

## 우리말 사업보고서 가독성 연구의 가능성에 대한 탐색적 연구\*

### An Explorative Study to Assess Feasibility of Annual Report Readability Research in Korea\*

(제1저자) 정 태 진 Taejin Jung\*\*

(교신저자) 임 승 연 Seung Yeon Lim\*\*\*

(공동저자) 이 우 종 Woo-Jong Lee\*\*\*\*

(공동저자) 조 미 옥 Meeok Cho\*\*\*\*\*

개요: 본 연구에서는 영문 기반의 사업보고서의 가독성을 측정한 기존 연구의 결과가 우리말 사업보고서에도 확장가능한지 가늠해보기로 한다. Li(2008)의 연구 이후, 회계와 재무 분야의 많은 연구에서 사업보고서의 가독성을 비롯한 기업공시의 언어적 특성(linguistic features)에 대한 연구가 폭넓게 이루어졌으나, 상이한 전자공시체계로 인한 분석의 용이성 차이 때문에 국내에서는 이와 관련된 연구가 미비한 상황이다. 본 연구는 Li(2008)의 방법론을 차용하여 국내 상장기업들을 대상으로 우리말 사업보고서의 가독성을 측정하고, 이익특성 변수들 간의 관계를 확인하고자 하였다. 주요 분석결과에 따르면, 상대적으로 낮

논문접수: 2018. 02. 19      1차수정: 2018. 06. 09      게재확정: 2018. 07. 03

\* 본 논문은 2016년 대한민국 교육부와 한국연구재단의 지원을 받아 수행된 연구입니다 (NRF-2016S1A5A8019273).

\* 심사과정에서 유익한 조언을 해주신 세 분의 심사위원들과 백복현 편집위원장님께 감사드립니다. 또한 본 연구에 귀한 조언을 주신 서울대학교 브라운백 세미나 참석자들과 2018년 한국회계학회 하계국제학술대회 참가자들에게도 감사드립니다.

\*\* 서울대학교 경영대학 박사과정 Ph.D. Candidate, School of Business Administration, Seoul National University (E-mail: taejin.jung@gmail.com)

\*\*\* 국민대학교 경영대학 부교수 Associate Professor, College of Business Administration, Kookmin University (E-mail: sylim11@kookmin.ac.kr)

\*\*\*\* 서울대학교 경영대학 부교수 Associate Professor, School of Business Administration, Seoul National University (E-mail: woojong@snu.ac.kr)

\*\*\*\*\* 서울대학교 경영대학 박사과정 Ph.D. Candidate, School of Business Administration, Seoul National University (E-mail: meeok.cho@gmail.com)

은 이익을 보고하는 기업의 경우 사업보고서의 가독성이 낮은 것으로 나타났다. 또한, 사업보고서의 가독성이 낮은 기업의 경우, 이익의 지속성 역시 떨어진다는 점을 확인하였다. 마지막으로 이익조정 유인이 큰 기업들의 사업보고서의 가독성이 떨어진다는 점을 확인하였다. 이는 영문 기반의 사업보고서 가독성을 분석한 Li(2008)의 연구 및 후속 연구들의 결과가 우리말 기반의 사업보고서에서도 동일하게 적용될 수 있는 가능성을 시사한다. 본 연구는 우리말 사업보고서의 가독성과 관련된 텍스트 분석의 확장 가능성을 체계적으로 제시하였다는 점에서 의미가 있다.

한글색인어: 사업보고서 가독성, 기업공시, 이익, 이익의 질

**ABSTRACT:** This paper assesses feasibility of annual report readability research in Korea. Using three different readability measures commonly used in accounting and finance literature (e.g., Flesch ease index, Fog index, and the length of annual report), we find that firms with lower earnings make their annual reports more complex and thus less readable. Additionally, we find that firms with less readable annual reports have less persistent positive earnings. We also present evidence consistent with prior studies that annual reports are less readable for firms just meeting or beating earnings benchmarks. Furthermore, the relation between annual report readability and corporate earnings is attenuated after IFRS adoption. Lastly, to check the validity of annual report readability measures estimated by using the Korean alphabets, we re-estimate our main analyses using alternative annual report readability measures (e.g., alternative Fog indices, average sentence length, and file size of annual reports); these results are robust to the alternative readability measures. In sum, we provide first large sample evidence of how annual report readability is related to attributes of corporate earnings in Korea, thereby being suggestive of feasibility of annual report readability research in Korea.

**Keywords:** Annual report readability, Corporate disclosure, Earnings, Earnings quality.

## I. 서 론

회계는 경영의 언어이다. 회계는 재무정보의 생산자(내부자)와 소비자(외부자)가 소통하는 프로토콜(protocol)을 제공하는 체계이기 때문이다. 회계정보는 원천적으로 비언어적 정보(즉, 숫자)로 요약됨에도 불구하고, 공시과정에서는 회계정보가 사회구성원들이 공유하는 언어를 매개로 전달된다. 따라서 한 사회가 공통적으로 사용하는 언어의 특성은 필연적으로 회계정보가 소비되는 양식을 결정한다.

그러므로 기업공시에 대한 연구에서 각 보고서가 구사하는 언어의 특성을 이해하는 것은 매우 중요한 일이다. 과거 미국 실증연구들에서 기업공시 중 Management Discussion and Analysis section(MD&A)나 주석에 초점을 두고 가독성을 평가하려는 다양한 시도가 있었으나(예: Jones and Shoemaker 1994; Bryan 1997), 사업보고서 가독성에 대한 보다 엄밀한 실증연구들은 Li(2008)가 언어학의 텍스트 분석 기법을 기반으로 가독성 측정치를 제시한 이후에 폭발적으로 증가하였다. Li(2008)는 이익의 수준이 낮거나 이익의 지속성이 낮아서 경영자가 정보를 투명하게 공시할 유인이 부족할수록 가독성이 낮다는 사실을 보고하였다. Li(2010)은 더 나아가 공시의 가독성이 투자나 재무활동에 반영된 경영자의 기회주의적인 특성과도 관련이 있다고 주장하였다. 이처럼 기업공시의 가독성은 비언어적 회계정보의 특성 뿐만 아니라, 이에 반영된 경영자의 유인을 이해하는데 의미있는 정보를 제공한다.

미국에서 가독성 연구가 폭발적으로 진행될 수 있었던 것은 전자공시시스템이 체계적으로 구성되어 있기 때문이었다. 미국의 전자공시시스템(Electronic Data Gathering, Analysis, and Retrieval system, 이하 “Edgar”)은 기업의 재무제표 뿐만 아니라 전화회의록(conference call transcript), 재무분석가 보고서 등 다양한 텍스트 정보를 일정한 법칙에 따라 체계적으로 디렉토리(directory)에 저장함으로써, 연구자들이 컴퓨터 코딩을 활용하여 대표본의 자료를 일괄적으로 처리할 수 있는 환경을 제공한다. 이를 이용하여 연구자들은 언어학계에서 발전시켜온 기법들을 적용하여 공시의 양(the amount of disclosure)과 어조(the tone), 그리고 가독성(readability)에 대한 측정치들을 구성해내었으며, 이를 전통적인 회계학 연구의 관심변수들과 관련시키며 활발한 연구를 진행하고 있다.

이처럼 미국에서 영문 기반의 가독성 연구들이 봇물 터지듯 생산되고 있는 반면, 국내에서는 가독성 연구가 활성화되지 못하고 있는 실정이다. 이는 영어와 한글의 언어적 차이 때문에 우리말 분석기법의 발달이 상대적으로 저조하기 때문이기도 하지만, 금융감독원의 전자공시시스템(Data Analysis, Retrieval and Transfer System, 이하 “DART”)이

제공하는 텍스트 분석환경이 미국의 시스템에 비하여 연구친화적이지 않기 때문이기도 하다.<sup>1)</sup> 본 연구는 국내 사업보고서 가독성 연구 활성화에 기여할 목적으로 우리말 기반의 사업보고서에서 주요 선행연구의 결과가 유지되는지 검토함으로써, 국내에서의 가독성 연구의 가능성을 타진하는 것을 그 목적으로 한다.

전자공시시스템(DART)에서 사업보고서를 제공하기 시작한 1998년부터 2015년까지 국내 유가증권시장 및 코스닥시장에 상장된 기업의 사업보고서를 대상으로 분석한 결과는 다음과 같다. 먼저, 회계와 재무 관련 연구에서 널리 사용되는 세 가지의 서로 다른 가독성 측정치를 사용하여 분석한 결과, 상대적으로 낮은 이익을 보고하는 기업의 경우 사업보고서의 가독성 역시 낮은 것으로 나타났다. 여기에 더하여, 사업보고서의 가독성이 낮은 기업의 이익은 그 지속성이 상대적으로 떨어진다는 점을 확인하였다. 이러한 결과는 Li(2008)의 연구 및 후속 연구들의 결과가 국내 기업들을 대상으로도 적용될 수 있다는 증거라고 해석할 수 있다. 뿐만 아니라, 이익조정 유인이 상대적으로 큰 기업들의 경우 사업보고서의 가독성이 떨어지는 것으로 나타나, 이러한 기업들의 이익의 질(earnings quality)이 낮다는 점을 확인하였다. 마지막으로 국제회계기준(K-IFRS)의 도입 이후, 사업보고서의 가독성과 이익 간의 관계는 약해지는 것으로 나타났다. 이러한 결과들은 영문 기반의 가독성 선행연구의 방법론이 우리말 기반의 텍스트 분석을 통하여 국내에서도 의미있게 확장될 수 있음을 시사하는 것이다.

본 연구의 가장 중요한 공헌점은 국내에서의 가독성 연구의 가능성을 타진하고 국내기업 대표본에서의 체계적인 증거를 제공하는 최초의 연구라는 점이다. 이는 Li(2008)에서 촉발된 미국에서의 가독성 연구를 국내에서 확장하는 의미있는 시도이며, 국내에서 가독성 측정치의 타당성(validity)과 적용가능성(feasibility)을 논의하는 것으로 이해할 수 있다.

물론 본 연구가 가독성과 관련된 연구흐름에 일차적으로 관련되어 있지만, 공시언어의 특성을 분석하는 독립적인 연구흐름에도 직접적인 관련이 있다. 예를 들어, 미국시장에 교차상장된 비영어권 기업들을 대상으로 언어적 특성의 유사도가 재무제표의 이해가능성에 미치는 영향을 검증하는 연구들은 공시언어의 특성이 국제회계기준의 도입과정에서 지니는 함의에 대하여 논의한다(Lundholm et al. 2014; Brochet et al. 2016). 또한 각 언어 간 특성의 차이가 국가별 재무보고에 미치는 영향을 검증하는 형태로 진행되는 최근의 연구들(예: Kim et al. 2017)도 언어적 특성이 회계정보의 질적 속성과 관련되어 있다는 점을 시사한다. 본 연구와 가장 직접적으로 관련되어 있는 Lang and Stice-Lawrence(2015)는 미국을 제외한 42개국의 15,000개 이상의 기업을 대상으로

1) 본 연구는 국내에서 가독성 연구가 적용가능한지 검토하는 것이 목적이므로, 왜 가독성 연구가 활성화되고 있지 못하고 있는지를 논의하는 것은 본 연구의 주제와 거리가 있다. 본 연구의 후속 장에서 미국과 한국 전자공시시스템의 구조적 차이에 대하여 더 자세히 논의할 것이다. 그러나 본 연구의 결과가 전자공시시스템을 특정 방식(예: 연구친화적)으로 개선해야한다는 점을 제안하는 것으로 해석될 필요는 없다. 전자공시시스템은 일차적으로 사용자 친화적으로 구성되는 것이 목표가 되어야할 것이기 때문이다.

사업보고서의 가독성을 분석하였다. 이는 사업보고서의 가독성을 포함한 공시의 언어적 특성과 관련된 연구가 미국 뿐만 아니라 서로 다른 언어를 사용하는 다른 국가들에서도 적용될 수 있음을 보여주는 실증적인 증거라고 할 것이다. 따라서 영문 기반의 공시분석기법을 한글 사업보고서에 적용해보는 본 연구는 Li(2008)과 Land and Stice-Lawrence(2015)의 연구결과가 국내 대표본에서 유지되는지 탐색하는데 그 의의가 있다.

본 연구에서는 후속 연구를 활성화하기 위한 목적으로 본 연구가 가독성 측정치를 구성하기 위하여 사용했던 절차와 방법론, 분석에 사용된 파이썬(Python)기반의 텍스트 분석코드와 이를 구동한 전산환경들에 대한 구체적인 논의들을 제공하고 있으나, 본 연구진의 기술적 미숙함으로 인하여 절차의 효율성 측면에서 상당한 한계가 있다는 점을 인지할 필요가 있다. 따라서 본 연구가 제공하는 분석의 절차는 한 가지 예시일 뿐이며, 그 효율성에 한계가 있을 수 있음을 밝힌다.

본 연구의 2장에서는 후속 연구들을 고찰하여 가설을 설정하고, 3장에서는 국내 사업보고서 가독성 측정치를 어떻게 구성했는지 구체적인 절차를 기술한다. 4장에서는 구성된 가독성 측정치를 이용하여 주요 선행연구의 결과를 재현해보는 방식으로 국내 가독성 연구의 확장가능성을 타진하며, 5장에서 본 연구를 요약하고 시사점을 정리하기로 한다.

## II. 선행연구의 검토 및 가설설정

### 2.1 기업공시의 언어적 특성(linguistic features)과 관련된 연구

기업 내부의 회계정보는 기업의 공시를 통해 정보의 소비자인 외부투자자나 정보이용자에게 전달된다. 이 과정에서 비언어적 정보인 회계정보는 각 기업이 정보를 공시하는 방식이나 특성에 따라 소비자들에게 전달되는 방식이나 이를 소비하는 방식이 달라진다고 할 수 있다. 이러한 측면에서 기업공시가 지니는 언어적 특성을 파악하는 것은 회계정보를 통해 생산자와 소비자가 소통하는 방식을 보다 폭넓게 이해하도록 한다는 점에서 중요한 역할을 한다고 할 것이다. 각 기업은 이러한 목적 하에 매년 사업보고서를 발행하고 이를 공시하고 있다. 비단 비언어적 정보인 회계정보 외에도, 사업보고서는 기업의 수익성이나 위험 등에 대해 비교적 자세하게 기술하고 있으며, 소비자(정보이용자)들은 이러한 사업보고서를 통해 각 기업이 처한 상황에 대해 보다 깊이있게 이해할 수 있게 된다. 이처럼 사업보고서의 공시는 회계정보의 소비자라고 할 수 있는 투자자와 정보이용자에게 지대한 영향을 미치고 있다. 이러한 관점에서 최근 회계와 재무 분야에서는 사업보고서를 비롯한 기업공시(corporate disclosure)의 언어적 특성(linguistic features)이 투자자와 정보이용자의

의사결정에 미치는 영향에 관한 연구가 활발하게 이루어지고 있다. 기존 연구에서는 기업이 공시를 하는 횟수나 공시의 양(the amount of disclosure)을 이용하여 투자자나 정보이용자의 반응을 연구하였다(Jones and Shoemaker 1994; Smith and Smith 1971; Healy and Palepu 2001). 한편, Core(2001)가 컴퓨터 언어(computational linguistics)를 활용하여 기업공시의 언어적 특성에 대한 연구의 가능성을 제시한 이래, 최근의 많은 연구는 텍스트분석(textual analysis) 기법을 활용하여 기업공시의 언어적 특성이 기업의 행동이나 투자자와 정보이용자에게 미치는 영향에 대해 분석하고 있다(Li 2008; Hoberg and Phillips 2010; Li et al. 2013; Hoberg et al. 2014; Loughran and McDonald 2014; Hoberg and Phillips 2016).

특히, Li(2008)은 1994년부터 2004년까지 미국 내 기업을 대상으로 이익이 상대적으로 낮거나 지속성이 낮은 기업의 경우, 사업보고서의 가독성이 떨어진다는 점을 확인하였다. 이는 기존의 연구가 1) 소표본(small-sample)을 활용하여 기업공시의 언어적 특성이 기업의 행동에 미치는 영향을 분석하거나 2) 사업보고서의 특정 부분(section)이 미치는 영향을 분석하는데 그쳤던 것에 비교해볼 때, 대표본을 활용하여 사업보고서의 전반적인 가독성이 미치는 영향을 확인한 최초의 연구라고 할 것이다.

Li(2008)의 연구 이후, 기업공시의 언어적 특성이 기업의 의사결정이나 투자자/정보이용자에게 미치는 영향과 관련된 많은 연구가 폭발적으로 증가하였다(Li 2010b; Loughran and McDonald 2016). 그 중, 기업공시의 가독성(readability)과 관련된 연구가 활발하게 이루어졌다. 먼저, Lee(2012)는 기업의 분기공시(10-Q)의 가독성과 주가와의 관계에 대하여 분석하였다. 그의 연구에 따르면 길이가 길고 내용이 복잡한 분기 보고서의 경우 주식시장의 투자자들이 이를 제대로 이해하지 못하며, 그 결과 기업의 이익과 관련된 정보를 보다 적시에 반영하지 못한다는 결과를 보고하였다. 비슷한 맥락에서 Kim et al.(2018)은 경영자가 가독성이 떨어지는 공시를 주식시장 투자자들로부터 부정적인 정보(adverse information)를 숨기는 수단으로 활용할 수 있고, 이러한 현상이 결국 주가폭락현상으로 연결된다는 실증적 증거를 제시하였다. 또한 You and Zhang(2009)은 연차 사업보고서(10-K)의 가독성에 대한 시장반응을 연구하였다. 연구에 따르면 사업보고서의 길이가 길수록, 기업의 이익에 대해 주식시장 투자자들이 과소반응(underreaction)하는 것으로 나타났다. 이를 개인투자자 측면에서 살펴본 연구도 존재한다. Lawrence(2013)는 개인투자자들이 보다 간결하고 읽기 쉬운 공시를 보고하는 기업들에 더 많이 투자한다는 사실을 보고하였다. 이는 주식시장의 개인투자자들 역시 사업보고서의 가독성에 영향을 받는다는 점을 밝혀냈다는 데 그 의의가 있다고 할 것이다. 뿐만 아니라, 사업보고서의 가독성은 주식시장 투자자 외에 채권시장 투자자들에게도 영향을 미치는 것으로 나타나고 있다. 사업보고서의 가독성이 낮은 경우, 신용등급 역시 낮은 것으로 나타나고 있으며, 이는 결국 높은 타인자본비용으로 연결되는 것으로 보고되고 있다(Bonsall and Miller 2017; Ertugrul et al. 2017).

한편, 가독성 외에도 공시에 담긴 경영자의 어조(tone)에 대한 연구도 점차 증가하고 있다. Li(2010a)는 미국 기업의 사업보고서 중 MD&A에 담긴 경영자의 미래 전망에 대한 어조를 분석하였다. 연구에 따르면, 미래 이익이 성장할 것으로 예상될 때, 경영자들이 보다 낙관적인 어조로 공시를 한다는 점을 확인할 수 있다. 뿐만 아니라, 현재의 실적이 좋고, 발생액이 작으며, 기업규모가 작은 기업의 경영자는 보다 긍정적인 어조로 MD&A를 작성한다는 점을 보고하였다. Kothari et al.(2009) 역시 경영자, 재무분석가, 그리고 언론의 다양한 공시 보고서를 활용하여 공시에 담긴 어조가 보다 호의적인(favorable) 경우, 자본비용이 감소하고, 주가의 변동성이 줄어들며, 재무분석가 예측치의 분산값(analyst forecast dispersion) 역시 감소한다는 점을 확인하였다. 이처럼 단순히 가독성 외에도 경영자의 어조를 비롯한 다양한 언어적 특성이 투자자 및 정보이용자의 의사결정에 영향을 미친다는 것을 확인할 수 있다.

또다른 연구들은 공시의 언어적 특성이 이익의 질이나 이익조정행위와 관련이 있다는 증거를 제시하였다. Lo et al.(2017)은 전기 사업연도의 이익(prior year's earnings)을 달성하기 위하여 이익을 조정하는 기업이 MD&A 공시를 보다 복잡하게 만든다는 실증적 증거를 제시하였다. 이는 기업의 실적이 좋은 경우 사업보고서의 가독성이 높다는 Li(2008)의 주장과 관련하여 이익조정과 사업보고서의 가독성 간의 관계를 고려해야 한다는 점을 보여주고 있다. 뿐만 아니라, Li et al.(2011)의 연구는 기업이 처한 경쟁환경에 대한 경영자의 공시내용이 이익의 질과 관련이 있다는 증거를 제시하였다.

보다 최근의 연구들은 사업보고서의 가독성을 비롯한 공시의 언어적 특성이 투자자 뿐만 아니라 기업의 여러 이해관계자들에게 미치는 영향에 대해서도 분석하고 있다. 먼저 Leheavy et al.(2011)은 사업보고서의 가독성이 떨어지는 경우, 기업을 담당하는 재무분석가의 수(analyst following)가 늘어날 뿐만 아니라 그들의 분석보고서가 갖는 정보의 유용성이 증가한다는 결과를 보고하였다. 이는 사업보고서의 가독성이 떨어지는 경우에 투자자들이 재무분석가의 분석에 보다 의존한다는 실증적인 증거라고 할 수 있다. 한편, Bao et al.(2015)의 연구에서는 산업전문감사인이 보다 복잡한 사업보고서를 공시하는 기업에 대하여 더 높은 감사보수 프리미엄(audit fee premium)을 청구한다는 사실을 확인하였다. 이를 통해 외부감사인 역시 기업의 사업보고서 공시에 영향을 받는다는 점을 확인할 수 있다.

비단 미국 뿐만 아니라 전세계 다른 국가들을 대상으로 공시의 언어적 특성을 분석하는 연구가 시작되고 있다. 특히 Lang and Stice-Lawrence(2015)는 미국을 제외한 42개국의 15,000개 이상의 기업을 대상으로 사업보고서의 가독성을 분석하였다. 이는 사업보고서의 가독성을 포함한 공시의 언어적 특성과 관련된 연구가 미국 뿐만 아니라 서로 다른 언어를 사용하는 다른 국가들에서도 적용될 수 있음을 보여주는 실증적인 증거라고 할 것이다. 이처럼 비영어권 국가의 공시가 갖는 언어적 특성에 대한 연구는 각 언어간 특성의 차이가 국가별 재무보고에 미치는 영향을 검증하는 최근의 연구흐름과 맞물려 그 중요성이

커지고 있다(Kim et al. 2017).

종합하면, 사업보고서의 가독성을 비롯한 공시의 언어적 특성은 기업의 의사결정 뿐만 아니라 다양한 이해관계자들의 행동에 영향을 미치고 있다. 특히 텍스트분석기법(textual analysis)의 활용이 널리 확대됨에 따라 대표본을 활용한 연구가 활발하게 이루어지고 있으며, 이는 국내에서도 텍스트 분석을 활용한 공시의 언어적 특성에 대한 연구의 필요성이 점차 증대되고 있음을 의미한다고 할 것이다.

## 2.2 기업공시의 언어적 특성과 관련된 국내연구

Section 2.1.에서 살펴본 바와 같이 최근 회계와 재무 분야에서는 텍스트분석(textual analysis)을 활용한 공시의 언어적 특성에 대한 분석이 널리 이루어지고 있다. 이와는 달리 국내에서는 텍스트분석을 활용한 연구가 상대적으로 미비한 실정이다. 이는 영어와 한글의 언어적 차이 때문에 분석기법의 발달이 상대적으로 저조하기 때문이기도 하지만, 현행 전자공시시스템 자체가 대표본을 활용한 텍스트분석 연구에 적합하지 않기 때문이기도 하다. 현행 전자공시시스템은 대량의 기업공시서류를 다운로드하는 것을 허용하지 않고 있다(김형준 등 2015). 뿐만 아니라, 김형준 등(2015)의 연구결과에 따르면, 웹 크롤링(web-crawling) 등의 우회적인 방법을 활용한다고 하더라도 DART에 공시된 모든 분기, 반기 및 사업보고서들 중 약 40%가 판독 자체가 불가능하며, 설령 판독이 가능한 보고서들의 경우에도 PDF 형식으로 제공되어 효과적인 텍스트 분석이 어려운 것으로 나타나고 있다. 이러한 이유 때문에 국내에서는 전반적인 사업보고서의 가독성보다는 사업보고서 내 특정부분(예: MD&A 공시)의 공시나 한국거래소의 공정공시정보 상 공시의 빈도와 관련된 연구들이 일부 존재하고 있다.

Ahn and Lee(2005)은 사업보고서 내 “영업의 개황(overview of operation)” 부분을 MD&A 공시에 해당하는 것으로 보고 이를 활용하여 경영자가 자발적으로 공시하는 동기를 조사한 바 있다. 또한, 이아영 등(2013)의 연구에서는 2009년 이래 금융감독원이 제시한 MD&A 작성 가이드라인이 기업별로 준수되고 있는 정도를 측정하여, 이를 MD&A 공시에 대한 품질로 활용하였다. 그 결과 기업규모가 크고, 부채비율이 높고, 기업성파가 좋은 기업일수록 MD&A 공시품질이 높은 것(즉, 금융감독원이 제시한 가이드라인을 준수하는 것)으로 나타났다. 한편, 권오진과 박하연(2017)의 연구에서는 2010년부터 2014년까지 사업보고서 중 “이사의 경영진단 및 분석의견”과 “그 밖의 투자자보호를 위하여 필요한 사항”의 공시를 MD&A 공시의 대응치로 보고 경쟁이 심한 산업에 속한 기업일수록 MD&A의 공시수준이 낮다는 점을 실증적으로 제시하였다. 이는 경쟁이 심한 산업의 경우, MD&A 공시로 인하여 경쟁기업들에게 내부 정보가 노출될 가능성이 높고, 이에 따라 산업 내에서 경쟁적 지위가 약화되는 것을 기피하려는 경향이 있음을 보여주는 결과라고



할 수 있다. 이와는 별개로 Hyeon et al.(2014)의 연구에서는 MD&A 공시에 담긴 비재무적 정보(nonfinancial information)를 공시하는 결정요인에 대한 실증적인 결과를 제시하였다. MD&A 공시에 비재무적 정보를 공시하는 빈도를 활용하여 분석한 결과, 무형자산이 많고, 이사회가 감시가 효과적이고(higher quality), 규모가 큰 기업일수록 MD&A 공시에 비재무적 정보를 공시한다는 점을 제시하였다.

또 다른 국내연구들은 한국거래소의 공정공시정보를 활용한 공시의 빈도에 초점을 맞추어 이루어졌다. 심호식 등(2010)은 지배주주의 의결권과 배당권의 차이인 지분율과 괴리도를 활용하여 기업지배구조의 수준을 측정하고, 기업지배구조의 수준과 한국거래소의 공정공시정보의 빈도 사이의 관련성을 조사하였다. 이를 통해 지분율 괴리도가 증가할수록 공시빈도가 감소한다는 점을 실증적으로 확인하였다. 또한 이문영 등(2012) 역시 대규모기업집단 소속기업들의 이사회 특성을 이용하여 측정한 기업지배구조 수준과 공정공시의 빈도 사이의 관련성을 확인하였고, 그 결과 이사회 중 사외이사의 비율이 높을수록, 회색이사의 비율이 높을수록 공정공시의 빈도가 높음을 확인하였다. 그러나 이와 같은 연구는 한국거래소를 통해 공시되는 공정공시의 빈도를 측정한 결과로, 사업보고서를 비롯한 공시의 언어적 특성을 측정하는 연구의 흐름과는 간접적인 관련이 있다고 할 것이다.

한편, 국내에서는 2011년부터 국제회계기준(K-IFRS)이 전면 도입되었다. K-IFRS 회계기준은 재무보고의 이해도를 높이고 회계기준의 복잡성을 감소시키기 위해 많은 노력을 기울인 것으로 알려져 있다. 이로 인해 K-IFRS 도입 이후 재무제표의 주석이 담고 있는 정보는 점차 증가하는 것으로 보인다. K-IFRS 도입이 재무제표의 주석의 가독성에 미치는 영향을 실증적으로 확인하기 위하여 Jang and Roh(2016)의 연구에서는 57개의 K-IFRS 도입 기업과 943개의 K-GAAP 기업 재무제표의 주석의 가독성을 분석하여 K-IFRS 재무제표의 주석이 K-GAAP 재무제표 주석에 비해 가독성이 떨어진다는 점을 확인하였다. 이는 IFRS 회계기준이 재무보고의 이해도를 향상시키고 회계기준의 복잡성을 감소시킨다는 주장과는 상반되는 결과로, IFRS의 도입이 비영어권 국가(non-English)의 언어로 변환되는 과정에서 오히려 낮은 가독성을 지니게 된다는 점을 확인할 수 있다.

이상에서 살펴본 바와 같이, 공시의 언어적 특성에 대한 연구가 활발히 이루어지는 다른 국가와 달리 국내에서는 관련연구가 비교적 미흡한 실정이다. 이에 본 연구에서는 국내 유가증권시장 및 코스닥 시장에 상장된 모든 상장기업의 사업보고서를 활용하여 공시의 언어적 특성 중 사업보고서의 가독성과 관련된 실증적 연구결과를 제시함으로써, 영문 기반의 사업보고서 가독성을 분석한 Li(2008)의 연구 및 후속연구들의 결과가 우리말 기반의 사업보고서에서도 동일하게 적용될 수 있는 가능성을 탐색해보고자 한다. 이를 통해 추후 국내에서 텍스트 분석을 활용한 연구의 확장 가능성을 제시하는 계기로 삼고자 한다.

### 2.3 가설의 설정

Li(2008)의 연구에서 경영자는 수익성이 좋지 않을 때 해당 정보가 주가에 보다 늦게 반영될 것이라며 투자자들의 해석에 시간이 소요되는 보다 복잡한 정보를 제공하게 되고 이로 인해 공시의 가독성이 떨어지는 것으로 나타난다(management obfuscation hypothesis). 이는 정보가 복잡한 경우 그 해석에 많은 시간과 비용이 소요되고 이로 인해 정보가 가격에 반영되는 속도와 정도가 상대적으로 낮다는 점에 근거한다(Bloomfield 2002). 결국 수익성이 좋은 경우 경영자들은 정보가 시장에 적시에 반영되기를 바라는 반면, 반대의 경우 정보가 시장가격에 지연되어 반영되기를 바라는 것이다. 이러한 주장에 따르면 기업이 공시하는 사업보고서의 가독성의 수준과 기업의 이익은 양(+ )의 관계가 있을 것이다. Li(2008)의 연구에서는 사업보고서의 가독성과 기업의 이익 간에 양(+ )의 관계를 제시한 바 있다. 이는 국내의 기업들을 대상으로 하더라도 비슷한 결과를 가져올 것으로 예상된다.

반면, 사업보고서의 가독성과 기업의 이익 간의 관계는 아래와 같은 이유로 쉽게 예측이 어려울 것으로 보인다. 먼저 사업보고서는 이익을 비롯한 재무정보 외에 다양한 비재무적 정보를 포함하고 있다. Hyeon et al.(2014)의 연구에서 확인할 수 있는 바와 같이 사업보고서에 포함되는 정보는 다양하며, 기업특성에 따라 이를 공시하는 수준이 다르다. 뿐만 아니라 사업보고서는 현재 시점의 정보 외에도 과거의 정보와 미래에 대한 예측정보를 모두 담고 있다. 이러한 사업보고서의 특성에 비추어 본다면, 단순히 경영자가 수익성이 좋지 않을 경우 가독성이 떨어지는 공시를 통해 투자자나 정보이용자에게 혼란을 준다고 예측할 수는 없을 것이다. 또한 기업이 공시하는 사업보고서는 표준 문안(boilerplate)을 반영하는 경향이 있다. 만약 경영자들이 법적 책임을 감소시키기 위하여 표준 문안에 맞게 사업보고서를 공시한다면 이는 기업의 수익성과는 별다른 관계가 없을 것으로 예측할 수 있다(Hoogervorst 2013). 미국을 제외한 42개국 15,000개 이상의 기업을 대상으로 사업보고서를 분석한 Lang and Stice-Lawrence(2015)의 연구에서도 회계기준이 취약한 국가에서는 사업보고서를 표준문안에 맞게 작성하는 현상이 강하게 나타난다고 보고하고 있다.

종합하면 사업보고서의 가독성과 이익 사이의 관계는 서로 다른 예측이 가능하다. 또한 국내에서는 이와 관련된 연구가 아직까지 이루어지지 않았으므로 아래와 같이 귀무가설의 형태로 첫 번째 가설을 설정하였다.

가설 1: 사업보고서의 가독성과 기업의 이익은 유의한 관계가 없다.

사업보고서의 가독성과 기업의 이익 간의 관계는 기업의 미래이익에 대한 분석으로 확장이 가능하다(Sloan 1996; Li 2008). 이 중 이익의 지속성과 관련하여 살펴보면, 기회주

의적인 경영자는 현재의 이익증가가 일시적이거나 혹은 현재의 이익감소가 영구적인 경우, 사업보고서의 가독성을 떨어뜨려 투자자나 정보이용자가 정보를 해석하는데 보다 많은 시간과 비용을 투입하도록 할 유인이 있다. 따라서 이익의 지속성이 낮다면 기업의 사업보고서는 보다 복잡한 형태를 나타낼 것이다. 마찬가지로 경영자가 이익을 조정하는 경우 이익의 지속성이 낮아지게 된다. 만약 이익조정을 하는 경영자가 사업보고서를 보다 복잡하게 구성한다면 결국 이익의 지속성 역시 낮아질 것으로 예측할 수 있다. 이와 같은 관점에서 사업보고서의 가독성과 이익의 지속성 간에 양(+)의 관계를 추론할 수 있다.

반면 가설 1에서 언급한 바와 같이 사업보고서의 가독성과 이익 간의 관계는 사업보고서가 담고 있는 방대한 정보와 표준문안 등의 특성으로 인해 단순히 예측할 수 없고, 이는 이익의 지속성에 대한 분석에서도 마찬가지로 적용될 수 있다. 이상의 논의에 따라 아래와 같이 두 번째 가설을 설정하였다.

가설 2: 사업보고서의 가독성과 기업의 이익의 지속성은 유의한 관계가 없다.

추가적으로, 본 연구에서는 사업보고서의 가독성과 이익의 질 간의 관계를 검증해보고자 한다. Li(2008)의 연구에서 사업보고서의 가독성과 이익 및 이익의 지속성 간의 관계를 살펴본 반면, 최근의 연구에서는 사업보고서의 가독성과 이익조정행위 간의 관계를 실증적으로 분석하였다(Lo et al. 2017). 본 연구에서는 Lo et al.(2017)의 방법론을 차용하여 이익조정유인이 큰 기업들의 경우(즉, 전기의 이익에 비해 소폭의 이익증가를 보인 기업들의 경우), 사업보고서의 가독성이 낮은지 검증해보고자 한다. 이익조정 유인이 이익의 품질을 결정하는데 결정적인 역할을 한다는 점을 감안하면, 이 분석은 사업보고서의 가독성을 이익의 품질의 맥락으로 확장하는 의미를 가진다고 할 것이다.

먼저, 경영자가 이익조정행위를 하는 경우 경영자는 이익조정행위라는 부정적인 정보가 시장가격에 보다 늦게 반영되기를 바라는 유인이 있을 수 있다. 이는 경영자가 이익의 부정적인 정보가 시장에 보다 늦게 반영되기를 바라는 것과 비슷한 맥락이라고 할 수 있다. 따라서 경영자는 사업보고서의 해석에 필요한 시간과 비용을 높이기 위해 가독성이 떨어지는 사업보고서 공시를 제공할 것이다(Jo and Kim 2007). 이와는 달리, 경영자가 이익조정행위를 하는 경우 오히려 사업보고서를 단순하게 만들 가능성도 존재한다. 이는 사업보고서가 단순할수록 외부투자자나 정보이용자가 사업보고서 안에서 이익조정행위에 대한 단서를 찾기 어렵다는 것에서 기인한다(Lo et al. 2017). 여기에 가설 1 및 가설 2와 마찬가지로 사업보고서가 담고 있는 정보의 성격과 표준문안의 경향 등을 고려한다면 사업보고서와 이익의 질 간의 관계를 추론하는 것은 단순하지 않다. 따라서 아래와 같이 귀무가설의 형태로 가설 3을 설정하였다.

가설 3: 사업보고서의 가독성과 이익의 질은 유의한 관계가 없다.

### III. 표본의 설정 및 가독성의 측정

#### 3.1 표본의 설정

본 연구의 표본은 금융감독원의 전자공시시스템(DART)이 전자공시의 형태로 사업보고서를 제공하는 1998년부터 2015년까지 국내 유가증권시장 및 코스닥시장에 상장된 기업을 대상으로 한다. 본 연구를 수행하기 위해서는 우선 분석의 대상이 되는 사업보고서로 구성된 데이터를 구축해야 하는데, 현재의 전자공시시스템에서는 대량의 보고서들을 한 번에 내려 받을 수 있는 별도의 익명 FTP서버를 운용하지 않고 있다. 따라서 보고서의 원본 파일을 내려 받기 위한 우회적인 방안들이 필요한 실정이다. 현행 DART 시스템 하에서는 10초 미만의 간격으로 빈번하게 내려 받기 요청<sup>2)</sup>을 보낼 경우, 이를 DART 서버에 대한 공격으로 간주하여 해당 IP 주소를 차단하도록 되어 있다(김형준 등 2015). 따라서 이러한 웹 크롤링(web-crawling) 작업을 IP 주소에 대한 차단 없이 수행하기 위해서는 DART 웹사이트를 담당하는 금융감독원 내 담당자와의 협조를 필요로 한다. 이와 같이 DART 시스템이 제공하는 제한적인 정보환경으로 인한 문제를 해결하고, 향후 다양한 연구로의 확장을 위하여 본 연구에서는 국내 모든 유가증권시장 및 코스닥시장 상장기업의 사업보고서를 Portable Document Format(이하 “PDF”) 형식과 마이크로소프트사의 WORD 프로그램의 DOCX 형식으로 다운로드 받아 분석을 진행하였다. PDF 혹은 DOCX 형식의 사업보고서를 다운로드한 후에, 표본기업의 사업보고서 가독성(Readability)을 산출하기 위하여 파이썬(Python)을 이용하였다. 가독성 측정을 위한 과정은 Section 3.2. 및 <Appendix B>를 참고하길 바란다.

가독성 정보를 제외한 각 표본기업의 재무정보는 FnGuide가 제공하는 DataGuide에서 추출하였다. 결산시점의 차이로 인해 비교가능성이 저하되어 연구결과가 왜곡되는 것을 방지하기 위해 12월 결산법인이 아닌 기업은 표본에서 제외하였고, 금융업의 경우 영업상의 특성상 재무제표 계정 과목의 특성이 다른 업종들과 상이하기 때문에 표본기업에서 제외하였다. 보고기업이 회계오류 수정 등의 이유로 보고서를 기재 정정한 경우에는 최종적으로 공시된 보고서가 표본에 포함되도록 하였다. 더불어, 필요한 재무정보가 누락된 기업 역시 표본에서 제외하였으며, 마지막으로 분석에 사용된 변수 중 연속변수는 이상치(outlier)가 분석 결과에 미칠 수 있는 효과를 최소화하기 위해 각 변수의 상하 1% 이내에서 조정(winsorization)하였다. 상기 요건을 충족하는 기업-연 자료(firm-year data)의 재무정보를 추출한 결과, 최종표본은 21,917개로 구성되었다.

2) 본 연구에서는 이처럼 모든 원본파일을 내려받기 위한 작업들을 통칭하여 “웹 크롤링(web-crawling)”이라는 용어를 사용하고 있다.

### 3.2 사업보고서 가독성(annual report readability)의 측정

정보의 가독성이 낮을수록 비전문가의 접근성이 떨어질 뿐만 아니라, 전문가들이 정보를 활용하여 재생산함에 있어서도 제약이 존재할 수 밖에 없다.<sup>3)</sup> 이러한 관점을 반영하여 공시의 언어적 특성에 대한 연구들은 사업보고서의 가독성을 얼마나 글을 읽기 어려운가로 정의해 왔다. 특히 최근의 선행연구에서는 가독성을 정량적인 지표로 평가하기 위하여 단어당 음절의 수, 문장당 단어의 수, 난해한 단어의 비중 등을 추출해 평가하는 특정 기법들을 사용한 가독성 척도(Reading Ease Formulas)를 사용해왔다(Li 2008; Jang and Roh 2016). 단어의 길이는 개별 단어가 갖는 의미를 인지하는 속도와 연관이 있고 문장의 길이는 기억력과 연관이 있기 때문에, 음절의 수로 측정된 단어의 길이와 단어의 수로 측정된 문장의 길이는 다양한 언어의 가독성을 측정하는 데 널리 사용되어 왔다(Subramanian et al. 1993).

본 연구에서는 사업보고서 가독성을 측정하기 위하여 두 가지 가독성 척도(Flesch Reading Ease Formula 및 Fog Index)와 사업보고서의 길이 등 총 3가지의 대응치를 사용하였다. 우선, Flesch(1948)의 Flesch Reading Ease formula(*FLESCH*)는 가독성을 측정함에 있어 가장 널리 사용되어 왔고, 가장 강력한 것으로 알려져 있다(Lewis et al. 1986). Flesch 가독성 지수의 논거(rationale)는 길이가 짧은 단어와 문장이 더 읽기 쉽다는 것이다. 구체적으로 이는 평균 단어 길이(average word length, 이하 “*AWL*”)와 평균 문장 길이(average sentence length, 이하 “*ASL*”)의 두 언어학적 변수로 구성되며, 그 측정 공식은 다음과 같다.

$$FLESCH = 206.835 - 84.6 * AWL - 1.015 * ASL$$

평균 단어 길이(*AWL*)는 글의 음절수를 단어수로 나누어 측정하며, 평균 문장 길이(*ASL*)는 글의 단어수를 문장수로 나누어 측정한다.<sup>4)</sup> 본 연구에서는 하나의 사업보고서를 개별문장으로 분절하기 위하여 온점(“.”)을 기준으로 문장을 나누었다. 또한 낱자, 소수 등을 문장으로 분류하지 않기 위하여 온점 “.” 뒤에 숫자가 이어져 나오는 경우 및 사업보고서

3) 사이언스온, 2017.04.03.“난해해지는 논문, 가독성 점점 낮아져”  
(<http://scienceon.hani.co.kr/504841>)

4) 또 다른 유사한 가독성 측정방식으로 Flesch - Kincaid formula(*Kincaid*)가 있으며, 이는  $Kincaid = (11.8 * AWL) + (0.39 * ASL) - 15.59$ 의 수식을 통하여 계산되고, 그 결과치가 미국의 학년체계와 연계된다. 예를 들어 Kincaid가 8.0이라면, 미국의 8학년 학생들이 평균적으로 이해할 수 있는 글이라는 의미이다. 이와 다르게 Flesch Reading Ease는 100점 체계로 산출된다. Flesch - Kincaid formula와 Flesch Reading Ease의 구성요소는 평균 단어 길이(*AWL*)와 평균 문장 길이(*ASL*)로 같고 그 가중치와 부호가 해석의 차이를 발생시킨다.

내 표로 작성된 부분은 제외하고 *AWL* 및 *ASL*을 측정하였다(Li 2008; Leheavy et al. 2011). 이상의 내용을 바탕으로 사업보고서의 *FLESCH* 값을 측정한다면, 측정값이 높을 수록, 해당 사업보고서는 보다 읽기 쉬운 것으로 볼 수 있다. 그러나 Flesch 가독성 지수가 비영어권 언어 특히, 한글의 가독성을 측정하기 위해 만들어진 지수가 아니기 때문에, 이를 연구에 사용함에 있어 주의를 요한다. 구체적으로, 한글로 작성된 문서의 가독성을 Flesch 가독성 지수를 이용하여 측정할 때에는 평균 단어 길이(*AWL*)와 평균 문장 길이(*ASL*)의 가중치가 다르게 적용되어야 할 수도 있을 것이다. 하지만 한글 자모의 경우 단어를 음절로 구분하는 것이 영어를 포함한 다른 언어에 비하여 자연스럽다는 점을 고려한다면, Flesch 가독성 지수와 같은 음절기반(syllable-based)의 가독성 측정공식이 철자기반(spelling-based)의 가독성 측정공식에 비해 한글로 작성된 문서의 가독성을 측정하는데 좀 더 적합하고 편리하다고 할 수 있겠다. Jang and Roh(2016)의 연구에서도 Flesch 가독성 지수를 사용하여 재무제표의 주석(footnotes)에 대한 가독성을 측정한 바 있다.<sup>5)</sup> 본 연구에서는 연구결과가 Flesch 가독성 지수 외에 여러가지 가독성 측정방식에도 강건하게 나타난다는 점을 보여주기 위하여 Flesch 가독성 지수 외에도 Fog Index(*FOG*)와 사업보고서의 전체길이(*LENGTH*)의 두 가지 추가 대용치를 활용하여 사업보고서의 가독성을 측정하였다.

먼저, 기존 연구에서 가독성의 측정방식으로 가장 널리 이용되어 온 Gunning(1952)의 Gunning Fog Index(*FOG*)를 사용하였다. Fog Index는 글의 단어수를 문장수로 나누어 측정된 평균 문장 길이(*ASL*)와 3음절 이상의 단어의 비중을 사용하여 구한 단어의 복잡성(*complexity*)으로 측정하는데, 그 측정 공식은 다음과 같다.

$$FOG = (ASL + \text{percentage of complex words}) * 0.4$$

평균 문장 길이(*ASL*)의 경우 *FLESCH* 측정의 구성요소와 맥락이 같지만, 복잡한 단어의 비중(*percentage of complex words*)을 가독성의 구성요소로 포함시킨 부분은 Fog Index만의 특징이라고 할 수 있겠다. Fog Index는 문장을 읽고 이해하는 데 필요한 교육년수를 나타내는데, 19 이상이면 읽을 수 없는(*unreadable*) 수준의 글을 의미하고, 14-18이면 읽기 어려운(*difficult*) 수준, 12-14이면 이상적인(*ideal*) 수준, 10-12이면 받아들일 수 있는(*acceptable*) 수준, 마지막으로 8-10이면 어린아이(*childish*) 수준의 글을 의미한다. 이러한 방식의 해석은 측정된 Fog Index를 한글문서에 그대로 적용할 수 있는가에 대한 의문이 있을 수 있겠으나, 문장을 읽고 이해하는 데 필요한 교육년수를 그대로 의미할 수는 없어도 상대적인 척도로서의 Fog Index 정의를 그대로 이용할 수 있을

5) 미국 증권거래위원회(Securities and Exchange Commission, 이하 "SEC") 역시 정보이용자들이 여러 가독성 측정공식 각각의 특징 및 장단점들에 대하여 주의를 기울여야 한다고 언급하고 있다(SEC, 1998).

것이다 (Li 2008). 반면, 본질적으로 비즈니스 관련 문서들은 3음절 이상의 복잡한 단어들로 구성이 되어 있으나, Fog Index 측정상 복잡하다고 규정된 단어들이 투자자들이나 재무분석가들에게 그다지 어렵지 않은 단어일 수도 있을 것이다(Loughran and McDonald 2014). 게다가 한글로 된 문서에서 3음절 이상 길이를 가진 단어가 복잡한 단어인지를 별도로 고려해 보아야 할 것이다. 이러한 Fog Index의 측정상 오류를 최소화하기 위하여 Section 4.5에서는 복잡한 단어(complex word)의 정의를 달리하여 결과의 강건성을 검증하고자 하였다.

마지막으로 사업보고서 가독성의 대응치로 사업보고서의 전체길이(*LENGTH*)를 다음과 같이 측정하였다.

$$LENGTH = \log(Word\_num),$$

*Word\_num*는 문서에 사용된 단어의 수를 의미하여, *Word\_num*에 자연로그를 취하여 문서의 길이(*LENGTH*)를 계산하였다. 긴 문서인 경우 이를 해석하는데 드는 시간과 비용이 더 크다는 점을 고려한다면, 문서가 길수록 읽기 어렵다고 할 수 있을 것이다. 만약 경영진이 사업보고서의 이용자들에게 기업에 불리한 정보를 드러내지 않기를 원하는 경우, 사업보고서의 길이를 전략적으로 사용하여 사업보고서를 이해하기 어렵게 만들 유인이 있다. 이러한 관점에서 미국 증권거래위원회(*SEC*)는 기업들이 길이가 긴 문장과 문서들을 공시하는 것을 피해야한다고 지속적으로 주장해 왔다(*SEC* 1998). 다른 가독성 측정공식들에 비하여 문서의 길이(*LENGTH*)는 계산하기 쉽고 이해하기도 쉽지만, 엄밀하게 말하면 가독성보다는 공시의 양과 오히려 관련이 있다고 할 수 있어 다른 가독성 측정방법과 마찬가지로 결과 해석에 있어 유의해야 할 것이다.

종합하면, 본 연구에서는 사업보고서 가독성을 측정하기 위하여 *FLESCH*, *FOG*, *LENGTH* 등 3가지의 대응치를 사용하였다. 사업보고서를 읽기 쉬울수록, 또는 사업보고서의 가독성이 높아질수록, *FLESCH*는 커지고, *FOG*와 *LENGTH*는 작아진다. 본 연구의 분석결과의 해석함에 있어, 사업보고서의 가독성 대응치들을 모두 같은 방향으로 해석하기 위하여, *FLESCH*에 (-1)을 곱하였으며, 결국 3가지 대응치의 값이 커질수록 사업보고서는 가독성이 낮아짐(*unreadability*)을 의미할 것이다.<sup>6)</sup>

6) 본 연구의 주요 관심변수(variable of interests)인 사업보고서의 가독성 대응치, *FLESCH*, *FOG*, *LENGTH* 등 3가지는 그 값이 커질수록 사업보고서의 가독성이 낮아짐을 의미하고, 이에 대한 변수명으로 '*Unreadability*'를 사용하였다.

## IV. 분석결과

### 4.1 사업보고서 가독성의 결정요인

사업보고서의 가독성은 경영자의 의도와 관계가 있을 것으로 보이지만, 경영자의 의도와 관계없는 기업 고유의 특성에도 영향을 받을 수 있다. 경영자의 의도와 무관한 사업보고서 가독성의 결정요인을 분리하여 사전적으로 분석해보고, 향후 가설 검증시 사업보고서 가독성이 기업의 성과인 이익과 그 지속성에 미치는 영향을 보다 정치하게 분석하기 위하여 본 연구에서는 사업보고서 가독성의 결정요인들을 통제변수들로 활용하여 다변량분석을 수행하고자 한다. 이에 앞서, 사업보고서 가독성의 결정요인들을 확인하기 위하여 아래와 같은 연구모형을 설정하였다(Li 2008; Jang and Rho 2016).

$$\begin{aligned} Unreadability_{jt} = & \beta_0 + \beta_1 SIZE_{jt} + \beta_2 MTB_{jt} + \beta_3 LEVERATGE_{jt} + \beta_4 AGE_{jt} \\ & + \beta_5 SI_{jt} + \beta_6 RET\_Vol_{jt} + \beta_7 EARN\_Vol_{jt} + \beta_8 NBSEG_{jt} + \beta_9 SEO_{jt} \\ & + \beta_{10} BIGN_j + \beta_{11} IFRS_{jt} + Year\ Fixed\ Effects \\ & + Industry\ Fixed\ Effects + \varepsilon_{jt} \end{aligned} \quad (1)$$

식 (1)에서 종속변수인 *Unreadability*는 *FLESCH*, *FOG*, *LENGTH*를 활용하였다. 선행연구에 따라 일반적인 기업의 특성들이 *Unreadability*를 결정하는 요인들로 활용될 것으로 보고, 아래의 변수들을 식 (1)에 포함하였다(Li 2008).

규모(*SIZE*) : 기업의 규모는 해당 기업이 직면하게 되는 여러 측면의 환경을 반영하므로, 기업규모의 크고 작음은 사업보고서의 가독성에 영향을 미치는 결정요인 중 하나일 것이다. 예를 들어, 상대적으로 규모가 큰 기업은 사업보고서의 ‘Ⅲ. 재무에 관한 사항’ 뿐만 아니라, ‘Ⅰ. 회사의 개요’ 및 ‘Ⅱ. 사업의 내용’ 등에 기재하여 투자자들에게 공시하여야 하는 사항들이 자연히 많을 수 밖에 없을 것이다. 예측하건대, 규모가 큰 기업은 더 길고 복잡한 사업보고서를 공시할 것이다. 선행연구들 역시 기업의 규모(*SIZE*)가 클수록 많은 이해관계자들로부터 공시에 대한 압력이 크기 때문에, 더 많은 양의 공시를 할 것으로 보고하였다(Botosan 1997; 최순재 와 황국재 2001; 이문영 등 2012).

시장·장부가비율(*MTB*) : 높은 시장·장부가비율을 가진 기업들은 투자기회나 잠재된 성장동력 등과 관련하여 낮은 시장·장부가비율을 가진 기업들과 다른 특성들을 가질 것이다. 성장기업의 경우 좀 더 복잡하고 불확실한 사업모형(business model)을 가지고 있을 가



능성이 크기 때문에, 이들의 사업보고서는 보다 가독성이 낮을 가능성이 있다. 하지만 시장·장부가비율(*MTB*)을 성장동력으로 해석하게 되는 경우 투자자들에게 보다 가독성이 높은 사업보고서를 공시하여 자금조달 등의 측면에서 혜택을 보려고 할 수도 있다.

부채비율(*LEVERAGE*) : 부채비율(*LEVERAGE*)이 높을수록 채권자들로부터 정보공개에 대한 요구가 증가하고(Hope and Thomas 2008), 기업의 입장에서 소송에 대한 위험을 줄이기 위해 공시를 더 많이 할 가능성이 있다(Bradbury 1992). 다시 말해, 부채비율(*LEVERAGE*)이 높은 기업의 사업보고서는 가독성이 낮을 것으로 예상된다(Barnett and Leoffler 1979).

기업 연수(*AGE*) : 기업연수는 기업공개 이후 각 기업-연 자료의 사업연도까지의 연수로 측정하였다. 기업연수가 오래될수록 투자자들이 해당기업이 공시하는 정보에 친숙해지고, 사업보고서를 공시하는 횟수가 늘어날수록 사업보고서의 정보비대칭(*information asymmetry*) 및 정보불확실성(*information uncertainty*)이 줄어든다고 예측해볼 수 있겠다(박순홍 2013). 이와 반대로, 상장직후에 다양한 이해관계자로부터의 소송위험(*litigation risk*)이 급격히 커진다면(Hanley and Hoberg 2012), 상장직후 시점에 사업보고서 작성시 가독성이 높은 사업보고서를 공시할 유인 또한 있을 것이다.

비경상영업외손익(*SI*) : 비경상영업외손익(*SI*)은 영업외수익 중 이자수익, 배당금수익, 임대료수익, 투자자산처분이익, 매입할인 등을 차감하고, 이자비용 및 투자자산처분손실을 제외한 영업외비용을 차감하여 측정하였다. 비경상영업외손익 항목의 비중이 큰 기업일수록, 특히 비경상항목이 음(-)의 큰 값을 갖는 경향이 클수록, 해당 기업이 특별한 사건(event)에 직면할 가능성이 높을 것이고, 이러한 경우 사업보고서의 가독성이 낮아질 경향을 보이려 할 것이다.

사업 또는 영업의 변동성(*volatility of business or operations*) : 사업 또는 영업의 변동성의 대용치로는 해당 사업연도 동안의 주가변동성(*RET\_VOL*) 및 과거 5개년의 이익변동성(*EARN\_VOL*)을 활용하였다. 주가나 이익의 변동성이 큰 기업의 경우 투자자들과의 커뮤니케이션은 복잡할 가능성이 높다. 기업은 공시, IR 등의 활동들을 통해 투자자들을 포함한 이해관계자들과 커뮤니케이션을 하는데, 그 중 가장 주요한 예로 사업보고서를 들 수 있겠으며, 주가나 이익의 변동성이 큰 기업의 사업보고서는 가독성이 낮아질 것이다.

복잡성(*complexity*) : 많은 사업부를 가진 기업일수록 그 기업의 영업활동은 복잡할 것이며, 그에 따른 사업보고서의 가독성은 낮아질 것이다. 본 연구에서는 영업활동의 복잡성(*complexity of operations*)의 대용치로 사업부의 수(*NBSEG*)를 사용하였다. 또한, 공시정보의 높은 감사품질을 누리하고자 하는 유인이 있는 기업들의 경우 복잡한 영업 및 재무활동이 복잡할 가능성이 높으므로 외부감사인의 대형회계법인 여부(*BIGM*)를 그 대용치로 활용하였다. 감사인은 외부감사를 수행하면서 기업의 회계정보 및 공시수준에 영향을 미칠 수 있으며, 일반적으로 대형회계법인은 전문성과 독립성을 갖춘 감사품질이 높은 감사인으로 알려져 있다(심호식 등 2010; 이문영 등 2012).

기업에 영향을 미치는 주요 사건(*firm events*) : 유상증자(*SEO*)나 K-IFRS의 도입(*IFRS*)과 같은 사건들은 비정상적으로 발생하는 사건들임에도 불구하고, 일시에 기업의 사업보고서 공시의 내용과 양에 모두 영향을 미치는 사건들임에 틀림이 없다. 따라서 유상증자(*SEO*)나 K-IFRS의 도입(*IFRS*)의 더미 변수로 이러한 사건들을 통제하였다. 그 밖의 사업보고서 가독성의 잠재적인 결정요인들을 통제하기 위하여, 본 연구에서는 한국표준산업분류에서 중분류(2자리)를 기준으로 한 산업더미(*Industry Fixed Effects*)와 연도더미(*Year Fixed Effects*)를 포함하였다.

〈Table 1〉은 사업보고서 가독성(*Unreadability*) 결정요인에 대한 연구모형의 회귀분석 결과를 나타낸 것이다. 앞서 언급한 것과 같이, 사업보고서 가독성 측정치인 *FLESCH*, *FOG*, *LENGTH*의 해석을 동일한 방식으로 하기 위하여, *FLESCH*에 (-1)을 곱하여 측정치가 높을수록 사업보고서의 가독성이 낮음(*obfuscation*, *desultory reading*, *difficulties to read*)을 의미하게 하였다. 〈Table 1〉의 열 (1), (2), (3)에서는 사업보고서 가독성(*Unreadability*) 측정치인 *FLESCH*, *FOG*, *LENGTH*를 각각 종속변수로 사용하여 분석하였다. 사업보고서는 규모(*SIZE*)가 클수록, 부채비율(*LEVERAGE*) 및 이익변동성(*EARN\_VOL*)이 높을수록, 그리고 유상증자(*SEO*)를 하였거나, K-IFRS의 도입(*IFRS*)이 있는 경우에 가독성이 낮아지고, 시장·장부가비율(*MTB*)이 낮을수록, 기업 연수(*AGE*)가 짧을수록, 비정상영업외손익(*SI*)의 크기가 적을수록 가독성이 낮아진다. 또한, 사업보고서 가독성 측정치를 *LENGTH*로 둔 열 (3)에서만 사업부의 수(*NBSEG*)가 많을수록, 그리고 외부감사인이 대형회계법인 경우(*BIGN*)에 사업보고서의 가독성이 낮음을 확인할 수 있다. 이러한 회귀분석 결과는 예측 및 선행연구들의 결과들과 대체로 일치하였다.<sup>7)</sup>

7) 앞서 논의한 가독성 측정의 결정요인 계수들의 예상되는 부호(pred. sign)들을 〈Table 1〉에 보고하였다. 식(1)의 회귀분석 결과는 규모(*SIZE*)가 크고, 기업 연수(*AGE*)가 짧으며, 비정상영업외손익(*SI*)의 크기가 적고, 이익변동성(*EARN\_VOL*)이 높고, 사업부의 수(*NBSEG*)가 많으며, 유상증자(*SEO*)를 하는 기업의 사업보고서의 가독성이 낮아짐을 보고한 Li(2008)의 결과와 일치한다.

〈Table 1〉 Determinants of Unreadability Index

<i>Dep. Variables</i>		<i>FLESCH</i>	<i>FOG</i>	<i>LENGTH</i>
	pred.sign	(1)	(2)	(3)
<i>Intercept</i>		55.421*** (10.88)	23.641*** (29.37)	7.027*** (54.03)
<i>SIZE</i>	+	1.165*** (8.89)	0.107*** (6.14)	0.064*** (14.86)
<i>MTB</i>	+/-	-0.272*** (-3.25)	-0.032*** (-2.99)	-0.015*** (-5.59)
<i>LEVERAGE</i>	+	4.877*** (7.34)	0.339*** (4.05)	0.204*** (10.37)
<i>AGE</i>	-	-0.061*** (-4.24)	-0.016*** (-8.60)	-0.003*** (-6.23)
<i>SI</i>	-	-2.902** (-2.24)	-0.374** (-2.27)	-0.122*** (-2.93)
<i>RET_Vol</i>	+	-0.453 (-0.46)	0.057 (0.47)	0.035 (1.30)
<i>EARN_VOL</i>	+	5.177*** (2.69)	1.630*** (6.73)	0.242*** (4.16)
<i>NBSEG</i>	+	-0.126 (-0.48)	0.011 (0.34)	0.043*** (5.86)
<i>SEO</i>	+	2.724*** (10.78)	0.372*** (11.62)	0.088*** (11.11)
<i>BIGN</i>	+	0.368 (1.15)	-0.004 (-0.10)	0.028*** (2.87)
<i>IFRS</i>	+	1.628*** (2.65)	0.174** (2.32)	0.050*** (3.20)
<i>Industry F.E.</i>		Included	Included	Included
<i>Year F.E.</i>		Included	Included	Included
<i>Clustered by</i>		Firm	Firm	Firm
N		21,917	21,917	21,917
Adjusted R <sup>2</sup>		0.480	0.666	0.569

Notes:

This table shows the determinants of Unreadability indices. Specifically, we estimate the following regression model:

$$\begin{aligned}
 \text{Unreadability}_{jt} = & \beta_0 + \beta_1 \text{SIZE}_{jt} + \beta_2 \text{MTB}_{jt} + \beta_3 \text{LEVERAGE}_{jt} + \beta_4 \text{AGE}_{jt} + \beta_5 \text{SI}_{jt} \\
 & + \beta_6 \text{RET\_Vol}_{jt} + \beta_7 \text{EARN\_Vol}_{jt} + \beta_8 \text{NBSEG}_{jt} + \beta_9 \text{SEO}_{jt} + \beta_{10} \text{BIGN}_{jt} \\
 & + \beta_{11} \text{IFRS}_{jt} + \text{Year Fixed Effects} + \text{Industry Fixed Effects} + \varepsilon_{jt}
 \end{aligned} \quad (1)$$

All variables are defined in Appendix A. Robust t-statistics are reported in parentheses. \*, \*\*, and \*\*\* denote the significance at 10%, 5%, and 1% level, respectively.

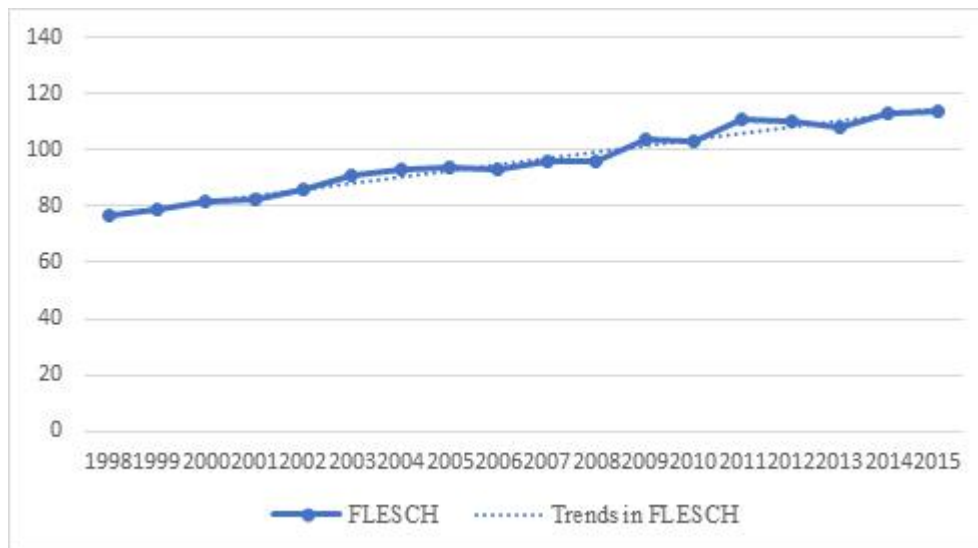
## 4.2 기술통계분석

〈Table 2〉는 본 연구의 4.1 사업보고서 가독성의 결정요인 및 이후 논의에서 사용된 변수들의 기술통계량을 보여주고 있으며, 해당 주요 변수들의 정의는 〈Appendix A〉를 참고 하길 바란다. 사업보고서의 가독성을 측정하는 *FLESCH*, *FOG*, *LENGTH*의 평균(중위수)은 각각 99.030 (101.713), 29.387 (29.723), 그리고 8.909 (8.858)이었으며, 사업보고서의 일부인 MD&A만을 분석의 대상으로 한 권오진과 박하연(2017)의 기초 통계량과 비교해 보았을 때 그 결과와 맥락을 같이 한다고 볼 수 있겠다.<sup>8)</sup> 이를 보다 구체적으로 분석하기 위하여 각 연도별 가독성 측정치의 평균값을 〈Figure 1〉에 표시하였다. 〈Figure 1〉의 Panel (A), (B), (C)에서 가독성 측정치인 *FLESCH*, *FOG*, *LENGTH*는 모두 지속적으로 증가하는 추세를 보이고 있다. Panel (B)에서 *FOG*의 경우 변수의 정의에 의거, 단순히 음절, 단어, 문장의 길이 이외에도 단어의 복잡성(*complexity*)을 고려한다고 본다면, 전반적으로 사업보고서는 시간이 지날수록 복잡해져 왔음을 추론해 볼 수 있겠다. 또한, Panel (C)에서, *LENGTH*의 추세분석을 통하여, 2011년과 2014년에 점진적인 증가추세에 특이사항을 발견할 수 있다. 2011년은 한국채택국제회계기준(K-IFRS)이 도입된 해이며, 2014년은 기업공시서식 작성기준의 개정<sup>9)</sup>에 따라 사업보고서 양식의 상당부분을 차지하는 ‘Ⅲ. 재무에 관한 사항’과 관련한 사업보고서 편제 개편이 이루어진 시기임을 감안하면 위와 같은 증가추세를 해석할 수 있을 것으로 보인다.

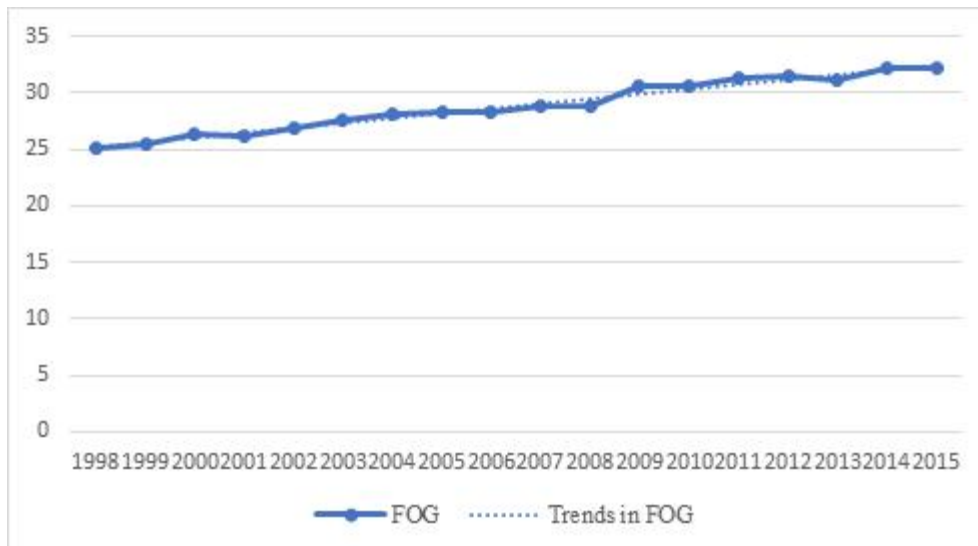
8) 권오진과 박하연(2017)은 산업 내 경쟁과 MD&A 공시수준 간의 관계를 확인하기 위해 2010년부터 2014년까지 유가증권시장에 상장된 기업을 대상으로 수행한 분석에서 MD&A에 공시된 모든 낱말수의 자연로그 값(*WORD*)의 평균값을 6.818로 보고하였다. 권오진과 박하연(2017)의 *WORD* 변수는 본 연구에서 *LENGTH*에 해당되며, 사업보고서 전체의 *LENGTH* 평균값은 8.909로 더 큰 값을 갖는다. 이는 표본의 기간이 다름에도 불구하고, MD&A가 사업보고서의 일부분임을 감안할 때 당연한 결과라 할 수 있겠다.

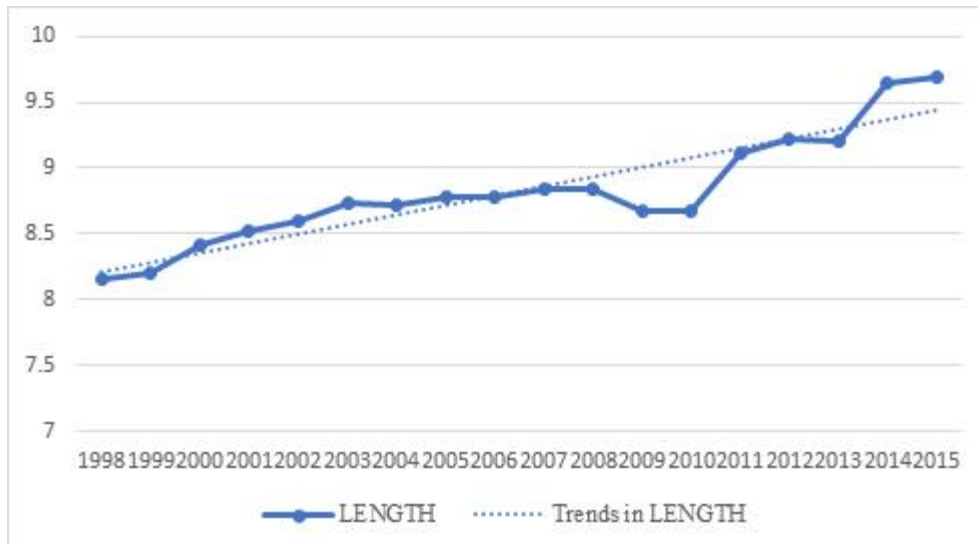
〈Figure 1〉 Trends in Annual report readability

Panel A. Trends in Flesch ease index(*FLESCH*)



Panel B. Trends in Fog index(*FOG*)



Panel C. Trends in Length of document(*LENGTH*)

Notes: This figure shows the trends of various annual report readability index. In Panel A, we report the trends of mean value of *FLESCH*, while in Panel B, and Panel C, we estimate the mean value of *FOG* and *LENGTH* in each year, respectively.

본 연구에서 사용된 기타 통제변수들의 기술통계량을 살펴보면 표본기업들의 기업규모(*SIZE*)는 평균이 17.899이고, 시장·장부가비율(*MTB*)과 부채비율(*LEVERAGE*)은 평균이 각각 1.393, 그리고 0.281로 나타났다. 또한, 자본시장에서 해당 기업이 추가로 주식을 발행했는지를 나타내는 유상증자여부더미(*SEO*)는 평균이 0.216로, 표본기업의 약 21.6%가 자본시장으로부터 자본을 조달하였음을 알 수 있다. 그리고 주가변동성(*RET\_VOL*) 및 이익변동성(*EARN\_VOL*)은 평균이 0.178과 0.073이었으며, 외부감사인의 대형회계법인 여부(*BIGN*)는 평균이 0.469로 표본기업의 약 47%가 대형회계법인에게 감사를 받고 있는 것으로 나타났다.

〈Table 2〉 Descriptive Statistics

Variables	N	Mean	Std.Dev	Q1	Median	Q3
<i>FLESCH</i>	21,917	99.030	16.447	88.650	101.713	111.742
<i>FOG</i>	21,917	29.387	2.657	27.549	29.723	31.677
<i>LENGTH</i>	21,917	8.909	0.557	8.500	8.858	9.289
<i>SIZE</i>	21,917	17.899	1.519	16.886	17.682	18.632
<i>MTB</i>	21,917	1.393	1.570	0.537	0.911	1.621
<i>LEVERAGE</i>	21,917	0.281	0.240	0.077	0.247	0.430
<i>AGE</i>	21,917	26.962	14.581	15.000	25.000	37.000
<i>SI</i>	21,917	-0.003	0.075	-0.014	0.003	0.020
<i>RET_VOL</i>	21,917	0.178	0.122	0.100	0.144	0.215
<i>EARN_VOL</i>	21,917	0.073	0.077	0.026	0.049	0.088
<i>NBSEG</i>	21,917	0.459	0.618	0.000	0.000	0.693
<i>SEO</i>	21,917	0.216	0.412	0.000	0.000	0.000
<i>EARNINGS</i>	21,917	-0.015	0.193	-0.032	0.026	0.072
<i>OCF</i>	21,896	0.036	0.132	-0.018	0.043	0.105
<i>ANNRET</i>	21,917	0.191	0.831	-0.292	0.000	0.400
<i>BIGN</i>	21,917	0.469	0.499	0.000	0.000	1.000

Notes:

This table shows the descriptive statistics of variables used in our main analysis. All variables are defined in Appendix A.

〈Table 3〉은 본 연구에서 사용된 변수 간의 상관관계를 나타낸다. 가독성의 대용치(proxy)인 *FLESCH*, *FOG*, *LENGTH*는 1% 유의수준에서 높은 양(+)의 상관관계를 갖는 것으로 나타났다. 또한, 통제변수들과 사업보고서의 가독성(*Unreadability*)과의 상관관계는 선행연구들의 결과와 크게 다르지 않았다(Li 2008; Lo et al. 2017; 권오진과 박하연 2017)

〈Table 3〉 (Pearson) Correlation Matrix

	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]	[13]	[14]	[15]
[1] <i>FLESCH</i>															
[2] <i>FOG</i>	<b>0.898</b>														
[3] <i>LENGTH</i>	<b>0.686</b>	<b>0.759</b>													
[4] <i>SIZE</i>	<b>0.256</b>	<b>0.251</b>	<b>0.310</b>												
[5] <i>MTB</i>	<b>0.122</b>	<b>0.145</b>	<b>0.141</b>	<b>0.279</b>											
[6] <i>LEVERAGE</i>	-0.027	-0.079	-0.014	-0.158	0.005										
[7] <i>AGE</i>	-0.046	-0.079	-0.039	0.183	-0.209	<b>0.077</b>									
[8] <i>SI</i>	-0.060	-0.069	-0.057	0.059	-0.127	<b>0.031</b>	<b>0.140</b>								
[9] <i>Ret_Vol</i>	-0.194	-0.226	-0.195	-0.202	<b>0.197</b>	<b>0.189</b>	-0.145	-0.073							
[10] <i>EARN_VOL</i>	<b>0.024</b>	<b>0.046</b>	<b>0.016</b>	-0.131	<b>0.217</b>	-0.008	-0.379	-0.197	<b>0.192</b>						
[11] <i>NBSEG</i>	-0.200	-0.249	-0.109	<b>0.088</b>	-0.094	<b>0.062</b>	<b>0.095</b>	<b>0.087</b>	<b>0.018</b>	-0.111					
[12] <i>SEO</i>	<b>0.052</b>	<b>0.053</b>	<b>0.040</b>	-0.072	<b>0.203</b>	<b>0.089</b>	-0.208	-0.156	<b>0.273</b>	<b>0.251</b>	-0.065				
[13] <i>EARNINGS</i>	-0.043	-0.013	-0.027	<b>0.277</b>	-0.141	-0.371	<b>0.090</b>	<b>0.552</b>	-0.259	-0.258	<b>0.098</b>	-0.278			
[14] <i>OCF</i>	-0.022	-0.010	-0.027	<b>0.232</b>	-0.061	-0.268	<b>0.025</b>	<b>0.178</b>	-0.216	-0.160	<b>0.089</b>	-0.248	<b>0.559</b>		
[15] <i>ANNRET</i>	-0.030	-0.029	-0.049	<b>0.204</b>	<b>0.287</b>	-0.086	<b>0.024</b>	<b>0.104</b>	<b>0.339</b>	-0.083	<b>0.030</b>	-0.010	<b>0.192</b>	<b>0.128</b>	
[16] <i>BIGN</i>	<b>0.148</b>	<b>0.146</b>	<b>0.168</b>	<b>0.311</b>	<b>0.008</b>	-0.084	<b>0.039</b>	<b>0.030</b>	-0.211	-0.036	<b>0.007</b>	-0.099	<b>0.134</b>	<b>0.117</b>	-0.017

Notes:

This table shows the Pearson correlation matrix among variables. The Pearson correlation coefficient in bold is significant at 1% level. All variables are defined in Appendix A.



### 4.3 가설 검증 결과

본 연구에서는 당기의 이익이 사업보고서 가독성(*Unreadability*)에 어떠한 관계가 있는지 살펴보고자 아래와 같은 연구모형을 설정하였다(Li 2008).

$$\begin{aligned} Unreadability_{jt} = & \beta_0 + \beta_1 EARNINGS(or Profit\_Dummy)_{jt} + \beta_2 SIZE_{jt} + \beta_3 MTB_{jt} \\ & + \beta_4 LEVERAGE_{jt} + \beta_5 AGE_{jt} + \beta_6 SI_{jt} + \beta_7 RET\_Vol_{jt} + \beta_8 EARN\_Vol_{jt} \\ & + \beta_9 NBSEC_{jt} + \beta_{10} SEO_{jt} + \beta_{11} BIGN_{jt} + \beta_{12} IFRS_{jt} \\ & + Year\ Fixed\ Effects + Industry\ Fixed\ Effects + \varepsilon_{jt} \end{aligned} \quad (2)$$

식 (2)는 사업보고서 가독성과 당기의 이익 간의 관계에 대한 가설 1을 검증하는 연구모형이다. 종속변수는 사업보고서 가독성 측정치인 *FLESCH*, *FOG*, *LENGTH*이며, 관심변수(*Variable of interests*)는 당기이익(*EARNINGS*) 또는 당기이익발생 더미(*Profit\_Dummy*)이다. 앞선 식(1)에서 사용된 사업보고서 가독성의 결정요인들이 식(2)의 통제변수로 모두 포함되었다. 당기의 이익이 낮을수록 사업보고서의 가독성이 낮아진다면, 위의 식(2)의 당기이익(*EARNINGS*) 또는 당기이익 더미(*Profit\_Dummy*)의 계수인  $\beta_1$ 는 0보다 작을 것으로 예상된다.

〈Table 4〉는 사업보고서의 가독성과 당기의 이익의 연관성에 대한 연구모형의 회귀분석 결과를 나타내고 있다. 열 (1), (2), (3)에서 사업보고서 가독성 측정치인 *FLESCH*, *FOG*, *LENGTH*를 종속변수로 사용하는 경우, 당기이익(*EARNINGS*)의 계수,  $\beta_1$ 은 각각 -5.798 ( $t=-6.77$ ), -0.346 ( $t=-3.35$ ), -0.173 ( $t=-6.61$ )이며, 당기이익 더미(*Profit\_Dummy*)의 계수  $\beta_1$ 은 각각 -1.184 ( $t=-4.41$ ), -0.066 ( $t=-1.90$ ), -0.050 ( $t=-5.83$ )이다. 다시 말해, 당기이익(*EARNINGS*) 또는 당기이익 더미(*Profit\_Dummy*)의 계수인  $\beta_1$ 은 예상과 마찬가지로 음(-)의 유의한 관계를 갖는다. 즉, 낮은 이익을 보고하는 기업의 경우 사업보고서의 가독성이 낮아지는 것을 알 수 있고, 이익이 존재하거나 커질수록 표본기업의 사업보고서 가독성은 높아지는 것을 알 수 있으며, 손실기업의 사업보고서가 이익기업의 사업보고서보다 가독성이 낮은 경향을 보이는데, 이는 가설1의 귀무가설인 기업의 당기이익이 사업보고서 가독성과 유의한 관계가 없음을 기각한다.<sup>9), 10)</sup>

9) 이러한 결과는 사업보고서 가독성의 결정요인들과 관련하여 기업-연 자료(firm-year data)의 특성을 생각해 볼 때 특정기업의 연도별 결정요인들이 크게 다르지 않을 수도 있음을 고려하지 못하는 바, 연구모형식(2)의 변화량 분석(changes analysis)을 하였다. 하지만 사업보고서 가독성 측정치를 *FOG*로 사용하여 변화량 분석을 하는 경우의 이익증가더미(*Earn\_Inc\_Dummy*)의 계수 ( $t=-2.18$ )를 제외하고는 대부분 유의하지 않은 결과를 보였다. 이는 우리나라의 많은 기업들이 사업보고서 표준문안(boilerplate)을 작성한 뒤 크게 바꾸지 않은 채 매년 거의 같은 내용을 공시하는 행태를 생각해 볼 때, 예상되는 결과라고 할 수 있겠다.

(Table 4) Current earnings and annual report unreadability

<i>Dep. Variables</i>		<i>FLESCH</i>		<i>FOG</i>		<i>LENGTH</i>	
	pred. sign	(1)		(2)		(3)	
<i>Intercept</i>		53.890*** (10.54)	55.476*** (10.87)	23.549*** (29.19)	23.644*** (29.35)	6.982*** (53.54)	7.030*** (54.31)
<i>EARNINGS</i>	-	-5.798*** (-6.77)		-0.346*** (-3.35)		-0.173*** (-6.61)	
<i>Profit_Dummy</i>	-		-1.184*** (-4.41)		-0.066* (-1.90)		-0.050*** (-5.83)
<i>SIZE</i>	+	1.322*** (9.85)	1.251*** (9.33)	0.116*** (6.52)	0.112*** (6.29)	0.069*** (15.71)	0.067*** (15.51)
<i>MTB</i>	+/-	-0.324*** (-3.88)	-0.299*** (-3.57)	-0.035*** (-3.27)	-0.033*** (-3.12)	-0.016*** (-6.20)	-0.016*** (-6.06)
<i>LEVERAGE</i>	+	3.431*** (4.93)	4.176*** (6.15)	0.253*** (2.87)	0.300*** (3.50)	0.161*** (7.58)	0.175*** (8.61)
<i>AGE</i>	-	-0.067*** (-4.70)	-0.063*** (-4.42)	-0.016*** (-8.79)	-0.016*** (-8.67)	-0.003*** (-6.66)	-0.003*** (-6.48)
<i>SI</i>	-	4.301*** (2.68)	-1.539 (-1.17)	0.056 (0.28)	-0.297* (-1.78)	0.093* (1.85)	-0.065 (-1.53)
<i>RET_VOL</i>	+	-1.169 (-1.19)	-0.735 (-0.75)	0.015 (0.12)	0.042 (0.34)	0.014 (0.51)	0.023 (0.87)
<i>EARN_VOL</i>	+	3.716* (1.92)	4.330** (2.25)	1.543*** (6.38)	1.583*** (6.56)	0.199*** (3.39)	0.207*** (3.56)
<i>NBSEG</i>	+	-0.122 (-0.47)	-0.123 (-0.47)	0.011 (0.34)	0.011 (0.34)	0.043*** (5.89)	0.043*** (5.88)
<i>SEO</i>	+	2.414*** (9.68)	2.592*** (10.33)	0.353*** (11.16)	0.365*** (11.45)	0.079*** (9.99)	0.082*** (10.47)
<i>BIGN</i>	+	0.408 (1.27)	0.362 (1.13)	-0.002 (-0.04)	-0.005 (-0.11)	0.029*** (2.99)	0.028*** (2.84)
<i>IFRS</i>	+	1.178* (1.90)	1.528** (2.49)	0.147* (1.93)	0.168** (2.24)	0.036** (2.35)	0.046*** (2.94)
<i>Industry F.E.</i>		Included	Included	Included	Included	Included	Included
<i>Year F.E.</i>		Included	Included	Included	Included	Included	Included

10) 기업의 공시는 기업 특유의 특성이 반영된 결과일 수 있다. 관찰불가능한(unobservable) 기업 특유의 속성을 반영하기 위하여 추가적으로 기업고정효과(firm-fixed effect)를 포함한 분석을 수행하였다. 그 결과 *EARNINGS*는 여전히 사업보고서의 가독성(*Unreadability*)와 음(-)의 관계를 갖는 것으로 나타났다. *Profit\_Dummy*의 경우 유의하지는 않지만 역시 전반적으로 음(-)의 관계를 갖는 것으로 나타났다.

Clustered by	Firm	Firm	Firm	Firm	Firm	Firm
N	21,917	21,917	21,917	21,917	21,917	21,917
Adjusted R <sup>2</sup>	0.482	0.481	0.666	0.666	0.570	0.570

(Continued on the next page)

Notes:

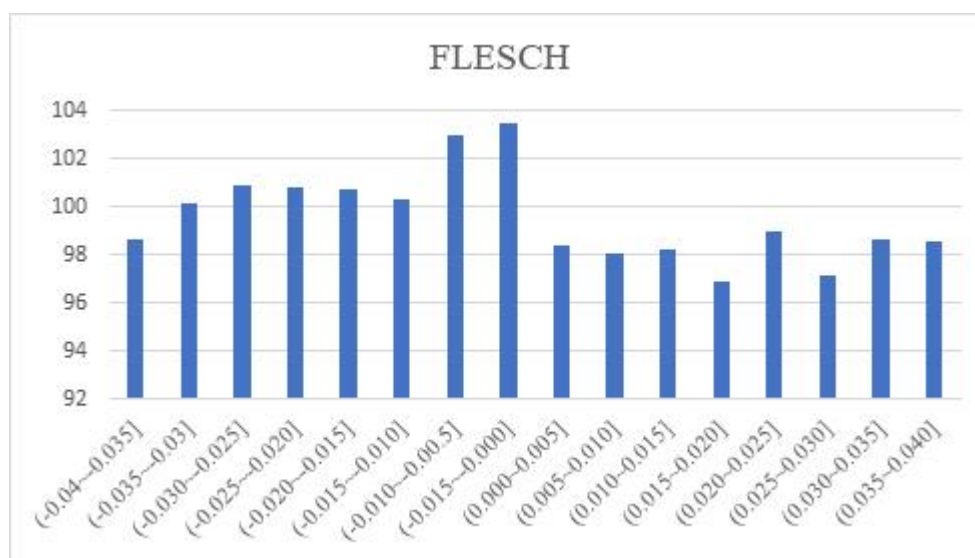
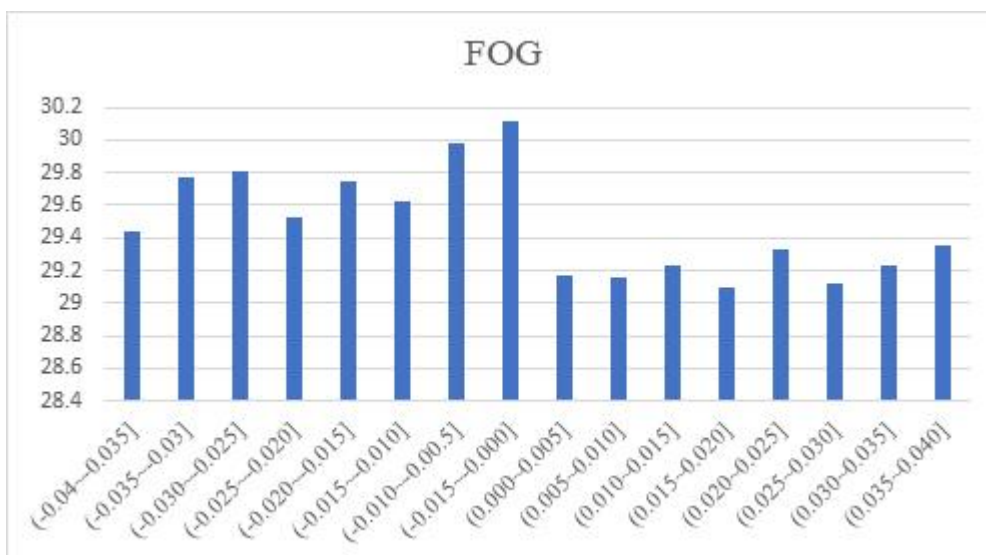
This table shows the results of equation (2):

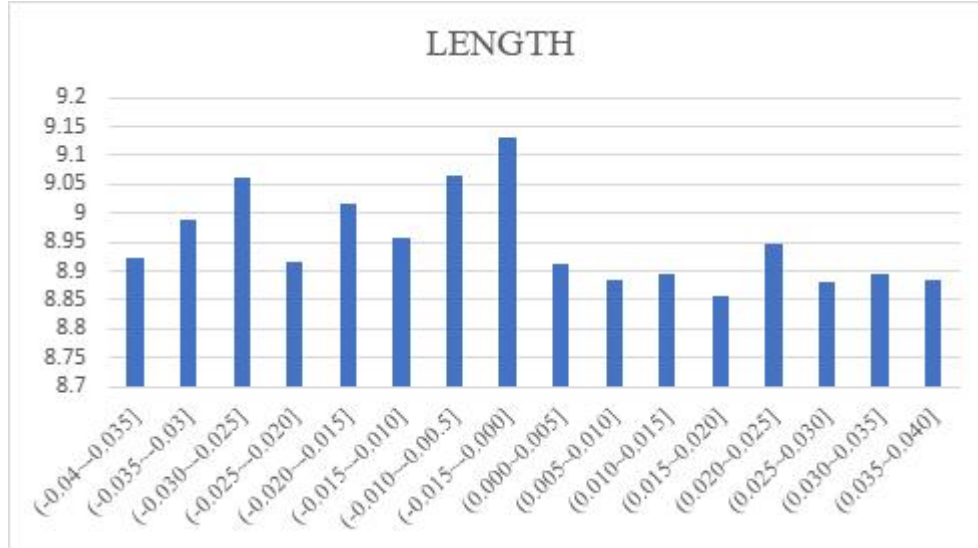
$$Unreadability_{jt} = \beta_0 + \beta_1 EARNINGS(or Profit\_Dummy)_{jt} + \beta_2 SIZE_{jt} + \beta_3 MTB_{jt} + \beta_4 LEVERAGE_{jt} + \beta_5 AGE_{jt} + \beta_6 SI_{jt} + \beta_7 RET\_Vol_{jt} + \beta_8 EARN\_Vol_{jt} + \beta_9 NBSEG_{jt} + \beta_{10} SEO_{jt} + \beta_{11} BIGN_{jt} + \beta_{12} IFRS_{jt} + Year\ Fixed\ Effects + Industry\ Fixed\ Effects + \varepsilon_{jt} \quad (2)$$

*Profit Dummy* equals to 1 if net income is not negative, and 0 otherwise; other variables are defined in Appendix A. Robust t-statistics are reported in parentheses. \*, \*\*, and \*\*\* denote the significance at 10%, 5%, and 1% level, respectively.

한편, 이와 같은 사업보고서의 가독성과 이익 간의 관계를 보다 자세히 알아보기 위하여 이익수준의 분포에 따른 사업보고서의 가독성을 측정하였다(Burstahler and Dichev 1997). 〈Figure 2〉는 당기순이익을 전기 총자산으로 나눈 이익수준의 분포를 영(0)을 기준으로 -0.40에서 +0.40까지 0.005씩의 구간으로 구분한 뒤 사업보고서 가독성 측정치인 *FLESCH*, *FOG*, *LENGTH*의 각 구간별 평균을 막대그래프로 나타낸 것이다. 〈Figure 2〉에서 확인할 수 있듯이, 적자구간과 흑자구간의 사업보고서 가독성 평균치가 다른 양상을 보이고 있으며, 특히 영(0)을 중심으로 한 양측 부분에서 상이한 분포를 보이고 있다. 전반적으로 적자구간에서 가독성 대응치들의 평균치가 흑자구간의 가독성 대응치들의 평균치보다 높아, 적자기업의 사업보고서가 이익기업의 사업보고서보다 가독성이 낮으며, 이익이 존재하거나 커질수록 표본기업의 사업보고서 가독성은 높아지는 〈Table 4〉의 결과와 일맥상통함을 알 수 있다. 더불어, 적자구간 중 당기 이익이 영(0)에 약간 미달하는 기업들의 경우 가장 가독성이 낮은 모습을 나타내고 있다. 당기 이익이 영(0)에 약간 미달하는 기업의 사업보고서의 경우 그 외 적자기업들의 사업보고서에 비하여 상대적으로 가독성이 떨어지는 것을 확인할 수 있고, 이러한 차이는 영(0)보다 약간 큰 이익을 보고한 기업들과도 몹시 상이한 결과이다. Burstahler and Dichev(1997)의 연구방법을 이용하여 우리나라 기업이 이익조정을 얼마나 행하고 있는지를 검증한 송인만 등(2004)의 연구 결과에 의하면, 영(0)에 약간 미달하는 기업들은 이익을 상향조정하여 흑자로 보고하는 경향이 있음을 알 수 있다. 이러한 실무적인 관행과 선행연구의 결과에도 불구하고 당기 이익이 영(0)에 약간 미달하게 보고할 수 밖에 없었던 기업들은 적자보고의 사유에 대하여 사업보고서 공시 내용에 설명함으로써 기업의 이해관계자들에게 기업의 손실로 인한 위험을 불식하고자 하는 유인이 있음을 추론해 볼 수 있겠다. 종합해보면 흑자 기업의 사업보고서 가독성은 높은 경향을 띄며, 이는 사업보고서의 가독성과 당기의 이익간에 양(+)의 관계가 있음을 보여준다.

〈Figure 2〉 Loss avoidance and Annual report readability

Panel A. Loss avoidance and Flesch ease index (*FLESCH*)Panel B. Loss avoidance and Fog Index (*FOG*)

Panel C. Loss avoidance and Length of document (*LENGTH*)

Notes: This figure shows the association between loss avoidance behavior and annual report readability. In Panel A, we report the mean value of *FLESCH* in each profit/loss section. In Panel B and Panel C, we estimate the mean value of *FOG* and *LENGTH* in each profit/loss section, respectively.

앞선 가설1의 당기이익분석에서 흑자기업은 사업보고서를 짧거나 혹은 이해하기 쉽게 공시하는 경향을 확인할 수 있었는데, 이러한 기업의 사업보고서의 가독성과 기업의 이익 간의 관계는 기업의 미래이익에 대한 분석으로 확장이 가능하다(Li 2008). 이익지속성(earnings persistence)은 당기의 이익이 미래기간에도 계속적으로 유지되는 정도를 의미하며, 선행연구들에서는 일반적으로 이익지속성을 당기의 이익( $EARNINGS_{jt}$ )과 차기의 이익( $EARNINGS_{jt+1}$ )간의 1차 자기상관관계(first-order autocorrelation)로 측정하고 있다(Sloan, 1996; 백원선, 2008). 본 연구는 당기( $t$ 기)의 이익 및 사업보고서의 가독성이 차기의 이익( $EARNINGS_{jt+1}$ )을 포함한 당기로부터 3년간( $t+1 \sim t+3$ 기)의 이익지속성에 어떠한 영향을 미치는지를 살펴보고자 아래와 같은 연구모형을 설정하였다.

$$\begin{aligned}
 & EARNINGS_{jt+1,2,3} \\
 &= \beta_0 + \beta_1 EARNINGS_{jt} + \beta_2 Unreadability_{ijt} + \beta_3 EARNINGS_{jt} * Unreadability_{ijt} \\
 &+ \beta_4 SIZE_{jt} + \beta_5 MTB_{jt} + \beta_6 LEVERAGE_{jt} + \beta_7 AGE_{jt} + \beta_8 SI_{jt} + \beta_9 RET\_Vol_{jt} \\
 &+ \beta_{10} EARN\_Vol_{jt} + \beta_{11} NBSEG_{jt} + \beta_{12} SEO_{jt} + \beta_{13} BIGN_{jt} + \beta_{14} IFRS_{jt} \\
 &+ \beta_{15} ABS\_ACC_{jt} + Year\ Fixed\ Effects + Industry\ Fixed\ Effects + \epsilon_j \quad (3)
 \end{aligned}$$

위의 연구모형 식(3)에서는, 앞선 사업보고서의 결정요인 및 식(2)와 마찬가지로 사업보

고서의 가독성 및 이익에 영향을 미칠 수 있는 변수들을 고려하여 통제변수로 포함하였다. 특히, Sloan(1996)은 발생액의 낮은 이익지속성에 대하여 연구하였는데, 본 연구에서는 정치한 이익지속성 분석을 위하여 발생액의 절대값(*ABS\_ACC*)을 이익지속성의 통제변수로 포함하였다. 식(3)에서  $\beta_1$ 은 그 추정값이 이익지속성을 나타내주는 계수이며, 그 값이 양(+)의 값을 가지며 그 크기가 클수록 이익이 지속됨을 나타낸다. 앞선 추측과 같이, 경영자가 미래 이익이 지속적일 것으로 예측하는 경우, 정보를 보다 투명하고 이해하기 쉽게 공시할 가능성이 있다면, 식(3)에서 가독성과 이익의 교차항의 계수인  $\beta_3$ 은 그 추정값이 음(-)의 값을 가질 것이다. 다시 말해, 흑자기업이라도 기업의 미래 이익의 지속성이 낮다면, 사업보고서가 길거나 이해하기 어렵게 작성될 가능성이 높다. 기업이 낮은 이익지속성에 대한 추가적인 설명으로 인하여 사업보고서를 길게 작성하거나, 정보를 모호하게 기술하여 사업보고서의 정보이용자들이 이해하기 어렵도록 작성할 가능성이 있기 때문이다. 반대로 귀무가설 2와 같이 사업보고서 가독성과 이익지속성이 유의한 관계가 없다면  $\beta_3$ 은 통계적으로 유의한 값을 갖지 못할 것이다.

〈Table 5〉의 Panel A는 사업보고서의 가독성과 이익지속성의 연관성에 대한 연구모형 (3)의 회귀분석 결과를 나타내고 있다. 열 (1)-(3), 열 (4)-(6), 열 (7)-(9)에서 종속변수는 각각  $EARNINGS_{jt+1}$ ,  $EARNINGS_{jt+2}$ ,  $EARNINGS_{jt+3}$ 이며, 이는 각각  $t+1$ 기,  $t+2$ 기,  $t+3$ 기의 이익을 의미한다. 사업보고서 가독성 측정치는 열 (1), (4), (7)에서 *FLESCH*, 열 (2), (5), (8)에서 *FOG*, 열 (3), (6), (9)에서 *LENGTH*를 사용하여 분석하였다. 우선, 이익지속성을 나타내는 계수  $\beta_1$ 은 9개의 열 모두에서 양(+)의 값을 가지며, 각각의 가독성 측정치를 사용한 분석에서  $t+1$ 기,  $t+2$ 기, 그리고  $t+3$ 기의 계수  $\beta_1$ 을 서로 비교할 때, 그 크기가 시간이 갈수록 작아짐을 알 수 있다. 예를 들어, 가독성 측정치로 *FLESCH*를 사용하는 경우, 계수  $\beta_1$ 은 열 (1)에서 0.580 ( $t=21.74$ ), 열 (4)에서 0.413 ( $t=16.66$ ), 그리고 열 (7)에서 0.339 ( $t=13.14$ )의 값을 나타내고 있다. 한편, 사업보고서의 가독성과 당기이익 간의 교차항의 계수인  $\beta_3$ 의 회귀분석 결과를 살펴보면, 9개의 열 모두에서 음(-)의 값을 갖는다. 종합해보면, 사업보고서 및 이익지속성에 영향을 미칠 가능성이 있는 요인들을 통제한 후에 이익은 지속성을 나타내지만, 이익지속성이 낮을수록 사업보고서 가독성 역시 낮아지는 결과를 보인다.

다만, 이러한 분석은 당기이익의 값이 양수인 경우(흑자기업)와 음수인 경우(적자기업)에 서로 다른 양상을 나타낼 수 있다. 따라서, 본 연구의 표본을 흑자기업과 적자기업으로 나누어 연구모형 (3)의 하위집단분석(sub-sample analysis)을 다시 수행하였으며 그 결과는 각각 〈Table 5〉의 Panel B와 Panel C에 보고하였다. 〈Table 5〉의 Panel B에서 확인할 수 있듯이, 흑자기업 표본의 경우 앞선 Panel A의 전체 표본의 분석결과와 비슷하다.  $\beta_1$ 은 9개의 열 모두에서 양(+)의 값으로 이익지속성을 보여주며, 각각의 가독성 측정치를 사용한 분석에서  $t+1$ 기,  $t+2$ 기, 그리고  $t+3$ 기의 계수  $\beta_1$ 을 서로 비교하면, 그 크기가 시간이 갈수록 작아짐을 확인할 수 있었다. 특히, 흑자기업만으로 구성된 표본의 분석

이기 때문에,  $\beta_1$ 은 <Table 5>의 Panel A의 계수 추정치보다 더 큰 양(+)의 값을 갖는다. 또한, 교차항의 계수인  $\beta_3$ 의 회귀분석 결과는 열 (4)를 제외한 8개의 열에서 음(-)의 값을 갖는다. 한편, <Table 5>의 Panel C와 같이 적자기업만으로 구성된 표본의 분석 결과는  $\beta_1$ 은 9개의 열 모두에서 양(+)의 값으로 이익(손실)지속성을 보여주며,  $t+1$ 기,  $t+2$ 기, 그리고  $t+3$ 기의 계수  $\beta_1$ 을 서로 비교해보면, 앞선 Panel A와 Panel B의 결과와 동일하게 그 크기가 시간이 갈수록 작아짐을 확인할 수 있었다. 즉, 적자기업인 경우 그 손실이 시간이 갈수록 지속되며, 그 손실의 지속성이 시간이 갈수록 약해지는 양상을 보이는 데, 이는 흑자기업의 이익지속성에 대한 분석과 다르지 않음을 알 수 있다. 하지만, Hayn(1995)은 흑자기업과 적자기업의 이익반응계수 및 수익률-이익 연관관계를 비교하여 이익기업의 이익반응계수 및 설명력이 손실기업의 그것보다 크다는 것을 보고하였는데, 이처럼 본 연구의 결과에서도 <Table 5>의 Panel B와 Panel C의  $\beta_1$ 의 크기를 비교해 본다면, Panel B의 이익지속성 계수값이 Panel C의 이익지속성 계수값보다 큰 것을 확인할 수 있었다. 반면, Panel C의 교차항의 계수인  $\beta_3$ 의 회귀분석 결과는 사업보고서의 대용치로 *LENGTH*를 사용한 열 (3),(6),(9)을 제외하고 유의하지 않았으며, 이는 선행연구인 Li(2008)의 결과와 유사하다.<sup>11)</sup> 요약하건대, 대체로 가독성이 낮은 사업보고서를 제출하는 기업의 이익지속성은 낮으며, 지속성이 낮은 긍정적인 사건(*good news*)을 공시하기 위하여 흑자기업은 가독성이 낮은 사업보고서를 공시하나, 기업가치에 지속적으로 부정적인 사건(*bad news*)을 공시하는 적자기업은 사업보고서의 가독성과 관련이 없음을 발견하였다. 다시 말해, 기업가치에 긍정적인 사건이라 할지라도 그 사건의 지속성이 낮다면 그 사건으로 인한 이익에 대하여 사업보고서의 독자인 이해관계자에게 설명하기 위하여 사업보고서에 부연적인 설명을 담아 사업보고서의 길이가 길어지거나 복잡해 질 가능성이 높다.

11) *LENGTH*를 사용하는 경우에도 교차항의 계수값이 10% 내에서 유의함을 알 수 있으며, 이는 기존의 전체표본 분석이나 흑자기업 하위집단을 대상으로 분석한 경우와 비교해 보면 그 유의성이 떨어짐을 확인할 수 있다.

〈Table 5〉 Earnings Persistence and annual report unreadability

Panel A : Full Sample										
<i>Dep. Variables</i>		<i>Earnings<sub>t+1</sub></i>			<i>Earnings<sub>t+2</sub></i>			<i>Earnings<sub>t+3</sub></i>		
Unreadability		<i>FLESCH</i>	<i>FOG</i>	<i>LENGTH</i>	<i>FLESCH</i>	<i>FOG</i>	<i>LENGTH</i>	<i>FLESCH</i>	<i>FOG</i>	<i>LENGTH</i>
	pred. sign	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
<i>Intercept</i>		0.163** (2.06)	0.373** (2.38)	0.321** (2.06)	-0.032 (-0.29)	-0.120 (-0.69)	-0.102 (-0.67)	-0.001 (-0.01)	0.078 (0.37)	-0.020 (-0.13)
<i>EARNINGS</i>	+	0.580*** (21.74)	0.581*** (23.54)	0.590*** (22.51)	0.413*** (16.66)	0.417*** (17.49)	0.424*** (16.86)	0.339*** (13.14)	0.343*** (14.43)	0.351*** (13.52)
<i>Unreadability</i>	-	-0.002*** (-3.37)	-0.016*** (-2.96)	-0.046*** (-3.06)	-0.000 (-0.46)	0.001 (0.26)	0.003 (0.18)	-0.001 (-0.86)	-0.006 (-0.83)	-0.007 (-0.41)
<i>EARNINGS</i> <i>*Unreadability</i>	+/-	-0.001*** (-3.44)	-0.002*** (-4.00)	-0.007*** (-3.94)	-0.000*** (-3.25)	-0.001*** (-3.72)	-0.005*** (-3.57)	-0.000*** (-3.65)	-0.002*** (-4.54)	-0.006*** (-4.08)
<i>SIZE</i>	+	-0.007 (-1.57)	-0.021** (-2.50)	-0.019** (-2.49)	0.003 (0.59)	0.004 (0.45)	0.006 (0.73)	0.002 (0.39)	-0.002 (-0.19)	0.006 (0.67)
<i>MTB</i>	+/-	-0.011 (-1.17)	-0.022 (-1.37)	-0.010 (-0.55)	-0.016 (-1.61)	-0.018 (-1.13)	-0.040** (-1.97)	-0.005 (-0.60)	-0.009 (-0.51)	-0.033 (-1.56)
<i>LEVERAGE</i>	+	0.016 (0.42)	0.006 (0.09)	-0.094 (-1.04)	-0.026 (-0.65)	-0.029*** (-3.22)	-0.197* (-1.92)	-0.035 (-0.78)	-0.083 (-0.96)	-0.114 (-0.94)
<i>AGE</i>	-	0.000 (1.13)	0.001 (1.62)	0.001 (1.59)	0.001* (1.71)	0.002** (2.21)	0.000 (0.35)	0.001 (1.11)	0.002 (1.28)	0.000 (0.14)
<i>SI</i>	-	-0.246	-0.059	-0.324	-0.082	0.121	0.739	-0.019	0.170	-0.202



		(-1.10)	(-0.14)	(-0.58)	(-0.39)	(0.27)	(1.00)	(-0.08)	(0.34)	(-0.27)
										(Continued on the next page)
RET_VOL	+	0.062	0.168	0.283	0.097	0.351***	0.412**	0.137**	0.447***	0.444**
		(1.12)	(1.48)	(1.55)	(1.48)	(2.81)	(2.00)	(2.18)	(3.30)	(2.06)
EARN_VOL	+	-0.123	-0.306	0.124	-0.205	-0.648	-0.115	-0.257	-0.903*	-0.468
		(-0.66)	(-0.92)	(0.30)	(-0.91)	(-1.60)	(-0.23)	(-1.00)	(-1.93)	(-0.78)
NBSEG	+	-0.011	-0.006	-0.003	-0.010	-0.013	-0.031	-0.000	0.001	-0.034
		(-1.29)	(-0.39)	(-0.15)	(-0.96)	(-0.63)	(-1.01)	(-0.00)	(0.03)	(-0.90)
SEO	+	-0.027	-0.016	-0.063	-0.011	-0.001	0.024	-0.021	-0.028	0.008
		(-1.21)	(-0.40)	(-1.33)	(-0.46)	(-0.02)	(0.41)	(-0.86)	(-0.60)	(0.13)
BIGN	+	0.023*	0.038*	0.062**	0.028*	0.037	-0.015	0.022	0.023	-0.011
		(1.93)	(1.69)	(2.02)	(1.77)	(1.31)	(-0.39)	(1.28)	(0.73)	(-0.22)
IFRS	+	-0.046**	0.086**	0.050	-0.110***	0.018	0.055	-0.150***	-0.053	-0.051
		(-2.32)	(2.07)	(1.03)	(-4.26)	(0.36)	(0.88)	(-4.63)	(-0.82)	(-0.69)
ABS_ACC	-	0.074	0.189	0.316	0.248**	0.391**	1.108***	0.026	-0.137	0.330
		(0.68)	(0.85)	(1.06)	(2.50)	(2.11)	(3.57)	(0.20)	(-0.54)	(0.98)
Industry F.E.	Included	Included	Included	Included	Included	Included	Included	Included	Included	Included
Year F.E.	Included	Included	Included	Included	Included	Included	Included	Included	Included	Included
Clustered by	Firm	Firm	Firm	Firm	Firm	Firm	Firm	Firm	Firm	Firm
N		21,063	21,063	21,063	19,659	19,659	19,659	17,781	17,781	17,781
Adjusted R <sup>2</sup>		0.418	0.419	0.418	0.341	0.341	0.342	0.304	0.305	0.304

(Continued on the next page)

Panel B : Profit Firm-Year Sample

<i>Dep. Variables</i>		<i>Earnings<sub>t+1</sub></i>			<i>Earnings<sub>t+2</sub></i>			<i>Earnings<sub>t+3</sub></i>		
Unreadability		<i>FLESCH</i>	<i>FOG</i>	<i>LENGTH</i>	<i>FLESCH</i>	<i>FOG</i>	<i>LENGTH</i>	<i>FLESCH</i>	<i>FOG</i>	<i>LENGTH</i>
	pred. sign	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
<i>EARNINGS</i>	+	0.641*** (21.68)	0.641*** (23.93)	0.646*** (22.35)	0.477*** (14.88)	0.474*** (15.18)	0.479*** (14.86)	0.407*** (12.57)	0.406*** (13.23)	0.406*** (12.73)
<i>Unreadability</i>	-	-0.001 (-1.51)	-0.001 (-0.30)	-0.009 (-0.83)	-0.000 (-0.01)	0.001 (0.27)	-0.007 (-0.63)	-0.001 (-0.71)	-0.05 (-0.90)	0.001 (0.007)
<i>EARNINGS</i> <i>*Unreadability</i>	+/-	-0.000* (-1.86)	-0.001** (-2.36)	-0.004** (-2.15)	-0.000 (-1.63)	-0.001* (-1.95)	-0.002* (-1.65)	-0.000*** (-2.90)	-0.001*** (-3.68)	-0.04*** (-3.06)
Control Variables		Included	Included	Included	Included	Included	Included	Included	Included	Included
Industry F.E.		Included	Included	Included	Included	Included	Included	Included	Included	Included
Year F.E.		Included	Included	Included	Included	Included	Included	Included	Included	Included
<i>Clustered by</i>		Firm	Firm	Firm	Firm	Firm	Firm	Firm	Firm	Firm
N		14,931	14,931	14,931	13,737	13,737	13,737	12,611	12,611	12,611
Adjusted R <sup>2</sup>		0.241	0.242	0.243	0.198	0.199	0.202	0.205	0.206	0.206

(Continued on the next page)

Panel C : Loss Firm-Year Sample

<i>Dep. Variables</i>		<i>Earnings<sub>t+1</sub></i>			<i>Earnings<sub>t+2</sub></i>			<i>Earnings<sub>t+3</sub></i>		
<i>Unreadability</i>		<i>FLESCH</i>	<i>FOG</i>	<i>LENGTH</i>	<i>FLESCH</i>	<i>FOG</i>	<i>LENGTH</i>	<i>FLESCH</i>	<i>FOG</i>	<i>LENGTH</i>
	pred. sign	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
<i>EARNINGS</i>	+	0.452*** (9.36)	0.456*** (9.30)	0.457*** (9.29)	0.347*** (7.07)	0.357*** (7.09)	0.363*** (7.19)	0.223*** (4.10)	0.227*** (4.14)	0.237*** (4.26)
<i>Unreadability</i>	-	-0.006*** (-3.11)	-0.041*** (-3.14)	-0.184*** (-3.93)	-0.000 (-0.10)	0.011 (0.75)	0.018 (0.31)	0.001 (0.44)	0.009 (0.52)	-0.052 (-0.93)
<i>EARNINGS</i> <i>*Unreadability</i>	+/-	-0.000 (-1.49)	-0.001 (-1.58)	-0.005* (-1.78)	-0.000 (-1.45)	-0.001 (-1.59)	-0.004* (-1.69)	-0.000 (-1.49)	-0.001 (-1.59)	-0.005* (-1.75)
<i>Control Variables</i>		Included	Included	Included	Included	Included	Included	Included	Included	Included
<i>Industry F.E.</i>		Included	Included	Included	Included	Included	Included	Included	Included	Included
<i>Year F.E.</i>		Included	Included	Included	Included	Included	Included	Included	Included	Included
<i>Clustered by</i>		Firm	Firm	Firm	Firm	Firm	Firm	Firm	Firm	Firm
N		6,672	6,672	6,672	5,922	5,922	5,922	5,170	5,170	5,170
Adjusted R <sup>2</sup>		0.329	0.330	0.331	0.300	0.299	0.299	0.271	0.272	0.270

Notes:

This table shows the regression results of hypothesis 2 using the following equation (3):

$EARNINGS_{jt}+1,2,3$

$$= \beta_0 + \beta_1 EARNINGS_{jt} + \beta_2 Unreadability_{jt} + \beta_3 EARNINGS * Unreadability_{jt} + \beta_4 SIZE_{jt} + \beta_5 MTB_{jt} + \beta_6 LEVERAGE_{jt} + \beta_7 AGE_{jt} + \beta_8 SI_{jt} + \beta_9 RET\_Vol_{jt} \\ + \beta_{10} EARN\_Vol_{jt} + \beta_{11} NBSEG_{jt} + \beta_{12} SEO_{jt} + \beta_{13} BIGN_{jt} + \beta_{14} IFRS_{jt} + \beta_{15} ABS\_ACC_{jt} + Year\ Fixed\ Effects + Industry\ Fixed\ Effects + \varepsilon_j \quad (3)$$

In Panel A, we report the results of full sample, whereas Panel B and C show the results of firms with profits and loss, respectively.  $ABS\_ACC$  is denoted as an absolute value of accruals, which is estimated from the difference between net income and the cash flows from operations deflated by lagged total assets: Other variables are defined in Appendix A. Robust t-statistics are reported in parentheses. \*, \*\*, and \*\*\* denote the significance at 10%, 5%, and 1% level, respectively.

이익의 지속성 외에도 이익의 질(*earnings quality*) 역시 이익의 중요한 속성 중 하나이다. 특히, 가설 1과 가설 2에서 사업보고서의 가독성과 당기의 이익 및 이익지속성과의 관계에 대한 가설검증을 하였는데, 당기 이익 및 미래 이익의 지속성에는 이익의 질이 영향을 미칠 수 있다. 현실적으로 이익이나 시장의 기대수준을 과하게 조정하는 것이 용이하지 않기 때문에, 기업들은 실제 주당순이익이 기준점(benchmark)인 전기 주당순이익과 동일하거나 초과하는 정도(meet or beat)가 적은 주당순이익을 보고하려는 유인이 클 것이다. 즉, 보고이익이 시장의 이익예측치보다 월등히 높다면 이는 이익조정 또는 시장의 기대수준의 조정에 따른 것이라기보다는 실제 기업의 영업이 탁월했을 가능성이 대부분일 것이다(남혜정과 최종학 2009). 또한, 경영자는 이익기준점을 달성하지 못할 경우, 추가락, 자본조달비용증가 등이 예상되기 때문에, 이를 이익조정 혹은 시장의 기대수준의 조정 등을 통해 이익기준점을 달성하고자 하는 유인이 존재할 수 있다(Kasznik and McNichols 2002; 남혜정과 최종학 2009). 본 연구에서는 이와 같은 이익기준점에 대한 달성유인을 이익의 질에 대한 대용치로 보고 사업보고서의 가독성(*Unreadability*)과 이익의 질 간의 관계에 대한 가설 3을 아래와 같은 연구모형을 통해 검증하였다.

$$\begin{aligned}
 &Unreadability_{jt} \\
 &= \beta_0 + \beta_1 MBE_{jt} + \beta_2 SIZE_{jt} + \beta_3 MTB_{jt} + \beta_4 LEVERAGE_{jt} + \beta_5 AGE_{jt} \\
 &\quad + \beta_6 SI_{jt} + \beta_7 RET\_VOL_{jt} + \beta_8 EARN\_VOL_{jt} + \beta_9 SALE\_VOL_{jt} \\
 &\quad + \beta_{10} OCF\_VOL_{jt} + \beta_{11} EARNINGS_{jt} + \beta_{12} OCF_{jt} + \beta_{13} LOSS_{jt} \\
 &\quad + \beta_{14} NBSEG_{jt} + \beta_{15} SEO_{jt} + \beta_{16} IFRS_{jt} + \beta_{17} BIGN_{jt} \\
 &\quad + Year\ Fixed\ Effects + Industry\ Fixed\ Effects + \varepsilon_{jt}
 \end{aligned} \tag{4}$$

*MBE*는 음(-)의 이익예측오차 회피정도를 나타내는 변수로서, 기업이 발표한 실제 주당순이익이 기준점(benchmark)과 같거나 초과하는 정도(meet or beat)가 적은 경우 1의 값을, 그 외에는 0의 값을 갖는 더미변수이다. <Table 6>은 연구모형 (4)에 대한 다중회귀분석 결과를 나타내고 있다. 실제 주당순이익이 기준점(benchmark)인 전기 주당순이익과 동일하거나 초과하는 정도(meet or beat)를 측정함에 있어 열 (1), (4), (7)에서는 주당 100원( $MBE = 1$  when  $\Delta EPS \in [0, 100]$ ), 열 (2), (5), (8)에서는 주당 300원( $MBE = 1$  when  $\Delta EPS \in [0, 300]$ ), 열 (3), (6), (9)에서는 주당 500원( $MBE = 1$  when  $\Delta EPS \in [0, 500]$ )의 세 가지 범위를 통하여 분석하였다. *MBE*의 계수,  $\beta_1$ 은 열 (1)과 (7)을 제외하고 모든 분석에서 통계적으로 유의한 양(+)의 값을 갖는다. 다시 말해, 이익조정 유인이 있는 기업의 경우 사업보고서의 가독성이 낮아지는 것을 알 수 있고, 이는 가설 3의 귀무가설인 사업보고서의 가독성과 이익의 질은 유의한 관계가 없음을 기각한다. 결국 이익조정 유인이 큰 기업들의 경우 사업보고서의 가독성이 낮아지는 경향이 있으며, 이는 이익의 질과 사업보고서의 가독성 간에 음(-)의 관계가 있음을 나타내는 것이라고 할 것이다.

〈Table 6〉 Earnings Quality and annual report unreadability

Dep. Variables		FLESCH			FOG		LENGTH			
MBE = 1 when $\Delta$ EPS $\in$ pred. sign		[0, 100] (1)	[0, 300] (2)	[0,500] (3)	[0, 100] (4)	[0, 300] (5)	[0,500] (6)	[0, 100] (7)	[0, 300] (8)	[0,500] (9)
Intercept		55.927*** (9.48)	55.773*** (9.44)	55.708*** (9.42)	24.374*** (27.34)	24.356*** (27.28)	24.336*** (27.22)	7.385*** (48.33)	7.373*** (48.19)	7.370*** (48.20)
MBE	+	0.483 (1.40)	0.475* (1.94)	0.520** (2.22)	0.079* (1.74)	0.069** (2.11)	0.087*** (2.84)	0.005 (0.48)	0.016** (2.10)	0.019*** (2.68)
SIZE	+	1.399*** (7.23)	1.407*** (7.24)	1.410*** (7.25)	0.100*** (4.18)	0.101*** (4.21)	0.102*** (4.25)	0.051*** (8.96)	0.052*** (9.06)	0.052*** (9.08)
MTB	+/-	-0.439*** (-3.93)	-0.443*** (-3.96)	-0.446*** (-3.99)	-0.037*** (-2.62)	-0.037*** (-2.66)	-0.038*** (-2.71)	-0.012*** (-3.39)	-0.012*** (-3.48)	-0.012*** (-3.51)
LEVERAGE	+	4.132*** (4.67)	4.135*** (4.67)	4.125*** (4.66)	0.289** (2.54)	0.289** (2.54)	0.288** (2.53)	0.178*** (6.27)	0.179*** (6.30)	0.179*** (6.29)
AGE	-	-0.066*** (-4.15)	-0.066*** (-4.14)	-0.066*** (-4.14)	-0.015*** (-7.23)	-0.015*** (-7.22)	-0.015*** (-7.22)	-0.003*** (-5.51)	-0.003*** (-5.48)	-0.003*** (-5.48)
SI	-	4.804** (2.02)	4.798** (2.02)	4.766** (2.01)	0.006 (0.02)	0.005 (0.02)	-0.000 (-0.00)	-0.012 (-0.16)	-0.012 (-0.16)	-0.013 (-0.18)
RET_VOL	+	1.211 (1.02)	1.196 (1.01)	1.186 (1.00)	0.208 (1.36)	0.206 (1.34)	0.204 (1.33)	0.033 (0.93)	0.033 (0.91)	0.032 (0.90)
EARN_VOL	+	2.183 (0.68)	2.294 (0.71)	2.308 (0.72)	1.231*** (3.05)	1.245*** (3.08)	1.253*** (3.10)	0.155 (1.59)	0.163* (1.66)	0.164* (1.67)

(Continued on the next page)

<i>SALE_VOL</i>	+	0.000 (0.12)	0.000 (0.13)	0.000 (0.15)	0.000 (1.54)	0.000 (1.55)	0.000 (1.57)	0.000*** (7.51)	0.000*** (7.50)	0.000*** (7.52)
<i>OCF_VOL</i>	+	2.052 (0.81)	2.057 (0.81)	2.077 (0.82)	0.297 (0.95)	0.297 (0.95)	0.301 (0.96)	0.034 (0.42)	0.034 (0.42)	0.035 (0.43)
<i>EARNINGS</i>	-	-5.304*** (-3.70)	-5.355*** (-3.73)	-5.395*** (-3.76)	-0.422** (-2.47)	-0.430** (-2.52)	-0.438** (-2.56)	-0.124*** (-2.93)	-0.126*** (-2.97)	-0.128*** (-3.00)
<i>OCF</i>	-	-0.114 (-0.10)	-0.135 (-0.12)	-0.141 (-0.13)	-0.136 (-0.97)	-0.139 (-0.99)	-0.140 (-1.00)	-0.146*** (-4.24)	-0.147*** (-4.25)	-0.147*** (-4.26)
<i>LOSS</i>	+	0.292 (0.86)	0.318 (0.93)	0.343 (1.00)	-0.015 (-0.34)	-0.012 (-0.27)	-0.006 (-0.14)	0.014 (1.26)	0.016 (1.42)	0.017 (1.51)
<i>NBSEG</i>	+	-0.210 (-0.69)	-0.210 (-0.69)	-0.209 (-0.68)	0.019 (0.49)	0.019 (0.49)	0.019 (0.50)	0.043*** (4.87)	0.043*** (4.88)	0.043*** (4.89)
<i>SEO</i>	+	2.064*** (6.74)	2.053*** (6.70)	2.050*** (6.69)	0.230*** (5.80)	0.229*** (5.76)	0.228*** (5.74)	0.057*** (5.66)	0.057*** (5.62)	0.057*** (5.60)
<i>IFRS</i>	+	0.505 (1.37)	0.508 (1.37)	0.513 (1.39)	-0.008 (-0.18)	-0.008 (-0.17)	-0.007 (-0.15)	0.034*** (2.95)	0.034*** (2.98)	0.034*** (3.00)
<i>BIGN</i>	+	4.523** (2.13)	4.548** (2.14)	4.560** (2.14)	0.534* (1.92)	0.539* (1.93)	0.540* (1.93)	0.429*** (5.83)	0.429*** (5.80)	0.429*** (5.81)
<i>Industry F.E.</i>		Included	Included	Included	Included	Included	Included	Included	Included	Included
<i>Year F.E.</i>		Included	Included	Included	Included	Included	Included	Included	Included	Included
<i>Clustered by</i>		Firm	Firm	Firm	Firm	Firm	Firm	Firm	Firm	Firm
N		16,290	16,290	16,290	16,290	16,290	16,290	16,290	16,290	16,290
Adjusted R <sup>2</sup>		0.493	0.493	0.493	0.666	0.666	0.666	0.585	0.585	0.585

Notes:

This table shows the regression results of hypothesis 3 using the following equation (4):

$$\begin{aligned}
 &Unreadability_{jt} \\
 &= \beta_0 + \beta_1 MBE_{jt} + \beta_2 SIZE_{jt} + \beta_3 MTB_{jt} + \beta_4 LEVERAGE_{jt} + \beta_5 AGE_{jt} + \beta_6 SI_{jt} + \beta_7 RET\_VOL_{jt} + \beta_8 EARN\_VOL_{jt} + \beta_9 SALE\_VOL_{jt} + \beta_{10} OCF\_VOL_{jt} \\
 &\quad + \beta_{11} EARNINGS_{jt} + \beta_{12} OCF_{jt} + \beta_{13} LOSS_{jt} + \beta_{14} NBSEG_{jt} + \beta_{15} SEO_{jt} + \beta_{16} IFRS_{jt} + \beta_{17} BIGN_{jt} + Year\ Fixed\ Effects + Industry\ Fixed\ Effects + \varepsilon_{jt}
 \end{aligned} \tag{4}$$

Robust t-statistics are reported in parentheses. \*, \*\*, and \*\*\* denote the significance at 10%, 5%, and 1% level, respectively. *MBE* equals to 1 if firms just meet or beat past years earnings by [100, 300, 500] KRW, and 0 otherwise; *OCF* = is cash flows from operations deflated by lagged total asset; *SALE\_VOL* is sales volatility during the past five years; *OCF\_VOL* is volatility of cash flows from operations during the past five years; *LOSS* equals to 1 if net income is negative, and 0 otherwise; Other variables are defined in Appendix A.



또한 연구모형 (4)에서 사용된 음의 이익예측오차 회피정도(*MBE*)외에도, 이익의 질의 또 다른 대용치인 재량적 발생액(discretionary accruals)을 사용하여 사업보고서의 가독성과 이익의 질 간의 관계에 대하여 추가적으로 분석하였다. 경영자가 가독성이 낮은 사업보고서를 공시함으로써 기업의 이해관계자들이 사업보고서의 해석에 필요한 시간과 비용을 높여(Jo and Kim 2007), 이익조정행위를 감추거나 쉽게 드러나지 않도록 노력한다면 이익의 질과 사업보고서의 가독성 간에는 음(-)의 관계가 존재할 것으로 예상된다. 이를 위해 Kothari et al. (2005)의 성과조정 재량적 발생액 모형을 차용하여 표본 기업의 재량적 발생액(*PMDA*)을 측정하였으며, 그 측정치를 이익의 질로 해석하기 위하여 (-1)을 곱하였다.<sup>12)</sup> 먼저, 재량적 발생액과 사업보고서의 가독성 간의 관계를 확인하기 위하여 사업보고서의 가독성 측정치인 *FLESCHE*, *FOG*, *LENGTH*를 10분위수(decile)로 나누어 각 분위수별 재량적 발생액((-1)\**PMDA*)의 평균값을 <Figure 3>에 표시하였다. 그 결과, 각 변수(*FLESCHE*, *FOG*, *LENGTH*)의 값이 커질수록 이익의 질이 낮아지는 경향을 <Figure 3>에서 확인할 수 있다.

연구모형 (4)의 관심변수를 음(-)의 이익예측오차 회피정도(*MBE*)에서 성과조정 재량적 발생액((-1)\**PMDA*)으로 바꾼 경우의 결과는 별도로 보고하지는 않았으나, 가독성 측정치로 사업보고서의 전체길이(*LENGTH*)를 사용한 결과를 제외하고, 전반적으로 이익의 질이 낮은 기업의 경우 가독성이 낮은 사업보고서를 공시한다는 점을 확인할 수 있었다.<sup>13), 14)</sup> 종합하면, 이익의 질이 낮은 기업의 경우 사업보고서를 공시함에 있어 사업보고서의 독자인 이해관계자에게 길거나 복잡한 서술을 하는 경향을 보인다는 점을 알 수 있다.

12) 연구모형 (4)를 다음과 같이 변형하여 추가분석을 진행하였다.

$$\begin{aligned} Unreadability_{jt} = & \beta_0 + \beta_1 \text{Earnings Quality}((-1)*PMDA_{jt})_{jt} + \beta_2 SIZE_{jt} + \beta_3 MTB_{jt} \\ & + \beta_4 LEVERAGE_{jt} + \beta_5 AGE_{jt} + \beta_6 SI_{jt} + \beta_7 RET\_Vol_{jt} + \beta_8 EARN\_Vol_{jt} \\ & + \beta_9 SALE\_Vol_{jt} + \beta_{10} OCF\_Vol_{jt} + \beta_{11} EARNINGS_{jt} + \beta_{12} OCF_{jt} + \beta_{13} LOSS_{jt} \\ & + \beta_{14} NBSEG_{jt} + \beta_{15} SEO_{jt} + \beta_{16} BIGN_{jt} + \beta_{17} IFRS_{jt} + Year\ Fixed\ Effects \\ & + Industry\ Fixed\ Effects + \varepsilon_{jt} \end{aligned}$$

13) 가독성 측정치로 사업보고서의 전체길이(*LENGTH*)를 사용한 경우에는 유의한 결과를 확인할 수 없었다.

14) 본문에서 행한 분석의 경우 IFRS와 Year 더미변수를 같이 고려한 결과를 보여주고 있어, 다중공선성(multicollinearity) 문제를 야기할 수 있다. 따라서 가설 1부터 3까지의 주된 연구모형식 (2), (3), (4)에서 다중공선성 문제를 확인해 보고자 분산팽창계수(Variance Inflation Factor, 이하 VIF)를 확인하였다. 통상적으로 VIF가 10을 넘지 않는 경우 다중공선성 문제로부터 비교적 자유로울 수 있는데(Christensen and Nikolaev 2017), 각 모델에 사용된 IFRS 변수의 VIF 결과값(3.41~3.69)에 따라 다중공선성 문제는 크게 걱정되지 않는다. 이와 더불어, (1) 전체표본을 이용한 경우에 Year 더미변수만을 고려한 분석과 (2) Year 더미변수를 제외하고, IFRS 변수만을 고려한 분석, (3) IFRS 전후로 나눈 표본에서 Year 더미변수만을 고려하는 분석을 추가로 진행하였다. 분석결과, 전반적으로 현 논문의 결과를 뒷받침한다.

〈Figure 3〉 Annual Report Unreadability and Discretionary Accruals



## Notes:

In this figure, we estimate the average value of discretionary accruals in each unreadability decile. We use performance-matched discretionary accruals following Kothari et al.(2005). We multiply by (-1) to show that higher value of discretionary accruals indicate higher accounting quality.

## 4.4 추가 분석

## 4.4.1. 미래 주식 수익률(future stock returns)과 사업보고서 가독성

4.3에서는 사업보고서의 가독성과 당기의 이익의 연관성, 이익의 지속성과의 연관성, 그리고 이익의 질과의 연관성에 대한 분석을 하였다. 한편, 기업의 경영진은 기업가치에 부정적인 영향을 주는 사건들을 최대한 늦게 주가에 반영시키기 위하여 더 복잡한 사업보고서를 작성하여 공시할 유인이 있다고 할 것이다(Li 2006). 이는 곧 사업보고서의 가독성과 미래의 주식수익률(future stock returns) 간에 일정한 관계가 있음을 암시하는 것이다. 국내 주식시장에서 주식수익률과 사업보고서 가독성 간의 관계를 분석하기 위하여, 〈Table 7〉의 Panel A에서는 Fama-MacBeth(1973)의 횡단면 회귀분석을 이용하여 주식수익률과 사업보고서 가독성과의 연관성을 살펴보았다. Panel A의 열 (1), (4), (7)에서는 각각 사업보고서 가독성 측정치로 *FLESCH*, *FOG*, *LENGTH*를 사용하였으며, 그 계수는 음(-)의 유의한 값을 갖는다. 이는 사업보고서의 가독성이 높아질수록 주가수익률이 높아짐을 의미한다. 열 (2), (5), (8)에 당기이익 및 당기이익과 사업보고서 가독성의 교차항을

포함하게 되면 사업보고서 가독성 계수의 크기는 작아지며 그 통계적 유의성도 대체로 잃게 된다. 또한, 당기이익과 사업보고서 가독성의 교차항의 계수는 유의하지 않는 값을 나타내며, 이는 흑자기업의 사업보고서가 길거나 이해하기 어려운 경우 차기의 주식수익률에 아무 영향을 주지 않음을 시사하는데, 이러한 결과는 Li(2008)의 결과와 동일하다. 열 (3), (6), (9)에는 사업보고서 가독성의 변화량을 포함시켜 미래 주식수익률과의 관계를 살펴보았는데, <Table 7>의 Panel A에서 확인할 수 있듯이 그 관계는 통계적으로 유의하지 않다. 한편, Panel B의 열 (1)-(3)에서는 실제 차기의 연간 주식수익률( $RETURN_{t+1}$ )을, 열 (4)-(6)에서는 미래 주식수익률의 변동성( $RET\_VOL_{t+1}$ )을 종속변수로 사용하여 사업보고서 가독성과의 연관성을 분석하였다. 열 (1)-(3)에서 알 수 있듯이, 사업보고서 가독성은 차기 주식수익률( $RETURN_{t+1}$ )과 통계적으로 유의한 관계를 갖지 않음을 보고하였다. 반면, 열 (4)-(6)에서 사업보고서 가독성이 미래 주식수익률의 변동성( $RET\_VOL_{t+1}$ )에 미치는 영향을 살펴보았을 때, 사업보고서의 가독성이 낮을 때 미래 주식수익률의 변동성이 커짐을 보여주고 있다. 결국, 시장에서 사업보고서의 미래성과(future performance)에 대한 기술내용과 기업가치에 부정적인 영향을 주는 사건들을 최대한 늦게 주가에 반영시키고자 하는 경영진의 의도를 충분히 이해하지 못하고 있고 이는 높은 주가변동성으로 나타남을 확인할 수 있다.

〈Table 7〉 Future stock returns and annual report unreadability

Panel A : Fama-MacBeth regressions of future returns on annual report unreadability

<i>Dep. Variables</i>	<i>Return<sub>t+1</sub></i>								
	<i>Unreadability</i>	<i>FLESCH</i>			<i>FOG</i>			<i>LENGTH</i>	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
<i>Intercept</i>	0.293*** (3.79)	0.241*** (3.03)	0.106 (1.61)	0.456*** (3.48)	0.344** (2.53)	0.105 (1.58)	0.702** (2.72)	0.556** (2.19)	0.112 (1.70)
<i>Unreadability</i>	-0.002** (-2.85)	-0.001* (-1.95)		-0.012** (-2.36)	-0.008 (-1.52)		-0.066** (-2.11)	-0.049 (-1.60)	
<i>EARNINGS</i>		0.226 (1.15)			0.384* (1.89)			0.379* (1.97)	
<i>Unreadability</i> * <i>EARNINGS</i>		0.002 (1.31)			0.001 (0.11)			0.002 (0.11)	
$\Delta$ <i>Unreadability</i>			-0.001 (-0.90)			-0.003 (-0.64)			-0.005 (-0.25)
No. of Years	18	18	17	18	18	17	18	18	17
N	14,059	14,059	12,974	14,059	14,059	12,974	14,059	14,059	12,974
Adjusted R <sup>2</sup>	0.005	0.038	0.002	0.005	0.038	0.002	0.007	0.039	0.002

(Continued on the next page)

Panel B : Future return and Return Volatility on annual report unreadability						
<i>Dep. Variables</i>	<i>Return<sub>t+1</sub></i>			<i>RET_VOL<sub>t+1</sub></i>		
<i>Unreadability</i>	<i>FLESCH</i>	<i>FOG</i>	<i>LENGTH</i>	<i>FLESCH</i>	<i>FOG</i>	<i>LENGTH</i>
	(1)	(2)	(3)	(4)	(5)	(6)
<i>Intercept</i>	1.149*** (9.13)	1.065*** (7.37)	1.174*** (7.59)	0.473*** (23.30)	0.444*** (18.14)	0.412*** (15.75)
<i>Unreadability</i>	0.000 (1.09)	0.005 (1.42)	-0.000 (-0.00)	0.000*** (2.72)	0.002*** (2.94)	0.010*** (4.15)
<i>SIZE</i>	-0.043*** (-8.08)	-0.043*** (-8.18)	-0.042*** (-7.91)	-0.017*** (-18.41)	-0.017*** (-18.50)	-0.018*** (-18.88)
<i>MTB</i>	-0.027*** (-6.52)	-0.027*** (-6.53)	-0.027*** (-6.54)	0.005*** (6.55)	0.005*** (6.53)	0.005*** (6.62)
<i>LEVERAGE</i>	-0.132*** (-5.25)	-0.131*** (-5.24)	-0.130*** (-5.13)	0.069*** (14.44)	0.069*** (14.61)	0.068*** (14.31)
<i>AGE</i>	-0.000 (-0.24)	-0.000 (-0.11)	-0.000 (-0.32)	-0.000* (-1.91)	-0.000* (-1.73)	-0.000* (-1.79)
<i>SI</i>	0.396*** (4.36)	0.396*** (4.36)	0.395*** (4.35)	-0.068*** (-4.06)	-0.068*** (-4.06)	-0.067*** (-4.00)
<i>RET_VOL</i>	-0.556*** (-10.03)	-0.557*** (-10.03)	-0.557*** (-10.03)			
<i>RETURN</i>				0.004*** (3.26)	0.004*** (3.26)	0.004*** (3.39)
<i>EARN_VOL</i>	-0.370*** (-4.22)	-0.376*** (-4.27)	-0.369*** (-4.20)	0.109*** (7.24)	0.107*** (7.10)	0.107*** (7.13)
<i>NBSEG</i>	0.034*** (3.82)	0.034*** (3.82)	0.034*** (3.80)	-0.001 (-0.60)	-0.001 (-0.63)	-0.001 (-0.87)
<i>SEO</i>	-0.096*** (-6.13)	-0.097*** (-6.14)	-0.095*** (-6.01)	0.019*** (7.72)	0.019*** (7.69)	0.018*** (7.59)
<i>BIGN</i>	0.003 (0.25)	0.003 (0.27)	0.003 (0.25)	-0.003 (-1.34)	-0.003 (-1.30)	-0.003 (-1.43)
<i>IFRS</i>	-0.197*** (-9.02)	-0.198*** (-9.04)	-0.196*** (-8.98)	0.041*** (9.92)	0.041*** (9.95)	0.041*** (9.97)
<i>DIV_Dummy</i>	1.149*** (9.13)	1.065*** (7.37)	1.174*** (7.59)	0.473*** (23.30)	0.444*** (18.14)	0.412*** (15.75)
<i>Industry F.E.</i>	Included	Included	Included	Included	Included	Included
<i>Year F.E.</i>	Included	Included	Included	Included	Included	Included
<i>Clustered by</i>	Firm	Firm	Firm	Firm	Firm	Firm
N	13,692	13,692	13,692	13,608	13,608	13,608
Adjusted R <sup>2</sup>	0.198	0.198	0.198	0.277	0.277	0.277

(Continued on the next page)

Notes:

This table reports future stock returns and annual report unreadability. Panel A shows Fama-MacBeth regression results of future returns on annual report unreadability, and Panel B estimates future return and return volatility on annual report unreadability, as follows:

$Return_{t+1}$  (or  $RET\_VOL_{t+1}$ )

$$= \beta_0 + \beta_1 Unreadability_{jt} + \beta_2 SIZE_{jt} + \beta_3 MTB_{jt} + \beta_4 LEVERAGE_{jt} + \beta_5 AGE_{jt} + \beta_6 SI_{jt} \\ + \beta_7 RET\_Vol_{jt} \text{ (or } Return_{t+1}) + \beta_8 EARN\_Vol_{jt} + \beta_9 NBSEC_{jt} + \beta_{10} SEO_{jt} + \beta_{11} BIGN_{jt} \\ + \beta_{12} IFRS_{jt} + \beta_{13} DIV\_Dummy_{jt} + Year\ Fixed\ Effects + Industry\ Fixed\ Effects + \varepsilon_{jt} \quad (5)$$

$RETURN$  is denoted as annualized monthly return during the fiscal year;  $DIV\_Dummy$  equals to 1 if a firm paid dividends, and 0 otherwise; other variables are defined in Appendix A. Robust t-statistics are reported in parentheses. \*, \*\*, and \*\*\* denote the significance at 10%, 5%, and 1% level, respectively.

#### 4.4.2. 국제회계기준(K-IFRS) 도입과 사업보고서 가독성

국내에서는 2011년부터 국제회계기준(K-IFRS)이 전면 도입되었는데, K-IFRS 도입의 가장 큰 특징 중 하나는 IFRS 회계기준이 재무보고의 이해도(understandability)를 높이기 위해 많은 노력을 기울였다는 점이다. K-IFRS 도입 이후 재무제표의 주석이 담고 있는 정보가 방대해진 것으로 널리 알려져 있듯이, K-IFRS 도입과 같은 사건은 일시에 기업의 사업보고서 공시의 양과 내용에 모두 영향을 미치게 된다. Jang and Roh(2016)는 K-IFRS 도입의 결과를 실증적으로 분석하였고, K-IFRS 재무제표의 주석이 K-GAAP 재무제표 주석에 비해 그 가독성이 떨어진다는 점을 확인하였으며, 이는 IFRS 회계기준이 재무보고의 이해도를 향상시키고 회계기준의 복잡성을 감소시킨다는 주장과는 상반되는 결과라 주장하였다. 본 연구에서도 사업보고서 가독성 결정요인분석을 위한 연구모형(1)에서 기업에 영향을 미치는 주요 사건(Firm events)으로 K-IFRS의 도입(IFRS)을 통제한 바 있다. 여기에 더하여 추가적으로 국제회계기준(K-IFRS) 도입여부가 본 연구의 가설검증에 미치는 영향을 살펴보고자 한다. 우선, 국제회계기준(K-IFRS) 도입여부와 당기이익의 교차항(interaction term)을 연구모형(2)에 추가하여 재수행하였으며, 그 결과는 <Table 8>의 Panel A에 보고하였다. 사업보고서 가독성 측정치인 *FLESCH*, *FOG*, *LENGTH*를 종속변수로 사용하는 경우, 열 (1), (3), (5)에서 보고된 당기이익(*EARNINGS*)의 계수 및 열 (2), (4), (6)에서 보고된 당기이익 더미(*Profit\_Dummy*)의 계수  $\beta_1$ 은 모든 회귀분석에서 음(-)의 값을 가진다. 즉, 낮은 이익을 보고하는 기업의 경우 사업보고서의 가독성이 낮아지는 것을 알 수 있으며, 이는 <Table 4>의 결과와 동일하다. 여기에 열 (1), (3), (5)에는 국제회계기준의 도입여부(*IFRS*)와 당기이익(*EARNINGS*)의 교차항, *IFRS\*EARNINGS*를, 열 (2), (4), (6)에는 국제회계기준의 도입여부(*IFRS*)와 당기이익 더미(*Profit\_Dummy*)의 교차항, *IFRS\*Profit\_Dummy*를, 각각 추가한 회귀분석에서 각 교차항의 계수는 양(+)의 값을 갖는다. 이는 이익이 존재하거나 커질수록 표본기업의 사업보고서 가독성은 높아지는 관계가 국제회계기준의 도입 이후 약화되는 것을 의미한다. 즉, 혹자기업임에도 K-IFRS 도입 후 작성된 사업보고서의 가독성은 낮아지는 경향을 보인다.

또한 전체 표본을 국제회계기준이 도입되기 전(*Pre\_IFRS*)과 후(*Post\_IFRS*)로 나누어 사업보고서의 가독성과 이익지속성의 관계에 대한 연구모형(3)의 하위집단분석(sub-sample analysis)을 시행하였으며, 그 종속변수를  $EARNINGS_{jt+1}$ 로 하는 경우의 결과를 <Table 8>의 Panel B에 보고하였다. 열 (1), (3), (5)에는 국제회계기준이 도입되기 전(*Pre\_IFRS*) 기간의 사업보고서의 가독성과 이익지속성의 연관성에 대한 결과를 나타냈으며, 열 (2), (4), (6)에는 국제회계기준이 도입된 이후(*Post\_IFRS*) 기간의 사업보고서의 가독성과 이익지속성의 연관성에 대한 결과를 나타내고 있다. 사업보고서 가독성 측정치로 열 (1), (2)에서 *FLESCH*, 열 (3), (4)에서 *FOG*, 열 (5), (6)에서 *LENGT H*를 사용하여 분석하는 경우, 이익지속성을 나타내주는  $EARNINGS_{jt}$ 의 계수  $\beta_1$ 은 6개의 열 모두에서 양(+)의 값을 가지며, 교차항의 계수인  $\beta_3$ 의 회귀분석 결과를 살펴보면, 6개의 열 모두에서 통계적으로 유의한 음(-)의 값을 갖는다. 이러한 결과는 <Table 5>의 Panel A의 결과와 동일하다<sup>15)</sup>. 이는 사업보고서 가독성과 이익의 지속성의 관련성 측면에서는 K-IFRS의 도입이 별다른 영향을 미치지 못했음을 보여주는 결과이다.<sup>16)</sup>

- 
- 15) 추가적으로 당기이익의 값이 양수인 경우(흑자기업)만으로 구성된 표본의 분석결과를 <Table 8>의 Panel C에서 살펴보면,  $\beta_1$ 은 6개의 열 모두에서 양(+)의 값을 가지나 교차항의 계수인  $\beta_3$ 은 국제회계기준이 도입되기 전(*Pre\_IFRS*)에만 음(-)의 유의한 결과를 보인다. 이는 국제회계기준의 도입에 따라 사업보고서의 가독성과 이익 간의 관계가 상대적으로 약화된 <Table 8>의 Panel A의 결과와 그 맥락을 같이한다고 할 수 있다. 반면, 당기이익의 값이 음수인 경우(적자기업)에는 교차항의 계수인  $\beta_3$ 이 국제회계기준이 도입된 이후(*Post\_IFRS*)에만 음(-)의 유의한 결과를 보인다.
- 16) Jang and Roh(2016)에서는 K-IFRS 재무제표의 주석이 K-GAAP 재무제표 주석에 비해 가독성이 떨어진다는 점을 확인하고 이를 IFRS 회계기준이 재무보고의 이해도를 향상시키고 회계기준의 복잡성을 감소시킨다는 해석과는 상반되는 결과라고 주장하였다. 본 연구에서는 이와 같은 연구를 확장하여, 사업보고서의 가독성과 기업의 이익(가설 1) 및 이익의 지속성(가설 2)의 관계에 대한 검증을 IFRS 도입 전후로 나누어 파악하고자 하였다. 이러한 분석은 IFRS의 도입이 회계의 질에 미친 영향을 살펴본 연구들(Barth et al. 2008; Christensen et al. 2008; Ahmed et al. 2013; 최관 등 2013)과는 상이하게 해석될 여지가 있다. 즉, 기존의 연구에서 IFRS 도입에 따른 주석의 증가가 관련정보의 이해가능성을 제고하기 위한 목적에서 기인된 것으로 본 반면, 본 연구의 결과는 이러한 주석의 증가가 문구의 간결성이나 길이로 측정되는 가독성 측정치 자체의 상승을 가져오게 되므로, IFRS 도입의 전체적인 효과성을 파악하기에는 어려울 것으로 판단된다. 이러한 면에서 <Table 8>의 결과해석에 있어 유의하여야 할 것으로 보인다.

〈Table 8〉 Cross-sectional Test: IFRS adoption

Panel A : Current Performance						
<i>Dep. Variables</i>	<i>FLESCH</i>		<i>FOG</i>		<i>LENGTH</i>	
	(1)	(2)	(3)	(4)	(5)	(6)
<i>EARNINGS</i>	-6.930*** (-6.70)		-0.477*** (-3.83)		-0.213*** (-6.97)	
<i>IFRS</i> <i>X EARNINGS</i>	2.421** (2.06)		0.280* (1.94)		0.086** (2.39)	
<i>Profit_Dummy</i>		-1.810*** (-4.66)		-0.147*** (-3.04)		-0.065*** (-6.47)
<i>IFRS</i> <i>X Profit_Dummy</i>		1.327*** (2.83)		0.172*** (2.83)		0.032** (2.26)
<i>Control Variables</i>	Included	Included	Included	Included	Included	Included
<i>Industry F.E.</i>	Included	Included	Included	Included	Included	Included
<i>Year F.E.</i>	Included	Included	Included	Included	Included	Included
<i>Clustered by</i>	Firm	Firm	Firm	Firm	Firm	Firm
N	21,917	21,917	21,917	21,917	21,917	21,917
Adjusted R <sup>2</sup>	0.482	0.481	0.666	0.666	0.571	0.570
Panel B : Earnings Persistence(Full Sample)						
<i>Dep. Variables</i>	<i>Earnings<sub>t+1</sub></i>					
<i>Unreadability</i>	<i>FLESCH</i>		<i>FOG</i>		<i>LENGTH</i>	
	Pre-IFRS	Post-IFRS	Pre-IFRS	Post-IFRS	Pre-IFRS	Post-IFRS
	(1)	(2)	(3)	(4)	(5)	(6)
<i>EARNINGS</i>	0.555*** (15.72)	0.584*** (16.79)	0.559*** (17.26)	0.587*** (16.89)	0.568*** (15.74)	0.588*** (17.31)
<i>Unreadability</i>	-0.002*** (-2.70)	-0.003* (-1.76)	-0.018*** (-2.76)	-0.011 (-1.08)	-0.044** (-2.53)	-0.061** (-2.19)
<i>EARNINGS</i> <i>* Unreadability</i>	-0.000** (-2.21)	-0.001*** (-4.55)	-0.002*** (-2.68)	-0.003*** (-4.71)	-0.006** (-2.49)	-0.009*** (-4.92)
<i>Control Variables</i>	Included	Included	Included	Included	Included	Included
<i>Industry F.E.</i>	Included	Included	Included	Included	Included	Included
<i>Year F.E.</i>	Included	Included	Included	Included	Included	Included
<i>Clustered by</i>	Firm	Firm	Firm	Firm	Firm	Firm
N	12,214	9,391	12,214	9,391	12,214	9,391
Adjusted R <sup>2</sup>	0.346	0.475	0.346	0.475	0.346	0.476

(Continued on the next page)



Panel C : Earnings Persistence(Profit Sample)

<i>Dep. Variables</i>	<i>Earnings<sub>t+1</sub></i>					
	<i>Unreadability</i>		<i>FOG</i>		<i>LENGTH</i>	
	Pre-IFRS	Post-IFRS	Pre-IFRS	Post-IFRS	Pre-IFRS	Post-IFRS
	(1)	(2)	(3)	(4)	(5)	(6)
<i>EARNINGS</i>	0.610*** (17.97)	0.673*** (15.72)	0.612*** (19.84)	0.665*** (14.91)	0.617*** (18.45)	0.670*** (15.07)
<i>Unreadability</i>	-0.001* (-1.68)	-0.001 (-0.48)	-0.001 (-0.35)	0.004 (0.44)	-0.012 (-1.33)	-0.018 (-0.65)
<i>EARNINGS</i> * <i>Unreadability</i>	-0.000* (-1.75)	-0.000 (-1.43)	-0.001** (-2.31)	-0.001 (-0.94)	-0.005** (-2.03)	-0.003 (-1.03)
<i>Control Variables</i>	Included	Included	Included	Included	Included	Included
<i>Industry F.E.</i>	Included	Included	Included	Included	Included	Included
<i>Year F.E.</i>	Included	Included	Included	Included	Included	Included
<i>Clustered by</i>	Firm	Firm	Firm	Firm	Firm	Firm
N	9,014	5,917	9,014	5,917	9,014	5,917
Adjusted R <sup>2</sup>	0.226	0.279	0.227	0.277	0.227	0.277

Panel D : Earnings Persistence(Loss Sample)

<i>Dep. Variables</i>	<i>Earnings<sub>t+1</sub></i>					
	<i>Unreadability</i>		<i>FOG</i>		<i>LENGTH</i>	
	Pre-IFRS	Post-IFRS	Pre-IFRS	Post-IFRS	Pre-IFRS	Post-IFRS
	(1)	(2)	(3)	(4)	(5)	(6)
<i>EARNINGS</i>	0.410*** (5.28)	0.451*** (8.19)	0.416*** (5.25)	0.456*** (8.33)	0.428*** (5.27)	0.448*** (8.16)
<i>Unreadability</i>	-0.006** (-2.13)	-0.006* (-1.87)	-0.066*** (-3.01)	-0.027 (-1.39)	-0.268*** (-2.81)	-0.148*** (-2.66)
<i>EARNINGS</i> * <i>Unreadability</i>	-0.000 (-0.62)	-0.001*** (-2.78)	-0.001 (-0.59)	-0.002*** (-2.93)	-0.004 (-0.72)	-0.007*** (-3.21)
<i>Control Variables</i>	Included	Included	Included	Included	Included	Included
<i>Industry F.E.</i>	Included	Included	Included	Included	Included	Included
<i>Year F.E.</i>	Included	Included	Included	Included	Included	Included
<i>Clustered by</i>	Firm	Firm	Firm	Firm	Firm	Firm
N	3,198	3,474	3,198	3,474	3,198	3,474
Adjusted R <sup>2</sup>	0.234	0.399	0.235	0.399	0.235	0.400

Notes: This table shows the impact of IFRS adoption on the properties of annual report readability. Panel A shows the regression results of the equation (2) which includes the interaction term between IFRS adoption and *EARNINGS* and *Profit Dummy*. The dependent variable in Panel B, C, and D is *Earnings<sub>t+1</sub>*. In Panel B, C, and D, we re-estimate the association of earnings persistence and annual report readability using the respective sub-sample: pre-IFRS adoption period sub-sample and post-IFRS adoption period sub-sample. Panel C only use firms with profits, whereas Panel D only use firms with loss. *Profit Dummy* equals to 1 if net income is not negative, and 0 otherwise; other variables are defined in Appendix A. Robust t-statistics are reported in parentheses. \*, \*\*, and \*\*\* denote the significance at 10%, 5%, and 1% level, respectively.

#### 4.5 우리말 가독성 지수의 활용가능성에 대한 검토

본 연구는 영어권 국가에서 사용되는 가독성 지수인 Flesch Reading Ease formula(*FLESCH*), Fog Index(*FOG*) 등의 가독성 지수가 우리말에도 적용가능하다는 전제 하에 이루어졌다. 그러나 우리말을 포함하여 한자를 기반으로 하는 언어의 경우, 단어의 길이가 길다거나 음절이 길다고 해서 반드시 단어의 가독성이 떨어진다고 보기는 어려울 수 있다. 뿐만 아니라 미국을 비롯한 영어권 국가에서도 학생들을 대상으로 하는 교육현장에서 개발된 Fog Index가 사업보고서와 같은 기업이 발행하는 문서의 가독성 지수로 적합하지 않다는 연구결과 역시 존재하고 있다(Loughran and McDonald 2014, 2016). 여기서는 기존의 가독성 지수를 비영어권 국가의 언어에서 사용하는 것의 한계에 대해 검토하고, 이를 개선하기 위해 본 연구에서 추가적으로 실시한 분석에 대해 논의하고자 한다.

##### 4.5.1. 가독성 지수의 한계점에 대한 논의

기존의 가독성 지수(*FLESCH*, *FOG*)가 지니는 한계와 관련하여 국내 및 해외의 언어학 및 교육학계에서는 비영어권 언어의 가독성 지수(*readability formula*)를 새롭게 개발하여 텍스트의 가독성에 대한 객관적인 수치를 제공하고자 하는 노력을 기울이고 있다. 일반적으로 가독성 지수는 어휘 차원의 변수인 의미적(*semantic*) 변수와 문장 차원의 변수인 통사적(*syntactic*) 변수로 이루어진다. 이 중, 후자인 통사적 변수로 가장 널리 사용되고 있는 ‘평균 문장 길이(ASL)’의 경우, 해당 값이 클수록 문장이 길고 어렵다는 것을 나타내기 때문에 언어의 종류에 상관없이 텍스트를 활용한 분석에서 공통적으로 활용할 수 있는 변수로 보인다. 이와 비슷한 관점에서, 언어학의 다양한 선행연구에서도 ‘평균 문장 길이(ASL)’를 가독성 지수 측정시 주요 변수로 활용하고 있다(Flesch 1946; Dale and Chall 1948; Gunning 1952; Kincaid et al. 1975; Stenner et al. 1988; 최인숙 2005; 홍정하 등 2011; 조용구 2016).

통사적 변수가 언어의 종류에 관계없이 공통적으로 활용될 수 있는 변수인 것과 달리, 어휘 차원의 의미적 변수는 사용되는 언어에 따라 다양한 분석 방법이 존재할 수 있다. 여기에는 ‘단어의 길이’, ‘단어의 빈도’, ‘어려운(복잡한) 단어의 수’, ‘단어의 등급’ 등이 고려될 수 있을 것이다. 예를 들어, 본 연구에서 사용하고 있는 Flesch Reading Ease formula(*FLESCH*)는 ‘평균 문장 길이’(통사적 변수)와 ‘단어의 길이’(의미적 변수)의 조합으로 이루어져 있으며, Fog Index(*FOG*)의 경우, ‘평균 문장 길이’(통사적 변수)와 ‘어려운(복잡한) 단어의 수’(의미적 변수)의 조합으로 구성되어 있다. 언어학계에서의 선행연구에 따르면, 우리말 환경에서 어휘 차원의 의미적 변수를 유의미하게 분석하기 위해서는 한자어를 포함한 외국어의 사용 뿐만 아니라, 존비어의 구조, 복합어의 사용에 대한 분석이 선행되어야 하며, 이를 위해 관련 어휘 목록의 작성 및 지속적인 어휘 목록의 업데이트, 그리고 이를 위한 분석 시스템의 개발이 필요한 실정이다(홍정하 등 2011). 다만, 국내 언어학계의 연구에서조차 의미적 변수와 관련하여 어휘목록의 정확성 확보 문제와 시스템 개발의 고비용 문제로 인하여 여전히 초기 단계의 연구에 머물러 있는 것이 현실이다.

이러한 관점에서의 국내 언어학 연구의 현실 및 비영어권 언어의 가독성 지수 분석에 대한 제약조건을 고려한다면, 본 연구에서 사용하고 있는 *FLESCH*와 *FOG* 등의 가독성 지수는 우리말 가독성 지수 분석을 위해 필요한 여러 가지 속성들을 충분히 반영하지 못한다고 볼 수 있다. 다

만, 위의 한계점에도 불구하고 *FLESCH*와 *FOG*변수가 국내 언어학계에서 연구 대상으로 삼고 있는 ‘단어의 길이’와 ‘어려운 단어의 수’ 등을 반영하고 있으므로, 기존의 가독성 지수를 사용하더라도 추가적인 공헌점이 있을 것으로 보인다.

#### 4.5.2. 가독성 지수의 한계에 따른 추가분석

4.5.1.에서 논의된 바와 같이, 영어권 언어를 대상으로 하는 기존의 가독성 지수(*FLESCH*와 *FOG*)를 우리말 사업보고서에 적용하는 것은 일견 한계점이 있는 것으로 보인다. 이와 같은 한계점을 완화하기 위하여 본 연구에서는 가독성 지수의 한계에 따른 다양한 추가분석을 실시하였다.

먼저 Fog Index의 구성요소인 ‘어려운 단어의 수’가 지니는 의미적 변수로서의 성격을 고려하여, 연구결과의 강건성을 위해 ‘어려운 단어’의 정의를 3음절 뿐만 아니라 5음절, 7음절, 10음절 이상의 단어로 대체하여 Fog Index를 다시 측정하였다. 이를 통해, 한자 기반의 단어가 많은 한글의 특성상 3음절 이상의 언어들이 특별히 어렵고 복잡한 의미를 전달하는 것으로 볼 수 없다는 주장을 반영할 수 있을 것으로 보인다. 이는 사업보고서에 포함된 비즈니스 용어들의 경우 일반적으로 3음절 이상인 단어들이 특별히 어렵고 복잡한 의미를 전달하지 않는다는 선행연구의 주장에 대한 강건성 검증이기도 하다(Loughran and McDonald 2014). 이처럼 다시 측정된 Fog Index를 사용하여 기존의 가설검증을 재수행하였으며, 그 결과는 <Table 9>에 보고하였다. <Table 9>에 보고된 사업보고서의 가독성과 당기의 이익의 연관성에 대한 연구모형(2)의 회귀분석 결과는 가설 1의 검증 결과를 지지하며, 별도로 보고하지는 않았으나 사업보고서의 가독성과 이익지속성의 연관성에 대한 연구모형(3)의 회귀분석 결과 역시 가설 2의 검증 결과를 지지한다.

다음으로 기존의 가독성 지수 중 의미적 변수(예: 단어의 길이, 어려운 단어의 수 등)가 지니는 영향을 최소화하기 위하여 기존의 가독성 지수에서 통사적 변수로 사용되는 ‘평균 문장 길이(ASL)’를 하나의 독자적인 가독성 지수로 정의하고 기존의 분석을 다시 실시하였다. 이는 국내 언어학 연구에서 ‘평균 문장 길이’를 주요 가독성 지수로 보고 있다는 점을 고려한 강건성 검증이라고 할 것이다. 별도로 보고하지는 않았으나, ‘평균 문장 길이(ASL)’를 하나의 독자적인 가독성 지수로 정의하더라도 본 연구의 결과는 강건하게 유지됨을 확인하였다. 즉, 기존의 분석과 동일하게 기업의 이익이 낮을수록 평균문장길이가 길어진다는 점을 확인할 수 있었다. 이는 우리말 사업보고서에서도 영어와 유사한 방식으로 측정한 지표들이 가독성 지수로서의 유의미한 역할을 한다는 점을 시사한다고 할 것이다.

또한 Loughran and McDonald(2014)에서는 Fog Index의 문제점을 지적하면서 문서의 파일크기(file size)와 같은 보다 직접적인 측정치를 사용하는 것이 적합하다고 보았다. 이와 관련하여 Bonsall et al.(2017)의 연구에서는 파일크기를 측정함에 있어 사업보고서 내에 담긴 HTML, XML, PDF, 그림파일 등으로 인해 그 측정치가 왜곡될 수 있다는 점을 주장하였다. 본 연구의 표본인 우리말 사업보고서 역시 Microsoft Word 상의 DOCX 형식의 파일로 저장되어 있고, 여기에는 조직도, 제품군 구성, 중소기업기준검토표 등 다양한 내용이 그림파일의 형식으로 포함되어 있다. Loughran and McDonald(2014)와 Bonsall et al.(2017)의 연구를 종합적으로 고려하여 본 연구에서는 각 기업-연별(firm-year) 사업보고서를 DOCX 형식에서 텍스트파일 형식(TXT)으로 변환하여, 사업보고서 작성에 적용된 서식 및 그림을 제외하고 문서의 파일크기(file size)를 측정하고자 하였다. 이러한 과정을 통해 사업보고서 내용만의 파일크기

기에 자연로그를 취한 값(*TXTSIZE*)을 가독성 지수의 추가적인 대용치로 보고 본 연구의 분석을 다시 한번 수행하였다. 별도로 보고하지는 않았으나, 기존의 분석과 동일하게 기업의 이익이 낮을수록 파일크기(*TXTSIZE*)가 커진다는 점을 확인하였다. 이는 Fog Index가 갖는 한계점을 어느 정도 보완하는 결과라고 해석할 수 있으며, FLESCH나 FOG와 같은 가독성 지수 외에도 파일크기와 같은 보다 직접적인 측정치 역시 가독성 지수로서의 역할을 한다는 점을 보여주고 있다.

이와 같은 추가적인 분석에도 불구하고, 본 연구가 우리말 사업보고서의 가독성에 대한 가장 편의없고 효율적인(unbiased and efficient) 측정치를 사용했다고 볼 수는 없을 것이다. 그러나 이와 같은 잠재적인 측정오차(measurement error)에도 불구하고, Li(2008) 등의 기본적인 결과가 우리말 사업보고서에서도 재현가능하다는 점은 주로 미국 등에서 연구가 이루어지는 가독성 측정치가 국내 연구에서도 의미있게 사용될 수 있다는 점을 시사한다고 볼 수 있다.

〈Table 9〉 Fog Index Variants

<i>Dep. Variables</i>	<i>FOG5</i>		<i>FOG7</i>		<i>FOG10</i>	
	(1)	(2)	(3)	(4)	(5)	(6)
<i>EARNINGS</i>	-0.479*** (-5.76)		-0.488*** (-8.77)		-0.305*** (-8.17)	
<i>Profit_Dummy</i>		-0.095*** (-3.50)		-0.109*** (-6.45)		-0.069*** (-5.94)
<i>Control Variables</i>	Included	Included	Included	Included	Included	Included
<i>Industry F.E.</i>	Included	Included	Included	Included	Included	Included
<i>Year F.E.</i>	Included	Included	Included	Included	Included	Included
<i>Clustered by</i>	Firm	Firm	Firm	Firm	Firm	Firm
N	21,917	21,917	21,917	21,917	21,917	21,917
Adjusted R <sup>2</sup>	0.584	0.583	0.508	0.506	0.626	0.625

Notes:

This table shows the regression results of hypothesis 1 using the equations (2) with alternative FOG indices. *FOG* is estimated as follows:  $FOG = (\text{words per sentence} + \text{percent of complex words}) * 0.4$ ; complex words in *FOG* index are originally defined as words with three syllables or more. Alternative *FOG* indices, complex words of *FOG5*, *FOG7*, and *FOG10* are defined as words with five, seven, and ten syllables or more, respectively; *Profit\_Dummy* equals to 1 if net income is not negative, and 0 otherwise; other variables are defined in Appendix A. Robust t-statistics are reported in parentheses. \*, \*\*, and \*\*\* denote the significance at 10%, 5%, and 1% level, respectively.

#### 4.6. 추가적인 강건성 검증

본 연구에서는 선행연구에 따라 *FLESCH*, *FOG*, *LENGTH* 등의 가독성 대용치를 측정할 때, 사업보고서 중 표로 작성된 부분은 제외하고 측정하였다 (Li, 2008; Leheavy et al. 2011). 하지만 우리나라 기업들의 경우 ‘기업공시서식 작성기준’에 의거하여 사업보고서 작성서식이 표로 규정된 내용 이외에도, 투자자들이 사업보고서를 한눈에 보고 이해하기 쉽게 표를 이용하여 많은 정보들을 요약하여 나타내는 경우가 많다. 따라서 위와 같은 국내 사업보고서의 특성을 반영하고 가설검증의 강건성을 검증하기 위하여 사업보고서 가독성 측정시 사업보고서 상의 표를 포함하여, *FLESCH*, *FOG*, *LENGTH*를 새로이 측정한 뒤, 가설 1과 가설 2의 분석

을 반복하여 수행하였다. 별도로 결과를 보고하지는 않았으나, 새로 측정된 *FLESCH*, *FOG*, *LENGTH*를 각각 종속변수로 사용하여 <Table 4>를 재검증하는 경우, 당기이익(*EARNINGS*)의 계수 및 당기이익 더미(*Profit\_Dummy*)의 계수  $\beta_1$ 은 모든 회귀분석에서 통계적으로 음(-)의 유의한 값을 갖는다. 이는 <Table 4>의 사업보고서 내용 중 표를 제외한 가독성 대용치인 *FLESCH*, *FOG*, *LENGTH*를 활용한 분석에 따른 결과와 동일하다. 또한, 사업보고서의 가독성과 이익지속성의 관계에 대한 연구모형(3)의 회귀분석 결과, 이익지속성을 나타내주는 *EARNINGS<sub>it</sub>*의 계수  $\beta_1$ 은 가독성 대용치인 *FLESCH*, *FOG*, *LENGTH*를 활용한 모든 분석에서 양(+)의 값을 가지며, 각각의 가독성 측정치를 사용한 분석에서  $t+1$ 기,  $t+2$ 기, 그리고  $t+3$ 기의 계수  $\beta_1$ 을 서로 비교할 때, 그 크기가 시간이 갈수록 작아졌다. 또한, 교차항의 계수인  $\beta_3$ 의 회귀분석 결과를 살펴보면, 모든 분석에서 통계적으로 유의한 음(-)의 값을 갖는다. 이러한 결과는 <Table 5>의 Panel A의 결과와 동일하다. 더불어 <Table 5>의 Panel B와 C의 분석과 마찬가지로, 본 연구의 표본을 흑자기업과 적자기업으로 나누어 연구모형(3)의 분석을 다시 수행하였으며, 흑자기업만으로 구성된 표본의 분석결과는 앞선 전체 표본의 분석결과와 유사하다.  $\beta_1$ 은 모든 분석에서 양(+)의 값으로 이익지속성을 보여주며, 교차항의 계수인  $\beta_3$ 은 가독성 대용치를 *LENGTH*로 사용하였을 때,  $t+2$ 기의  $\beta_3$ 을 제외한 모든 분석에서 음(-)의 값을 갖는다. 다시 말해, 이익지속성이 낮을수록 사업보고서 가독성이 낮아지는 결과를 보인다. 한편, 적자기업표본 분석결과를 살펴보면,  $\beta_1$ 은 가독성 대용치인 *FLESCH*, *FOG*, *LENGTH*를 활용한 모든 분석에서 양(+)의 값으로 이익(손실)지속성을 보여주며, 교차항의 계수인  $\beta_3$ 의 회귀분석 결과는 가독성 대용치를 *FLESCH*와 *LENGTH*로 사용하였을 때,  $t+1$ 기의  $\beta_3$ 과 가독성 대용치를 *LENGTH*로 사용하였을 때,  $t+3$ 기의  $\beta_3$ 에서만 통계적으로 유의한 음(-)의 결과를 보인다. 즉 적자기업인 경우 그 손실이 지속되며, 그 손실의 지속성이 시간이 갈수록 작아지는 양상을 보이는 데, 이는 흑자기업의 이익에 대한 분석과 다르지 않다. 또한 대체로 가독성이 낮은 사업보고서를 제출하는 기업의 이익지속성은 낮으며, 지속성이 낮은 긍정적인 사건을 공시하기 위하여 흑자기업은 가독성이 낮은 사업보고서를 공시하나, 기업가치에 지속적으로 부정적인 사건을 공시하는 적자기업은 사업보고서의 가독성과 관련이 없음을 발견한 <Table 5>의 결과와 동일하다. 종합해보면, 사업보고서 상의 표를 포함한 사업보고서 가독성 측정치, *FLESCH*, *FOG*, *LENGTH*를 활용하여 본 연구의 가설1과 2를 재수행한 결과는 기존 분석 결과인 <Table 4>와 <Table 5>의 결과와 다르지 않다고 할 수 있다.

## V. 결 론

본 연구는 영문기반의 가독성 연구들이 제시한 텍스트 분석기법들이 우리말 기반의 사업보고서에도 적용가능한지를 검증함으로써, 국내에서 가독성 연구의 가능성을 타진해보는 것을 그 목적으로 한다. 금융감독원 전자공시시스템에서 1998년과 2015년 사이에 공시된 사업보고서를 전수 분석한 결과, 기존 선행연구의 결과가 우리말 기반의 사업보고서에서도 의미있게 재현될 수 있음을 보고하였다.

본 연구는 Li(2008)에서 촉발된 가독성 연구를 국내에 적용하고 확장하기 위하여, 대표본의 실증분석 결과를 보고하였다는 것을 그 공헌으로 삼는다. 국내 뿐만 아니라 비영어권 국가들에

서 전반적으로 현지어 기반의 가독성 연구들이 활성화되고 있지 못한 상황을 고려한다면, 본 연구에서 보고한 분석결과들을 후속 연구에서 더 발전시키고 확장할 필요가 있다. 이를 위하여 본 연구는 실증분석을 진행하면서 직면했던 절차를 구체적으로 기술하고, 진행과정에서의 기술적 어려움을 논의하였으며, 분석시 사용했던 파이썬 코드(Appendix B. 참조)를 공개하고 이를 구현할 수 있는 전산환경에 대한 기초적인 논의를 제공하고 있다. 이를 통하여 국내에서 가독성 연구를 진행할 때의 제약사항에 대한 논의의 단초를 제공하고, 후속 연구자들이 더 효율적인 방식으로 실증분석을 진행할 수 있기를 기대한다. 그러나 본 연구진의 기술적 미숙함으로 인하여 절차의 효율성 측면에서 상당한 한계가 있을 수 있다는 점을 인지할 필요가 있으며, 본 연구의 연구절차와 방법론은 한 가지 예시로만 이해되어야만 할 것이다.

“본 논문은 다른 학술지 또는 간행물에 게재되었거나 게재 신청되지 않았으며,  
한국회계학회 연구윤리규정을 준수하여 작성되었음을 확인함”

## REFERENCES

- Ahmed, A., M. Neel, and D. Wang. 2013. Does mandatory adoption of IFRS improve accounting quality? Preliminary evidence. *Contemporary Accounting Research* 30 (4): 1344-1372.
- Ahn, T. S, and J. H. Lee. 2005. Determinants of Voluntary Disclosures in Overview of Operations - Korean Evidence. *Korean Accounting Review* 30 (2): 33-78.
- Barth, M., W. Landsman, and M. Lang. 2008. International accounting standards and accounting quality. *Journal of Accounting Research* 46 (3): 467-498.
- Bao, D., R. Files, and S. Radhakrishnan. 2015. Industry-specialist Audit Fee Premium and Financial Statement Complexity. *Working Paper*.
- Barnett, A., and K. Leoffler. 1979. Readability of accounting and auditing messages. *The Journal of Business Communication* 16 (3): 49-59.
- Bloomfield, R. J. 2002. The “incomplete revelation hypothesis” and financial reporting. *Accounting Horizons* 16 (3): 233-243.
- Bonsall, S. B., and B. P. Miller. 2017. The impact of narrative disclosure readability on bond ratings and the cost of debt. *Review of Accounting Studies* 22 (2): 608-643.
- Botosan, C. A. 1997. Disclosure Level and Cost of Equity Capital. *The Accounting Review* 72 (3): 323-349.
- Bradbury, M. E. 1992. Voluntary disclosure of financial segment data. *Accounting and Finance* 32 (1): 15-26.
- Bryan, S. H. 1997. Incremental Information Content of Required Disclosures Contained in Management Discussion and Analysis. *The Accounting Review*

- 72: 285-301.
- Brochet, F., P. Naranjo, and G. Yu. 2016. The capital market consequences of language barriers in the conference calls of non-U.S. firms. *The Accounting Review* 91 (4): 1023-1049.
- Burgstahler, D. and I. Dichev. 1997. Earnings Management to Avoid Earnings Decreases and Losses, *Journal of Accounting and Economics* 24 (1): 99-126.
- Choe, I. S. 2005. A Comparative Study on Modelling Readability Formulas : Focus on Primary and Secondary Textbooks. *Journal of the Korean Society for Information Management* 22 (4): 173-195. [Printed in Korean]
- Choi, K., J. I. Park, and S. H. Choi. 2013. The Effect of IFRS Adoption on Accounting Quality and Economic Consequences: A Literature Review and Suggestions for Future Research. *Korean Accounting Review* 38 (2): 525-591. [Printed in Korean]
- Choi, S. J. and K. J. Hwang. 2001. Analysis on Characteristics of INternet Financial Reporting of Korean Firms. *Korean Accounting Journal* 10 (3): 1-20. [Printed in Korean]
- Christensen, H. B., E. Lee, M. Walker, and C. Zeng. 2008. Incentives or standards: what determines accounting quality changes around IFRS adoption? *European Accounting Review* 24 (1): 31-61.
- Christensen, H. B., and V. V. Nikolaev (2017). Contracting on GAAP changes: large sample evidence. *Journal of Accounting Research* 55 (5): 1021-1050.
- Core, J. E. 2001. A review of the empirical disclosure literature: discussion. *Journal of Accounting and Economics* 31 (1-3): 441-456.
- Dale, E., and J. S. Chall. 1948. A formula for predicting readability. *Educational Research Bulletin* 27 (2): 37-54.
- Ertugrul, M., J. Lei, J. Qiu, and C. Wan. 2017. Annual Report Readability, Tone Ambiguity, and the Cost of Borrowing. *Journal of Financial and Quantitative Analysis* 52 (2): 811-836.
- Fama, E. F., and J. D. MacBeth. 1973. Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy* 81 (3): 607-636.
- Flesch, R. 1948. A New Readability Yardstick. *Journal of Applied Psychology* 32 (3): 221-233.
- Gunning, Robert, 1952, *The Technique of Clear Writing*, McGraw-Hill, New York.
- Hanley, K. W., and G. Hoberg. 2012. Litigation risk, strategic disclosure and the underpricing of initial public offerings. *Journal of Financial Economics* 103 (2): 235-254.
- Hayn, C. 1995. The information content of losses. *Journal of Accounting and Economics* 20 (2): 125-153.
- Healy, P. M., and K. G. Palepu. 2001. Information asymmetry, corporate disclosure, and the capital markets: A review of the empirical disclosure literature. *Journal of Accounting and Economics* 31 (1): 405-440.

- Hoberg, G., and G. Phillips. 2010. Product market synergies and competition in mergers and acquisitions: A text-based analysis. *Review of Financial Studies* 23 (10): 3773-3811.
- Hoberg, G., and G. Phillips. 2016. Text-based network industries and endogenous product differentiation. *Journal of Political Economy* 124 (5): 1423-1465.
- Hoberg, G., G. Phillips, and N. Prabhala. 2014. Product market threats, payouts, and financial flexibility. *The Journal of Finance* 69 (1): 293-324.
- Hong, J. H., J. W. Choe, and S. H. You. 2011. A Verification and Application of a Correlation between Text Levels and Readability Using Korean Learning Material. *Language Information* 12: 111-148. [Printed in Korean]
- Hoogervorst, H. 2013. Breaking the Boilerplate, Remarks to the IFRS Conference. Amsterdam, June 27.
- Hope, O. K., and W. B. Thomas. 2008. Managerial Empire Building and Firm Disclosure. *Journal of Accounting Research* 46 (3): 591-626.
- Hyeon, J., T. S. Ahn, Y. S. Choi and J. H. Hyun. 2014. The Determinants of Nonfinancial Information Disclosure in Management Discussion and Analysis (MD&A). *Korean Accounting Review* 39 (2): 179-219.
- Jang, M. H., and J. H. Rho. 2016. IFRS adoption and financial statement readability: Korean evidence. *Asia-Pacific Journal of Accounting & Economics* 23 (1): 22-42.
- Jo, H, and Y. Kim. 2007. Disclosure frequency and earnings management. *Journal of Financial Economics* 84: 561-590.
- Jo, Y. G. 2016. Korean Language Readability Formula. *Journal of Reading Research* 41: 73-92. [Printed in Korean]
- Jones, M. J., and P. A. Shoemaker. 1994. Accounting narratives: A review of empirical studies of content and readability. *Journal of Accounting Literature* 13: 142-162.
- Kasznik, R., and M. McNichols. 2002. Does meeting earning expectation matter? Evidence from analyst forecast revisions and share prices. *Journal of Accounting Research* 40 (3): 727-759.
- Kim, C. F., K. Wang, and L. Zhang. 2018. Readability of 10-K Reports and Stock Price Crash Risk. *Contemporary Accounting Research*. Forthcoming.
- Kim, H. J., J. W. Park., and J. W. Lee. 2015. A Study on the Textual Analysis Research Environment using the DART System in Korea. *Korean Accounting Journal* 24 (4): 199-221. [Printed in Korean]
- Kim, J., Y. Kim, and J. Zhou. 2017. Languages and Earnings Management. *Journal of Accounting and Economics* 63: 288-306.
- Kincaid, J. P., R. P. Fishburne Jr, R. L. Rogers, and B. S. Chissom. 1975. Derivation of new readability formulas for navy enlisted personnel. Naval Technical Training Command Millington TN Research Branch.
- Kothari, S. P., A. J. Leone, and C. E. Wasley. 2005. Performance matched



- discretionary accrual measures. *Journal of Accounting and Economics* 39 (1): 163-197.
- Kothari, S. P., Li, X., & Short, J. E. 2009. The effect of disclosures by management, analysts, and business press on cost of capital, return volatility, and analyst forecasts: A study using content analysis. *The Accounting Review* 84 (5): 1639-1670.
- Kwon, O. J., and H. Y. Park. 2017. The Association between Industry Competition and MD&A Disclosure. *Korean Accounting Journal* 26 (3): 141-173. [Printed in Korean]
- Lang, M., and L. Stice-Lawrence. 2015. Textual analysis and international financial reporting: Large sample evidence. *Journal of Accounting and Economics* 60 (2-3): 110-135.
- Lawrence. 2013. Individual investors and financial disclosure. *Journal of Accounting and Economics* 56 (1): 130-147.
- Lee, A. Y., S. H. Kim, and S. B. Chun. 2013. The disclosure quality of management discussion and analysis and the firms' characteristics. *Korean Accounting Journal* 22 (1): 175-201. [Printed in Korean]
- Lee, M. Y., H. S. Shim, and J. H. Choi. 2012. Board Characteristics and the Frequency of Disclosure of the Firms Belonging to Large Business Groups. *Korean Accounting Review* 37 (2): 297-320. [Printed in Korean]
- Lee, Y. 2012. The effect of quarterly report readability on information efficiency of stock prices. *Contemporary Accounting Research* 29 (4): 1137-1170.
- Lehavy, R., F. Li, and K. Merkley, 2011. The effect of annual report readability on analyst following and the properties of their earnings forecasts. *The Accounting Review* 86 (3): 1087-1115.
- Li, F. 2006. Do stock market investors understand the risk sentiment of corporate annual reports?. University of Michigan, *Working Paper*
- Li, F. 2008. Annual report readability, current earnings, and earnings persistence. *Journal of Accounting and Economics* 45 (2): 221-247.
- Li, F. 2010a. The information content of forward looking statements in corporate filings—A naïve Bayesian machine learning approach. *Journal of Accounting Research* 48 (5): 1049-1102.
- Li, F. 2010b. Textual Analysis of Corporate Disclosures : A Survey of the Literature. *Journal of Accounting Literature* 29: 143-165.
- Li, F., R. Lundholm, and M. Minnis. 2013. A measure of competition based on 10-K filings. *Journal of Accounting Research* 51 (2): 399-436.
- Li, X. 2010. The impacts of product market competition on the quantity and quality of voluntary disclosures. *Review of Accounting Studies* 15 (3): 663-711.
- Lo, K., F. Ramos, and R. Rogo. 2017. Earnings management and annual report readability. *Journal of Accounting and Economics* 63 (1): 1-25.

- Loughran, T., and B. McDonald. 2014. Measuring readability in financial disclosures. *The Journal of Finance* 69 (4): 1643-1671.
- Loughran, T., and B. McDonald. 2016. Textual Analysis in Accounting and Finance: A Survey. *Journal of Accounting Research* 54 (4): 1187-1230.
- Lundholm, R. J., R. Rogo, and J.L. Zhang. 2014. Restoring the tower of Babel: how foreign firms communicate with U.S. investors. *The Accounting Review* 89: 1453-1485.
- Nam, H. J. and J. H. Choi. 2009. The Association between Firm Location and the Tendency to Meet or Beat Analysts' Earnings Forecasts. *Korean Management Review* 38 (2): 611-639. [Printed in Korean]
- Paek, W. S. 2008. Fourth-quarter earnings management and earnings persistence. *Korean Accounting Review* 33 (2): 111-144. [Printed in Korean]
- Park, S. H. 2013. The Effects of Controlling Shareholders on Cash Holdings, Investment, and Dividend Policy. *The Korean Journal of Financial Management* 30 (3): 133-169. [Printed in Korean]
- SEC, 1998. A Plain English Handbook: How to Create Clear SEC Disclosure Documents. U.S. Securities and Exchange Commission, Washington, DC.
- Shim, H. S., M. Y. Lee, and J. H. Choi. 2010. Ownership Divergence and the Frequency of Information Disclosure. *Korean Accounting Review* 35 (4): 39-74. [Printed in Korean]
- Sloan, R. G. 1996. Do Stock Prices Fully Reflect Information in Accruals and Cash Flows about Future Earnings? *The Accounting Review* 71 (3): 289-315.
- Smith, J. E., and N. P. Smith. 1971. Readability: A measure of the performance of the communication function of financial reporting. *The Accounting Review* 46 (3): 552-561.
- Song, I. M., W. S. Paek., and H. S. Park. 2004. Earnings Management for the Avoidance of Losses. *Korean Accounting Journal* 13 (2): 29-51. [Printed in Korean]
- Stenner, A. J., I. Horabin, D. R. Smith, and M. Smith. 1988. Most comprehension test do measure reading comprehension: A response to McLean and Goldstein. *Phi Delta Kappan*: 765-767.
- Subramanian, R., Insley, R. G., and Blackwell, R. D. 1993. Performance and readability: A comparison of annual reports of profitable and unprofitable corporations. *The Journal of Business Communication* 30 (1): 49-61.
- You, H., and X. Zhang. 2009. Financial Reporting Complexity and Investor Underreaction to 10-K Information. *Review of Accounting Studies* 14 : 559-86.

## 국내 참고 문헌

- 권오진, 박하연. 2017. 산업 내 경쟁과 경영진단의견서 (MD&A) 공시수준 간의 관련성 연구. 회계저널 (제26권 제3호): 141-173.
- 김형준, 박종원, 이재원. 2015. 전자공시시스템 (DART)을 활용한 국내 텍스트 분석 (Textual Analysis) 환경에 관한 연구. 회계저널 (제24권 제4호): 199-221.
- 남혜정, 최종학. 2009. 기업의 위치와 음의 이익예측오차 회피 성향 사이의 관계. 경영학연구 (제38권 제2호): 611-639.
- 박순홍. 2013. 지배주주가 보유현금, 투자, 배당에 미치는 영향. 재무관리연구 (제30권 제3호): 133-169.
- 백원선. 2008. 4분기 이익조정과 이익지속성. 회계학연구 (제33권 제2호): 111-144.
- 송인만, 백원선, 박현섭. 2004. 적자보고를 회피하기 위한 이익조정. 회계저널 (제13권 제2호): 29-51.
- 심호식, 이문영, 최종학. 2010. 지분을 괴리도와 공시 정보의 빈도 사이의 관계. 회계학연구 (제35권 제4호): 39-74.
- 이문영, 심호식, 최종학. 2012. 대규모기업집단 소속기업의 이사회 특성과 공시 빈도 사이의 관계. 회계학연구 (제37권 제2호): 279-320.
- 이아영, 김성혜, 전성빈. 2013. 경영진단의견서 (Management Discussion and Analysis: MD&A) 의 공시품질과 기업특성. 회계저널 (제22권 제1호): 175-201.
- 조용구. 2016. 글의 수준을 평가하는 국어 이독성 공식. 독서연구 (제41권): 73-92.
- 최관, 박종일, 최성호. 2013. IFRS 도입이 회계의 질에 미친 영향과 경제적 효과: 문헌검토와 미래연구에 대한 제안. 회계학연구 (제38권 제2호): 525-591.
- 최순재, 황국재. 2001. 인터넷을 통한 자발적 재무공시기업의 특성 분석. 회계저널 (제10권 제3호): 1-20.
- 최인숙. 2005. 텍스트의 언어적 난이도 측정 공식 비교연구-초중고 교과서를 중심으로. 정보관리학회지 (제22권 제4호): 173-195.
- 홍정하, 최재웅, 유석훈. 2011. 텍스트 수준과 가독성: 한국어 학습교재를 이용한 검증과 응용. 언어정보 (제12권): 111-148.

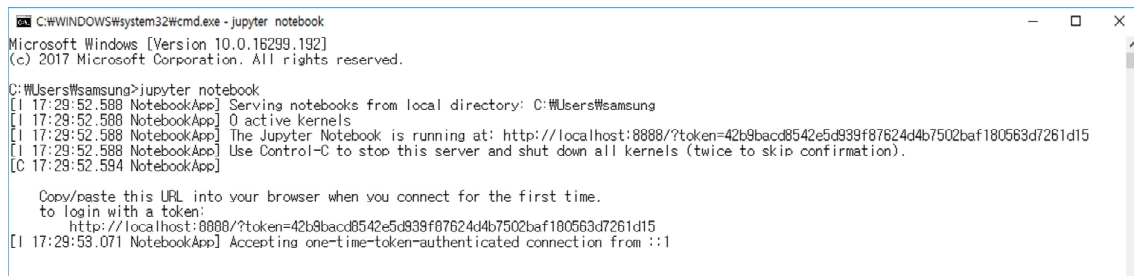
## 〈Appendix A〉 Variable Definitions

Variables	Definition
<b><i>Unreadability Variables</i></b>	
<i>FLESCH</i>	: FLESCH Index. $FLESCH = (-1) * (206.835 - 84.6 * \text{Average word length}(AWL) - 1.015 * \text{Average sentence length}(ASL))$
<i>FOG</i>	: Fog Index. $FOG = (ASL + \text{percentage of complex words}) * 0.4$
<i>LENGTH</i>	: the natural logarithm of number of words in annual report( <i>Word_num</i> ).
<b><i>Regression Variables</i></b>	
<i>SIZE</i>	: the natural logarithm of the market value of equity at the end of the fiscal year <i>t</i> .
<i>MTB</i>	: Market-to-Book ratio calculated as market value of equity divided by book value of equity.
<i>LEVERAGE</i>	: Financial leverage, calculated as short-term debt plus long term debt divided by book value of total assets.
<i>AGE</i>	: Years after going public (IPO)
<i>SI</i>	: Special item calculated as special gain minus special loss divided by lagged total asset
<i>RET_Vol</i>	: Standard deviation of monthly stock return during the fiscal year <i>t</i>
<i>EARN_VOL</i>	: Standard deviation of operating earnings from year <i>t-4</i> to year <i>t</i> .
<i>NBSEG</i>	: Natural logarithm of number of business segments firm operates.
<i>SEO</i>	: Indicator variable equals to 1 if firm issue equity in fiscal year <i>t</i> and 0 otherwise..
<i>EARNINGS</i>	: Net income divided by lagged total assets.
<i>OCF</i>	: Operating cash flow divided by lagged total assets.
<i>RETURN</i>	: Average of monthly return during the fiscal year <i>t</i> .
<i>BIG4</i>	: Indicator variable equals to 1 if a firm's auditor is one of the Big 4(5 or 6) auditors in fiscal year <i>t</i> .
<i>IFRS</i>	: Indicator variable equals to 1 if a firm adopted IFRS, and 0 otherwise.

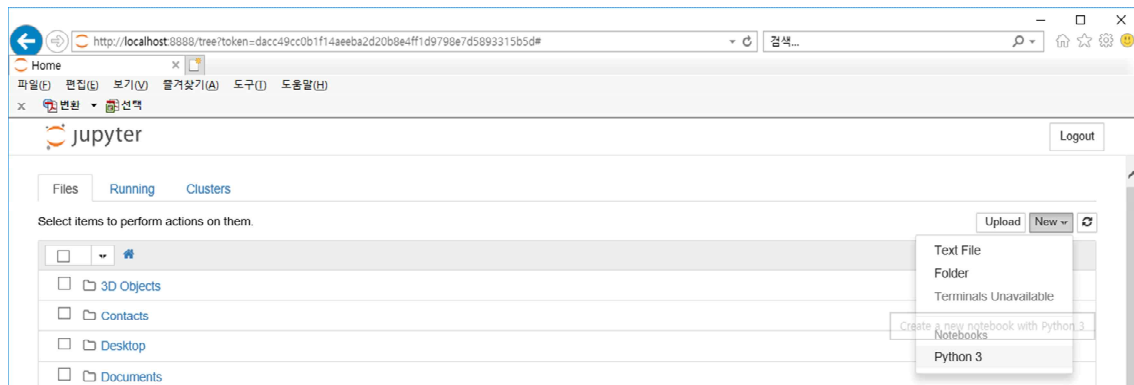
## 〈Appendix B〉 How to construct readability index

Section 3.1에서 언급한 것과 같이, 본 연구의 표본은 금융감독원의 전자공시시스템(DART)에서 열람가능한 1998년부터 2015년까지 국내 유가증권시장 및 코스닥시장에 상장된 모든 기업을 대상으로 한다. 본 연구를 수행하기 위해서 분석의 대상이 되는 사업보고서들로 구성된 데이터를 웹 크롤링(web-crawling) 방식이 아닌, PDF 형식 혹은 Microsoft word 프로그램의 DOCX 형식으로 다운로드 받아 분석을 진행하였다. 표본기업의 사업보고서 가독성(Readability)을 산출하기 위하여, 파이썬 프로그램<sup>17)</sup>을 사용하였으며, 특히 효율적인 코딩을 위하여 사업보고서 파일명을 “회사명+사업년도”방식으로 부여하였다.<sup>18)</sup> 가독성 측정을 위한 상세한 절차는 다음과 같다.

1. 윈도우에서 실행(command)창을 열고, “jupyter notebook”을 입력한 후 “enter”로 실행하면 파이썬을 시작할 수 있다.<sup>19)</sup>

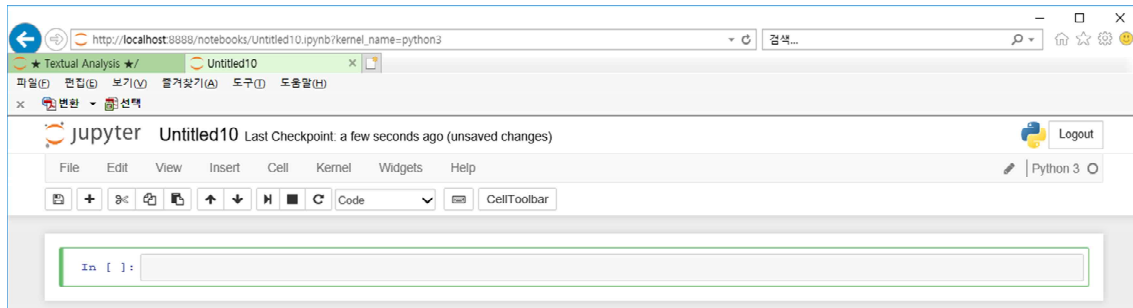


2. 파이썬 화면에서 아래의 “New”를 누르면 나오는 선택사항 중 “Python 3”을 선택하여 새 파일을 생성한다.



- 17) 파이썬 프로그램은 <https://www.python.org/downloads/> 에서 무료로 다운로드 가능하다. 또한 윈도우에서 실행(command)창에서 “pip install jupyter”을 입력 후 “enter”로 실행하면 “jupyter notebook”을 설치할 수 있다. 여기서, jupyter notebook이란, 파이썬(Python)이나 R 파일을 작성하고 실행하는 개발 환경을 제공하는 일종의 웹 애플리케이션으로, 길고 복잡한 코드도 셀 단위로 나누어 작성하고 실행하면서 인터랙티브한 동작이 가능한 것이 그 장점이라 할 수 있겠다.
- 18) 예를 들어, “삼성전자”의 “2015”년도 사업보고서의 경우 파일명을 “삼성전자2015”로 부여하였다.
- 19) 사용자의 환경에 따라 파이썬을 처음 사용하는 경우, 혹은 파이썬에 docx package가 설치되어 있지 않을 경우 분석을 위하여 윈도우에서 실행(command)창에서, “pip install python-docx”를 입력 및 실행하여 docx package를 설치 후 진행하여야 한다.

3. 효율적인 코딩을 위하여 사업보고서 파일명을 “회사명+사업년도”방식으로 부여하였는데, 아래의 명령어 창에 사업년도별로 가독성을 측정하기 위한 코드를 다음과 같이 입력한다. 또한, Python 코드에 “#”뒤로 상세한 주석을 덧붙였다.<sup>20)</sup>



```

from docx import Document
import xlswriter
import os

savename = "D:\\Fiscal Year.xlsx" # 연도별 가독성 측정 결과 저장경로21)
workbook = xlswriter.Workbook(savename)
worksheet = workbook.add_worksheet()
row_w = 0

for root, directories, files in os.walk("D:\\Annual Report"):# 사업보고서 저장경로
    # 가독성지수의 구성요소별 집계 시작점을 "0"으로 지정
    for filename in files:
        file_len = len(os.path.splitext(filename)[0]) # 파일명의 글자수 세기
        if os.path.splitext(filename)[0][file_len-4:file_len] == "2015": # 회계연도(예시)
            letter_num = 0 # 글자수
            word3_num = 0 # 3음절이상 단어수
            word5_num = 0 # 5음절이상 단어수
            word7_num = 0 # 7음절이상 단어수
            word10_num = 0 # 10음절이상 단어수
            word_num = 0 # 단어수
            sentence_num = 0 # 문장수
            document = Document(root + "\\\" + filename)

```

20) 본 연구의 표본인 금융감독원의 전자공시시스템(DART)에서 열람가능한 1998년부터 2015년까지 국내 유가증권시장 및 코스닥시장에 상장된 기업을 대상으로 위의 Python Code를 실행하면, 본 연구에 사용된 기업-연 자료(firm-year data)의 가독성 지수를 측정하는 데 약 2일의 시간이 소요되는데, 표본의 수가 21,917개 기업-연자료임을 고려할 때, 상당한 속도라 할 수 있다. 다만 가독성 지수 측정 시의 소요시간은 연구자들의 연구 환경에 따라 상이할 수 있다.

21) “\”문자는 python의 문자열 안에서 특수한 의미를 갖는 escape 문자로, 파일경로를 구분하는 “\”를 그 의미 그대로 사용하기 위하여 “\” 앞에 escape 문자인 “\”를 붙여 표현하여야 한다. 즉 첫 번째 “\”는 escape 문자를 의미하고, 두 번째 “\”는 파일경로의 구분자를 표현한다.

```

# 단락 내 각 구성요소 집계과정
for para in document.paragraphs:
    for i in range(0, len(para.text.split())):
        # 단어수는 빈칸(" ")으로 구분
        # 글자수 또는 음절수는 단어를 구성하는 요소로 판단함.
        letter_num = letter_num + len(para.text.split()[i])
        word_num = word_num + 1
        if len(para.text.split()[i]) > 2:
            word3_num = word3_num + 1
        if len(para.text.split()[i]) > 4:
            word5_num = word5_num + 1
        if len(para.text.split()[i]) > 6:
            word7_num = word7_num + 1
        if len(para.text.split()[i]) > 9:
            word10_num = word10_num + 1
    for i in range(0, len(para.text.split("."))): # 문장수는 마침표로 구분
        split_length = len(para.text.split(".")[i])
        if split_length == 0:
            sentence_num = sentence_num
        elif para.text.split(".")[i][split_length - 1].isdigit():
            sentence_num = sentence_num
            # 마침표 앞에 숫자면 제외, 예 : 10.31%, 17.08.11 등
        else:
            sentence_num = sentence_num + 1
            # 위의 경우를 제외하고, 문장수 집계

# 엑셀 시트에 행단위로 출력 (제일 첫 번째 열에 회사명으로 시작)
worksheet.write(row_w, 0, os.path.splitext(filename)[0][0:file_len - 4])
worksheet.write(row_w, 1, letter_num) # 글자수
worksheet.write(row_w, 2, word_num) # 단어수
worksheet.write(row_w, 3, word3_num) # 3음절이상 단어수
worksheet.write(row_w, 4, sentence_num) # 문장수
AWL = letter_num/word_num # Flesch 구성요소 : 평균 단어 길이
ASL = word_num/sentence_num # Flesch 구성요소 : 평균 문장 길이
flesch = 206.835 - 84.6 * AWL - 1.015 * ASL # Flesch 측정
worksheet.write(row_w, 5, AWL) # AWL
worksheet.write(row_w, 6, ASL) # ASL
worksheet.write(row_w, 7, flesch) # Flesch Index
worksheet.write(row_w, 8, word5_num) # 5음절이상 단어수
worksheet.write(row_w, 9, word7_num) # 7음절이상 단어수
worksheet.write(row_w, 10, word10_num) # 10음절이상 단어수

```

```

print(os.path.splitext(filename)[0] + "is collected:", letter_num,
word_num, word3_num, sentence_num, AWL, ASL, flesch, word5_num,
word7_num, word10_num) # 프로그램 화면에 출력하여 결과 진행 확인가능
row_w = row_w + 1 # 다음 행으로 넘어가기

```

```

workbook.close() # 엑셀파일 워크시트 닫고 저장
print("end") # 해당 회계연도 전부 수행시 "end" 출력

```

4. 사업보고서 중 표에 기재된 내용을 가독성 지수 측정에 포함한다면, "document = Document(root + "\\\" + filename)" 뒤로 다음의 코드를 추가하여 수행한다.

```

for table in document.tables:
    for row in table.rows:
        for cell in row.cells:
            for paragraph in cell.paragraphs:

```

5. Flesch reading ease formula는 평균 단어 길이(*AWL*)와 평균 문장 길이(*ASL*)의 두 언어학적 변수로 구성되며, 앞서 언급한 Python 코드에서 " $FLESCH = (-1) * (206.835 - 84.6 * AWL - 1.015 * ASL)$ "를 계산하여 측정하였다. 또한, Robert Gunning(1952)의 Gunning Fog Index는 Python 코드에서 측정된 평균 문장 길이(*ASL*)와 3음절 이상 단어수(*word3\_num*)를 사업보고서의 전체 단어수(*word\_num*)로 나눈 복잡성 비중(*percentage of complex words*)를 산출공식 " $FOG = (ASL + percentage\ of\ complex\ words) * 0.4$ "을 통하여 측정하였다. 4.5.2.의 <Table 9>에서와 같이 한글로 작성된 사업보고서의 가독성 분석에서 Fog Index의 강건성을 확인하기 위하여, 5음절 이상 단어수(*word5\_num*), 7음절 이상 단어수(*word7\_num*), 10음절 이상 단어수(*word10\_num*)를 별도로 측정하여 복잡성의 정의를 다양하게 분석에 활용하였다. 마지막으로 사업보고서의 전체길이(*LENGTH*)는 사업보고서의 전체 단어수(*word\_num*)에 자연로그를 취하여 측정하였다.