요구자본과 연계하여 기업가치 관련성을 분석하였거나 비록 보험회사가 아니더라도 순자산과 순이익의 상대적 가치관련성을 기업특성 측면에서 분석한 연구로는 Barth et al.(1998), 장지인 외(2010), 김정옥(2010), 오창수·조석희(2013) 등이 있다.

5) 2016년 3분기말 상장 손해보험회사의 지급여력비율의 평균은 230%를 상회하였으나, 2016년 4분기말에는 190%를 하회하는 수준이 되었다. 대부분 연결기준 지급여력제도의 도입관련 효과로 보여진다.

Barth et al.(1998)은 순자산과 순이익의 기업가치 관련성이 해당기업의 재무건전성에 따라 차별화될 수 있음을 실증연구결과로 보고하고 있다. 이 연구에서는 재무건전성 정도가 높은 표본에서 그렇지 않은 경우보다 순이익의 상대적 가치관련성이 높게 나타나고 있음을 보였다. 이 연구는 본 연구와 같이 보험회사만을 대상으로 한 연구는 아니지만 기업의 재무적 특성에 따라 회계정보의 가치관련성에 차이가 있음을 보고하였다는 점에서 본 연구와 관련이 있다.

장지인 외(2010)에서는 1989년부터 2008년까지 미국의 60개 무부채기업과 이에 대응되는 184개의 부채기업을 대상으로 기업 자본구조상의 특성이 순이익과 순자산의 가치관련성에 영향을 미치는지 여부에 대하여 분석하였다. 구체적으로 이 연구에서는 순이익과 순자산의 주가 반응계수와 순이익과 순자산의 상대적 주가설명력의 차이를 관찰하는 방식으로 분석하였다. 분석결과, 순자산의 주가 반응계수는 부채기업에서 높게 나타나고, 순이익의 반응계수는 무부채기업에서 높게 나타났다. 그리고 순자산과 순이익의 주가에 대한 총설명력은 무부채기업에서 높게 나타났지만 순자산만의 증분설명력은 부채기업에서 더 높게 나타나는 것으로 보고하였다.

김정옥(2010)에서는 부채의 구성에 따라 순자산과 순이익의 가치관련성에 차이가 발생 하는지 여부에 대하여 분석하였다. 부채의 구성은 일반부채와 금융기관차입금으로 구분하였고, 연구결과 금융기관차입금의 비율이 높아질수록 순자산의 상대적 가치관련성이 높아지고, 일반부채의 비율이 높아질수록 순이익의 상대적 가치관련성이 높아짐을 보고하였다.

오창수·조석희(2013)에서는 손해보험회사 지급여력비율의 증감에 따른 순자산과 순이익의 상대적 가치관련성에 변화가 있는지에 대해 분석하였다. 연구결과 지급여력비율이 적절한 수준(너무 높거나 낮지 않은 수준)일 때 순이익의 가치관련성이 가장 높게 나타나고 있음을 보고하였다. 이 연구는 재무건전성의 높고 낮음에 따라 회계정보의 가치관련성이 달라진다는 이분법적 주장의 오류가능성에 대해 언급하였다는 점에서 의의가 있다.

5. 보험회사 회계정보의 전반적 가치관련성

회계정보의 전반적 가치관련성은 정보위험의 측면에서 생각해 볼 수 있다. 정보위험이 발생하는 이유는 경영자와 외부투자자 간의 정보비대칭에 있는 것으로서 비교적 회계정보가

복잡하고 회계정보의 산출에 많은 추정과정이 개입되는 보험업의 회계정보에서 정보위험은 더욱 높아질 것으로 판단한다. Ohlson(1995) 모형에서 주가로 측정되는 기업가치는 기본적으로 회계정보와 기타정보의 함수로 간주되는데, 이러한 가치평가 관점에서는 정보위험이 낮을수록 기업가치를 결정하는데 필요한 정보가 회계정보에 많이 반영되어 회계정보를 이용한 가치평가의 정확성이 높아질 수 있다. 김경순·이진훤(2016)에서는 투자자들이 기업의 정보위험이 낮다고 판단하면, 기업가치 평가과정에서 회계정보를 보다 중요하게 반영시킬 것이기 때문에 긍정적인 이익특성은 회계정보의 가치관련성을 높이는 요소로 보았다. 왜냐하면 높은 이익특성은 정보위험에 대한 평가를 반영하는 것이라고 보았기 때문이다.

정보위험을 보험회사의 리스크 특성과 연관하여 생각해보면 다음과 같다. 보험업 회계정보의 정보위험이 높아지는 주요 원인으로는 회계처리의 복잡성과 낮은 이해가능성 그리고 회계정보 산출에 많은 추정이 필요하고 주관적 판단의 개입가능성이 높다는 점을 꼽을 수 있다. 특히, 고도의 보험수리적 기법과 가정이 개입되는 책임준비금 결산 등과 관련된 영역이 회계정보의 정보위험을 높이는 주요부문으로 판단된다. 즉, 보험회사의 활동을 보험영업과 투자영업으로 대별하는 경우, 정보위험을 높이는 영역은 투자영업부문보다는 보험영업부문으로 판단된다. 따라서 보험영업부문의 중요성이 상대적으로 높아질 때 정보위험은 높아지고, 시장참여자들은 정보위험이 높은 회계정보를 기업가치 평가에 잘 반영하지 않을 것이므로 해당 회계정보의 가치관련성은 자연스럽게 하락할 가능성이 높다(김경순·이진훤 2016; Easley and O'hara 2004; Francis et al. 2004; Hwang and Lim 2012). 그리고 보험영업부문의 중요성이 상대적으로 큰 경우 보험리스크와 운영리스크의 상대적 비중은 높아질 것이므로 보험리스크나 운영리스크의 상대적 비중이 높을수록 회계정보의 전반적 가치관련성은 하락할 것이다. 이와 반대로 보험영업부문의 상대적 비중이 낮아질수록 투자영업부문의 상대적 비중은 커지게 되고, 이러한 상황에서는 신용리스크와 시장리스크의 상대적 비중이 높아지게 된다. 따라서 신용리스크와 시장리스크의 상대적 비중이 높아질수록 회계정보의 전반적 가치관련성은 높아질 수 있다. 다만, 금리리스크의 경우에는 보험영업과 투자영업부문 모두와 관련되어 있으므로 금리리스크의 상대적 비중이 높아지는 경우 정보위험의 크기가 어떻게 변화하는지 또는 회계정보의 전반적 가치관련성에 어떤 영향이 있을지를 예상하기는 어렵다. 이러한 논의를 종합적으로 반영하여 본

논문에서는 다음과 같은 연구가설을 설정한다.

[가설 1] 손해보험회사 회계정보의 전반적 가치관련성은 지급여력 요구자본의 구성내역에 따라 차별화된다.

6. 보험회사 회계이익의 상대적 가치관련성

Ohlson(1995)의 회계정보를 이용한 기업가치 평가모형에 따르면, 기업가치 중 회계이익으로 설명되는 부분은 기업의 미래 초과이익에 대한 현재가치를 의미하게 된다. 여기서 미래 초과이익은 예상되는 이익의 수준에서 정상이익을 차감한 것으로서 정상이익은 통상 기업의 자본비용을 반영하여 결정된다. 즉, 높은 자본비용은 정상이익의 규모를 증가시키게 되고, 높아진 정상이익은 미래 초과이익을 감소시키게 된다. 따라서 미래 초과이익의 감소는 자연스럽게 회계이익의 가치관련성을 낮게 한다. 이와 관련하여, 오창수·조석희(2013)에서는 높은 지급여력비용은 기업의 파산확률을 감소시키게 되어 계속기업으로서의 존속가능성을 높이고 회계이익의 지속가능성을 높인다고 보았다. 하지만 너무 과다한 수준으로 자본을 보유하게 되면 높은 지급여력비용을 유지할 수는 있지만 자본비용의 증가와 함께 미래 초과이익의 감소가능성이 있음을 주장하였다.

회계정보(순자산과 순이익)의 상대적 가치관련성과 관련하여 본 연구에서는 자본비용보다는 이익의 지속가능성에 초점을 맞추고자 한다. 지속가능한 이익은 안정적인 영업활동에서 나오는 것이며 보험회사 영업의 안정성은 투자영업부문보다는 보험영업부문에서 나오는 것이다. 왜냐하면, 투자영업이익의 경우 일반적으로 이익조정 가능성이 높고, 이익이 일시적으로 나타나거나 그 변동성이 크다(조석희·박규서 2016)는 특징이 있으며, 투자자산의 손익실현주기는 보험계약의 그것보다 짧아 이익의 지속성이 낮아질 수 있기 때문이다. 보험회사 회계정보의 가치관련성에 관한 선행연구에서는 순이익보다는 순자산의 가치관련성이 높게 나타나고 있다. 이와 관련하여, 김호중 외(1997)에서는 손해보험회사의 회계정보를 이용한 가치평가에서는 주당순이익이나 주가수익비율(PER)의 큰 변동성으로 인하여 이러한 이익지표들의 기업가치 설명력에 한계가 있을 수밖에 없으므로 손해보험회사의 기업가치 분석에서는 주당순자산에 대한 정보가 보다 큰 의미를 가질 수 있다고 주장

하였다. 이러한 논리에 따르는 경우, 보험영업부문과 관련이 깊은 보험리스크나 운영리스크의 상대적 비중이 높은 경우 순이익의 상대적 가치관련성은 높아질 수 있으며,⁶⁾ 신용이나 시장리스크에 대해서는 이와 반대의 결과가 나타날 가능성이 있다. 다만, 금리리스크에 대해서는 보험과 투자부문 모두와 관련이 있기 때문에 상대적 가치관련성의 변화방향에 대해서 예측하기 어려운 면이 있고, 시장리스크의 경우에는 전체 지급여력 요구자본에서 차지하는 비중이 미미하기 때문에 예상된 결과를 확인하지 못할 가능성 또한 존재한다. 이러한 논의를 종합적으로 반영하여 본 논문에서는 다음과 같은 연구가설을 설정한다.

[가설 2] 손해보험회사 회계이익의 상대적 가치관련성은 지급여력 요구자본의 구성내역에 따라 차별화된다.

III. 실증연구 설계

1. 표본의 선정

본 논문의 실증분석대상이 되는 표본은 국내 상장손해보험회사의 2011사업연도 1분기부터 2021사업연도 3분기까지 총 42개 분기의 별도 및 연결재무제표의 자료와 추가자료를 이용하여 구성하였다. 연구대상 기간의 시작점을 2011사업연도부터로 한 이유는 2011사업연도부터 국내 보험회사에 국제회계기준이 적용되기 시작하였는바, 그 이전의 기간까지 연구 표본에 포함하는 경우 회계기준 변경의 효과가 분석결과에 포함됨으로 인하여 발생할 수 있는 분석결과의 왜곡가능성을 제거하기 위함이다. 표본의 추출은 기본적으로 NICE신용평가(주)의 Kis-Value 데이터베이스를 이용하였지만, 동 데이터베이스상의 자료가 불완전한 경우에는 개별 보험회사가 공시한 사업보고서상의 재무자료를 통하여 검증 및 보완하였다.

6) 물론 최근 국내 손보사 영업이익의 평균적인 수준은 보험영업이익보다 투자영업이익이 높은 것이 사실이지만, 회계이익의 가치관련성은 이익의 규모보다는 이익의 지속성이나 변동가능성에 따라 더욱 많은 영향을 받게 된다(백원선·최관 1999; 김정옥 2010; 오창수·조석희 2013; 김경순·이진원 2016). 따라서 보험부문의 비중이 높아질 때 회계이익의 상대적 가치관련성은 높아질 수 있는 것이다.

〈Table 2〉 Sample Selection

Sample selection criteria	Number of firm-quarters
Initial samples with listed non-life insurers for 2011-2021	336
Less : Samples with re-insurers	(42)
Less : Outlier	(12)
Final samples	282

연구대상 기간 중 국내 상장손해보험사의 분기별 총 표본수는 336개이지만, 이중에서 재보험회사의 표본 42개를 제외하였고, 극단치(outlier)의 포함으로 인한 분석결과와 왜곡을 방지하기 위하여 기초 자산총액으로 표준화한 순자산 장부금액 및 순이익이 상하위 1%에 해당하는 표본 역시 제외하였다. 이러한 과정을 통하여 최종적으로 결정된 연구표본은 국내 상장된 7개 보험회사의 총 282개 기업-분기이다.

2. 리스크 특성치의 측정

손해보험회사의 리스크 특성에 따른 회계정보의 가치관련성 차이를 확인하기 위해서는 우선 리스크 특성에 대한 정의가 필요하며 이 정의에 따라 리스크 특성별 해당 특성의 고저를 결정하여야 한다. 본 연구에서의 리스크 특성은 손해보험회사의 지급여력 요구자본을 구성하는 보험, 금리, 신용, 시장 및 운영리스크로 구분하여 파악하기로 한다. 파악된 개별 리스크량이 회사 전체의 지급여력 요구자본에서 차지하는 비율을 해당 손해보험회사의 개별 리스크 특성치로 정의한다. 다만, 손해보험회사의 지급여력 요구자본은 각각의 리스크량을 개별적으로 산출하여 이들 간의 분산효과를 고려하여 최종합산하는 방식을 사용하는바 개별 리스크량의 총합과 회사 전체의 지급여력요구자본과는 차이가 있다. 따라서 본 연구에서는 각각의 개별 리스크가 반영되지 않았을 경우에 산출될 전체 지급여력 요구자본과 원래의 전체 지급여력 요구자본과의 차이가 원래의 전체 지급여력 요구자본에서 차지하는 비율을 해당 리스크 항목에 대한 리스크 특성치로 정의하기로 한다. 예를 들어, 메리츠화재해상보험주식회사(이하, ‘메리츠화재’라 함)의 2017년 1분기 회사 전체의 지급여력 요구자본은 10,724억 원이지만, 개별 리스크에 대한 리스크량은 보험, 금리, 신용, 시장 그리고 운영리스크에 대하여 각각 5,904억 원, 2,327억 원, 4,912억 원, 482억 원

그리고 608억 원이며 이들의 산술적 합계는 14,233억 원으로서 회사 전체의 지급여력 요구자본과 다르다. 왜냐하면 회사 전체의 지급여력 요구자본(RC)은 보험업감독업무시행세칙 제5-7조의 3 및 [별표 22]에 따라 다음과 같이 산출하기 때문이다.

$$\text{회사전체 요구자본(RC)} = \sqrt{\sum_i \sum_j (RA_i \times RA_j) \times Cor.coef_{ij}} + ORA$$

위의 식에서 RA 는 개별 리스크량을 의미하고, ORA 는 운영리스크량을 의미하며, $Cor.coef.$ 는 개별리스크 간 상관관계수를 그리고 i, j 는 보험, 금리, 신용, 시장리스크를 의미한다. 개별 리스크 간의 상관관계수는 보험리스크와 다른 리스크 항목에 대해서는 0.25를 적용하고, 금리, 신용 및 시장리스크 간에는 0.5를 적용하여 산출한다. 이에 따라 메리츠 화재의 2017년 1분기 회사전체의 지급여력 요구자본(TRC)은 다음과 같이 계산된다.

$$\begin{aligned} TRC &= \sqrt{5,904^2 + 5,904 \times 2,327 \times 0.25 + 5,904 \times 4,912 \times 0.25 + 5,904 \times 482 \times 0.25} \\ &\quad + 2,327 \times 5,904 \times 0.25 + 2,327^2 + 2,327 \times 4,912 \times 0.5 + 2,327 \times 482 \times 0.5 \\ &\quad + 4,912 \times 5,904 \times 0.25 + 4,912 \times 2,327 \times 0.5 + 4,912^2 + 4,912 \times 482 \times 0.5 \\ &\quad + 482 \times 5,904 \times 0.25 + 482 \times 2,327 \times 0.5 + 482 \times 4,912 \times 0.5 + 482^2} + 608 \\ &= 10,724 \text{ (억원)} \end{aligned}$$

여기서 보험리스크에 대한 리스크 특성치는 보험리스크가 존재하지 않았을 경우 즉, 보험리스크량이 0(영)인 경우의 지급여력 요구자본(FARC)과 원래 요구자본(TRC)과의 차이가 원래 요구자본에서 차지하는 비율로 계산한다.

$$\begin{aligned} FARC &= \sqrt{0^2 + 0 \times 2,327 \times 0.25 + 0 \times 4,912 \times 0.25 + 0 \times 482 \times 0.25} \\ &\quad + 2,327 \times 0 \times 0.25 + 2,327^2 + 2,327 \times 4,912 \times 0.5 + 2,327 \times 482 \times 0.5 \\ &\quad + 4,912 \times 0 \times 0.25 + 4,912 \times 2,327 \times 0.5 + 4,912^2 + 4,912 \times 482 \times 0.5 \\ &\quad + 482 \times 0 \times 0.25 + 482 \times 2,327 \times 0.5 + 482 \times 4,912 \times 0.5 + 482^2} + 608 \\ &= 7,293 \text{ (억원)} \end{aligned}$$

그리고 보험리스크가 존재하지 않았을 경우의 지급여력 요구자본과 원래 요구자본과의 차이(DARC) 및 보험리스크의 기여도(PARC)는 각각 다음과 같이 계산된다.

$$\bullet \text{ DARC} = \text{TRC} - \text{FARC} = 10,724 \text{억 원} - 7,293 \text{억 원} = 3,431 \text{억 원}$$

$$\bullet \text{ PARC} = \frac{\text{DARC}}{\text{TRC}} = \frac{3,431}{10,724} = 31.99\%$$

이렇게 계산된 개별 리스크 특성치는 동일한 기간(분기)의 다른 보험회사들의 개별 리스크 특성치의 기여도와 비교하여 평균적인 수준보다 높으면 보험리스크 특성치가 높은 표본으로 구분하고, 그렇지 않은 경우에는 보험리스크 특성치가 낮은 표본으로 구분한다. 예를 들어, 2017년 1분기의 국내 상장손해보험회사들의 평균적인 보험리스크 특성치는 24.24%이었지만 같은 분기에 대한 메리츠화재의 보험리스크 특성치는 이보다 높은 31.99%이므로 메리츠화재의 2017년 1분기 표본은 높은 보험리스크 특성치를 갖는 집단(집단 H)으로 분류된다. 이러한 방식으로 동일한 기간에 대한 메리츠화재의 개별 위험특성치를 산출하고 업계의 평균과 비교하여 집단을 구분하면 다음과 같다.

〈Table 3〉 Individual Risk Characteristics

Risk Components	Meritz	Average (Industry)	Group classification
Insurance Risk	31.99%	24.24%	High
Interest Rate Risk	12.24%	17.11%	Low
Credit Risk	28.45%	32.23%	Low
Market Risk	2.40%	0.85%	High
Operating Risk	5.67%	7.02%	Low

국내 보험회사들의 지급여력 요구자본의 산출방식은 지속적인 변화가 있었는데, 본 연구의 연구대상 기간중에도 한 차례의 변화가 있었다. 사실상 위의 요구자본 산출공식은 2014년 4분기부터 적용되는 개별 리스크 간 상관관계에 대한 규정을 적용한 것으로 2014년 3분기까지는 다음과 같은 합산공식을 사용하였다.

$$TRC = \sqrt{ARA^2 + (IRA + CRA)^2 + MRA^2} + ORA$$

위의 식에서 ARA, IRA, CRA, MRA 그리고 ORA는 각각 보험, 금리, 신용, 시장 그리고 운영리스크 요구자본을 의미한다. 이러한 계산식은 금리리스크와 신용리스크의 상관관계를 1로 가정하고 나머지 개별 리스크 간의 상관관계는 없는 것으로 가정하는 방식으로서 본 연구의 2014년 3분기 이전의 표본들에 대한 개별 리스크의 특성치를 계산할 때 사용하였다.

3. 기업가치 평가모형

본 논문에서의 실증분석에는 다음과 같은 두 가지의 검정을 사용하기로 한다.

(검정 1) 회계정보(순자산과 순이익)의 전반적 가치관련성 차이 검정

(검정 2) 순자산과 순이익의 상대적 가치관련성 차이검정

(검정 1)과 (검정 2)는 각각 [가설 1]과 [가설 2]의 검정을 위한 것으로서, (검정 1)은 설정된 회귀모형을 이용하여 추정된 회귀식의 동질성 여부를 Chow(1960) 검정을 통하여 확인하는 것이고, (검정 2)는 모든 개별 리스크 각각에 대하여 리스크 특성치의 고저를 기준으로 구분된 집단 간 회귀계수 차이의 유의성을 확인할 수 있는 평행성 검정을 사용하기로 한다.

한편, 이들 분석을 위해서는 기업가치 평가모형이 필요한데, 본 논문에서는 기본적으로 Ohlson(1995)의 회계변수를 이용한 기업가치 평가모형을 사용하기로 한다. 사용될 평가모형은 다음과 같다.

$$P_{i,t} = \alpha_0 + \alpha_1 BV_{i,t} + \alpha_2 E_{i,t} + e_{i,t} \quad (1)$$

$$P_{i,t} = \beta_0 + \beta_1 BV_{i,t} + e_{i,t} \quad (2)$$

$$P_{i,t} = \gamma_0 + \gamma_2 E_{i,t} + e_{i,t} \quad (3)$$

$$P_{i,t} = \delta_0 + \delta_1 BV_{i,t} + \delta_2 E_{i,t} + \delta_3 NGE_{i,t} + e_{i,t} \quad (4)$$

$$P_{i,t} = \theta_0 + \theta_1 BV_{i,t} + \theta_2 E_{i,t} + \theta_3 NGE_{i,t} + \theta_X \sum QR_t + e_{i,t} \quad (5)$$

여기서, P 는 분기말 시가총액(보통주와 우선주의 합계)을 의미하고, BV 와 E 는 각각 분기말 자기자본 장부금액과 해당분기의 분기순이익을 의미한다. NGE 는 분기순이익이 음(-)인 경우는 분기순이익으로 정의되고 양(+)인 경우에는 영(0)으로 정의되는 손실더미변수를 의미하며 $\sum QR$ 은 분기더미변수를 의미한다. 한편, 분기더미변수를 제외한 모든 변수는 분기초 총자산으로 표준화하여 사용하였다.

식 (1)은 기업의 순자산과 순이익이 기업가치를 설명하는 Ohlson(1995)의 단순모형을 나타낸 것이고, 식 (2)와 식 (3)은 각각 식 (1)에서 순이익 변수와 순자산 변수를 제외하고 구성한 식들이다. 식 (1)내지 (3)을 이용하여 기업가치에 대한 회계정보의 전체설명력(Total Explanation Power; TEP)과 증분설명력(Incremental Explanation Power;

IEP) 그리고 공통설명력(Common Explanation Power; CEP)을 측정하고자 한다.

식 (4)는 식 (1)에 분기순손실의 발생에 따른 효과를 통제하기 위하여 손실 통제변수 NGE를 추가한 것이다. 손실의 발생과 손익정보의 가치관련성 간의 관계에 대해서는 이미 여러 선행연구를 통하여 몇 가지 사실이 밝혀진 바 있다. 가장 기본적인 것은 손실정보와 이익정보의 가치관련성에는 차이가 있다는 것이다. 손실정보에 대해서는 투자자들이 일시적인 현상으로 해석할 수 있어 이익의 지속성에 부정적 평가를 할 수 있고, 이에 따라 순이익보다는 순자산 정보를 기업가치평가에 많이 활용할 수 있게 되는 것이다(Ohlson 1995; Burgstahler and Dichev 1997; Collins, Pincus and Xie 1999). 이러한 효과를 통제하기 위하여 선행연구(Hayn 1995; Collins et al. 1997; Collins et al. 1999 등)에서 일반적으로 사용되고 있는 것처럼 손실이 발생한 경우에만 비대칭적으로 반응하는 통제변수 NGE를 추가한 것이다.

식 (5)는 식 (4)에 각 분기별 더미변수를 추가한 것이다. 이는 본 연구표본이 기업-분기로 이루어진 패널자료에 해당하므로 종속변수(P)의 시계열 자기상관의 존재가능성이 있고 이러한 경우 잔차항의 자기상관이 유발되어 추정된 회귀계수와 그 표준오차에 편의가 있을 수 있다(Beaver et al. 1989; 윤재원·최현돌 2005). 분기 더미변수의 추가는 이러한 현상을 통제하는 효과가 있다.

한편, 국내 보험회사들의 지급여력비율은 2016년 10월부터 연결재무제표를 기준으로 산출되고 있으므로 지급여력 요구자본을 통해 추정되는 리스크 특성이 회계정보의 가치관련성에 미치는 영향을 분석하기 위한 회계정보는 당연히 연결재무제표상의 회계정보이어야 한다. 따라서 본 연구에서도 2016년 4분기 이후의 표본에서는 손해보험회사의 연결재무제표상의 회계정보를 이용하여 분석하기로 한다. 구체적으로 2016년 4분기 이후의 표본부터 자기자본 장부금액은 연결재무상태표상의 연결자본 중 지배기업소유주에게 귀속되는 부분을 사용하고, 분기순이익은 연결분기순이익 중 지배기업소유주귀속분을 사용하여 분석하기로 한다. 국내 상장 손해보험회사들 중에는 연결대상 종속회사가 없는 경우도 있고, 있더라도 비지배지분이 없거나 그 비중이 극히 미미한 경우가 대부분이다. 물론 연결자본 중 지배기업소유주지분 및 연결분기순이익 중 지배기업소유주귀속분 대신 별도재무제표상의 자본총계 및 별도기준 분기순이익을 그대로 이용하여 분석하는 경우에도 큰 차이가 없을 것으로 예상되지만 보다 정확한 분석을 위해서는 연결재무제표상의 정보를 이용하는 것이 바람직하다고 판단된다.

4. 회계정보의 전반적 가치관련성 차이에 대한 분석방법

본 논문에서는 개별 리스크의 특성에 따라 회계정보의 가치관련성에 차이가 발생하는 여부를 검정하기 위하여 식 (4)를 기업가치 평가모형으로 사용하여 개별 리스크 특성치의 고저(高低)에 따라 구분된 집단별로 회귀식을 추정한 후 각각의 리스크별로 대응되는 두 개 집단을 통하여 추정된 회귀식의 동질성을 Chow(1960) 검정을 이용하여 확인하고자 한다.

Chow(1960) 검정은 Gregory Chow에 의해 제안된 방법으로 선형 회귀모형에서 절편과 기울기의 동질성 또는 구조적 단절(structural break)이 존재하는지 여부를 검정하기 위하여 사용되는 방법들 중의 하나이다. 본 연구에서 Chow(1960) 검정을 위한 귀무가설은 다음과 같다.

$$H_0 : \delta_{0,low} = \delta_{0,high}, \delta_{1,low} = \delta_{1,high}, \delta_{2,low} = \delta_{2,high}$$

귀무가설(H_0)에 표현된 δ_{low} 와 δ_{high} 는 식 (4)에서 설정된 회귀계수 δ_0 , δ_1 , δ_2 를 각각 개별 리스크 특성치의 고저(高低)에 따라 구분된 리스크 특성치가 낮은 집단(집단 L)과 높은 집단(집단 H)에 대하여 추정한 것이다. Chow(1960) 검정을 위한 F-통계량은 다음과 같이 산출한다.

$$F-value = \frac{[SSE_{total} - (SSE_{low} + SSE_{high})]/k}{(SSE_{low} + SSE_{high})/(n_1 + n_2 - 2k)}$$

여기서, SSE_{total} , SSE_{low} 그리고 SSE_{high} 은 각각 전체 표본, 개별 리스크 특성치가 낮은 집단(집단 L) 그리고 개별 리스크 특성치가 높은 집단(집단 H)에 대한 잔차제곱합(sum of squares error)을 의미하고, n_1 과 n_2 는 각각 집단 L과 집단 H에 포함된 표본의 수를 의미한다. 그리고 k 는 독립변수의 개수를 의미한다. 이렇게 산출된 $F-value$ 는 분자의 자유도가 k 이고 분모의 자유도가 $n_1 + n_2 - 2k$ 인 F-분포를 따르게 된다.

개별 리스크별로 귀무가설의 기각여부를 검정한 결과, 귀무가설을 기각하는 경우 회계정보의 전반적인 가치관련성은 개별 리스크 특성치의 고저(高低)에 따라 차이가 있음을 의미하게 되어 [가설 1]을 채택할 수 있게 된다.

5. 순자산과 순이익의 상대적 가치관련성 차이에 대한 분석방법

현행 RBC체제상 각각의 리스크 특성에 따라 순이익의 상대적 가치관련성이 차별화되는지 여부를 분석하기 위하여 본 논문에서는 상호보완적인 두 가지 검정모형을 사용하기로 한다. 첫째는 증분설명력의 상호비교방법을 사용하는 것이고 둘째는 평행성 검정(parallelism test)을 사용하는 것이다.⁷⁾

첫 번째 분석방법인 증분설명력의 상호비교방법은 다음과 같다. 특정 회계정보의 증분설명력이라 함으로 해당 정보를 활용하여 가치평가를 하는 경우와 해당 정보를 활용하지 않고 가치평가를 하는 경우 간에 나타나는 가치평가 정확성의 차이로 볼 수 있다. 다만 본 연구에서는 회귀분석의 방법을 사용하는 것이고, 가치평가를 위한 정보로 두 가지 회계정보 즉 순자산 정보와 순이익 정보를 이용하는 것이며, 회귀모형의 설명력은 결정계수(R^2)의 크기로 판단한다. 따라서 순자산의 증분설명력은 순자산과 순이익 정보가 모두 포함된 회귀식 (1)의 설명력(R^2_1 , 전체설명력)에서 순자산 정보가 생략된 회귀식 (3)의 설명력(R^2_3)을 차감하여 측정하고, 순이익의 증분설명력은 전체설명력(R^2_1)에서 순이익 정보가 생략된 회귀식 (2)의 설명력(R^2_2)을 차감하여 측정한다. 한편, 증분설명력 비교연구에서는 일반적으로 공통설명력의 개념을 함께 이용한다. 공통설명력은 전체설명력 중에서 증분설명력의 합으로 설명되지 않는 부분을 말하는 것으로서 구체적으로는 전체설명력에서 순자산과 순이익의 증분설명력 합계를 차감하여 측정한다. 이상에서 설명한 설명력의 종류별 측정방법을 요약하면 다음과 같다.

전체설명력(TEP) = 회귀식 (1)의 결정계수(R^2_1)

순자산의 증분설명력(IEP_{BV}) = $R^2_1 - R^2_3$

7) 다만, 본 연구와 유사한 상황의 회계학 연구에서는 일반적으로 더미변수를 포함한 상호작용 효과 분석방법이 더 많이 사용되기도 한다. 이 방법을 사용하는 경우 회귀식 (5)에 개별 리스크 특성에 대한 더미변수와 등 더미변수와 순자산 및 순이익 간의 상호작용항이 포함된 다중회귀분석을 수행하여 상호작용항의 회귀계수가 유의적인지 여부를 확인하는 방식으로 분석이 진행된다. 그러나 이 방법을 사용하는 경우에는 더미변수와 이를 이용한 상호작용변수의 투입으로 인하여 다중공선성이 심각해질 수 있고, 이렇게 되는 경우 회귀식의 추정결과가 왜곡될 가능성이 존재한다. 따라서 본 연구에서는 평행성 검정을 주된 분석방법으로 이용하고 더미변수를 활용한 상호작용효과 분석은 보조적인 방법으로 활용하고자 한다.

$$\text{순이익의 증분설명력(IEP}_E\text{)} = R^2_1 - R^2_2$$

$$\text{순자산과 순이익의 공통설명력(CEP)} = \text{TEP} - (\text{IEP}_{BV} + \text{IEP}_E)$$

두 번째 분석방법인 평행성 검정은 기본적으로 순자산과 순이익의 상대적 가치관련성에 관한 연구가설 [가설 2]의 채택여부에 대한 통계적 판단을 위한 것이다. 구체적으로 식 (2)를 기업가치 평가모형으로 사용하여 각각의 개별 리스크 특성치의 고저에 따라 구분된 집단별로 회귀식을 추정한 후 각각의 리스크별로 대응되는 두 개 집단을 통하여 추정된 순자산과 순이익의 회귀계수가 유의적인 차이를 보이는지 여부를 검정하는 것이다.

평행성 검정은 동일한 회귀모형에 대하여 서로 다른 집단을 통하여 추정된 회귀식의 회귀계수간에 통계적으로 유의적인 차이가 있는지 여부를 분석하기 위한 검정방법이다. 본 연구에서 평행성 검정을 위한 귀무가설은 다음과 같다.

$$H_0(1) : \theta_{1,low} = \theta_{1,high}$$

$$H_0(2) : \theta_{2,low} = \theta_{2,high}$$

귀무가설(H_0)에 표현된 θ_{low} 와 θ_{high} 는 식 (5)에서 설정된 회귀계수인 θ_1 과 θ_2 를 각각 개별 리스크 특성치의 고저(高低)에 따라 구분된 집단(집단 L, 집단 H)에 대하여 추정한 것이다. 귀무가설 $H_0(1)$ 에 대한 평행성 검정의 검정통계량(t-value)은 다음과 같이 측정한다 (전규안 외 2004; Hollander 1970).

$$t-value = \frac{\hat{\theta}_{1,low} - \hat{\theta}_{1,high}}{\sqrt{Var(\hat{\theta}_{1,low}) + Var(\hat{\theta}_{1,high})}}$$

여기서, $\hat{\theta}_{1,low}$ 과 $\hat{\theta}_{1,high}$ 은 각각 개별 리스크 특성치가 낮은 집단(집단 L)과 높은 집단(집단 H)을 대상으로 추정한 θ_1 이며, θ_1 는 회귀식 (5)로 추정한 순자산의 회귀계수를 의미한다. $Var(\hat{\theta}_{1,low})$ 과 $Var(\hat{\theta}_{1,high})$ 은 각각 $\hat{\theta}_{1,low}$ 과 $\hat{\theta}_{1,high}$ 의 분산을 의미한다.

개별 리스크별로 귀무가설의 기각여부를 검정한 결과, 귀무가설을 기각하는 경우 회계이익의 상대적 가치관련성은 개별 리스크 특성치의 고저(高低)에 따라 차이가 있음을 의미하게 되어 [가설 2]를 채택할 수 있게 된다.

IV. 실증분석결과

1. 기술통계 분석

본 논문의 실증분석에 사용된 주요 변수의 기술통계량은 다음과 같다.

〈Table 4〉 Descriptive Statistics

The number of observations used in the estimation is 282. Sample selection criteria is presented in Table 3, and all variables are defined in equation (1)–(2). PARC refers to the contribution of insurance risks to the total required capital. PIRC refers to the contribution of interest rate risks to the total required capital. PCRC refers to the contribution of credit risks to the total required capital. PMRC refers to the contribution of market risks to the total required capital. PORC refers to the contribution of operating risks to the total required capital. Other variables are defined in equation (1) – (5).

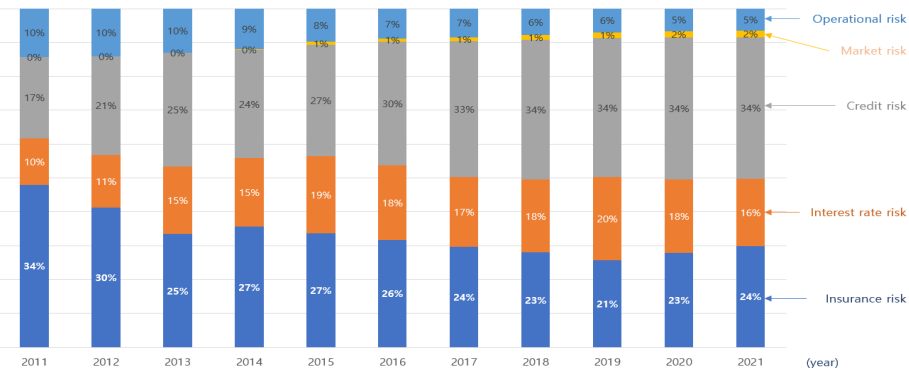
Variables	Mean	Std. dev.	Minimum	Median	Maximum
P	0.0952	0.0655	0.0108	0.0795	0.3118
BV	0.0962	0.0391	0.0422	0.0918	0.2042
E	0.0025	0.0020	-0.0032	0.0025	0.0079
NGE	-0.0001	0.0004	-0.0032	0.0000	0.0000
PARC	0.2612	0.0676	0.1094	0.2516	0.4554
PIRC	0.1616	0.0570	0.0663	0.1473	0.3849
PCRC	0.2809	0.0877	0.0969	0.2772	0.5983
PMRC	0.0074	0.0087	0.0000	0.0043	0.0480
PORC	0.0762	0.0218	0.0444	0.0696	0.1307

〈Table 4〉에 따르면 분기초 자산총액으로 나눈 분기말 시가총액은 평균적으로 9.52%로 나타났고, 분기초 자산총액으로 나눈 자기자본 장부금액은 이보다 약간 높은 수준으로 나타나, 국내 손해보험회사의 MTB 비율(market to book ratio)은 98.96%(=0.0952/0.0962)인 것으로 나타났다. 분기순이익의 평균적인 수준은 기초 자산총액 대비 0.25%이고, 이는 분기말 자기자본의 2.60%(= 0.0025/0.0962)로 나타나 선행연구와 유사하게 순이익의 비중이 자기자본에 비하여 매우 낮은 수준으로 나타났다.

한편 본 논문에서 중요한 변수라 할 수 있는 개별 리스크의 리스크 특성치는 리스크 항

목에 따라 큰 차이가 나타나고 있다. 평균적인 특성치가 가장 높은 리스크 항목은 신용리스크로서 전체 요구자본의 28.09%를 차지하고 있고, 그 다음으로는 보험리스크가 전체 요구자본의 26.12%를 차지하고 있다. 시장리스크의 리스크 특성치는 평균적으로 1% 수준에도 미치지 못하여 회사 전체의 요구자본에 미치는 영향이 중요하지 않은 것으로 나타났다. 그러나 국내 감독제도의 개편으로 인하여 연구대상 기간 중 지급여력 요구자본의 산출방식이 수차례 변경되었기 때문에 모든 시점을 통합하여 개별 리스크 항목의 리스크 특성치를 살펴보는 것은 그 의미가 제한적일 수 있다. 따라서 연도별 개별 리스크 특성치의 변화추이를 살펴보면 <Figure 1>과 같다.

<Figure 1> Trend of the Proportion of Individual Risk Component



<Figure 1>에 따르면, 개별 리스크 항목 중 그 기여도가 가장 큰 신용리스크의 경우 시간이 갈수록 기여도가 상승하고 있으며, 보험리스크의 경우는 그 기여도 점점 감소하다가 다시 증가하는 현상을 보이고 있다. 금리리스크의 경우는 그 비중이 다소 증가하다가 최근 들어 다시 줄어드는 추세에 있고, 시장리스크의 경우는 다소 증가하는 추세로 보이긴 하지만 보험회사의 지급여력 요구자본에서 차지하는 비중이 미미하다.⁸⁾ 따라서 보험회사 회계정보의 가치관련성은 시장리스크의 고저에 따라 차별화될 가능성은 높지 않을 것으로 예상되고, 그 중

8) 이태기(2021)에 따르면, 최근 손해보험회사의 보험리스크는 보험가격리스크를 중심으로 그 절대적 금액이 증가하였고, 신용리스크 역시 개별 손보사들의 수익성 중시성향이 높아짐에 따라 금액증가가 이어지고 있다. 이와 함께 손보사들의 장기상품 판매증가와 함께 금리리스크 역시 증가하였지만, 그 절대적인 수준이 보험 및 신용리스크에 비해 낮은 수준이므로, 개별 리스크의 상대적 비중을 나타내는 <Figure 1>에서는 신용리스크 비중의 변화는 거의 없고, 보험리스크의 비중은 다시 높아지고 있으며, 금리리스크의 비중은 다시 낮아지는 것으로 나타나고 있다.

요성이 높은 보험과 금리 및 신용리스크 특성치의 고저에 따라 영향을 받을 가능성이 높다. 그러나 앞서 논의한 바와 같이 금리리스크의 경우에는 보험회사의 투자영업부문과 보험영업 부문 모두와 관련성이 높다는 점에서 가치관련성의 차별화를 유발할 가능성이 높지 않고, 운영리스크의 경우에는 그 상대적 중요성이 낮지만 리스크량의 산출방식을 고려할 때 보험리스크와 그 변화방향이 유사할 것으로 판단되는바 가치관련성의 차별화가 예상된다.

2. 상관관계 분석

본 논문의 실증분석에 사용되는 주요 변수들의 피어슨 상관관계는 <Table 5>와 같다.

<Table 5> Pearson Correlation Matrix of Variables

The number of observations used in the estimation is 282. All variables are defined in equation (1)-(5), and the text in the parenthesis represents the p-value of the statistics.

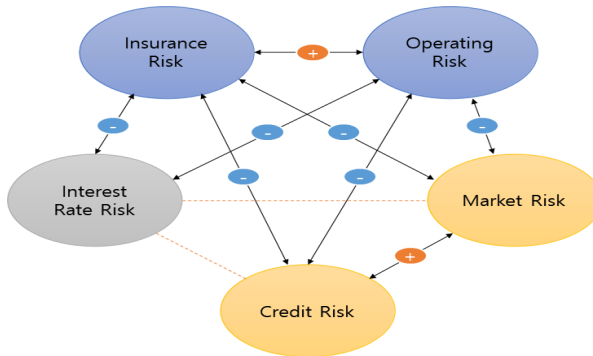
	P	BV	E	NGE	PARC	PIRC	PCRC	PMRC	PORC
P	1.000								
BV	0.834 (<.0001)	1.000							
E	0.581 (<.0001)	0.481 (<.0001)	1.000						
NGE	0.168 (0.005)	0.179 (0.003)	0.488 (<.0001)	1.000					
PARC	0.058 (0.328)	-0.141 (0.018)	0.222 (0.001)	0.045 (0.456)	1.000				
PIRC	-0.296 (<.0001)	-0.083 (0.163)	-0.293 (<.0001)	-0.058 (0.334)	-0.434 (<.0001)	1.000			
PCRC	-0.009 (0.875)	0.215 (0.001)	-0.057 (0.343)	0.075 (0.210)	-0.811 (<.0001)	0.064 (0.286)	1.000		
PMRC	-0.038 (0.521)	0.261 (<.0001)	0.100 (0.095)	0.127 (0.033)	-0.297 (<.0001)	0.068 (0.258)	0.508 (<.0001)	1.000	
PORC	0.031 (0.609)	-0.328 (<.0001)	-0.113 (0.058)	-0.180 (0.002)	0.324 (<.0001)	-0.434 (<.0001)	-0.487 (<.0001)	-0.641 (<.0001)	1.000

〈Table 5〉에 따르면, 기업가치(P)는 순자산(BV)과 가장 높은 상관관계를 보이고 있으며 순이익(E)과도 유의적인 상관관계가 있는 것으로 나타났다. 이와 같은 결과는 제조업 등 비금융업을 대상으로 하는 일반적인 회계학 연구의 결과와는 다소 차이가 있는 것이지만, 보험산업에 대한 선행연구의 결과와는 대체로 일치하는 결과이다. 그러나 본 논문에서는 이러한 보험업에서의 회계이익 정보가 낮은 가치관련성을 가지는 것이 보험업 전반에 걸친 일반적인 현상인지 아니면 보험회사의 리스크 특성에 따라 차별화 될 수 있는 것인지를 확인하려고 한다.

한편, 보험회사의 리스크 특성을 나타내는 변수들 중 금리리스크 특성치(PIRC)의 경우에는 기업가치와 유의적인 음(-)의 상관관계를 가지나, 이를 제외한 나머지 리스크 특성치(PARC, PCRC, PMRC, PORC)는 기업가치와 유의적인 상관관계를 가지지 않는 것으로 나타났다. 본 연구의 목적은 이러한 리스크 특성치들의 기업가치 설명력을 확인하려 하는 것이 아니라 회계정보 특히 회계이익의 가치관련성에 영향을 미치는지 여부를 확인하는 것이므로 가설의 검정은 다중회귀분석을 기초로 수행되는 평행성 검정과 Chow(1960) 검정을 통해 살펴보기로 한다.

또한, 개별 리스크 특성치들 간의 상관관계를 살펴보면, 보험리스크는 운영리스크와 유의적인 정(+)의 상관관계가 있고, 금리, 신용 그리고 시장리스크 특성치와는 유의적인 부(-)의 상관관계를 가지는 것으로 나타났다. 반면에 금리리스크의 경우에는 보험 및 운영리스크와 유의적인 부(-)의 상관관계가 있지만, 신용 및 시장리스크와는 유의적인 상관성이 없는 것으로 나타났다. 그리고 신용리스크의 경우에는 시장리스크와 유의적인 정(+)의 상관관계를 가지는 것으로 나타났다.

〈Figure 2〉 Correlation of Individual Risk Component



이러한 개별 리스크 특성치들의 상관관계를 종합하면, 본 논문의 제2장에서 언급한 개별 리스크 항목의 경제적 특성과 일관된 측면이 있는 것을 알 수 있다. 왜냐하면, 보험 및 운영리스크는 모두 보험영업과 관련된 리스크를 반영하지만 신용 및 시장리스크는 투자영업과 관련된 리스크를 반영하기 때문에 보험과 운영리스크 상호간 그리고 신용 및 시장리스크 상호간에는 정(+)의 상관관계가 있지만 보험 및 운영리스크와 신용 및 시장리스크 특성치와는 부(-)의 상관관계가 있기 때문이다. 다만, 투자 및 보험영업의 리스크를 모두 반영하는 금리리스크의 경우에는 투자부문의 리스크(신용 및 시장리스크)와 상관관계가 없어 예상한 바와 다르지 않지만, 보험부문의 리스크(보험 및 운영리스크)와는 부(-)의 상관관계가 있다는 점은 예상과 다르게 나타났다는 측면에서 향후의 분석결과를 통하여 보다 자세히 살펴볼 필요가 있을 것이다.

3. 리스크 특성에 따른 회계정보의 전반적 가치관련성 변화

손해보험회사의 개별 리스크 특성치에 따른 회계정보의 전반적 가치관련성은 〈Table 6〉의 회귀분석 결과 및 이를 기초한 Chow(1960) 검정을 통해 분석한다.

〈Table 6〉 Results of Chow Test

Estimation model: $P_{i,t} = \delta_0 + \delta_1 BV_{i,t} + \delta_2 E_{i,t} + \delta_3 NGE_{i,t} + e_{i,t}$

All variables are defined in equation (4). ***, ** and * indicate the significance at 1%, 5% and 10% level, respectively.

Panel A. Insurance Risk Characteristics

Variables	Coefficient (t-statistics)				Adj. R ²	Chow's F (p-value)
	Intercept	BV _{i,t}	E _{i,t}	NGE _{i,t}		
Low	-0.051*** (-7.53)	1.274*** (17.60)	8.838*** (4.35)	-20.619** (-2.00)	0.821	3.193*** (0.014)
High	-0.005 (-0.36)	0.690*** (4.77)	10.877*** (6.43)	-13.734* (-1.73)	0.433	

Panel B. Interest Rate Risk Characteristics

Variables	Coefficient (t-statistics)				Adj. R ²	Chow's F (p-value)
	Intercept	BV _{i,t}	E _{i,t}	NGE _{i,t}		
Low	-0.025** (-3.31)	1.048*** (13.58)	8.378*** (5.52)	-18.833* (-1.87)	0.661	3.164*** (0.015)
High	-0.062*** (-7.71)	1.294*** (14.52)	10.888*** (4.68)	-20.041** (-2.28)	0.802	

Panel C. Credit Risk Characteristics

Variables	Coefficient (t-statistics)				Adj. R ²	Chow's F (p-value)
	Intercept	BV _{i,t}	E _{i,t}	NGE _{i,t}		
Low	-0.016* (-1.77)	0.775*** (6.52)	10.744*** (7.36)	-16.218*** (-2.68)	0.526	3.838*** (0.005)
High	-0.053*** (-4.88)	1.242*** (13.45)	11.135*** (4.28)	-44.736* (-1.95)	0.779	

Panel D. Market Risk Characteristics

Variables	Coefficient (t-statistics)				Adj. R ²	Chow's F (p-value)
	Intercept	BV _{i,t}	E _{i,t}	NGE _{i,t}		
Low	-0.052*** (-7.72)	1.264*** (15.73)	9.456*** (6.43)	-17.896*** (-2.95)	0.721	0.550 (0.699)
High	-0.037*** (-3.03)	1.135*** (11.35)	9.421*** (3.77)	-30.096 (-1.05)	0.686	

Panel E. Operating Risk Characteristics

Variables	Coefficient (t-statistics)				Adj. R ²	Chow's F (p-value)
	Intercept	BV _{i,t}	E _{i,t}	NGE _{i,t}		
Low	-0.056** (-6.53)	1.296*** (17.23)	9.395*** (5.46)	-17.929 (-1.16)	0.740	4.782*** (0.001)
High	0.001 (0.11)	0.510*** (3.60)	10.842*** (5.93)	-16.906*** (-2.66)	0.392	

분석결과, 시장리스크를 제외한 모든 리스크 항목에서 리스크 특성치의 고저에 따라 회계정보의 전반적 가치관련성에 차이가 발생하고 있다. 보험리스크와 운영리스크의 경우 리스크 특성치가 높은 집단이 낮은 집단보다 수정 결정계수(Adj. R²)로 측정한 회계정보의 설명력이 더 낮게 나타났으나, 금리리스크와 신용리스크 항목에 대해서는 이와 반대로 나타났다.

보험부문의 리스크를 나타내는 보험 및 운영리스크의 상대적 중요성이 증가하는 경우 회계정보의 전반적 가치관련성은 낮게 나타났는데, 이는 가설설정 단계에서의 예상과 일치하는 결과이다. 다만, 투자부문의 리스크를 나타내는 신용리스크와 시장리스크에 대해서는 신용리스크의 상대적 중요성이 증가하는 경우에만 회계정보의 전반적 가치관련성이 증가하였고, 시장리스크의 고저에 대해서는 전반적 가치관련성이 차별화되지 않았다. 이러한 결과는 시장리스크가 전체 지급여력 요구자본에서 차지하는 비중이 미미하기 때문에 나타난 결과로 추측된다.

금리리스크의 경우에는 보험 및 투자영업부문의 리스크가 모두 포함되어 있기 때문에 가설설정 단계에서는 회계정보의 전반적 가치관련성의 변화방향을 예측할 수 없었다. 그러나 분석결과에 따르면 신용리스크 특성치가 높은 집단의 전반적 가치관련성이 더 높게 나타났고, 이는 투자부문의 리스크 특성치에 대한 예상과 동일한 것이다. 이러한 결과는 금리리스크가 투자부문의 리스크와 성격이 유사하기 때문인 것으로 해석하기 보다는 금리리스크 특성치가 보험 및 운영리스크와 유의한 부(-)의 상관관계를 가지기 때문인 것으로 해석하는 것이 바람직하다고 판단된다. 왜냐하면, 앞서 수행된 상관관계 분석의 결과에서 금리리스크 특성치는 신용 및 시장리스크 특성치와 유의적인 상관관계가 없었으나 보험 및 운영리스크 특성치와는 유의적인 부(-)의 상관관계가 있는 것으로 나타났기 때문이다.

4. 리스크 특성에 따른 순자산과 순이익의 증분적 및 상대적 가치관련성

〈Table 7〉은 개별 리스크 특성치의 고저에 따라 순자산과 순이익의 증분적 가치관련성이 어떻게 변화되는지를 분석한 결과이다.

〈Table 7〉 Explanation Power Based on the Risk Characteristics

This table reports R square measures obtained from the following linear model:

$$P_{i,t} = \alpha_0 + \alpha_1 BV_{i,t} + \alpha_2 E_{i,t} + e_{i,t} \quad (1)$$

$$P_{i,t} = \beta_0 + \beta_1 BV_{i,t} + e_{i,t} \quad (2)$$

$$P_{i,t} = \gamma_0 + \gamma_1 E_{i,t} + e_{i,t} \quad (3)$$

TEP and CEP denotes total and common explanation power of BV and E. IEP(BV) and IEP(E) denotes increment explanation power of BV and E. Each variables are measured as follows:

TEP = R square of equation (1)

IEP(BV) = R square of equation (1) - R square of equation (3)

IEP(E) = R square of equation (1) - R square of equation (2)

CEP = TEP-IEP(BV)-IEP(E)

Panel A. Insurance risk characteristics

Variables		EP(BV)	EP(E)	EP(CO)	EP(TO)	BV vs E
Low	R square	0.3554	0.0198	0.4464	0.8215	BV ≫ E
	Proportion	43%	2%	54%	100%	
High	R square	0.0728	0.1967	0.1938	0.4632	BV ≪ E
	Proportion	16%	42%	42%	100%	

Panel B. Interest risk characteristics

Variables		EP(BV)	EP(E)	EP(CO)	EP(TO)	BV vs E
Low	R square	0.3371	0.0657	0.2686	0.6713	BV ≫ E
	Proportion	50%	10%	40%	100%	
High	R square	0.3746	0.0310	0.3949	0.8005	BV ≫ E
	Proportion	47%	4%	49%	100%	

Panel C. Credit risk characteristics

Variables		EP(BV)	EP(E)	EP(CO)	EP(TO)	BV vs E
Low	R square	0.1026	0.1464	0.2871	0.5361	BV < E
	Proportion	19%	27%	54%	100%	
High	R square	0.3568	0.0365	0.3895	0.7827	BV ≫ E
	Proportion	46%	5%	50%	100%	

Panel D. Market risk characteristics

Variables		EP(BV)	EP(E)	EP(CO)	EP(TO)	BV vs E
Low	R square	0.2870	0.0666	0.3669	0.7205	BV ≫ E
	Proportion	40%	9%	51%	100%	
High	R square	0.3506	0.0459	0.3033	0.6998	BV ≫ E
	Proportion	50%	7%	43%	100%	

Panel E. Operational risk characteristics

Variables		EP(BV)	EP(E)	EP(CO)	EP(TO)	BV vs E
Low	R square	0.4217	0.0578	0.2757	0.7552	BV ≫ E
	Proportion	56%	8%	37%	100%	
High	R square	0.0689	0.1638	0.1477	0.3804	BV ≪ E
	Proportion	18%	43%	39%	100%	

분석결과, 보험부문의 리스크를 나타내는 보험리스크 및 운영리스크 특성치가 높은 경우에는 낮은 경우에 비하여 회계정보의 전체설명력이 큰 폭으로 감소하는 것으로 나타났고, 금리 및 신용리스크에 대해서는 이와 반대로 나타났으며, 시장리스크의 경우에는 리스크 특성의 고저로 구분한 집단 간에 큰 차이가 없었다. 이와 같은 결과는 [가설 1]의 검정과 관련된 Chow(1960)검정의 결과와도 일관된다.

개별 리스크별로 그 특성치의 고저에 따라 순자산과 순이익의 상대적 가치관련성에 차이가 있는지를 살펴보면 다음과 같다. 보험리스크의 경우 리스크 특성치가 낮은 집단의 경우 순자산의 증분설명력이 43%이었으나 리스크 특성치가 높은 집단의 경우에는 16%에 불과한 것으로 나타났고, 순이익의 증분설명력은 이와 반대로 리스크 특성치가 낮은 집단에서는 2%에 불과하였으나, 리스크 특성치가 높은 집단에서는 42% 수준으로 나타나 리스크 특성치의 변화에 따라 상대적 가치관련성에 차이가 있었다. 금리리스크의 경우에는 리스크 특성치가 낮은 집단의 경우 순자산의 증분설명력이 50%이었으나 리스크 특성치가 높은 집단의 경우에는 47%로 소폭 감소하였고, 순이익의 증분설명력 역시 리스크 특성치가 낮은 집단에서는 10%이었으나, 리스크 특성치가 높은 집단에서는 4% 수준으로 나타나 상대적 가치관련성에 차이가 있는지 유무를 판단할 수 없다. 신용리스크의 경우에는 리스크 특성치가 낮은 집단의 경우 순자산의 증분설명력이 19%이었으나 리스크 특성치가 높은 집단의 경우에는 46% 수준으로서 큰 차이가 있었고, 순이익의 증분설명력은 이와 반대로 리스크 특성치가 낮은 집단에서는 27% 수준이었으나, 리스크 특성치가 높은 집단에서는 5%에 불과하여 리스크 특성치의 변화에 따라 상대적 가치관련성에 차이가 있었다. 시장리스크의 경우에는 리스크 특성치가 낮은 집단의 경우 순자산의 증분설명력이 40%이었으나 리스크 특성치가 높은 집단의 경우에는 50%로 증가하였고, 순이익의 증분설명력은 리스크 특성치가 낮은 집단에서는 9%이었으나, 리스크 특성치가 높은 집단에서는 7%로서

소폭의 감소되긴 하였으나 그 차이가 미미하여 차이의 유의성이 의심되는 상황이다. 마지막으로 운영리스크의 경우에는 보험리스크와 유사한 형태로 나타났는데, 리스크 특성치가 낮은 집단의 경우 순자산의 증분설명력이 56%이었으나 리스크 특성치가 높은 집단의 경우에는 18%에 불과한 것으로 나타났고, 순이익의 증분설명력은 이와 반대로 리스크 특성치가 낮은 집단에서는 8%에 불과하였으나, 리스크 특성치가 높은 집단에서는 43% 수준으로 나타나 리스크 특성치의 변화에 따라 상대적 가치관련성에 차이가 있었다.

분석결과를 요약하면, 보험부문의 리스크를 반영하는 보험 및 운영리스크의 경우 그 특성치가 낮은 집단에서는 순자산의 가치관련성이 상대적으로 높지만 특성치가 높아지는 경우 순이익의 가치관련성이 높아지고 있다. 그리고 투자부문의 리스크를 반영하는 신용 및 시장리스크의 경우에는 이와 반대로 리스크 특성치가 낮은 집단에서는 순이익의 가치관련성이 상대적으로 높지만 리스크 특성치가 높아지는 경우 순자산의 가치관련성이 높아지고 있다. 다만, 여기서 시장리스크의 경우 순자산 및 순이익의 증분설명력 차이가 미미하여 통계적 유의성에는 의문이 있다. 그리고 금리리스크의 경우에는 리스크 특성치의 증가에 따라 순자산 및 순이익 모두의 증분설명력이 감소하는 것으로 나타나 상대적 가치관련성에 변화가 있다고 볼 수 없다.

하지만, 증분설명력을 이용한 순자산과 순이익의 상대적 설명력(가치관련성)에 대한 해석은 어디까지나 대략적인 증감의 방향을 확인하는 것에 불과한 것으로서 그 통계적 유의성을 단정할 수는 없다. 따라서 본 연구에서는 개별 리스크별 특성치의 고저로 구분한 집단 간에 순자산과 순이익의 주가반응계수(식 (5)에서 순자산과 순이익의 회귀계수)에 차이가 있는지 있다면 그 차이에 통계적 유의성이 있는지 여부를 평행성 검정의 방법으로 확인해야 한다.

주가반응계수 및 평행성 검정의 결과는 <Table 8>와 같다.

〈Table 8〉 Parallelism Test

Estimation model: $P_{i,t} = \theta_0 + \theta_1 BV_{i,t} + \theta_2 E_{i,t} + \theta_3 NGE_{i,t} + \theta_X \sum QR_t + e_{i,t}$

All variables are defined in equation (5). ***, ** and * indicate the significance at 1%, 5% and 10% level, respectively.

Panel A: Insurance risk characteristics

Variables	Low – Insurance		High – Insurance		Parallelism test
	coefficient	t-statistics	coefficient	t-statistics	
Intercept	-0.0606	-5.29***	-0.0504	-5.03***	-
BV	1.2784	22.38***	1.0759	10.04***	-1.67
E	5.0780	2.67***	11.3099	8.47***	2.68***
NGE	1.6471	0.20	-12.4470	-2.45***	-
QT dummy	Included				-
F-value	37.78***		16.34***		
Adj.R ²	0.9150		0.8322		
Max(VIF)	3.7254		3.1558		

Panel B: Interest rate risk characteristics

Variables	Low – Interest		High – Interest		Parallelism test
	coefficient	t-value	coefficient	t-value	
Intercept	-0.0526	-5.80***	-0.0464	-3.44***	-
BV	1.1209	19.85***	1.2785	17.54***	1.71
E	8.4223	6.72***	8.5109	3.96***	0.04
NGE	-11.2942	-1.51	-3.2623	-0.45	-
QT dummy	Included				-
F-value	25.74***		27.74***		
Adj.R ²	0.8643		0.9105		
Max(VIF)	2.5657		3.6957		

Panel C: Credit risk characteristics

Variables	Low – Credit		High – Credit		Parallelism test
	coefficient	t-value	coefficient	t-value	
Intercept	-0.0411	-4.88***	-0.0746	-4.02***	-
BV	1.0157	10.98***	1.2641	18.93***	2.18**
E	10.1911	8.00***	4.1334	1.57	-2.08**
NGE	-7.7609	-1.74*	0.8617	0.04	-
QT dummy	Included				-
F-value	17.91***		25.04***		
Adj.R ²	0.8034		0.9102		
Max(VIF)	3.2964		4.2321		

Panel D: Market risk characteristics

Variables	Low – Market		High – Market		Parallelism test
	coefficient	t-value	coefficient	t-value	
Intercept	-0.0426	-4.96***	-0.0908	-6.51***	-
BV	1.2237	19.80***	1.3182	19.42***	1.03
E	7.8857	6.23***	7.3077	3.15***	-0.22
NGE	-5.9129	-1.28	-9.1473	-0.50***	-
QT dummy	Included				-
F-value	29.96***		25.23***		
Adj.R ²	0.8756		0.9026		
Max(VIF)	2.7909		4.3521		

Panel E: Operating Risk Characteristics

Variables	Low – Operational		High – Operational		Parallelism test
	coefficient	t-value	coefficient	t-value	
Intercept	-0.0642	-5.63***	-0.0254	-2.13***	-
BV	1.2631	24.81***	0.8858	7.22***	-2.84***
E	7.4695	5.21***	9.1654	5.44***	0.77
NGE	-4.1967	-0.4015	-7.1303	-1.29	-
QT dummy	Included				-
F-value	37.21***		6.57***		
Adj.R ²	0.9026		0.6815		
Max(VIF)	2.5196		2.8599		

〈Table 8〉은 개별 리스크 특성치의 고저를 기준으로 구분된 두 집단에 포함된 표본을 이용하여 다중회귀분석의 방법으로 앞서 정의된 식 (5)를 추정하고, 추정된 순자산과 순이익의 주가 반응계수(회귀계수)가 두 집단 간에 서로 유의적으로 다른지 여부를 검정한 결과이다. 분석결과 모든 회귀식의 추정과정에서 분산팽창계수의 최대값이 5 이하로 나타나 회귀계수의 왜곡을 초래할 수 있는 다중공선성의 문제는 없는 것으로 나타났으며, 모형의 F-값 역시 충분한 수준으로 모든 회귀식이 적합하게 추정된 것으로 판단된다.

평행성 검정⁹⁾의 결과, 보험리스크 특성치의 고저로 구분한 두 집단 간 순자산의 반응계

9) 동일한 회귀모형이 서로 다른 표본집단을 통해 추정될 때, 추정된 회귀계수의 동질성 또는 차이의 유사성을 검증하기 위한 연구방법으로서 집단에 따라 구분되는 더미변수를 추가하고 이 더미변수와 관찰대상 변수와의 상호작용항의 회귀계수의 유의성을 확인하는 방법을 사용할 수도 있으므로 실제로 본 연구데이터를 이용하여 다음과 같은 더미변수를 이용한 분석을 수행한 결과 모든 회귀식의 추정과정에서 분산팽창계수의 최대값이 10을 초과하여 다중공선성이 문제로 회귀분석 결과의 왜곡가능성이 높음을 발견하였다.

수(회귀계수)에는 유의적인 차이가 없었으나 순이익의 반응계수는 리스크 특성치가 높은 집단에서 유의적으로 더 크게 나타났다. 이러한 결과는 증분설명력의 비교결과와 일관된 결과로서 보험리스크 특성치가 높아질수록 순이익의 상대적 가치관련성이 높아진다는 것을 나타낸다. 그러나 금리리스크에 대해서는 그 특성치의 고저에 따라 순자산 및 순이익 모두의 반응계수에 유의적인 차이가 없었다. 신용리스크의 경우에는 리스크 특성치가 높아질수록 순자산의 반응계수는 유의적으로 커지고 순이익의 반응계수는 유의적으로 낮아지는 것으로 나타났다. 이 역시 증분설명력의 비교결과와 일관된 결과로서 신용리스크 특성치가 높아질수록 순이익(순자산)의 상대적 가치관련성이 낮아(높아)진다는 것을 나타낸다. 한편 시장리스크에 대해서는 가설 설정단계에서의 예상과 증분설명력의 비교결과와 일관되게 순자산 및 순이익의 가치관련성 차이가 유의적이지 않았다. 마지막으로 운영리스크에 대해서는 보험리스크의 경우와 유사하게 리스크 특성치가 높아짐에 따라 순자산의 반응계수가 유의적으로 낮게 나타나 순이익의 상대적 가치관련성은 높아지는 것으로 해석할 수 있다.

순자산과 순이익의 상대적 가치관련성에 대한 증분설명력의 비교분석과 평행성 검정결과를 종합하면, 손해보험회사의 보험부문 리스크(보험 및 운영리스크)가 커질수록 순이익의 상대적 가치관련성은 높아지는 것으로 전체 기업가치를 순이익이 설명하는 비중이 더 높아지게 된다. 그러나 투자부문 리스크(신용 및 시장리스크)가 커질수록 순이익보다는 순자산의 기업가치 설명력이 높아지는 것으로 분석되었으나 시장리스크의 경우에는 보험산업에서의 중요성이 크지 않아 상대적 가치관련성의 유의적인 차이를 보이지는 않았으며, 금리리스크의 경우에도 순자산 및 순이익의 설명력 차이가 유의적으로 나타나지 않았다.

$$P_{i,t} = \eta_0 + \eta_1 BV_{i,t} + \eta_2 E_{i,t} + \eta_3 NGE_{i,t} + \eta_4 DA_{i,t} + \eta_5 DA_{i,t} * BV_{i,t} + \eta_6 DA_{i,t} * E_{i,t} + \eta_X \sum QR_t + e_{i,t}$$

$$P_{i,t} = \kappa_0 + \kappa_1 BV_{i,t} + \kappa_2 E_{i,t} + \kappa_3 NGE_{i,t} + \kappa_4 DI_{i,t} + \kappa_5 DI_{i,t} * BV_{i,t} + \kappa_6 DI_{i,t} * E_{i,t} + \kappa_X \sum QR_t + e_{i,t}$$

$$P_{i,t} = \lambda_0 + \lambda_1 BV_{i,t} + \lambda_2 E_{i,t} + \lambda_3 NGE_{i,t} + \lambda_4 DC_{i,t} + \lambda_5 DC_{i,t} * BV_{i,t} + \lambda_6 DC_{i,t} * E_{i,t} + \lambda_X \sum QR_t + e_{i,t}$$

$$P_{i,t} = \mu_0 + \mu_1 BV_{i,t} + \mu_2 E_{i,t} + \mu_3 NGE_{i,t} + \mu_4 DM_{i,t} + \mu_5 DM_{i,t} * BV_{i,t} + \mu_6 DM_{i,t} * E_{i,t} + \mu_X \sum QR_t + e_{i,t}$$

$$P_{i,t} = \nu_0 + \nu_1 BV_{i,t} + \nu_2 E_{i,t} + \nu_3 NGE_{i,t} + \nu_4 DO_{i,t} + \nu_5 DO_{i,t} * BV_{i,t} + \nu_6 DO_{i,t} * E_{i,t} + \nu_X \sum QR_t + e_{i,t}$$

여기서, $DA_{i,t}$, $DI_{i,t}$, $DC_{i,t}$, $DM_{i,t}$, $DO_{i,t}$ 는 각각 기업 i 의 t 분기말 보험, 금리, 신용, 시장 그리고 운영리스크 특성 더미변수로서 보험(금리, 신용, 시장 그리고 운영)리스크 기여도가 업계평균보다 높으면 1, 그렇지 않으면 0(영)의 값을 가진다. 기타 다른 변수들의 정의는 식 (1) 및 식 (2)와 동일하다.

5. 추가 분석

손해보험회사의 리스크 특성에 따른 보험회사 회계정보의 전반적 가치관련성과 회계이익의 상대적 가치관련성 차이에 대한 앞선 분석은 개별 리스크 특성의 수준을 평균값을 중심으로 구분하여 분석한 결과이다. 그러나 이러한 방법으로 리스크 특성의 높낮이를 구분하는 것은 다소 자의적인 측면이 있다. 따라서 본 추가분석에서는 개별 리스크 특성치가 상위 25%에 해당하는 집단과 하위 25%에 해당하는 집단으로 구분하여 양 집단 간의 차이를 살펴봄으로서 분석결과의 강건성을 높이고자 한다. <Table 9>는 이와 같은 분석방법을 이용하여 리스크 특성에 따른 회계정보의 전반적 가치관련성 변화에 대한 분석결과이다.

<Table 9> Results of Additional Chow Test

Estimation model: $P_{i,t} = \delta_0 + \delta_1 BV_{i,t} + \delta_2 E_{i,t} + \delta_3 NGE_{i,t} + e_{i,t}$

All variables are defined in equation (4). ***, ** and * indicate the significance at 1%, 5% and 10% level, respectively.

Risk Characteristics	Adj. R ²		Chow's F (p-value)
	Low	High	
Insurance Risk	0.963	0.236	8.076*** (<0.001)
Interest rate Risk	0.754	0.900	2.702** (0.038)
Credit Risk	0.473	0.971	8.326*** (<0.001)
Market Risk	0.559	0.644	1.525 (0.204)
Operating Risk	0.648	0.386	4.156*** (0.005)

분석결과, 시장리스크를 제외한 모든 리스크 특성치의 높낮이로 구분한 집단 간 차이가 1% 또는 5% 수준에서 유의적인 것으로 나타났다. 보험리스크와 운영리스크의 경우 리스크 특성치가 높은 집단이 낮은 집단보다 회계정보의 설명력이 더 낮게 나타났고, 금리리스크와 신용리스크에 대해서는 이와 반대로 나타나 <Table 6>에 제시된 분석결과와 일관성을 보였다.

한편, 리스크 특성에 따른 순자산과 순이익의 상대적 가치관련성의 변화여부에 대한 추가분석 결과는 <Table 10>과 같다.

〈Table 10〉 Additional Parallelism Test

Estimation model: $P_{i,t} = \theta_0 + \theta_1 BV_{i,t} + \theta_2 E_{i,t} + \theta_3 NGE_{i,t} + e_{i,t}$

All variables are defined in equation (5). ***, ** and * indicate the significance at 1%, 5% and 10% level, respectively.

Risk Characteristics	Variables	Standardized coefficient		Parallelism test
		Low	High	
Insurance Risk	BV	0.9107***	0.3010**	-4.54***
	E	0.1056*	0.3970***	0.36
Interest rate Risk	BV	0.7074***	0.7068***	0.39
	E	0.2460**	0.4087***	2.39**
Credit Risk	BV	0.3028**	0.9828***	4.12***
	E	0.7928***	0.0095	-3.99***
Market Risk	BV	0.5974***	0.7595***	2.18**
	E	0.3736**	0.2575**	0.63
Operating Risk	BV	0.6746***	-0.2947**	-5.67***
	E	0.3146***	0.7338***	-0.30

〈Table 10〉은 개별 리스크의 평균값을 구분으로 구분한 집단 간 차이분석 결과를 나타내는 〈Table 8〉의 분석결과에 대한 강건성을 검증하기 위한 추가분석 결과를 나타낸다.

추가분석결과, 신용리스크와 운영리스크 특성치의 차이는 순자산 및 순이익 모두에서 〈Table 8〉에서의 분석결과와 일치하였다. 즉, 신용리스크 특성치가 높아질수록 순자산의 반응계수는 유의적으로 증가하는 반면 순이익의 회귀계수는 유의적으로 감소하였고, 운영리스크 특성치가 높아질수록 순자산의 반응계수는 유의적으로 낮아진 반면 순이익의 회귀계수는 유의적인 차이를 보이지 않았다. 이러한 결과는 신용리스크의 증가에 따라 순이익의 상대적 가치관련성은 감소하고, 운영리스크의 증가에 따라 순이익의 가치관련성은 증가한다는 주장을 뒷받침하는 것이므로 〈Table 8〉에 제시된 분석결과와 일관된 것이다.

한편, 보험리스크의 경우 리스크 특성치의 증가에 따라 순자산에 대한 반응계수의 차이는 유의적으로 감소한 반면 순이익에 대한 반응계수의 차이는 유의적이지 않았다. 이와 같은 결과를 해석하면 다음과 같다. 순자산과 순이익 각각에 대한 반응계수 차이의 유의성 측면에서는 〈Table 8〉에서의 분석결과와 다른 것이지만, 순이익의 상대적 가치관련성 측면에서는 동일한 의미를 내포하고 있다. 왜냐하면 순이익의 상대적 가치관련성 증가(감소)는 순자산 반응계수에 유의적인 변화가 없는 상황에서 순이익 반응계수가 유의적으로 증

가(감소)하는 경우에도 발생하지만, 순이익 반응계수에 유의적인 변화가 없는 상황에서 순자산 반응계수가 유의적으로 감소(증가)하는 경우에도 발생하기 때문이다.

또한 금리리스크와 시장리스크의 경우에는 <Table 8>에 제시된 분석결과를 통해서도 리스크 특성치가 높아짐에 따라 순이익의 가치관련성에 유의적인 변화가 없었으나, 추가 분석을 통해 <Table 10>에 제시된 결과에 따르면 금리리스크 특성치의 증가에 따라 순이익의 상대적 가치관련성은 증가하였고, 시장리스크 특성치가 높아짐에 따라 순이익(순자산)의 상대적 가치관련성은 감소(증가)하였다. 이와 같은 결과는 금리리스크의 특성이 보험부문과 투자부문의 특성을 모두 반영한다는 점과 손해보험회사의 전체 리스크량에 비하여 시장리스크량의 규모가 미미하다는 점에서 기인하는 결과로 해석된다.

V. 결론

본 연구에서는 국내 상장 손해보험회사의 회계정보의 품질을 가치관련성의 관점에서 살펴보았다. 구체적으로 회계정보의 전반적인 가치관련성 및 순자산과 순이익의 상대적 가치관련성이 어떤 상황에서 높아지거나 낮아질 수 있는지에 대하여 정보위험과의 관련성을 기초로 검토한 후 연구가설을 설정하고 이를 실증적으로 검증하였다.

본 연구에서는 보험회사 회계정보의 가치관련성을 변화시킬 수 있는 요인으로 보험회사의 부문별 활동을 반영하여 산출되는 지급여력 요구자본의 구성내역을 고려하였다. 현행 감독체계상 지급여력 요구자본은 다섯 가지 개별 리스크에 대해 리스크량을 산출하고 이의 상관관계 등을 고려하여 합산하여 회사 전체의 요구자본을 산출하는데, 본 연구에서는 이에 착안하여 개별 리스크 항목의 상대적 중요성을 의미하는 리스크 특성치를 고안하고, 각 개별 리스크별로 리스크 특성치의 고저에 따라 회계정보의 가치관련성이 전반적으로 변화하거나 순자산과 순이익이 기업가치 평가과정에서 기여하는 정도(상대적 가치관련성)가 달라질 것이라는 가설을 검정하였다.

가설의 검정결과, 주로 보험회사의 보험부문 리스크에 해당하는 보험리스크와 운영리스크 특성치가 커질수록 회계정보의 전반적 가치관련성은 낮아지지만, 순이익의 상대적 가

치관련성은 증가하는 현상을 발견하였다. 그리고 투자부문의 리스크 항목 중 그 상대적 중요성이 큰 신용리스크의 경우 리스크 특성치가 커질수록 보험부문 리스크와 반대되는 현상 즉 리스크 특성치가 커질수록 회계정보의 전반적 가치관련성은 높아지지만 이 과정에서 순이익의 상대적 가치관련성은 오히려 낮아지는 것으로 관찰되었다.

그간 국내 보험산업에 대한 연구결과에서는 보험산업의 이익품질이 낮은 것으로 보고되고 있으나 아직 이에 대한 명확한 진단은 없는 상황이다. 이러한 상황에서 본 논문의 결과는 보험산업의 낮은 이익품질에 대한 기초적인 진단과 원인분석에 참고 가능한 정보를 제공할 것으로 판단된다. 또한 본 연구결과는 회계이론적인 관점에서 이미 정립된 정보위험과 가치관련성 간의 관계를 보험회사의 영업구조와 연관시켰다는 점에서 남다른 의미를 가진다. 특히 보험회사의 영업부문(보험부문과 투자부문)의 구성내역을 리스크량으로 계량화한 지표를 기준으로 구분하고, 이를 기초로 회계정보의 전반적인 가치관련성 및 회계이익(순이익)의 상대적인 가치관련성이 차별적으로 나타날 수 있음을 보였다는 점에서 공헌점이 있다.

참고문헌

금융감독원 (2020), **(개정) 보험회사 RBC제도 해설서**.

(Translated in English) Korean Financial Supervisory Service (2020).
(Revised) Guide for RBC system for the insurers.

김경순·이진훤 (2016), “이익의 질이 회계정보를 이용한 가치평가에 미치는 차별적 효과: 이익특성변수들 간의 비교”, **회계저널**, 제25권 제2호, pp. 1-43.

(Translated in English) Kim, K., and J., Lee (2016). “The Effect of Earnings Quality on the Value Relevance of Accounting Information: A Comparison among the Earnings Attributes Variables”, *Korean accounting journal*, Vol. 25(2):1-43.

김정옥 (2010), “부채와 회계정보의 가치관련성: 금융기관차입금과 일반부채를 중심으로”, **회계학연구**, 제35권 제1호, pp. 63-93.

(Translated in English) Kim, J. (2010). “The Effect of Debt Characteristics on Value Relevance”, *Korean accounting review*, 35(1):63-93.

김호중·이석영·손광기 (1997), “국제 비교연구를 통한 우리나라 손해보험회계의 문제점 및 개선방안”, **회계저널**, 제6권 제2호, pp. 287-311.

(Translated in English) Kim, H., S., Lee and K., Son (1997). “Problems and Improvement of Korean Non-life insurance Accounting by International Comparison”, *Korean accounting journal*, 6(2): 287-311.

박경원·장지인 (2012), “연구논문: 생명보험사 자기자본제도가 유가증권 분류 회계선택에 미치는 영향”, **회계저널**, 제21권 제4호, pp. 163-183.

(Translated in English) Park, K., and J., Jang (2012). “The effect of capital regulation of life insurance companies on accounting choice about classification of securities”, *Korean accounting journal*, 21(4):163-183.

백원선·최관 (1999), “이익지속성이 순자산과 회계이익에 대한 주가배수에 미치는 영향”, **회계학연구**, 제24권 제4호, pp. 61-83.

(Translated in English) Paek, W., and K., Choi (1999). “Effects of Persistence of Abnormal Earnings on Relative Pricing Multiples on Book Value of Equity and Earnings”, *Korean Accounting Review*, 24(4):61-83.

오창수·변재웅 (2015), “손해보험회사 지급준비금 적립형태에 관한 실증연구”, **보험금융연구**, 제26권 제2호, pp. 33-58.

(Translated in English) Ouh, C., and J., Byon (2015). “A study on discretionary accounting for loss reserves of non-life insurance companies”, *Journal of Insurance and Finance*, 26(2):33-58.

오창수·조석희 (2013), “지급여력비율 수준에 따른 회계정보의 가치관련성에 관한 연구-손해보험산업을 중심으로”, **리스크 관리연구**, 제24권 제2호, pp. 41-67.

(Translated in English) Ouh, C., and S., Cho (2013). “A study on the value relevance of accounting information based on the level of solvency margin ratio - Focusing on the non-life insurance industry”, *The Journal of Risk Management*, 24(2):41-67.

윤재원·최현돌 (2005), “외환위기와 은행회계정보의 주가관련성 변화에 관한 연구”, **회계학연구**, 제30권 제1호, pp. 35-63.

(Translated in English) Yoon, J., and H., Choi (2005). “The 1997 Financial Crisis and the Changes in Value Relevance of Accounting Information”, *Korean Accounting Review*, 30(1):35-63.

이기영·이성관·정소영·허태영 (2017), “차우검정을 활용한 선형회귀모형간 유사성 검증” **한국 ITS 학회 논문지**, 제16권 제2호, pp. 73-82.

(Translated in English) Lee, K. et al. (2017). “A Stability Test of the Regression Coefficients for the Linear Models using Chow Test”, *The Korea Institute of Intelligent Transport Systems*, 16(2), 2:73-82.

이원아·심현우 (2018), “손해보험회사 경영정보의 가치관련성에 관한 실증연구”, **리스크 관리연구**, 제29권 제3호, pp. 107-138.

(Translated in English) Lee, W., and H., Shim (2018). “An Empirical Study on the Value Relevance of Non-life Insurer’s Management Information”, *The Journal of Risk Management*, 29(3):107-138.

이태기 (2021), “위험기준 자기자본제도(RBC)-운영경과 및 변화”, 보험연구원·리스크 관리학회 공동 세미나 자료.

(Translated in English) Lee, T. (2021). “Risk Based Capital - Operation Results and Change”, Korea Insurance Research Institute & Korea Insurance Research Institute Seminar Notes.

장지인·이경주·김봉현 (2010), “기업의 자본구조에 따른 회계이익과 장부가액의 가치관련성: 자기자본기업과 부채기업의 비교”, **유라시아연구**, 제7권 제3호, pp. 65-79.

(Translated in English) Jang, J., K., Lee and B., Kim (2010). “The Effect of the Firm’s Capital Structure on the Value Relevance of Earnings and Equity Book Value: All-equity versus Levered Firms”, *The Journal of Eurasian Studies*, 7(3):65-79.

전규안·김재준·오용락 (2004), “원가효율성의 가치관련성에 관한 연구”, **회계학연구**, 제29권 제3호, pp. 167-197.

(Translated in English) Jeon, K., J., Kim and O., O (2004). “A Study on the Value Relevance of Cost Efficiency”, *Korean Accounting Review*, Vol. 29(3):167-197.

정요섭 (1999), “생명보험 지급능력제도의 발전방향”, **리스크 관리연구**, 제10권 단일호, pp. 197-232.

(Translated in English) Chung, Y. (1999). “Suggestions for Improving the Current Solvency Systems of Life Insurance Companies”, *The Journal of Risk Management*, 10:197-232.

조석희 (2016), “손해보험회사의 지급여력비율과 경영효율성의 가치관련성에 관한 연구”,
보험금융연구, 제27권 제3호, pp. 117-154.

(Translated in English) Cho, S. (2016). “A Study on Value Relevance of Solvency Margin Ratio and Operational Efficiency in the Non-life Insurance Companies”, *Journal of Insurance and Finance*, 27(3):117-154.

조석희·박규서 (2016), “보험영업손익과 투자영업손익의 연관성에 관한 실증연구”,
Journal of The Korean Data Analysis Society, 제18권 제3호, pp. 1509-1521.

(Translated in English) Cho, S., and K., Park (2016). “A Study on the Relation between Insurance and Investment Operating Income”, *Journal of The Korean Data Analysis Society*, 18(3):1509-1521.

Barth, M., H., William and R., Wayne (1998). “Relative valuation roles of equity book value and net income as a function of financial health”, *Journal of Accounting and Economics*, 25.1:1-34.

Beaver, W., C., Eger, S., Ryan and M., Wolfson (1989). “Financial reporting, supplemental disclosures, and bank share prices”, *Journal of Accounting Research*, 27(2):157-178.

Burgstahler, D., and I., Dichev (1997). “Earnings, adaptation, and equity value”, *The Accounting Review*, 72(April):187-215.

Chow G. (1960). “Tests of equality between sets of coefficients in two linear regressions”, *Econometrica: Journal of the Econometric Society*, 28(3):591-605.

Collins, D., E., Maydew and I., Weiss (1997). “Changes in the Value-Relevance of Earnings and Book Values over the Past Forty Years”, *Journal of Accounting and Economics*, 24:39-67.

- Collins, D., M., Pincus and H., Xie (1999). "Equity valuation and negative earnings: The role of book value of equity", *The Accounting Review*, 74(January):29-61.
- Easley, D., and M., O'hara (2004). "Information and the cost of capital", *The journal of finance*, 59(4):1553-1583.
- Francis, J., R., LaFond, P., Olsson, and K., Schipper (2004). "Costs of equity and earnings attributes", *The accounting review*, 79(4):967-1010.
- Hayn, C. (1995). "The information content of losses", *Journal of accounting and economics*, 20(2):125-153.
- Hollander, M. (1970). "A Distribution-Free Test for Parallelism, *Journal of the American Statistical Association*, 65:387-394.
- Hwang, L., and S., Lim (2012). "Do investors price accruals quality? A reexamination in the implied cost of equity capital", *Asia-Pacific Journal of Financial Studies*, 41(4):458-490.
- Ohlson, J. (1995). "Earnings, book values, and dividends in equity valuation", *Contemporary Accounting Research*, 11(Spring):661-687.

Abstract

This study examines whether value relevance of insurance accounting information differs according to risk characteristics of insurance companies. The analysis results are as follows.

First, the higher insurance or operating risk proportion, the lower overall value relevance of accounting information, while the higher interest rate or credit risk proportion, the higher overall value relevance of accounting information.

Second, the higher insurance or operating risk proportion, the higher relative value relevance of net income, while the higher credit risk proportion, the lower relative value relevance of net income.

This study contributes to the literature on insurance accounting information quality by exploring conditions affecting overall and relative value relevance of insurance accounting information.

※ Key words: Risk Characteristics, RBC System, Required Capital, Value Relevance

국민건강보험 빅데이터를 활용한 치아보험 요율 차등화에 관한 연구: 치아우식증, 치주질환을 중심으로*

A Study on Differentiation of Dental Insurance Rates Using National Health Insurance Service Big Data: Periodontal Caries and Disease

전 희 주**·최 경 진***

Heuiju Chun·Kyung-jin Choi

최근 보험회사는 국민건강보험 빅데이터를 활용하여 고위험 집단 대상 보험상품 개발 및 신규사업 확대를 모색하고 있다. 이에 본 연구는 치아우식증, 치주질환을 보장하는 치아보험에 대해 국민건강보험 코호트2.0 DB를 이용하여 치주질환 치료의 관찰 기간을 기준년도(2007년) 직전 5년에서 1년으로 완화할 경우 기존 가입대상 집단(무병자A)대비 완화된 기준을 적용한 가입대상 집단(무병자B)의 상대위험도를 산출하였다. 일반화선형모형(GLM)을 이용하여 상대위험도를 산정한 결과 기존집단인 무병자A 대비 고위험 집단인 무병자B의 상대위험도는 치아우식과 치주질환에 대한 충전치료에서는 각각 1.0026과 1.0049로 매우 낮은 반면, 치아우식과 치주질환에 대한 보철치료의 상대위험도는 각각 1.053과 1.103으로 충전치료에 비해 높았다. 결국 치주질환 치료 관찰 기간을 5년에서 1년으로 줄여 치아보험의 가입대상을 확대할 경우 충전치료 보다는 보철치료의 보험료 할증이 높다는 점을 보여준다.

국문 색인어: 국민건강보험, 빅데이터, 상대위험도, 일반화선형모형, 치아보험

한국연구재단 분류 연구분야 코드: B051603, B051608

* 본 연구는 한국보험계리사회 2021 계리연구지원사업의 지원을 받아 연구되었음

** 동덕여자대학교 정보통계학과 부교수(hjchun@dongduk.ac.kr), 제1저자

*** 성균관대학교 수학과(보험계리전공) 강사/주택금융연구원 연구위원(choikj23@naver.com), 교신저자

논문 투고일: 2022. 2. 24, 논문 최종 수정일: 2022. 7. 22, 논문 게재 확정일: 2022. 8. 18

I. 서론

치아는 인간생존에 필수인 음식 섭취에 직접적으로 관여한다. 치아의 저작기능은 소화 과정의 첫 단계일 뿐만 아니라 위장 건강, 기억력, 면역력 등에도 관여하는 것으로 알려져 있다. 만약, 치아가 부실해 제대로 씹지 못한다면 우리 신체 여러 기능에 문제가 발생할 수 있다.

구강건강을 위협하는 대표적인 질환은 치아우식증¹⁾과 치주질환²⁾으로 생애 전반에 걸쳐 발생하는 다빈도 질병이다. 2020년 치과 외래 다빈도 질병결과를 살펴보면 ‘치은염 및 치주질환’(잇몸질환)이 1,627만 명(31.4%)으로 가장 많으며, ‘치아우식’(충치) 613만 명(11.8%), ‘치수 및 근단주위조직의 질환’(치수염) 389만 명(7.5%)의 순이었다. 또한, 연령대별 다빈도 질병을 살펴보면, 19세 이하에서는 ‘치아우식’(26.3%), ‘치아의 발육 및 맹출 장애’(20.0%), 20~50대는 ‘치은염 및 치주질환’(35.0%), ‘치아우식’(9.1%), 60대 이상은 ‘치은염 및 치주질환’(36.1%), ‘치수 및 근단주위조직의 질환’(10.1%)이었다.³⁾

구강질환은 치아 관리에 익숙하지 않은 어린이뿐만 아니라 스트레스로 인한 치주질환을 호소하는 청년 및 중·장년층, 당뇨, 고혈압, 골다공증과 같은 만성질환에 의해 치아상실 위험에 노출되어 있는 고령층에 이르기까지 전 세대에 걸쳐 발생하는 질환임을 알 수 있다. 또한, 구강질환은 만성적이고 누적되는 특성으로 인해 치과진료비는 총 진료비에서 차지하는 비중이 크고 매년 꾸준히 증가하는 추세에 있다. 실제, 2017년 기준 1인당 평균 총 진료비(치과포함)는 69만 원, 치과 총 진료비는 52만 원이며 치과진료비는 2011년부터 2016년까지 약 49%가 증가하였다(치과의료정책연구원 2020).

최근, 건강보험 보장성 확대로 치과진료비에 대한 본인부담금은 낮아지고 있으나 2019년 기준 건강보험 보장률 살펴보면, 전체 보장률은 64.2%인데 비해 치과의원은 36.9%,

-
- 1) 치아우식증은 구강 내 *Streptococcus mutans*, *Streptococcus sobrinus*와 같은 세균에 의해 발생되며(Loesche 1986), 치면세균막 내 세균의 대사산물인 젖산 등의 유기산 생성에 의해 치아의 법랑질과 상아질이 탈회되는 과정으로 치아를 파괴시키며, 구강건강에 대한 인식 부족, 구강관리 습관 등으로 인해 빈번하게 발생한다.
 - 2) 치주질환은 치아표면에 형성되는 치면세균막과 이에 대한 숙주의 반응에 의해 치은 출혈과 치은퇴축, 치주낭 형성 및 치아를 지지하는 결합 조직과 지지골의 치주조직을 파괴하여 치아 상실을 유발하는 만성 염증성 질환으로 성인에서 치아를 상실하는 주된 원인이다(Schatzle et al. 2004).
 - 3) 건강보험심사평가원 보도자료(2021), “건강보험치과 다빈도질병 및 진료행위 분석결과”.

치과병원은 27.4%에 불과하여 치과진료에 대한 의료비 부담은 여전하다.⁴⁾ 따라서 치과진료비에 대한 경제적 부담 완화를 위해 치아보험에 대한 소비자의 관심과 수요는 높은 상황이다. 그러나 최근 실손의료보험의 손해율 및 적자규모 증가로 인해 지난 10년간 12곳의 중소형 보험회사가 실손의료보험 판매를 중단하였다. 가장 최근에 출시된 4세대 실손의료보험⁵⁾을 판매하고 있는 보험회사도 2022년 1월 현재 생보사 5곳, 손보사 15곳에 불과한 상황으로 소비자의 선택권 위축이 우려되고 있다.⁶⁾ 보장성 보험에 해당하는 실손의료보험의 급격한 손해율 상승에도 불구하고 보험회사들은 새로운 국제회계기준(IFRS 17) 도입을 앞두고 있어 금리에 민감하지 않은 보장성 보험 판매 확대의 필요성은 여전히 높은 상황이다. 기존 실손의료보험 상품의 판매를 확대하기는 어렵지만 최근 충치 및 치주환자 증가로 인해 수요가 높고 소액보험 형태로 판매가 가능한 치아보험이 보험사의 새로운 틈새시장으로 부상하고 있다. 수익성 측면에서 치아보험은 언더라이팅 강화, 보험금 감액설계, 보험금 지급에 대한 면책기간 설정 등을 통해 손해율 관리가 가능한 장점이 있다. 실제로 출시 초기시점인 2008년 치아보험의 손해율은 180% 수준이었으나 점차 낮아져 한 때 50~60% 수준까지 하락하기도 하였다.⁷⁾

또한, 장기적으로 보험회사는 저출산 고령화로 인한 보험수요 변화에 대응하기 위해 빅데이터 활용방안을 모색하고 있다. 실제로 국민건강보험 빅데이터와 같은 공공의료 데이터를 이용한 정교한 위험분석을 통해 기존에 보장하지 않았던 질환을 보장하는 신규상품 개발 또는 기존상품의 가입대상 확대를 추진하고 있다. 그러나 공공의료 데이터 제공 시 개인정보 유출 우려 등으로 보험회사들은 신규상품 개발을 위한 데이터 확보에 난항을 겪고 있는 상황이다.

4) 국민건강보험 보도자료(2020), “2019년도 건강보험환자 진료비 실태조사”.

5) 실손의료보험은 판매 시기 및 보장한도 등에 따라 1세대(2009년 10월 이전), 2세대(2009년 10월~2017년 3월), 3세대(2017년 4월~2021년 6월), 4세대(2021년 7월 이후)로 구분하고 있다. 4세대 실손의료보험은 필수치료인 급여(주계약)에 대해서는 보장을 확대하되 환자의 선택사항인비급여(특약)에 대해서는 의료이용에 따라 보험료가 할인·할증 비급여 이용량에 따라 보험료가 상승하는 차등제를 도입한 실손보험을 말한다.

6) 이코노미스트(2021. 12. 18), “10년간 생보사 12곳 실손 판매중단 남은 5곳 “포기못한다””, 기사 참고.

7) 서울경제(2018. 3. 5), “치아보험 경쟁 불붙은 이유”, 기사 참고.

조선일보(2022. 4. 5), “충치·치주 환자 늘었다....치열해지는 치아보험 시장”, 기사 참고.

이러한 관점에서 본 연구는 국민건강보험공단에서 제공하는 국민건강 코호트2.0 DB 빅데이터를 이용하여 현재 판매되고 있는 치아우식증과 치주질환 관련 보험상품의 기대위험률을 일반화선형모형을 통해 산출하고, 새로운 가입조건을 허용할 경우 기존 가입집단 대비 신규 가입이 가능한 집단의 상대위험도와 보험료 할증 유무를 제안하고자 한다.

II. 이론적 배경 및 선행연구

일반적으로 보험회사는 장래 담보위험의 손실액(보험금)을 예측한 결과를 바탕으로 보험상품의 위험보험료를 산정한다. 만약, 담보위험에 대한 손실 분포를 정확히 알고 있다면 개인별 위험을 반영한 보험료 부과가 가능하지만, 이를 적용하기에는 보험회사가 보유하고 있는 데이터의 제약이 존재할 뿐만 아니라 가입자의 개별위험을 반영한 보험료를 부과하더라도 보험회사 관리측면의 어려움이 존재할 수 있다. 따라서 연령 및 성별 집단으로 구분하여 대수의 법칙에 의한 기대 손실액을 추정하여 표준화된 위험보험료를 적용하는 것이 일반적이다.

그러나 현실적으로 기대손실액은 실제로 발생한 손실액과 다르게 나타난다. 만약, 기대손실액 보다 실제발생 손실액이 커질 경우 보험회사의 손해액(울)이 증가하는 보험리스크가 발생한다. 실제와 기대 손실액의 차이에 대한 공통적 원인은 가입자와 보험회사 간 정보 비대칭에 의한 역선택과 도덕적 해이를 들 수 있다.

먼저, 역선택은 가입자의 개별 위험을 반영한 보험료를 적용하지 않으므로 예상과 달리 고위험군에 속하는 소비자가 보험에 가입하는 현상이다. 반면, 저위험군에 해당하는 소비자는 자신의 담보위험 수준에 비해 부과되는 보험료가 비싸다고 생각하여 가입하지 않게 되는 것이다. 역선택은 단순한 이론에 그치지 않고 국내외 다수 문헌에서도 실증된 바 있다(김대환·이봉주 2013; Akerlof 1970; Rothschild and Stiglitz 1976; Einav and Finkelstein 2011).

다음으로 도덕적 해이는 정보 비대칭으로 인해 보험 가입자 또는 이해관계자가 법적 또는 윤리적으로 최선의 의무를 다하지 않는 행위를 말한다. 실손의료보험에서는 도덕적 해

이를 사전과 사후적 측면에서 구분하고 있다.

사전적 도덕적 해이는 의료보험 가입자가 의료보험을 통한 의료비 보장으로 인해 의료 보험 비가입자에 비해 건강유지 노력을 게을리 하는 행위를 의미한다(Arrow 1963). 반면, 사후적 도덕적 해이는 보험가입자가 보험가입을 통해 부담하는 실질 의료서비스 가격이 하락함으로 인해 의료서비스 이용을 늘리는 행위를 말한다. 즉, 주어진 건강상태에서 의료 서비스 가격 변화에 따른 소비자 수요의 반응이다(Pauly 1968; Cutler and Zeckhauser 2000). 이러한 의미에서 사후적 도덕적 해이는 숨겨진 행동에 대한 일종의 남용이라고 볼 수 있다(Einav and Finkelstein 2018). 실손의료보험의 사후적 도덕적 해이는 가입자의 직접 행위뿐만 아니라 의료기관의 비급여 항목에 대한 과잉진료에도 원인이 있다. 실손의료보험의 사후적 도덕적 해이의 결과는 보험회사의 손해액 상승에 의한 보험료 인상으로 이어졌다. 실제로 지난 4년간(2017~2020년) 실손의료보험의 보험료는 연 평균 13.4% 증가하였으며 보험회사가 적자를 해소하기 위해서는 10년간 연 평균 19.3%의 보험료 인상이 필요한 것으로 나타났다.⁸⁾ 결국, 사후적 도덕적 해이로 인한 보험료 인상은 기존 선량한 가입자의 이탈과 함께 장기적으로 잠재 수요자의 보험가입 유인을 저하시켜 역선택을 더욱 심화시킬 수 있다. 이처럼 도덕적 해이와 역선택은 정보 비대칭에 의한 현상이라는 공통점뿐만 아니라 상호 연관성도 존재한다.

그 밖에 실손의료보험은 의료제도 변화 및 의료기술 발달에 따른 치료비 증가로 인한 손해를 증가 가능성도 존재하지만(조재린·정성희 2018) 실손의료보험의 지속가능성 확보를 위해서는 근본적으로 역선택 해소가 필요하다.

실손의료보험의 역선택 해소를 위한 방안으로 보험료의 차등화 또는 세분화 적용이 필요하다는 주장이 꾸준히 제기되어왔다(이경아·이항석 2016; 이항석 외 2017; 조재린·정성희 2018). 실손의료보험의 보험료 차등화 또는 세분화를 위한 통계적 모형은 다양하나 대부분의 실손의료보험 상품이 단기 갱신형 위주의 손해보험 상품임을 감안하면, 가장 널리 활용될 수 있는 모형은 일반화선형모형(Generalized Linear Model; GLM)이다.

일반적으로 GLM을 이용한 보험료 산정에는 계약자별 인적 및 계약관련 요인(Factors)

8) 조선일보(2021. 12. 15), “실손보험료, 매년 19.3%씩 올리지 않으면 적자...현 추세면 2031년 누적적자 122조 원”, 기사 참고.

의 선형결합이 체계적 성분에 해당하며 반응변수인 랜덤성분은 사고빈도(Claim Frequency) 또는 사고심도(Claim Severity)이다. 또한, 체계적 성분과 랜덤성분의 연결 함수는 로그함수 또는 지수족 분포가 주로 사용된다. 실무적으로 심도와 빈도는 독립으로 가정할 때, GLM 방식에 의해 추정된 계약자별 빈도와 심도의 곱을 통해 계약자의 특성을 반영한 최종 순 보험료가 도출된다.

해외에서는 이미 GLM을 이용한 보험료 산정방법이 여러 차례 소개되었다(Kaas et al. 2008; De Jong and Heller 2008; Ohlsson and Johansson 2010; Goldburd et al. 2019; Denuit et al. 2019). 우리나라도 이항석 외(2017)와 이항석·백혜연(2018)에 의해 GLM과 Generalized Linear Mixed Model(GLMM)⁹⁾을 이용하여 실손의료보험의 요율 차등화를 위한 실증분석이 이루어져 왔다.

그러나 기존 선행연구는 특정 보험회사의 실적자료에 근거하여 추정하였다는 한계가 존재한다. 또한, 실손의료보험에서 담보하는 질병종류별 발생 및 심도가 다름에도 이를 구분하지 못한 한계도 존재한다. 이에 본 연구는 건강보험 코호트 DB를 이용하여 치아우식에 대한 충전치료, 치주질환에 대한 충전치료, 치아우식에 대한 보철치료, 치주질환에 대한 보철치료 4개의 기존 치아보험 상품군인 표준집단 대비 가입기준 완화 시 가입이 가능한 집단의 상대위험도를 GLM모형을 적용하여 최초로 산정할 수 있음을 보이고자 한다.

III. 연구 방법

1. 연구 데이터

본 연구는 국민건강보험공단에서 제공하는 표본코호트 2.0DB를 사용하였다. 표본코호트 DB는 2006년 기준 1년간 건강보험 가입자 및 의료급여 수급권자 자격을 유지한 전 국민을 대상으로 무작위 추출된 100만 명(모집단의 2%)의 2002년부터 2015년까지의 출생 및 사망, 건강보험 자격유지 및 보험료 납부, 의료이용 현황 정보를 포함하고 있다. 본 연

9) GLMM은 GLM의 확장모형으로 고정효과에 해당하는 GLM 설명변수에 임의효과를 추가한 모형이다.

구는 표본코호트 2.0DB의 정보 중 출생 및 사망, 건강보험 자격유지 및 진료명세서, 치아우식증 또는 치주질환과 같은 구강진료내역 정보를 활용하였다.

분석대상자는 무작위로 추출된 표본 데이터 중 만 30세~79세(2007년 기준)인 자이다. 다만, 추적 관찰의 정확도를 위해 관찰기간(2002년~2005년)동안 사망 등으로 인한 중도탈퇴자와 2002~2005년 진료내역 정보가 존재하지 않는 의료급여 수급권자의 표본은 분석대상에서 제외하였다.

2. 연구설계 및 모형

가. 연구설계

먼저, 치아우식증 또는 치주질환의 치료 이력 확인을 위해 관찰기간(2002년~2015년)은 2007년을 기준으로 사전 관찰기간(2002년~2006년)과 사후 관찰기간(2007년~2015년)으로 구분하였다.

다음으로 분석대상자를 사전 관찰기간 동안 치아우식증 및 치주질환 치료 여부에 따라 유병자와 무병자 집단으로 분류하였다. 세부적으로 유병자와 무병자 집단은 치주질환 치료의 사전 관찰기간에 따라 유병자A와 무병자A, 유병자B와 무병자B로 구분된다. 유병자A와 무병자A는 기존 상품에서 적용되는 분류기준으로서 기준년도(2007년) 직전 1년간 치아우식증 또는 직전 5년간 치주질환 치료가 존재할 경우 유병자A로 분류된다. 반면, 유병자B는 유병자A 분류기준 중 치주질환 치료여부에 대한 사전 관찰기간을 기준연도 직전 5년에서 1년으로 단축한 경우이다. 결국, 무병자B 분류기준은 현재 치아보험 가입대상인 무병자A에 비해 완화된 기준으로 볼 수 있다. 또한, 유병자A 집단과 무병자A 집단의 중복성은 존재하지 않으며 마찬가지로 유병자B 집단과 무병자B 집단 역시 집단 간 중복성은 존재하지 않는다.

- 유병자A: 기준년도(2007년) 직전 1년(2006년)간 치아우식증 치료 또는 기준년도 직전 5년간(2002~2006년) 치주질환 치료가 있는 경우
- 무병자A: 기준년도(2007년) 직전 1년(2006년)간 치아우식증 치료 및 기준년도 직전

5년간(2002~2006년) 치주질환 치료가 없는 경우

- 유병자B: 기준년도(2007년) 직전 1년(2006년)간 치아우식증 또는 치주질환 치료가 있는 경우
- 무병자B: 기준년도(2007년) 직전 1년(2006년)간 치아우식증 및 치주질환 치료가 없는 경우

본 연구는 상기와 같이 설정된 4개 집단에 대해 각각 성별, 5세 연령대별로 구분하고 사후 관찰기간 치아우식증과 치주질환 원인별 충전 및 보철치료(발치술 포함)에 대한 평균 발생빈도 차이를 살펴보기 위한 코호트 분석을 수행하였다.

여기서 발생빈도 산정과 관련하여 보철치료의 경우 보철치료에 해당하는 수가코드 이외에 임플란트, 브릿지, 틀니 치료와 기본적으로 동반되는 발치술(자연치 발치)에 해당하는 수가코드를 포함하였다. 또한, 연 단위 기준질병 또는 행위 종류별 동일 인원이 여러 번 치료받은 경우에는 치료 횟수를 모두 빈도수에 반영하였다. 마지막으로 충전 및 보철치료는 치료 목적을 명확히 구분하기 위해 건강보험 급여 치료에 한정하였다.

나. 통계적 모형

본 연구의 관심대상인 반응변수는 사후 관찰기간(2007년~2015년도) 동안 발생한 치아우식과 치주질환에 대한 충전치료 건수와 치아우식과 치주질환에 대한 보철치료 건수의 총 4가지로 구성된다. 도수자료(count data)인 이들 반응변수를 예측하기 위한 통계적 모형은 GLM의 포아송회귀모형 또는 음이항회귀모형을 이용할 수 있다. 일반적으로 포아송회귀모형은 반응변수가 양(+)의 정수 값을 갖고 평균과 분산이 비슷할 때 적용할 수 있는 모형이다. 포아송회귀모형은 설명변수 벡터 $\mathbf{x}_i = (x_1, \dots, x_k)$ 가 주어질 때, 랜덤성분에 해당하는 발생건수 y_i 의 확률은 식 (2.1)과 같은 포아송분포로부터 구해진다.

$$P(Y_i = y_i | \mathbf{x}_i) = \frac{e^{-\mu_i} \mu_i^{y_i}}{y_i!}, \quad y_i = 0, 1, 2, \dots \quad (2.1)$$

선형예측식 $\mathbf{x}_i'\boldsymbol{\beta}$ 은 식 (2.2)와 같이 평균 모수 $\mu_i = E(y_i|\mathbf{x}_i)$ (i 번째 개체의 조건부 평균)의 로그함수로 표현된다.

$$\log(\mu_i) = \mathbf{x}_i'\boldsymbol{\beta}, \quad (2.2)$$

여기서 $\boldsymbol{\beta}$ 는 $(k+1) \times 1$ 모수 벡터를 나타내고, k 는 설명변수의 수이다.

반응변수에 포아송분포를 적용하기 위해서는 조건부 평균과 조건부 분산이 동일하다는 제약이 존재한다. 그러나 일반적으로 사용되는 도수 데이터는 평균이 증가할 때 분산 또한 증가하는 과대산포(over-dispersion)현상이 존재한다. 이러한 과대산포를 갖는 데이터에 포아송 분포를 적용하면, 포아송회귀모형은 모형 적합의 효율성이 떨어지게 된다(Gumu 1991; 전희주·안철경 2012). 따라서 이러한 포아송분포의 과대산포현상을 해소하기 위한 모형으로 음이항분포(negative binomial distribution)를 사용하게 된다.

음이항분포는 포아송분포와 같이 양(+)의 정수값을 갖는 확률변수의 분포로 이분산성(hetero-skedasticity)을 허용하는 분산함수로 정의되므로 과대산포가 발생하는 경우 사용할 수 있는 대체 수단이 된다. 음이항분포는 임의의 i 번째 개체의 발생 건수에 대한 평균과 분산은 식 (2.3)과 같이 표현될 수 있다.

$$E(y_i | \mathbf{x}_i) = \mu_i, \quad Var(y_i) = \mu_i + D\mu_i^2, \quad (2.3)$$

상기 식 (2.3)의 i 번째 개체의 발생 건수에 대한 음이항분포의 분산을 살펴보면, 관측되지 않은 비동질(heterogeneity)요소인 산포모수 $D(= \frac{1}{\tau})^{10)}$ 값에 의해 포아송분포의 확장이 되며 과대산포 문제를 해결하게 된다. 만약, 산포모수 D 값이 0보다 커질수록 포아송 분포보다 분산이 더 커지고, D 값이 0에 가까울수록 $Var(y_i)$ 는 $E(y_i|\mathbf{x}_i)$ 에 가까워지므로 음이항분포는 포아송분포로 수렴하게 된다.

음이항회귀모형 역시 반응변수가 양의 정수를 갖는 음이항분포를 갖고, 선형예측식

10) 여기서 τ 는 GLM모형에서 추정된 산포모수값으로 <Table 4>, <Table 6>, <Table 8>, <Table 10>에서 나타난다.

$x_i'\beta$ 은 식 (2.4)와 같이 로그함수를 통해 연결이 된다.

$$\log (\mu_i) = x_i'\beta \quad (2.4)$$

여기서 β 는 $(k+1) \times 1$ 모수 벡터를 나타내고, k 는 설명변수의 수이다.

따라서 본 연구는 사후 관찰기간에 발생하는 치아우식 및 치주질환별 충전 및 보철치료 건수의 조합인 4개의 반응변수에 대해 현재 보험상품에 적용하고 있는 성별, 연령, 사전 관찰집단의 3가지 설명변수와 오프셋(offset)을 포함한 음이형회귀모형을 이용하여 분석하고자 한다.

IV. 실증분석

1. 관찰대상 집단의 일반적인 특성

〈Table 1〉은 기존 분류기준을 적용한 유병자A와 무병자A 집단의 성별, 연령별 분포를 나타낸 표이다. 유병자A의 비율은 전체 분석대상 인원의 약 27.4%를 차지하고 있다. 남성의 유병자A 비율은 전체남성의 약 30%, 여성의 유병자A 비율은 전체여성의 약 26% 수준이다. 또한, 유병자A 집단 내 성별 비율은 남성 52%, 여성 48%인 반면, 무병자A 집단 내 성별 비중은 남성 48%, 여성 52%이다.

〈Table 1〉 Distribution by gender and age of diseased A and disease-free A in basis year 2007(unit: person, %)

variable		diseased A		disease-free A	
		number	ratio	number	ratio
gender	male	81,077	51.8	198,536	48.0
	female	75,465	48.2	215,296	52.0
age	30~34	16,362	10.5	63,715	15.4
	35~39	18,472	11.8	74,182	17.9
	40~44	19,146	12.2	66,435	16.1
	45~49	22,866	14.6	63,637	15.4
	50~54	22,165	14.2	48,799	11.8
	55~59	17,265	11.0	31,239	7.5
	60~64	14,691	9.4	23,907	5.8
	65~69	13,041	8.3	20,493	5.0
	70~74	8,529	5.4	14,001	3.4
	75~79	4,005	2.6	7,424	1.8
total		156,542	27.4	413,832	72.6

연령대별 인원 수 대비 유병자A의 비율이 가장 높은 연령대는 65~69세(38.9%), 60~64세(38.1%), 70~74세(37.9%), 55~59세(35.6%)의 순으로 나타났다. 반면, 유병자 A 집단 내 비중을 살펴보면, 45~49세(14.6%), 50~54세(14.2%), 40~44세(12.2%), 35~39세(11.8%), 30~34세(10.5%)의 순으로 나타나 60세 미만의 연령대가 전체 유병자 A 집단의 74.3%를 차지하고 있다.

〈Table 2〉는 완화된 분류기준을 적용한 유병자B와 무병자B의 성별, 연령별 분포이다. 전체적으로 유병자B의 비율은 13.5%로 유병자A 대비 약 2배 정도 낮은 비율임을 알 수 있다. 유병자B를 성별 기준으로 살펴보면 남성의 경우 유병자A 남성의 약 47.5%, 여성은 유병자A 여성 대비 50.8% 수준이다.

연령대별로 살펴보면 유병자B는 유병자A 대비 43~51% 수준으로 나타났다. 특히, 40~44세, 45~49세, 50~54세 연령집단은 유병자B 분류기준을 적용할 경우 감소폭이 51%로 다른 연령대에 비해 높게 나타났다.

〈Table 2〉 Distribution by gender and age of diseased B and disease-free B in basis year 2007(unit: person, %)

variable		diseased B		disease-free B	
		number	ratio	number	ratio
gender	male	38,559	50.1	241,054	48.0
	female	38,333	49.9	252,428	52.0
age	30~34	7,682	10.0	72,395	14.7
	35~39	8,976	11.7	83,678	17.0
	40~44	9,695	12.6	75,886	15.4
	45~49	11,738	15.3	74,765	15.2
	50~54	11,206	14.6	59,758	12.1
	55~59	8,513	11.1	39,991	8.1
	60~64	7,150	9.3	31,448	6.4
	65~69	6,179	8.0	27,355	5.5
	70~74	4,021	5.2	18,509	3.8
	75~79	1,732	2.3	9,697	2.0
total		76,892	13.5	493,482	86.5

본 연구의 코호트 분석 대상자는 2007년 기준 570,374명이 2015년까지 경과기간에 따라 매년 위험에 노출된 총 5,133,366명으로 남성 2,516,517명(49.0%), 여성 2,616,849명(51.0%)이고, 성별/연령대별/관찰기간에 따른 병력에 대한 분포는 〈Table 3〉과 같다.

〈Table 3〉 Sample characteristics of subjects exposed to risk each year from 2007 to 2015(unit: person, %)

	age	number	ratio	disease-free A	diseased A	disease-free B	diseased B
male	30~34	119,338	4.7	96,337	23,001	107,948	11,390
	35~39	337,288	13.4	272,372	64,916	305,854	31,434
	40~44	410,821	16.3	326,942	83,879	370,359	40,462
	45~49	394,865	15.7	298,554	96,311	348,368	46,497
	50~54	370,620	14.7	260,011	110,609	318,020	52,600
	55~59	294,339	11.7	190,399	103,940	245,447	48,892
	60~64	210,797	8.4	126,445	84,352	171,400	39,397
	65~69	159,752	6.3	91,261	68,491	127,824	31,928
	70~74	118,646	4.7	67,035	51,611	94,337	24,309
	75~79	100,051	4.0	57,468	42,583	79,929	20,122
	sum	2,516,517	49.0	1,786,824	729,693	2,169,486	347,031

female	30~34	115,299	4.4	89,740	25,559	103,766	11,533
	35~39	324,902	12.4	256,256	68,646	292,854	32,048
	40~44	395,712	15.1	313,530	82,182	354,600	41,112
	45~49	382,489	14.6	298,064	84,425	337,492	44,997
	50~54	366,650	14.0	276,158	90,492	316,765	49,885
	55~59	302,145	11.5	217,508	84,637	255,997	46,148
	60~64	223,028	8.5	152,507	70,521	185,961	37,067
	65~69	182,097	7.0	120,168	61,929	150,938	31,159
	70~74	154,814	5.9	100,632	54,182	128,973	25,841
	75~79	169,713	6.5	113,101	56,612	144,506	25,207
	sum	2,616,849	51.0	1,937,664	679,185	2,271,852	344,997
total		5,133,366	100.0	3,724,488	1,408,878	4,441,338	692,028

2. 분석 결과

앞에서 언급한 데이터 및 분석모형을 바탕으로 2007년~2015년도 사후기간에 발생하는 4개의 반응변수(치아우식에 대한 충전치료 건수, 치주질환에 대한 충전치료 건수, 치아우식에 대한 보철치료 건수, 치주질환에 대한 보철치료 건수)에 대한 분석결과를 제시하고자 한다.

가. 치아우식에 대한 충전치료

성별, 연령, 사전관찰집단을 설명변수로 하여 오프셋을 포함한 음이항회귀모형을 이용한 치아우식에 대한 충전치료 발생 건수를 적합한 결과, 사전관찰집단과 연령변수는 유의한 변수인 반면, 성별은 유의수준 0.05하에서 유의하지 않은 변수임을 알 수 있다(〈Table 4〉 참조).

〈Table 4〉의 결과에서 사전관찰집단 변수를 살펴보면, 유병자B 집단이 평균적으로 치아우식에 대한 충전치료 건수가 가장 많으며 다음으로 유병자A, 무병자B, 무병자A의 순서를 보였다. 유병자B의 치아우식에 대한 충전치료 건수는 유병자A 대비 $\exp(0.2547)=1.2900$ 배 높게 나타나 유병자A 대비 유병자B의 예측 상대위험도는 1.290임을 알 수 있다. 한편, 무병자B의 경우는 무병자A 대비 $\exp(0.0026)=1.0026$ 배 높게 나타났다.

〈Table 4〉 The result of model fitted the number of filling treatments for dental caries

variable		df	estimate	s.e.	P-value	LR χ^2	P-value
intercept		1	-1.1908	0.0336	<.0001	1254.75	<0.0001
gender	female	1	-0.0024	0.0183	0.8973	0.02	0.8971
	male	0	0.0000	0.0000	0.0000		
age	30~34	1	-0.2869	0.0410	<.0001	154.88	<0.0001
	35~39	1	-0.3500	0.0408	<.0001		
	40~44	1	-0.2917	0.0408	<.0001		
	45~49	1	-0.1730	0.0408	<.0001		
	50~54	1	-0.0395	0.0408	0.3284		
	55~59	1	0.0496	0.0408	0.2257		
	60~64	1	0.1438	0.0408	0.0004		
	65~69	1	0.2106	0.0408	<.0001		
	70~74	1	0.1945	0.0408	<.0001		
	75~79	0	0.0000	0.0000	.		
pre-observ ed group	disease-free A	1	-0.6056	0.0258	<.0001	192.51	<0.0001
	diseased A	1	-0.2547	0.0258	<.0001		
	disease-free B	1	-0.6030	0.0258	<.0001		
	diseased B	0	0.0000	0.0000	.		
dispersion(τ)		1	0.0066	0.0010	.		
Deviance			80.01 (df=66)				
Deviance/df			1.2123				
Pearson χ^2			79.35				
Full Log Likelihood			-696.88				
AIC			1423.76				

연령대별 치아우식에 대한 충전치료 건수를 보면, 연령이 높을수록 치아우식에 대한 충전치료 건수가 증가하는 경향을 보이나, 75세 이상 연령대에서는 감소하는 것으로 나타났다. 치아우식에 대한 충전치료 건수가 가장 높은 연령대인 65~69세는 가장 낮은 연령대인 35~39세 대비 $\exp(0.5606)=1.7517$ 배 높았다.

〈Table 5〉 Predicted value and relative risk of the number of filling treatments for dental caries

gender	age	incidence rate of diseased A	incidence rate of diseased B	diseased relative risk (B/A)	incidence rate of disease-free A	incidence rate of disease-free B	disease-free relative risk (B/A)
male	30~34	0.177	0.228	1.2900	0.1244	0.1247	1.0026
	35~39	0.166	0.214		0.1168	0.1171	
	40~44	0.176	0.227		0.1238	0.1241	
	45~49	0.198	0.255		0.1394	0.1397	
	50~54	0.226	0.292		0.1593	0.1597	
	55~59	0.247	0.319		0.1741	0.1746	
	60~64	0.272	0.351		0.1913	0.1918	
	65~69	0.291	0.375		0.2045	0.2051	
	70~74	0.286	0.369		0.2013	0.2018	
	75~79	0.235	0.304		0.1657	0.1662	
female	30~34	0.177	0.228		0.1244	0.1247	
	35~39	0.166	0.214		0.1168	0.1171	
	40~44	0.176	0.227		0.1238	0.1241	
	45~49	0.198	0.255		0.1394	0.1397	
	50~54	0.226	0.292		0.1593	0.1597	
	55~59	0.247	0.319		0.1741	0.1746	
	60~64	0.272	0.351		0.1913	0.1918	
	65~69	0.291	0.375		0.2045	0.2051	
	70~74	0.286	0.369		0.2013	0.2018	
	75~79	0.235	0.304		0.1657	0.1662	

〈Table 5〉는 음이항회귀모형을 이용한 유병자A와 유병자B의 예측된 치아우식에 대한 충전치료 건수를 발생비율로 나타낸 표이며, 유병자A 대비 유병자B의 치아우식 충전치료 건수의 상대위험도는 모두 동일하게 1.2900이며, 무병자A 대비 무병자B의 치아우식 충전치료 건수의 상대위험도는 모두 동일하게 1.0026으로 음이항회귀모형의 회귀계수에서 얻은 사전관찰집단의 상대위험도와 일치한다.

나. 치주질환에 대한 충전치료

〈Table 6〉은 치주질환에 대한 충전치료 건수에 대한 음이항회귀모형 적합 결과 성별, 연령, 사전관찰집단 모두 치주질환에 대한 충전치료건수에 매우 유의한 영향을 미치는 것으로 나타났다. 이들 변수 중 사전관찰집단이 가장 유의한 영향 변수이며, 다음으로 연령 대, 성별 순으로 나타났다.

〈Table 6〉 The result of model fitted the number of filling treatments for periodontal disease

variable		df	estimate	s.e.	P-value	LR χ^2	P-value
intercept		1	-3.5560	0.0353	<.0001		
gender	female	1	-0.0932	0.0187	<.0001	20.84	<0.0001
	male	0	0.0000	0.0000	.		
age	30~34	1	-1.1078	0.0465	<.0001	259.20	<0.0001
	35~39	1	-0.4866	0.0424	<.0001		
	40~44	1	-0.1598	0.0418	<.0001		
	45~49	1	0.1086	0.0416	0.0091		
	50~54	1	0.3143	0.0415	<.0001		
	55~59	1	0.4619	0.0415	<.0001		
	60~64	1	0.4726	0.0416	<.0001		
	65~69	1	0.3768	0.0418	<.0001		
	70~74	1	0.1830	0.0422	<.0001		
	75~79	0	0.0000	0.0000	.		
pre-observ ed group	disease-free A	1	-0.3789	0.0268	0.0001	115.04	<0.0001
	diseased A	1	-0.1616	0.0272	<.0001		
	disease-free B	1	-0.3740	0.0268	<.0001		
	diseased B	0	0.0000	0.0000	.		
dispersion(τ)		1	0.0063	0.0012			.
Deviance			87.67 (df=66)				
Deviance/df			1.3284				
Pearson χ^2			87.97				
Full Log Likelihood			-526.91				
AIC			1083.82				

〈Table 6〉의 사전관찰집단에 대한 적합결과를 살펴보면, 유병자B가 평균적으로 치주질환에 대한 충전치료 건수가 가장 많으며 다음으로 유병자A, 무병자B, 무병자A의 순서로 나

타났다. 유병자B의 치주질환에 대한 충전치료 건수는 유병자A 대비 $\exp(0.1616)=1.1753$ 배 높아 유병자A 대비 유병자B의 예측 상대위험도는 1.1753임을 알 수 있다. 또한, 무병자 B는 무병자A 대비 $\exp(0.0049)=1.0049$ 배 높았다.

연령대별 적합결과를 보면, 연령이 높을수록 치료 건수가 증가하는 현상을 보이며, 50~70세의 치료 건수가 타 연령대에 비해 매우 많으며, 70세 이상에서는 급격히 감소하는 경향을 보인다. 치주질환에 대한 충전치료 건수가 가장 높은 60~64세는 가장 낮은 30~34세 대비 치주질환에 대한 충전치료 건수가 $\exp(1.5804)=4.8569$ 배 높게 나타났다.

성별 결과를 살펴보면, 남성의 치주질환에 대한 충전치료 건수는 여성대비 $\exp(0.0932)=1.0977$ 배 높은 경향을 보인다.

〈Table 7〉 Predicted value and relative risk of the number of filling treatments for periodontal disease

gender	age	incidence rate of diseased A	incidence rate of diseased B	diseased relative risk (B/A)	incidence rate of disease-free A	incidence rate of disease-free B	disease-free relative risk (B/A)
male	30~34	0.008	0.009	1.1753	0.0065	0.0065	1.0049
	35~39	0.015	0.018		0.0120	0.0121	
	40~44	0.021	0.024		0.0167	0.0167	
	45~49	0.027	0.032		0.0218	0.0219	
	50~54	0.033	0.039		0.0268	0.0269	
	55~59	0.039	0.045		0.0310	0.0312	
	60~64	0.039	0.046		0.0314	0.0315	
	65~69	0.035	0.042		0.0285	0.0286	
	70~74	0.029	0.034		0.0235	0.0236	
	75~79	0.024	0.029		0.0195	0.0196	
female	30~34	0.007	0.009		0.0059	0.0059	
	35~39	0.014	0.016		0.0109	0.0110	
	40~44	0.019	0.022		0.0152	0.0153	
	45~49	0.025	0.029		0.0199	0.0199	
	50~54	0.030	0.036		0.0244	0.0245	
	55~59	0.035	0.041		0.0283	0.0284	
	60~64	0.036	0.042		0.0286	0.0287	
	65~69	0.032	0.038		0.0260	0.0261	
	70~74	0.027	0.031		0.0214	0.0215	
	75~79	0.022	0.026		0.0178	0.0179	

〈Table 7〉은 음이항회귀모형을 이용한 유병자A와 유병자B의 예측된 치주질환에 대한 충전치료 건수를 발생비율로 나타낸 표이며, 유병자A 대비 유병자 B의 치주질환에 대한 충전치료 건수 상대위험도는 모두 동일하게 1.1753가 된다. 무병자A 대비 무병자 B의 치주질환 충전치료 건수 상대위험도는 모두 동일하게 1.0049로 음이항회귀모형의 회귀계수에서 얻은 사전관찰집단의 상대위험도와 일치된다.

다. 치아우식에 대한 보철치료

〈Table 8〉은 치아우식에 대한 보철치료 건수 적합을 위한 음이항회귀모형 결과이다. 성별, 연령대, 사전관찰집단 모두 보철 치아우식 치료 건수에 매우 유의한 영향을 미치며, 이 중 성별이 가장 유의한 영향 변수이며, 다음으로 사전관찰집단, 연령대의 순서로 나타났다.

〈Table 8〉의 성별변수를 살펴보면, 남성의 치아우식에 대한 보철치료 건수는 여성대비 $\exp(0.2591)=1.2958$ 배 높은 결과를 보였다.

다음으로 사전관찰집단의 경우, 유병자B가 평균적으로 치아우식에 대한 보철치료 건수가 가장 많으며 다음으로 유병자A, 무병자B, 무병자A의 순서로 나타났다. 유병자B의 치아우식에 대한 보철치료 건수는 유병자A 대비 $\exp(0.0233)=1.0236$ 배, 무병자B의 치아우식에 대한 보철치료 건수는 무병자A 대비 $\exp(0.0516)=1.0530$ 배 많은 것으로 나타났다. 결국, 유병자A 대비 유병자 B의 치아우식에 대한 보철치료 건수의 예측 상대위험도는 1.0236임을 알 수 있다.

〈Table 9〉는 음이항회귀모형을 이용한 유병자A와 유병자B, 무병자와 무병자B의 예측된 치아우식에 대한 보철치료 건수를 발생비율로 나타낸 표이다. 유병자A 대비 유병자 B의 치아우식에 대한 보철치료 건수 상대위험도는 모두 동일하게 1.0236이 된다. 무병자A 대비 무병자 B의 치아우식에 대한 보철치료 건수 상대위험도 또한 모두 동일하게 1.053으로 음이항모형의 회귀계수에서 얻은 사전관찰집단의 상대위험도와 일치한다.

〈Table 8〉 The result of model fitted the number of prosthetic treatments for dental caries

variable		df	estimate	s.e.	P-value	LR χ^2	P-value
intercept		1	-2.8287	0.0229	<.0001		
gender	female	1	-0.2591	0.0121	<.0001	148.67	<0.0001
	male	0	0.0000	0.0000	.		
age	30~34	1	-0.8814	0.0292	<.0001	308.61	<0.0001
	35~39	1	-1.0617	0.0275	<.0001		
	40~44	1	-0.9263	0.0270	<.0001		
	45~49	1	-0.6848	0.0267	<.0001		
	50~54	1	-0.4812	0.0265	<.0001		
	55~59	1	-0.3299	0.0265	<.0001		
	60~64	1	-0.1745	0.0266	<.0001		
	65~69	1	-0.0633	0.0267	0.0175		
	70~74	1	0.0082	0.0268	0.7584		
	75~79	0	0.0000	0.0000	.		
pre-observ ed group	disease- free A	1	-0.2951	0.0177	<.0001	143.74	<0.0001
	diseased A	1	-0.0233	0.0179	0.0222		
	disease- free B	1	-0.2435	0.0175	<.0001		
	diseased B	0	0.0000	0.0000	.		
dispersion(τ)		1	0.0024	0.0005	.		
Deviance			86.65 (df=66)				
Deviance/df			1.3128				
Pearson χ^2			87.79				
Full Log Likelihood			-512.22				
AIC			1054.44				

〈Table 9〉 Predicted value and relative risk of the number of
prosthetic treatments for dental caries

gender	age	incidence rate of diseased A	incidence rate of diseased B	diseased relative risk (B/A)	incidence rate of disease- free A	incidence rate of disease- free B	disease- free relative risk (B/A)
male	30~34	0.024	0.024	1.0236	0.0182	0.0192	1.0530
	35~39	0.020	0.020		0.0152	0.0160	
	40~44	0.023	0.023		0.0174	0.0183	
	45~49	0.029	0.030		0.0222	0.0234	
	50~54	0.036	0.037		0.0272	0.0286	
	55~59	0.042	0.042		0.0316	0.0333	
	60~64	0.048	0.050		0.0369	0.0389	
	65~69	0.054	0.055		0.0413	0.0435	
	70~74	0.058	0.060		0.0444	0.0467	
	75~79	0.058	0.059		0.0440	0.0463	
female	30~34	0.018	0.019		0.0141	0.0148	
	35~39	0.015	0.016		0.0117	0.0124	
	40~44	0.018	0.018		0.0134	0.0142	
	45~49	0.022	0.023		0.0171	0.0180	
	50~54	0.028	0.028		0.0210	0.0221	
	55~59	0.032	0.033		0.0244	0.0257	
	60~64	0.037	0.038		0.0285	0.0300	
	65~69	0.042	0.043		0.0319	0.0336	
	70~74	0.045	0.046		0.0342	0.0360	
	75~79	0.045	0.046		0.0339	0.0357	

라. 치주질환에 대한 보철치료

〈Table 10〉는 치주질환에 대한 보철치료 건수 적합을 위한 음이항회귀모형 결과이다. 성별, 연령, 사전관찰집단 모두 치주질환에 대한 보철치료 건수에 매우 유의한 영향을 주며, 이 중 성별이 가장 유의한 변수이며, 다음으로 사전관찰집단, 연령순으로 나타났다.

〈Table 10〉의 음이항회귀모형 적합결과에서 성별변수는 남성의 치주질환에 대한 보철치료 건수는 여성대비 $\exp(0.4347)=1.5445$ 배 높은 경향을 보인다. 다음으로 사전관찰집단을 살펴보면, 유병자A의 치주질환에 대한 보철치료 건수가 가장 많으며 다음으로 유병

자B, 무병자B, 무병자A의 순서로 나타났다. 유병자B는 유병자A 대비 치주질환에 대한 보철 치료 발생건수는 $\exp(-0.0361)=0.9645$ 배로 상대위험도는 0.9645이나 그 차이에 대한 $p\text{-value}=0.3258$ 로 나타나 통계적으로 유의하지 않았다. 반면, 무병자B는 무병자A 대비 $\exp(0.0978)=1.1027$ 배 많은 것으로 나타났다.

연령대별 치주질환에 대한 보철치료 건수를 보면, 연령이 높을수록 증가하는 현상을 보였는데 치주질환에 대한 보철치료 건수가 가장 높은 75세 이상은 가장 낮은 30~34세 대비 무려 $\exp(1.4430)=4.2334$ 배 높게 나타났다.

〈Table 10〉 The result of model fitted the number of prosthetic treatments for periodontal disease

variable		df	estimate	s.e.	P-value	LR χ^2	P-value
intercept		1	-1.6403	0.0484	<.0001		
gender	female	1	-0.4347	0.0260	<.0001	119.68	<0.0001
	male	0	0.0000	0.0000	-		
age	30~34	1	-1.4430	0.0586	<.0001	237.84	<0.0001
	35~39	1	-1.3784	0.0579	<.0001		
	40~44	1	-1.0219	0.0577	<.0001		
	45~49	1	-0.6290	0.0577	<.0001		
	50~54	1	-0.3773	0.0576	<.0001		
	55~59	1	-0.2323	0.0576	<.0001		
	60~64	1	-0.1408	0.0576	0.0145		
	65~69	1	-0.0767	0.0576	0.1830		
	70~74	1	-0.0505	0.0576	0.3811		
pre-observed group	75~79	0	0.0000	0.0000	-	92.07	<0.0001
	disease-free A	1	-0.3673	0.0367	<.0001		
	diseased A	1	0.0361	0.0367	0.3258		
	disease-free B	1	-0.2695	0.0367	<.0001		
dispersion(τ)		0	0.0000	.0000	-		
Deviance			80.95 (df=66)				
Deviance/df			1.2265				
Pearson χ^2			84.53				
Full Log Likelihood			-650.61				
AIC			1331.21				

〈Table 11〉은 음이항회귀모형을 이용한 유병자A와 유병자B의 예측된 치아우식에 대한 보철치료 건수를 발생비율로 나타낸 표이다. 유병자A 대비 유병자B의 치아우식에 대한 보철치료 건수 상대위험도는 통계적으로 유의하지는 않지만 〈Table 10〉과 동일하게 0.9645로 나타났다. 무병자A 대비 무병자B의 치주질환 보철치료 건수의 상대위험도 또한 모두 동일하게 1.1027로 음이항회귀모형의 회귀계수에서 얻은 사전관찰집단의 상대위험도와 일치한다.

〈Table 11〉 Predicted value and relative risk of the number of prosthetic treatments for periodontal disease

gender	age	incidence rate of diseased A	incidence rate of diseased B	diseased relative risk (B/A)	incidence rate of disease-free A	incidence rate of disease-free B	disease-free relative risk (B/A)
male	30~34	0.047	0.046	0.9645	0.0317	0.0350	1.1027
	35~39	0.051	0.049		0.0338	0.0373	
	40~44	0.072	0.070		0.0483	0.0533	
	45~49	0.107	0.103		0.0716	0.0790	
	50~54	0.138	0.133		0.0921	0.1016	
	55~59	0.159	0.154		0.1065	0.1174	
	60~64	0.175	0.168		0.1167	0.1287	
	65~69	0.186	0.180		0.1244	0.1372	
	70~74	0.191	0.184		0.1277	0.1408	
	75~79	0.201	0.194		0.1343	0.1481	
female	30~34	0.031	0.030		0.0205	0.0227	
	35~39	0.033	0.032		0.0219	0.0242	
	40~44	0.047	0.045		0.0313	0.0345	
	45~49	0.069	0.067		0.0464	0.0511	
	50~54	0.089	0.086		0.0596	0.0658	
	55~59	0.103	0.100		0.0689	0.0760	
	60~64	0.113	0.109		0.0755	0.0833	
	65~69	0.121	0.116		0.0805	0.0888	
	70~74	0.124	0.119		0.0827	0.0912	
	75~79	0.130	0.126		0.0870	0.0959	

결론적으로 치아우식증에 의한 충전치료 건수를 제외하고 사전관찰집단, 연령, 성별변수는 사후 관찰기간 치아우식과 치주질환 원인에 의한 충전 및 보철 치료발생 건수에 모두 유의한 영향을 미치는 요인으로 나타났다.

사전관찰 집단변수의 경우 치주질환에 대한 보철치료를 제외하고 유병자B의 유병자A 대비 상대위험도가 1.024~1.290배 높게 나타난 반면, 가입기준 완화 시 가입대상인 무병자B는 기존 무병자A 대비 충전치료의 상대위험도는 치아우식과 치주질환에서는 각각 1.0026과 1.0049로 거의 크지 않은 반면, 치아우식과 치주질환에 대한 보철치료에서는 각각 1.053과 1.1027로 상대위험도는 크게 증가함을 보여준다.

다음으로 연령대 변수의 경우 충전치료에서 일부 초 고연령대에서 감소하는 현상을 제외하면 연령대가 증가할수록 상대위험도가 높아지는 경향을 보였다. 특히, 치주질환에 대한 보철치료의 경우 75세 이상은 30~34세 대비 무려 $\exp(1.4430)=4.2334$ 배 높게 나타났다.

마지막으로 성별은 치아우식에 대한 충전치료를 제외한 나머지 반응변수에서 남성의 상대위험도가 여성보다 1.0977~1.2958배 정도 높게 나타났다. 상기 결과를 종합해 볼 때, 고연령, 남성, 유병자B 집단의 위험도가 가장 높음을 알 수 있다.

V. 결론 및 시사점

실손의료보험 상품은 매년 손해율 및 손해액이 증가함에 따라 지속가능성이 위협받고 있다. 실손의료보험의 지속가능성 확보를 위해서는 건강보험 빅데이터를 활용한 정교한 위험분석에 기반한 보험료 차등화가 중요하다고 할 수 있다. 이러한 상황에서 본 연구는 국민건강보험 코호트2.0 DB를 이용하여 구강질환에 대한 일반화선형모형을 이용한 상대 위험도를 측정하여 고위험 상품인 치아보험의 시장확대 가능성을 모색하였다.

본 연구는 세부적으로 기준년도(2007년) 이전의 사전 관찰기간 1년 이내의 치아우식증 또는 과거 5년 이내에 적어도 치주질환 치료가 있는 유병자A와 무병자A, 새로운 제안 상품군인 고위험군인 사전 관찰기간 직전 1년 이내의 치아우식증 또는 1년 이내의 치주질환

치료가 있는 유병자B, 무병자B의 4집단으로 구분하였다. 2007년~2015년도 사후기간에 발생하는 4개의 반응변수인 치아우식 충전치료 치료 건수, 치주질환 충전치료 건수, 치아우식 보철치료 치료 건수와 치주질환 보철치료 건수의 4개 반응변수에 대해 오프셋(offset)을 포함한 음이항회귀모형을 적합하여 상대위험도를 산정하였다.

보험상품의 가입대상 확대 시 할증률을 산출하기 위한 대상인 사전관찰집단변수의 무병자A 대비 무병자B의 상대위험도는 치아우식과 치주질환에 대한 충전치료에서는 각각 1.0026과 1.0049로 매우 낮은 반면, 치아우식과 치주질환에 대한 보철치료에서는 각각 1.053과 1.103으로 상대위험도가 크게 증가하였다. 결국 상품가입의 범위를 치주질환 치료에 대한 사전 관찰기간을 5년에서 1년으로 단축함으로써 가입대상을 확대할 경우 충전치료보다는 보철치료에서 보험료 할증 폭이 더욱 크게 증가할 수 있음을 보여준다.

현재 보험업계에서는 상품개발을 할 때 GLM을 이용한 예측값을 사용하기보다 실제 측정된 값을 사용한다. 이는 성별, 연령별에 따라 상대위험도가 모두 일정하게 나오지 않을 수 있어 상품개발에 대한 전체적인 의사결정을 어렵게 한다. 그러나 GLM모형을 이용하게 되면 성별, 연령별 발생률은 다르더라도 사전관찰집단 간의 상대위험도는 성별, 연령별 오차를 반영하여 동일한 값이 제시되므로 통일되고 합리적인 결론을 내릴 수 있는 장점이 될 것이다.

본 연구의 결과를 통해 보험회사 및 가입자, 정부 측면에서의 기대효과 및 시사점은 다음과 같다.

먼저, 보험회사 측면에서 건강보험 빅데이터를 이용하여 질병과 사망에 영향을 미칠 수 있는 요인들을 설명변수로 포함한 일반화선형모형을 활용한다면 연령 및 성별 이외의 다양한 요인을 고려한 위험분석이 가능하므로 더욱 세밀하고 차등적인 보험료 산정이 가능할 것으로 사료된다. 세밀하고 차등적인 보험료 산정은 가입자의 역선택 위험을 감소시켜 보험회사의 재무건전성 향상에도 도움이 될 것이다. 또한, 기존에 가입이 불가능했던 고위험 집단의 보험상품 가입이 가능하므로 시장 확대 효과를 기대할 수 있을 것이다.

다음으로, 가입자 측면에서는 차등적이고 세분화된 보험료 적용을 통해 기존의 가입이 어려웠던 고위험 유병력자도 보험가입의 혜택을 볼 수 있다. 또한, 가입자 간 보험료 부담의 형평성이 제고됨으로써 선량한 가입자에 대한 부당한 보험료 부담 전가가 해소될 수 있을 것이다.

정부 차원에서도 실손의료보험 상품과 국민건강보험과의 효과적인 연계와 협력이 이뤄진다면 국민건강보험 지속가능성 제고에도 도움이 될 것이다.

향후 국민건강보험 빅데이터와 같은 공공의료데이터의 활용이 더욱 활성화된다면 기존 상품의 시장 확대뿐만 아니라 사회적 변화 및 질병패턴 변화에 대응한 보험상품 개발에도 기여 할 수 있다. 특히, 고령화 시대 당뇨, 고혈압, 치매 등 노인성 만성질환을 앓고 있는 고령자·유병력자 대상 보험상품 개발도 가능할 것이다. 또한, 난임 시술 등과 같이 기존에 보장하지 않았던 보장범위의 확대도 기대할 수 있다.

보험회사의 보험상품 시장 확대 노력과 다양한 보험상품 개발이 꾸준히 이뤄진다면, 보험소비자의 다양한 위험에 대한 보장 서비스가 가능한 동시에 보험상품 선택권이 늘어날 수 있을 것이다.

궁극적으로 빅데이터를 활용한 보험상품 개발은 대수의 법칙에 근거한 표준화된 상품에서 개인별 맞춤형 상품으로 탈바꿈하는 계기가 될 것이다. 4차 산업혁명 시대 보험산업의 발전과 소비자 보호 및 권리 증진을 위해서는 보험업계의 노력뿐만 아니라 관련 학계 및 감독당국과의 협력이 필요한 시점이다.

참고문헌

- 김대환·이봉주 (2013), “실손의료보험의 역선택 분석”, **보험학회지**, 제96호, pp. 25-50.
- (Translated In English) Kim, D., and B., Lee (2013). “An Analysis on Adverse Selection in Fee-For-Service Health Insurance”, *KOREAN JOURNAL OF INSURANCE*, (96):25-50.
- 이경아·이항석 (2016), “실손의료보험의 역선택과 보험료 차등화,” **리스크관리연구**, 제 27권 제3호, pp. 93-128.
- (Translated In English) Lee, K., and H., Lee (2016). “Adverse Selection and Rates Differentiation in the Korean Private Health Insurance”, *The Journal of Risk Management*, 27(3):93-128.
- 이항석·이가은·이경아 (2017), “실손의료보험에서 요율산정과 일반화선형모형(GLM)의 활용,” **리스크관리연구**, 제28권 제4호 pp. 37-71.
- (Translated In English) Lee, H., K., Lee and K., Lee (2017). “A Ratemaking of the Korean Private Health Insurance using GLM”, *The Journal of Risk Management*, 28(4):37-71.
- 이항석·백혜연 (2018), “실손의료보험 할인할증제도의 실증분석,” **보험학회지**, 제116호, pp. 33-66.
- (Translated In English) Lee, H., and H., Baek (2018). “An Empirical Analysis of Bonus-Malus System in Private Health Insurance”, *KOREAN JOURNAL OF INSURANCE*, (116):33-66.
- 전희주·안철경 (2012), “음이항회귀모형을 이용한 보험설계사들의 조직성과에 관한 연구”, **보험금융연구**, 제23권 4호, pp. 29-60.
- (Translated In English) Chun, H., and C., Ahn (2012). “A Study on Organizational Outcome of Life Insurance Solicitors Using Negative Binomial Regression Model”, *Journal of Insurance and Finance*,

23(4):29-60.

조재린·정성희 (2018), **계리적 관점에서 본 실손의료보험 개선 방안**, 보험연구원.

(Translated In English) Cho, J., and S., Chung (2018). *A Study on the Improvement of Private Health Insurance from a Actuarial Perspective*, Korea Insurance Research Institute.

치과의료정책연구원 (2020), “우리나라 국민의 외래 치과진료비 지출 경향: 2011-2017 한국의료패널 분석”, 이슈보고서

(Translated In English) Health Policy Institute (2020). “The Trends of Korean Outpatient Dental Expenditure: 2011-2017 Korea Medical Panel Analysis”, Issue Report.

Akerlof, G. (1970). “The Market for ‘Lemons’: Quality Uncertainty and the Market Mechanism”, *Quarterly Journal of Economics*, 84:488-500.

Arrow, K. (1963). “Uncertainty and the Welfare Economics of Medical Care”, *The American Economic Review*, 53(5):941-973.

Cutler, D., and R., Zeckhauser (2000). “The anatomy of health insurance”, *In Handbook of Health Economics*, 1:563-643.

De Jong, P., and G., Heller (2008). *Generalized linear models for insurance data*, Cambridge University Press.

Denuit, M., D., Hainaut and J., Trufin (2019). *Effective statistical learning methods for actuaries I: GLMs and extensions*. Springer Publishing.

Einav, L. and A. Finkelstein (2011). “Selection in Insurance Markets: Theory and Empirics in Pictures.” *Journal of Economic Perspectives*, 25(1):115-138.

_____ (2018). “Moral Hazard in Health Insurance: What We Know and How We Know It”, *Journal of the European Economic Association*, 16(4):957-982.

- Goldburd, M., A., Khare, D., Tevet and D., Gullder (2019). *Generalized linear models for insurance rating*, Casualty Actuarial Society, 2nd edition.
- Gurmu, S. (1991). Tests for Detecting Overdispersion in the Positive Poisson Regression Model, *Journal of Business and Economic Statistics*, 9(2):215-222.
- Loesche W. (1986). "Role of Streptococcus mutans in Human Dental Decay", *Microbiological Review*, 50(4):353-380.
- Kaas, R., M., Goovaerts, J., Dhaene and M., Denuit (2008). *Modern Actuarial Risk Theory Using R*, Springer Publishing, 2nd edition.
- Ohlsson, E. and B., Johansson (2010). *Non-Life Insurance Pricing with Generalized Linear Models*. Springer Publishing.
- Pauly, M. (1968). "The Economics of Moral Hazard: Comment", *American Economic Review*, 58(3):531-537.
- Rothschild, M., and J., Stiglitz (1976). "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information", *Quarterly Journal of Economics*, 90(4):629-649.
- Schatzle, M., et al. (2004). "The clinical course of chronic periodontitis", *Journal of Clinical Periodontology*, 31(12):1122-1127.

Abstract

The relative risks of disease-free groups with different periodontal disease treatment observation periods - five years and one year before the base year - with the National Health Insurance Cohort 2.0 DB for dental insurance that covers periodontal caries and disease were calculated using GLM. It was shown that the relative risk of the one-year disease-free group to the five-year disease-free group were 1.0026 and 1.0049 for filling treatments for periodontal caries and disease. However, the relative risk of prosthetic treatments for periodontal caries and disease were 1.053 and 1.103, which were higher than filling treatment cases. These results suggest that the additional premium for prosthetic treatment will be higher than that for filling treatment if the observation period for periodontal disease treatment shortens from 5 years to 1 year and the eligible group for dental insurance expands.

※ Key words: big data, dental insurance, generalized linear model, national health insurance, relative risk

