

# 프라이버시 염려가 디지털 콘텐츠 지출에 미치는 영향: 이부모형을 이용한 실증분석\*

## The Effects of Privacy Concerns on Digital Contents Spending: Empirical Analysis via a Two-Part Model

신 일 순 (Ilsoon Shin)\*\*  
김 현 수 (Hyun Soo Kim)\*\*\*

### 국문초록

본 연구는 온라인 소비자의 지출 데이터가 준연속적인(semi-continuous) 특징을 가진다는 점에 착안하여, 프라이버시 염려가 디지털 콘텐츠 지출에 미치는 영향을 디지털 콘텐츠 지출을 하지 않던 소비자가 디지털 콘텐츠 지출을 선택하는 데에 어떠한 영향을 미치는지와 이미 양(+의) 디지털 콘텐츠 지출을 하는 소비자가 디지털 콘텐츠 지출액을 조정하는 데에 어떠한 영향을 미치는지를 구분하여 분석하였다. 이를 위해 Cragg의 이부모형(two-part model)을 이용하여 참가 방정식(participation equation)과 수량 방정식(amount equation)을 추정하였다. 실증분석의 결과는 다음과 같다. 첫째, 디지털 콘텐츠 지출 여부와 지출액을 구분하지 않은 실증분석에서는 기존의 연구와 유사하게 프라이버시 역설이 성립한다. 둘째, 이러한 프라이버시 역설은 주로 참가 방정식에서의 효과, 즉 디지털 콘텐츠 지출 여부의 결정에 대한 효과에 기인한다. 셋째, 이에 비해 이미 디지털 콘텐츠 지출을 하고 있는 사람들을 대상으로 한 수량 방정식의 결과를 살펴보면, 비록 통계적 유의성은 떨어지지만 프라이버시 염려가 증가할 때 디지털 콘텐츠 지출액을 줄이는 식의 반응이 나타난다. 이러한 결과는 패널 모형 및 처치효과 모형에서 강건성이 검증된다. 논문에서는 프라이버시 역설이 디지털 콘텐츠 지출 여부와 디지털 콘텐츠 지출액의 측면에서 차별적으로 나타나는 결과에 대해 의미 및 시사점을 간략히 논의하였다.

**주제어:** 프라이버시 역설, 프라이버시 염려, 디지털 콘텐츠 지출, 이부모형

※ 논문접수일: 2021. 2. 24, 수정일: 2021. 5. 12, 게재확정일: 2021. 5. 25

\* 본 논문은 인하대학교의 지원으로 연구되었음.

\*\* 인하대학교 사회과학대학 경제학과 교수, 제1저자, E-mail: [ishin@inha.ac.kr](mailto:ishin@inha.ac.kr)

\*\*\* 인하대학교 사회과학대학 경제학과 박사과정, 교신저자, E-mail: [hskim150920@gmail.com](mailto:hskim150920@gmail.com)

**ABSTRACT**

Because digital contents spending data are semi-continuous in nature, this study analyzes two distinct effects of privacy concerns on digital contents spending: first, whether inexperienced consumers participate in digital contents spending, and second, how consumers with positive spending adjust their spending amounts. Using the Cragg two-part model, we estimate participation and amount equations. Our results show the following. First, without distinguishing between participation and adjustment, the privacy paradox appears to hold in the sense that the amount of digital contents spending is positively related to privacy concerns. Second, the privacy paradox is mainly due to the participation equation, that is, the decision to participate in digital contents spending. Third, the amount equation shows that people respond to rising privacy concerns by reducing their digital contents spending, although coefficient estimates are statistically insignificant. The robustness of these results is verified using panel data models and treatment effect models. The paper briefly discusses the implications of the consequences of the privacy paradox in relation to digital contents spending.

**Key words:** Privacy paradox, Privacy concern, Digital contents spending,  
Two-part model

## I. 서론

식별가능한 개인정보의 수집 및 공유를 기반으로 하는 온라인 활동은 전통적인 오프라인 활동에 비해 프라이버시 침해의 개연성을 증가시킨다. 특히 디지털 콘텐츠 지출 등의 온라인 거래 활동은 개인정보뿐만 아니라 지급결제 및 거래 내용에 대한 정보의 수집이 요구되기 때문에 소비자들은 프라이버시 침해 가능성에 더 민감할 것으로 예상할 수 있다. 그런데 최한솔 외 (2019)에서는 실증분석을 통해 프라이버시 염려(Privacy Concern)가 더 큰 소비자들이 오히려 온라인 구매 활동을 더 적극적으로 하는 현상이 관찰됨을 보이고 있다. 즉, 프라이버시에 대한 태도와 행동이 일관적이지 아니라는 의미의 프라이버시 역설(Privacy paradox)이 온라인 구매에서도 발견됨을 밝히고 있다.

본 연구에서는 최한솔 외 (2019)의 분석에서 한 걸음 더 나아가, 디지털 콘텐츠 지출 자료가 준연속적인(semi-continuous) 특징을 가진다는 점에 착안하여 프라이버시 염려가 디지털 콘텐츠 지출에 미치는 영향을 디지털 콘텐츠 지출 여부 - 디지털 콘텐츠 지출을 할지 말지 - 및 디지털 콘텐츠 지출액 조정 - 디지털 콘텐츠 지출을 하는 경우, 늘일지 줄일지 - 등 두 가지 서로 다른 결정에 미치는 영향으로 구분하여 분석하는 시도를 하였다. 여기서 준연속적인 데이터는 0의 값과 연속적인(continuous) 값이 혼합되어 나타나는 형태의 자료를 의미하는데, 실제 자료에서 온라인 소비자는 디지털 콘텐츠 지출을 전혀 하지 않는 상당수의 소비자와 양(+)의 지출을 하는 일부의 소비자로 구분된다. 이 경우 우리는 이부모형(two-part model)을 이용하여, 프라이버시 염려가 디지털 콘텐츠 지출을 하지 않던 소비자가 디지털 콘텐츠 지출을 선택하는 데에 어떠한 영향을 미치는지와 함께 이미 양(+)의 디지털 콘텐츠 지출을 하는 소비자가 디지털 콘텐츠 지출액을 조정하는 데에 어떠한 영향을 미치는지를 구분하여, 두 가지 측면에서 분석하는 것이 가능하다.

프라이버시 역설을 두 가지 측면에서 분석하는 것이 필요한 이유는 다음과 같다. 만일 프라이버시 염려가 증가할 때 디지털 콘텐츠 지출 여부와 디지털 콘텐츠 지출액 조정의 두 측면 모두에서 프라이버시 역설 현상이 관찰된다면, 전반적으로 소비자들이 디지털 콘텐츠 지출을 위해 프라이버시를 포기하는 것을 의미하므로 이 문제의 해결을 위한 정책적인 접근이 매우 시급하고 중요하게 됨을 시사한다. 그러나 만일 디지털 콘텐츠 지출 여부 및 디지털 콘텐츠 지출액 조정의 두 측면

중 어느 하나에서만 프라이버시 역설 현상이 관찰된다면, 소비자의 합리성이 어느 정도 존재한다는 것을 의미하므로 이를 감안하여 지나치지 않고 적절하게 타계팅된 정책적인 접근이 필요함을 시사한다. 다시 말하면, 본 연구는 디지털 콘텐츠 지출 및 거래와 관련한 프라이버시 보호 정책의 기본적인 방향 및 강도를 판단하기 위한 필요성 및 목적을 가지고 있다.

본 연구는 이부모형을 이용하여 프라이버시 염려가 디지털 콘텐츠 지출 여부 및 디지털 콘텐츠 지출액에 미치는 영향을 실증적으로 추정하였으며, 패널 모형 및 처치효과 모형을 이용하여 이부모형 결과의 강건성을 검증하였다. 이부모형, 패널 모형 및 처치효과 모형을 통한 실증분석에서는 모두 질적으로 유사한 결과가 도출되었는데, 프라이버시 역설은 주로 디지털 콘텐츠 지출을 하지 않던 소비자가 디지털 콘텐츠 지출을 선택하는 데에서 발생하며, 이미 디지털 콘텐츠 지출을 하고 있는 사람들을 대상으로 하는 경우는 비록 통계적 유의성은 떨어지지만 프라이버시 염려가 증가할 때 디지털 콘텐츠 지출액을 줄이는 식으로, 다시 말하면 프라이버시 역설이 성립하지 않는 것으로 관찰되었다.

본 연구의 구성은 다음과 같다. II 장에서는 프라이버시 염려 및 프라이버시 역설에 대한 설명과 기존연구를 제시하며, III 장에서는 본 연구에서 사용한 자료를 소개하고, 주요 변수인 프라이버시 염려와 디지털 콘텐츠 지출 변수를 설명한다. IV 장에서는 실증분석에 사용된 모형을 소개하고, 실증분석의 결과 및 설명이 이루어진다. 결론과 시사점은 V 장에서 정리한다.

## II. 이론적 배경

### 1. 프라이버시 염려 및 프라이버시 역설

먼저 본 연구에서 말하는 프라이버시란 정보 프라이버시(information privacy)를 의미한다. 정보 프라이버시란 개인정보의 수집이나 이용을 제한하거나 개인정보 자체에 대한 보호를 넘어서 개인정보가 침해되는 행위로부터의 보호와 통제를 의미하는 것으로, 개인정보가 디지털 형태로 정보시스템에 저장되고 전송되는 과정에서 다양한 형태로 복제될 수 있을 뿐만 아니라 개인정보 취급자의 관리에서 쉽게 벗어날 수 있다는 가능성을 염두에 두고 있다 (김종기 외, 2018).

프라이버시 역설(Privacy paradox)이란 인터넷과 디지털의 확산에 따라 자연스럽

게 프라이버시 침해에 대한 염려(privacy concern)가 증가하지만, 이에도 불구하고 개인정보(personal information)를 더 적극적으로 제공하는 현상을 의미하는 것으로, 태도와 행동이 일관적이지 아니라는 의미에서 ‘역설’이라는 단어로 표현된다 (신일순·김현수, 2018).

프라이버시 역설이 발생하는 이유로는 프라이버시 계산(Privacy Calculus) 이론 및 행동경제학(Behavioral Economics)적인 설명이 제시되고 있다.<sup>1)</sup> 두 이론은 개인의 합리성에 대한 전제에서 차이가 나는데, 먼저 ‘프라이버시 계산’ 이론에 따르면, 개인의 행동은 경제적인 관점의 위험-혜택 분석(Risk-Benefit Analysis)을 기반으로 위험과 혜택의 상충관계(trade-off)를 합리적으로 반영하여 결정된다고 상정된다 (Li et al., 2010). 이에 따르면 프라이버시 역설이 나타나는 것은 온라인 활동의 위험의 측면만을 보고 혜택을 간과하기 때문으로, 만일 적절하게 혜택의 측면을 고려했다면 역설적인 현상이 나타나지 않았을 것이라고 해석한다.

이에 비해 ‘행동경제학’적 설명은 개인의 선호체계에 대해 지나치게 합리적이고 제한적인 가정에 입각한 전통적인 설명 대신 심리학에 기반을 둔 인지적 요인을 통해 인간의 행동을 설명하고자 하는 이론이다 (Acquisti et al., 2015). 행동경제학에 따르면, 개인은 의사결정에 있어 효용을 극대화하기 위한 정보가 부족하고, 이러한 정보를 적절히 수집하기에는 시간적 제약이 따르며, 정보를 처리함에 있어서도 인지능력의 한계를 지닌다. 즉, 높은 불확실성, 불완전한 정보, 제한된 시간, 정보처리에 대한 인지적 한계 등의 현실적인 문제로 인해 인간의 의사결정이 완전히 합리적일 수 없다는 것이다 (Brecht et al., 2011). 따라서 프라이버시 염려가 증가하더라도 이를 합리적으로 행동에 반영하기에는 현실적으로 인지적인 요인을 포함한 여러 가지 제약이 존재하기 때문에, 반드시 합리적인 행동이 나타나는 것은 아니고 역설적인 현상이 관찰 가능하다고 주장한다.

## 2. 기존 연구

프라이버시 염려와 관련된 연구는 다양한 분야에서 상당히 활발하게 이루어지고

1) Kokolakis (2017)의 연구에 따르면, 프라이버시 역설이 발생하는 이유에 대해 프라이버시 계산 이론과 행동경제학적인 설명 외에도 사회적 이론(Social theory), 인지적 편향 및 휴리스틱(Cognitive biases and heuristics), 제한된 합리성(Bounded rationality), 불완전한 정보(incomplete information), 정보비대칭(information asymmetries), 양자이론(Quantum theory) 등 다양한 개념과 이론 등이 존재한다.

있으나 본 연구의 주제인 디지털 콘텐츠와 프라이버시 염려 사이의 연구는 상대적으로 미흡한 상황으로, 본 절에서는 온라인 구매(지출) 맥락에서 프라이버시 염려의 영향을 다룬 기존 연구들을 살펴보았다.<sup>2)</sup>

2000년 초반 연구 (George, 2004; Limayem et al., 2000; Khalifa & Limayem, 2003)를 제외한 대부분의 최근 연구들은 프라이버시 염려와 행동 간 일치를 지지하는 결과를 확인함으로써 프라이버시 역설이 일반적인 온라인 환경과 달리 온라인 구매의 맥락에서는 성립되지 않는 현상임을 나타내고 있다. Dinev & Hart (2005)는 프라이버시 염려가 구매 의도에 부정적 영향을 미친다는 것을 밝혔고, 온라인 상거래 맥락에서 기업의 명성과 프라이버시 염려, 몰입, 신뢰, 구매 의도 간의 관계를 연구한 Eastlick et al. (2006)은 프라이버시 염려와 신뢰가 구매 의도에 모두 부정적인 영향을 미친다는 것을 확인하였다. Suh & Han (2003)은 지각된 프라이버시 보호 수준이 소비자들의 전자상거래 신뢰에 긍정적인 영향을 미치고 신뢰가 매개변수로서 전자상거래 수용에 유의한 영향을 미치는 것을 확인하였다. 유일 외 (2008)도 프라이버시 염려가 기업의 신뢰에 부정적인 영향을 미치고, 신뢰는 거래 의도에 유의한 긍정적 영향을 미친다는 점을 밝혔다. 민병권·김이태 (2006) 역시 프라이버시 염려가 기업 명성에 부정적인 영향을 미치고, 기업 명성이 높을수록 구매 경험에 긍정적 영향을 미쳐, 프라이버시 염려와 구매 경험이 기업 명성을 통해 부적으로 상관이 있음을 나타낸 바 있다. 한편, 소수에 불과하지만 일부 연구들은 프라이버시 염려가 온라인 구매 행동에 유의미한 영향을 미치지 못한다는 점을 확인함으로써 온라인구매 맥락에서도 프라이버시 역설이 성립될 수 있음을 나타내었다. 계획된 행동이론에 근거하여 온라인 구매행동에 미치는 영향을 알아본 George (2004)는 기업이 소비자들의 승인 없이 개인정보를 2차적으로 사용하는 것에 대한 소비자들의 염려가 계획된 행동이론 주요요인 중 하나인 온라인 구매 태도에 영향을 미칠 것이라는 가설을 설정하였으나, 채택되지 못하였음을 확인한 바 있다. 국내 연구로서는 드물게 소비자들의 온라인 구매 행동에서의 프라이버시의 역설을 규명한 이동주·김명수 (2011)는 다소 복합적인 연구 결과를 도출하였는데, 전자상거래 이용자들의 경우, 프라이버시 염려 수준이 전자상거래 구매 빈도에는 영향을 미치지 않으나 프라이버시 보호에 소홀한 기업의 거래는 중단하고 프라이버시 보호에 노력하는 기업의 경우에는 거래를 증가시키려 하는 등 기업의 프라이버시 보호 행동에 따라 차별적 반응을 보인다는 것을 밝히고 있다.

2) 이하의 내용은 최한솔 외 (2019)를 참조하였다.

이처럼 온라인 구매 맥락에서 수행된 연구들 대부분은 프라이버시 염려와 소비자들의 행동 간 일치성을 주장하고 있고, 프라이버시 역설을 지지하는 연구들은 극히 일부에 불과한 것으로 나타났으며, 선행연구 대부분은 소비자들의 실제 구매 행동이나 경험이 아닌 행동 의도만을 측정하고 있어 프라이버시 염려의 영향력을 일반화시키기에는 다소 한계가 있다 (최한솔 외, 2019).

이에 본 연구에서는 온라인 구매 대신 디지털 콘텐츠 지출을 대상으로 프라이버시 염려와 소비자들의 구매 행동을 분석하였으며, 디지털 콘텐츠 지출 자료가 준연속적인(semi-continuous) 특징을 가진다는 점에 착안하여 프라이버시 염려가 디지털 콘텐츠 지출에 미치는 영향을 디지털 콘텐츠 지출 여부 - 디지털 콘텐츠 지출을 할지 말지 - 및 디지털 콘텐츠 지출액 조정 - 디지털 콘텐츠 지출을 하는 경우, 늘일지 줄일지 - 등 두 가지 서로 다른 결정에 미치는 영향으로 구분하여 분석하는 시도를 하였다.

### Ⅲ. 자료 설명

#### 1. KISDI 한국미디어패널 조사

본 연구에서는 프라이버시 염려와 디지털 콘텐츠 지출간의 상관관계 및 인과관계에 대한 실증 분석을 위해 KISDI(정보통신정책연구원)의 ‘한국미디어패널 조사’ 데이터를 사용하였다. ‘한국미디어패널 조사’는 동일 응답자, 동일 항목에 대해 매년 반복적으로 조사하는 패널(panel) 자료의 형태이고, 제공하는 기간이 장기간이며, 충실한 표본추출 및 방문 면접조사를 통해 매년 1만 명 이상의 응답자를 대상으로 하는 자료로서, 가구 설문, 개인 설문, 미디어 다이어리 등으로 구성되어 있다.<sup>3)</sup> 본 연구에서는 프라이버시 염려와 디지털 콘텐츠 지출에 대한 자료가 포함되어 있는 개인 설문 자료 중에서 2014~15년 2개년 자료를 균형 패널(balanced panel)로 구성하여 총 8,804명의 데이터를 사용하였다.<sup>4)5)</sup>

3) <https://stat.kisdi.re.kr/default.aspx>에서 KISDI 한국 미디어패널 원시자료, 설문지, 코드북, 유저가이드 등을 다운받을 수 있다.

4) 본 연구에서 2014년 이후의 자료를 사용한 이유는 2010년부터 시작된 KISDI 미디어패널 자료에서 프라이버시 관련 항목이 2014년부터 조사되었기 때문이다.

5) 본 연구에서 2015년까지의 자료만을 사용한 이유는 KISDI 미디어패널 조사 중 디지털 콘

## 2. 주요 변수의 기초통계량

먼저 본 연구 주제의 실증분석에서 가장 중요한 변수들인 프라이버시 염려 및 디지털 콘텐츠 지출은 한국미디어패널 조사에서 다음의 항목으로부터 추출하였다.

프라이버시 염려는 한국미디어패널 조사(개인용) 설문지의 (문 16) 항목인 ‘프라이버시 침해 우려(염려)’를 이용하였다. 이에는 7가지 세부 항목이 있는데, 구체적으로는 ‘① 모르는 사람이 나의 온라인 활동을 보고 나에 대한 개인 정보를 획득할까 걱정스럽다; ② 예전에 쓰던 기기(컴퓨터, 휴대폰) 등에 나에 대한 정보가 남아 있을까 걱정스럽다; ③ 온라인 사이트에 가입할 때 개인 정보를 너무 많이 요구하는 것이 걱정스럽다 ④ 내 온라인 아이디를 도용당할까 걱정스럽다; ⑤ 일반적으로 인터넷을 사용할 때 나의 프라이버시에 대해 걱정스럽다; ⑥ 온라인에서 자기가 누구인지 밝히지 않은 사람들은 의심스럽다’ 등이다. 이러한 여섯 가지 질문에 대해 응답자는 ‘① 전혀 그렇지 않다; ② 그렇지 않다; ③ 보통이다; ④ 그렇다; ⑤ 매우 그렇다’ 등의 5점 척도로 답하게 된다. 본 연구에서는 프라이버시 염려에 대한 변수(pvC)를 이에서 설명한 여섯 가지 세부 항목의 답변을 모두 더하여 사용하였다.<sup>6)</sup> 프라이버시 염려는 높은 수치가 프라이버시 염려가 큰 것을 의미하며, 척도 변수여서 등간격을 보장하지는 않지만 실증 분석을 위해 등간격을 가정하기로 한다.

<표 1> 프라이버시 염려(pvC) 기초통계량

변수	연도	관측치	평균	표준편차	최소값	최대값
프라이버시 염려 (pvC)	2014	8,804	18.79	7.12	6	30
	2015	8,804	17.87	6.35	6	30

텐츠 지출에 대한 자료의 특징 및 문제점 때문이다. 구체적으로 설명하면, 2014~18년의 5개년도의 균형 패널(5,409명)을 대상으로 2014~15년의 디지털 콘텐츠 지출의 자료는 결측치(missing value) 없이 모두 존재하지만, 2016, 17, 18년에는 각각 9, 48, 279명만이 디지털 콘텐츠 지출 자료가 있고, 나머지는 모두 결측치로 나타나 관측치 수의 부족이 심각하기 때문이다.

- 6) 2015년 조사에 프라이버시 침해 염려에 대한 세부 항목 중 “온라인상에 내가 기억하지 못하는 나에 관한 정보가 삭제되지 않은 채 남아 있을까 걱정스럽다”에 대한 질의가 추가되었다. 그러나 본 연구에서는 2014년과의 일관성 확보를 위해 이를 제외하고 여섯 가지 항목을 이용하였다.



<표 1>은 프라이버시 염려 변수의 기초통계량을 표현한 것이다. 2015년은 전년도에 비해 프라이버시 염려의 평균치가 약간 감소하였는데, ‘일반적으로 인터넷을 사용할 때 나의 프라이버시에 대해 걱정스럽다(세부 항목 ⑤)’에 대해 ‘그렇다 및 아주 그렇다’라고 답변한 사람을 프라이버시 고염려군으로 간주하면, 이들의 비중은 1년 사이 8.4%p (43.3% ⇨ 34.9%) 감소한 것으로 나타난다.

다음으로 디지털 콘텐츠 지출은 한국미디어패널 조사(개인용) 설문지의 (문 6) 항목인 ‘온라인 디지털 콘텐츠 서비스 월평균 지출요금’을 이용하였다. 이 항목은 “귀하께서 온라인 상에서 디지털 콘텐츠를 이용하기 위해 지출한 금액은 한 달에 어느 정도 입니까?”의 질문에 대해 TV 방송 프로그램, 동영상/영화, 신문/잡지/책, 음악, 게임 등의 5가지로 구분하여 지난 3달 동안 지출한 금액의 월평균액을 답하도록 하고 있는데, 본 연구에서는 이를 모두 더하여 디지털 콘텐츠 지출액(onTRS) 변수로 정의하였다.

그런데 디지털 콘텐츠 지출 변수는 준연속적(semi-continuous)인 특징을 가지고 있다. 여기서 준연속적인 자료란 0의 값과 양(+)의 연속적인 값이 혼합되어 나타나는 형태의 자료를 의미한다. 따라서 디지털 콘텐츠 지출 변수를 두 가지, 즉 ‘지출 여부’와 지출을 하는 사람의 경우 ‘지출액’으로 나누어 기초통계량을 <표 2>에 제시하였다.

<표 2> 디지털 콘텐츠 지출(onTRS) 기초통계량

변수	연도	관측치	평균	표준편차	최소값	최대값
디지털 콘텐츠 지출 여부 (bionTRS)	2014	8,804	0.07	0.26	0	1
	2015	8,804	0.04	0.19	0	1
변수	연도	관측치	평균	표준편차	최소값	최대값
디지털 콘텐츠 지출액 (onTRS   onTRS>0)	2014	618	10.19	10.15	1	85
	2015	323	9.27	9.05	1	50

<표 2>의 디지털 콘텐츠 지출 여부를 살펴보면, 2014년에는 전체의 7%인 618명만 양(+)의 디지털 콘텐츠 지출을 하고, 나머지는 디지털 콘텐츠 지출을 수행하지 않으며, 2015년도 질적으로 유사하게 전체의 4%인 323명만 양(+)의 디지털 콘텐츠 지출을 하는 것을 알 수 있다. 다시 말하면, 데이터는 소수의 사람들만 디지털

털 콘텐츠 지출을 하고 있으며, 다수는 전혀 디지털 콘텐츠 지출을 하지 않는다는 특징을 보여주고 있으며, 이러한 특징이 실증분석의 과정에서 적절하게 감안되어야 함을 시사한다. 다음으로 디지털 콘텐츠 지출을 하는 사람들만을 대상으로 디지털 콘텐츠 지출액(한 달 기준)을 살펴보면, 평균적으로 9.3천~10.2천 원 정도로 디지털콘텐츠에 대해 지출하는 것으로 나타난다.<sup>7)</sup>

실증분석에서 통제변수의 역할을 하는 기타변수는 대부분 설문자의 개인적인 특성을 나타내는 변수들이며, 기초통계량은 다음의 <표 3>과 같다. 나이 및 소득을 제외한 다른 변수는 0과 1의 값을 가지는 이항 변수(binary variable)이며, 소득은 ‘소득 없음’이 1의 값, ‘800만 원 이상의 소득’이 18의 값을 가지고, 50만원을 간격으로 구분된 변수이다.

<표 3> 기타 변수의 기초통계량

변수	관측치	평균	표준편차	최소값	최대값
특별/광역시 거주	17,608	0.46	0.50	0	1
성별	17,608	0.45	0.50	0	1
나이	17,608	45.99	20.52	6	105
결혼 여부	17,608	0.62	0.49	0	1
직업 유무	17,608	0.50	0.50	0	1
1인 가구	17,608	0.07	0.26	0	1
3세대 이상 가구	17,608	0.08	0.27	0	1
소득	17,608	3.56	3.10	1	18

7) 디지털 콘텐츠 지출액(>0)의 히스토그램을 그려본 결과 정규분포의 형태가 아닌 로그정규분포의 형태가 나타났다. <부록 그림 1> 참조.

## IV. 실증 분석 및 결과

### 1. 실증 모형

Ⅲ장에서 제시한 자료를 이용하여 본 연구에서는 프라이버시 염려와 디지털 콘텐츠 지출에 대한 실증분석을 수행하였다. 먼저 기본적인 이해를 위한 선형회귀(OLS) 모형을 제시하면 다음과 같다.

$$Y = \alpha + \beta X + \gamma pvC + \epsilon \quad (1)$$

식 (1)에서 종속변수  $Y$ 는 디지털 콘텐츠 지출이고, 우측항의  $X$ 는 통제변수 벡터,  $pvC$ 는 프라이버시 염려에 대한 변수이고,  $\epsilon$ 는 오차항이다. 통제변수로는 거주 지역, 소득, 나이, 성별, 결혼 여부, 가구원의 수 등을 포함하였다. 선형회귀 모형은 개인적인 특성의 통제변수와 함께 프라이버시 염려와 디지털 콘텐츠 지출의 인과관계를 가장 단순하게 모형화하는 시도로 이해할 수 있다.

그러나 앞서 설명한 것처럼 디지털 콘텐츠 지출에 대한 분포는 모서리해(corner solution)의 문제가 심각한 준연속적(semi-continuous) 특징을 가지고 있다. 2014~15년의 데이터를 풀링(pooling)할 경우, 5.3%의 소비자만이 양(+의) 디지털 콘텐츠 지출을 하고, 나머지 94.7%는 디지털 콘텐츠 지출을 전혀 하지 않는다. 이러한 준연속적인 종속변수의 경우 연속적인 형태를 상징하는 추정방법을 사용하면 편의성이 발생하게 된다.<sup>8)</sup> 이에 따라 준연속적인 종속변수를 추정하는 몇 가지 방법이 개발되었다.

먼저 준연속적인 종속변수를 명시적으로 반영하는 분석모형으로 토빗(Tobit) 모형을 들 수 있다. 이는 다음과 같이 표현된다.

$$Y = 1[\alpha + \beta X + \gamma pvC + \epsilon > 0] \cdot (\alpha + \beta X + \gamma pvC + \epsilon), \epsilon \sim N(0, \sigma^2) \quad (2)$$

여기서  $1[\cdot]$ 은 지시함수(indicator function)로 대괄호 내부가 참일 경우 1, 거짓을 경우 0의 값을 갖는 함수이고,  $\epsilon$ 는 평균 0, 분산  $\sigma^2$ 인 정규분포를 따

8) 정병철 (2018), p.117.

르는 확률변수이다. 따라서 토빗 모형은  $\alpha + \beta X + \gamma pvC + \epsilon > 0$  인 경우  $Y = \alpha + \beta X + \gamma pvC + \epsilon$  이 되지만,  $\alpha + \beta X + \gamma pvC + \epsilon \leq 0$  인 경우  $Y = 0$  이다. 따라서 토빗 모형은 일부 개인의 종속변수가 0의 값을 갖고 나머지 개인에 대해서는 연속적으로 분포하는 패턴을 효과적으로 모형에 반영할 수 있다.

그런데 토빗 모형에는 한 가지 제약적인 가정이 존재한다. 구체적으로는 어떤 개인이 디지털 콘텐츠 지출 여부에 영향을 미치는 요인과 디지털 콘텐츠 지출액에 영향을 미치는 요인의 부호 및 상대적 크기가 동일하다는 가정이다.<sup>9)</sup> 만일 어떤 개인이 디지털 콘텐츠 지출을 하도록 만드는 요인과 디지털 콘텐츠 지출을 선택한 이후 지출액을 늘리도록 만드는 요인이 상이하거나 또는 동일한 요인이라도 중요성이 달라진다면 토빗 모형의 가정은 타당하지 않다.

이에 따라 자연스럽게 토빗 모형의 제약적인 가정을 완화한 또 다른 실증모형을 생각할 수 있다. 즉, 실증모형을 두 개로 나누어 하나는 종속변수가 양(+)의 값을 갖는지 여부를 결정하는 참가 방정식(participation equation), 다른 하나는 종속변수가 양(+)일 경우 얼마의 값을 갖는지를 결정하는 수량 방정식(amount equation)으로 구성되는 것으로 이해하는 방법이다. 이 경우 다른 두 개의 식을 추정하므로 토빗 모형의 제약성은 완화될 수 있다. 이러한 방법을 이부모형(two-part model)이라고 하며, 대표적으로 Cragg (1971)의 모형을 들 수 있다.<sup>10)</sup> 이는 다음과 같이 표현된다.

$$Y = 1[\alpha_1 + \beta_1 X + \gamma_1 pvC + v > 0] \cdot (\alpha_2 + \beta_2 X + \gamma_2 pvC + \epsilon), \quad (3)$$

여기서  $v$  는 표준정규분포를 따르는 확률변수이고,  $\epsilon$  은 평균 0, 분산  $\sigma^2$ 이며,  $-\alpha - \beta X - \gamma pvC$  에서 절단된 표준정규분포를 따르는 확률변수이다. 이부모형은

9) 이 가정은 수식으로는  $\frac{\partial P(Y > 0) / \partial x_j}{\partial P(Y > 0) / \partial x_h} = \frac{\partial E(Y | Y > 0) / \partial x_j}{\partial E(Y | Y > 0) / \partial x_h}$  로 표현할 수 있다. 즉, 변수  $x_j$ 와  $x_h$ 가  $\Pr(Y > 0)$  - 본 연구에서는 디지털 콘텐츠 지출 여부 - 에 미치는 효과의 상대적 크기와  $E(Y | Y > 0)$  - 본 연구에서는 디지털 콘텐츠 지출을 하는 경우의 지출액 - 에 미치는 효과의 상대적 크기와 동일하다는 의미이다.

10) 이부모형을 이용한 연구로는 민희철 외 (2016)를 들 수 있다. 이 연구에서는 바이오 연구자의 특허수 데이터가 준연속적인 특징 - 대부분 바이오 연구자는 특허를 보유하지 않고, 일부 소수의 연구자만 특허를 보유하고 있는 - 을 반영하여 이부모형을 사용하였다.

$(\alpha_2, \beta_2, \gamma_2)$ 와  $(\alpha_1, \beta_1, \gamma_1)$ 을 별도로 추정하므로, 토빗 모형과 달리 설명변수가 지출 여부와 지출 조건부 지출액 결정에서 차별적인 효과를 가지게 된다. 편의상  $\epsilon$ 와  $v$ 는 서로 독립이라고 가정하였는데, 이 경우 참가 방정식은 프로빗(probit) 모형으로,<sup>11)</sup> 수량 방정식은 절단된 회귀분석(truncated regression) 방법 또는 로그정규분포(log-normal) 모형으로 추정할 수 있다.<sup>12)</sup>

이러한 이부모형을 보다 구체적으로 표기하면 다음과 같다.

$$Y = s h^* \quad (4)$$

여기서  $Y$ 는 종속변수인 디지털 콘텐츠 지출로 0 또는 0보다 큰 값으로 관찰되며, 참가 방정식을 나타내는 선택변수  $s$ 와 수량 방정식을 의미하는 잠재변수  $h^*$ 의 형태로 이루어져 있다. 선택변수  $s$ 는  $Y=0$ 일 때 0이고,  $Y>0$ 일 때 1의 값을 갖는다.

$$s = \begin{cases} 1 & \text{if } \alpha_1 + \beta_1 X + \gamma_1 pvC + v > 0 \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

식 (5)에서  $X$ 는 통제변수 벡터로 거주 지역, 소득, 나이, 성별, 결혼 여부, 가구원의 수 등을 포함하고 있으며,  $pvC$ 는 프라이버시 염려에 대한 변수이고,  $v$ 는 오차항으로 표준정규분포를 따르는 확률변수이다.

잠재변수  $h^*$ 는  $s = 1$ 인 경우에만 관찰되며, 수량 방정식은 식 (6)과 같이 절단 선형모형 또는 지수모형으로 표현된다.  $X$ 는 통제변수 벡터,  $pvC$ 는 프라이버시 염려에 대한 변수이고,  $\epsilon$ 는 오차항으로 절단된 표준정규분포를 따르는 확률변수이다.<sup>13)</sup>

11) 참가 방정식으로 해석의 관점에서 이점을 가지고 있는 로짓(logit) 모형을 생각해볼 수 있으나, 이부모형을 개발한 Cragg (1971)와 이를 stata에서 추정할 수 있게 제시한 Burke (2009)이 프로빗 모형만을 제시하고 있으므로 본 연구에서는 프로빗(probit) 모형을 사용하였다.

12) Cragg (1971)은 절단된 정규분포를 이용하는 대신 로그정규분포를 이용한 대안적 모형을 언급하였고, Duan et al. (1984)이 그 가능성을 자세히 연구하여 로그정규분포 모형을 제시하였다.

$$h^* = (\alpha_2 + \beta_2 X + \gamma_2 pvC + \epsilon) \text{ (절단된 선형 모형)}$$

$$h^* = \exp(\alpha_3 + \beta_3 X + \gamma_3 pvC + \epsilon) \text{ (로그 정규분포 모형)}$$
(6)

요약하면, 디지털 콘텐츠 지출 데이터의 준연속성의 특징에 따라 Cragg의 이부모형을 이용하여 추정하면, 프라이버시 염려가 증가할 때 디지털 콘텐츠 지출이 어떻게 변하는지를 두 가지로 구분하여, 즉 디지털 콘텐츠 지출 여부에 대한 결정과 디지털 콘텐츠 지출액 조정에 대한 결정에 미치는 차별적인 영향에 대한 분석이 가능하다.

본 연구에서는 이부모형을 이용한 실증분석의 강건성(robustness) 검증을 위하여 두 가지 실증방법을 추가로 사용하여 결과의 질적인 유사성을 비교하였다. 첫째, 패널 자료의 특성을 보다 적극적으로 반영하여 디지털 콘텐츠 지출 여부의 결정에 대해서는 패널 프로빗 모형, 디지털 콘텐츠 지출이 양인 경우 디지털 콘텐츠 지출액 결정에 대해서는 패널 선형 분석을 시도하였다. 둘째, 처치 효과(Treatment Effects) 모형을 이용하여 프라이버시 염려의 효과만을 추출하여 분석하는 시도를 하였다.<sup>14)</sup>

## 2. 실증 분석 결과

### 1) Cragg 이부모형(Two-Part Model)

종속변수를 디지털 콘텐츠 지출로 설정하고 프라이버시 염려에 어떠한 영향을 받는지를 살펴보기 위해 2014~15년 자료를 통합하여(pooling) 분석한 실증결과는 <표 4>에 제시하였다. 본 연구에서 사용한 프라이버시 염려의 변수가 척도 변수여서 한 단위 변화의 의미가 불명확하기 때문에, 추정계수의 크기보다는 부호 및 통계적 유의성에 집중하여 결과를 요약하기로 한다.

13) 지수모형의 양변에 로그를 취하면  $\log(h^*) = (\alpha + \beta X + \gamma pvC + \epsilon)$ 과 같이 쓸 수 있으며, 이때 오차항  $\epsilon$ 는 정규분포로 가정한다. 따라서  $h^*$ 는 로그정규분포(log-normal)를 따르게 된다.

14) 처치효과 모형에 대한 보다 자세한 설명은 신일순·김현수 (2018)를 참조할 것.

<표 4> 디지털 콘텐츠 지출에 대한 분석 결과: OLS, Tobit, 이부모형)

	(1)	(2)	(3)	(4)	(5)
	선형회귀 모형 (OLS)	토빗 모형 (Tobit)	이부 모형-1 프로빗 모형 (디지털 콘텐츠 지출 여부)	이부 모형-2 절단된 회귀 모형 (디지털 콘텐츠 지출액)	이부 모형-2 지수 모형 (log 디지털 콘텐츠 지출액)
<b>프라이버시 염려</b>	<b>0.030*** (0.004)</b>	<b>0.836*** (0.069)</b>	<b>0.039*** (0.003)</b>	<b>-1.103 (1.935)</b>	<b>-0.005 (0.005)</b>
특별/광역시 거주	0.141*** (0.048)	4.155*** (0.752)	0.199*** (0.034)	-12.057 (21.331)	-0.04 (0.059)
성별	0.007 (0.054)	-1.392* (0.814)	-0.087** (0.037)	54.119 (53.251)	0.137** (0.058)
나이	-0.010*** (0.002)	-0.382*** (0.034)	-0.017*** (0.001)	0.816 (1.208)	0.001 (0.004)
결혼 여부	-0.214*** (0.067)	-2.639** (1.114)	-0.148*** (0.050)	36.849 (44.016)	0.091 (0.098)
직업 유무	0.049 (0.069)	1.332 (1.176)	0.051 (0.053)	43.721 (49.341)	0.169* (0.092)
1인 가구	-0.109 (0.115)	-2.248 (2.213)	-0.090 (0.100)	-52.452 (80.346)	-0.092 (0.14)
3세대 이상 가구	0.259*** (0.089)	3.617*** (1.282)	0.148** (0.059)	40.017 (45.404)	0.203** (0.096)
소득	0.039*** (0.012)	1.008*** (0.181)	0.048*** (0.008)	-4.926 (6.261)	-0.011 (0.014)
상수항	0.311*** (0.108)	-42.918*** (2.191)	-1.911*** (0.081)	-299.246 (311.430)	1.863*** (0.148)
관측치	17,608	17,608	17,608	941	941
R-sq	0.0167				0.0241
Log Likelihood		-6394.303	-3269.171	-3083.158	

주: 괄호속의 수치는 추정계수의 표준오차이고, \*\*\*는 p<0.01, \*\*는 p<0.05, \*는 p<0.01를 각각 의미함.

먼저 (1)열은 선형회귀(OLS) 모형을 이용한 결과이다.<sup>15)</sup> 우리가 관심을 가지는 추정계수는 프라이버시 염려가 증가할 경우 디지털 콘텐츠 지출이 어떻게 변하는지에 대한 것으로, 여기서는 통계적으로 유의한 양(+)의 효과를 가지는 것으로 나타난다. 이 결과는 프라이버시 역설(privacy paradox) – 프라이버시 염려가 증가함에도 불구하고 개인정보를 적극적으로 제공하는 현상 – 이 개인정보의 누출을 수반하는 디지털 콘텐츠 지출 측면에서도 성립함을 의미한다. 프라이버시 염려 외의 다른 통제변수에 대한 실증 분석 결과는 특별/광역시에 거주할수록, 나이가 적을수록, 미혼일수록, 고소득일수록 디지털 콘텐츠 지출이 증가하는 것으로 나타나, 직관적인 설명과 부합하는 것을 알 수 있다.

다음으로 디지털 콘텐츠 지출 분포의 준연속적인 특성을 감안하여 토빗(Tobit) 모형을 추정방법으로 한 결과는 (2)열에 제시하였다. 선형회귀 모형의 결과와 비교하면 추정계수의 크기는 다르지만, 추정계수의 부호와 통계적 유의성은 동일하게 유지되는 것으로 나타난다. 즉, 토빗 모형을 사용한 실증분석도 프라이버시 역설 현상이 나타나는 것으로 결론지을 수 있다.

이상의 결과를 요약하면 다음과 같다. 디지털 콘텐츠 지출에 대한 개인의 선택은 디지털 콘텐츠 지출을 하는지/하지 않는지와 한다면 얼마나 하는지의 두 가지 선택으로 구분할 수 있다. 만일 연구자가 디지털 콘텐츠 지출 여부와 지출액을 구분하지 않고 실증적으로 분석하면, 준연속적인 특징을 반영함에 상관없이 OLS 및 토빗 모형을 통해 디지털 콘텐츠 지출 측면에서 프라이버시 역설이 성립하는 것으로 결론지을 수 있다. 이러한 결과는 SNS 등의 다른 온라인 활동에 비해 거래 및 지급결제의 정보의 누출 가능성이 추가적으로 존재하는 디지털 콘텐츠 지출에 대한 행동도 프라이버시 역설 현상에서 벗어나지 못하는 점을 시사하는 것으로, 프라이버시 보호에 대한 보다 강력한 정책적인 보호의 필요성을 제기하는 것으로 해석할 수 있다.

다음으로 (3), (4), (5)열에는 Cragg (1971)의 이부모형(two-part model)을 추정한 결과를 보고하였다. 이부모형은 참가 방정식과 수량 방정식의 오차항에 서로 상관관계가 없다는 가정 하에, 참가 방정식(디지털 콘텐츠 지출 여부의 결정 방정식)은

15) 선형회귀모형의 추정결과는 IV장 1절 ‘실증모형’에서 설명한 바와 같이 준연속적인 종속 변수의 특징으로 인해 편의성이 발생할 수 있으므로 적절한 추정결과는 아니나, 다른 모형과의 비교를 위해 제시하였다.



프로빗 모형으로, 수량 방정식(양의 디지털 콘텐츠 지출을 한다는 조건 하에서 디지털 콘텐츠 지출액의 결정 방정식)은 절단된 회귀모형과 지수모형을 이용하여 추정할 수 있다.<sup>16)</sup>

(3)열에서 프라이버시 염려의 효과는 선형회귀 모형의 결과와 유사하며, 여성이 디지털 콘텐츠 지출을 결정할 확률이 통계적으로 유의하게 높다는 점을 제외하면 추정계수의 부호 및 통계적 유의성의 측면에서 다른 통제변수도 대체로 선형회귀 모형의 결과와 유사하게 나타난다.<sup>17)</sup>

이에 비해 (4)열은 수량 방정식을 절단 회귀모형으로 추정한 결과로 관측치의 수가 941개인 것은 디지털 콘텐츠 지출이 양수(+)인 개인만을 대상으로 하였기 때문이다. 그런데 추정결과를 보면 통계적으로 유의한 변수가 전혀 없는 것으로 나타난다. 특히 프라이버시 염려에 대한 추정계수는 디지털 콘텐츠 지출 여부의 참여 방정식에서와 달리 통계적 유의성이 없어지는 것을 알 수 있다. 이 결과는 이미 디지털 콘텐츠 지출을 하고 있는 사람들을 대상으로 판단할 때 더 이상 강한 수준의 프라이버시 역설이 성립하지 않는다는 것을 의미한다.<sup>18)19)</sup>

(5)열은 수량 방정식을 지수모형으로 추정한 결과로 추정계수의 방향성은 (4)열의 결과와 크게 차이가 없으나 성별, 직업 유무, 3세대 이상 가구 등의 변수에서 통계적 유의성이 나타났다. 이러한 차이는 자료의 특성이 절단 회귀분포보다는 로그정규분포에 가깝기 때문인 것으로 판단된다.

이부모형의 실증결과는 디지털 콘텐츠 지출 측면에서 프라이버시 역설에 대한

16) Burke (2009)에서는 Cragg의 이부모형을 추정할 수 있는 stata의 craggit 명령어를 제시하고 있다.

17) 프로빗 모형은 비선형모형이기 때문에 추정계수가 해당 변수의 한계효과를 의미하지는 않는다는 점을 주의할 필요가 있다.

18) 프라이버시 염려가 증가할 때 디지털 콘텐츠 지출액이 통계적으로 유의하게 증가하는 경우 강한 수준의 프라이버시 역설, 통계적으로 유의하지 않는 경우에는 약한 수준의 프라이버시 역설이 성립한다고 볼 수 있으며, 디지털 콘텐츠 지출액이 통계적으로 유의하게 감소하는 경우 프라이버시 역설이 성립하지 않는 것으로 해석할 수 있다.

19) <부록 표 1> 개별 디지털 콘텐츠(음악)를 대상으로 분석한 결과를 보면, 프라이버시 염려는 통계적으로 유의하게 음(-)의 값이 나타나 더 이상 프라이버시 역설이 성립하지 않는 것으로 나타난다. <부록 표 1>에 5가지 디지털 콘텐츠 중 ‘음악’만을 제시한 이유는 나머지 디지털 콘텐츠의 경우 모형적합도(추정계수 중 적어도 하나는 0이 아니라는 귀무가설)가 유의하지 않았기 때문이다.

다음과 같은 흥미롭고 중요한 시사점을 제공한다.<sup>20)</sup> 첫째, 디지털 콘텐츠 지출 여부와 지출액을 구분하지 않은 실증분석에서 프라이버시 역설이 나타난 것은 주로 참가 방정식에서의 효과, 즉 디지털 콘텐츠 지출 여부에 대한 효과에 기인한다. 다른 말로 표현하면, 프라이버시 염려가 증가함에도 불구하고 디지털 콘텐츠 지출이 증가하는 것은 디지털 콘텐츠 지출을 하지 않던 사람들이 프라이버시 염려가 증가해도 디지털 콘텐츠 지출을 선택하기 때문이다. 둘째, 이미 디지털 콘텐츠 지출을 하고 있는 사람들을 대상으로 한 수량 방정식의 결과를 살펴보면, 프라이버시 역설의 정도가 강한 수준에서 약한 수준으로 줄어든 것으로 나타난다.<sup>21)</sup>

이 결과는 이부모형을 사용하지 않은 경우 도출되는 해석, 즉 프라이버시 역설의 발생이 현저하고, 따라서 그 해결의 필요성이 정책적으로 요구된다는 식의 결론에 대해 약간의 조정이 필요함을 의미한다. 이를 좀 더 자세히 설명하면 다음과 같다. 첫째, 프라이버시를 희생하면서 디지털 콘텐츠 지출을 하는 행동은 주로 디지털 콘텐츠 지출을 하지 않던 사람들이 디지털 콘텐츠 지출을 선택하기 때문에 발생하는 것인 바, 만일 이들이 프라이버시 희생의 비용과 디지털 콘텐츠 지출의 편익을 상호 비교하여 합리적으로 결정한다면, 프라이버시를 강화하는 정책의 수립 및 시행에도 불구하고 개인의 행동은 변함이 없고, 거래 당사자의 비용과 불편성만 가증될 가능성이 존재하다. 둘째, 이미 디지털 콘텐츠 지출에 적극적으로 참여하고 있는 사람들은 프라이버시 염려에 대해 불합리한 반응을 보이지 않는 것으로 해석할 수 있으므로, 굳이 이를 해결하기 위한 정책적인 접근이 필요한지에 대해서는 재고가 필요함을 의미한다.

20) 시사점을 언급하기에 앞서, 이부모형의 실증결과는 (3)열 참가 방정식과 (4), (5)열의 수량 방정식에서 같은 요인(변수)이더라도 추정계수의 방향이 다르거나 통계적 유의성이 달라지는 것을 통해 토빗 모형의 가정이 본 연구에서는 타당하지 않음을 보여준다.

21) 본 연구는 자료의 한계 때문에 이러한 일이 왜 발생하는지에 대한 이유를 적극적으로 설명하지는 못하고 있다. 그럼에도 불구하고 하나의 가능한 설명은 다음과 같다. 디지털 콘텐츠 지출을 하지 않던 사람이 디지털 콘텐츠 지출을 선택하는 경우, 그 이유는 디지털 콘텐츠 지출의 편익(benefit)이 크면서 동시에 그 편익이 프라이버시 침해의 가능성으로 인한 비용(cost)을 초과하기 때문이라고 볼 수 있으며, 이 때 참여 방정식에서 프라이버시 역설 현상이 관찰된다. 이에 비해 이미 디지털 콘텐츠 지출을 하고 있는 사람이 지출액을 증가시키는 편익은 앞서 설명한 디지털 콘텐츠 지출 선택의 편익에 비해 상대적으로 더 작기 때문에 수량 방정식에서 프라이버시 염려가 증가할 때 지출액을 줄이는 (또는 영향을 받지 않는) 반응을 보인다고 해석할 수 있다. 물론 이러한 해석의 정당성에 대한 입증은 전술한 것처럼 보다 세밀한 자료와 정직한 분석을 필요로 하는 것으로 차후의 연구로 미루기로 한다.

## 2) 강건성 검증: 패널(Panel) 모형

이번 절에서는 앞서 분석한 이부모형의 강건성을 검증하기 위해 패널 모형을 이용하였다. 패널 모형을 이용하면 개인의 성향은 시간에 따라 쉽게 변하지 않는다는 가정 하에서 개인의 이질성을 통제할 수 있는 장점을 가진다.<sup>22)</sup> 본 연구에서는 디지털 콘텐츠 지출 여부에 대해서는 확률효과 패널 프로빗(random effect panel probit) 모형을 이용하고, 디지털 콘텐츠 지출이 양(+인) 개인의 디지털 콘텐츠 지출액 선택은 확률효과 패널 선형(random effect panel regression) 모형을 이용하였다.<sup>23)</sup>

먼저 <표 5>의 (1)열은 종속변수를 디지털 콘텐츠 지출 여부로 하고 확률효과 패널 프로빗 모형을 이용한 결과이다. 프라이버시 염려 및 다른 통제변수에 대한 추정계수의 부호 및 통계적 유의성이 <표 4> (3)열의 참가 방정식의 결과와 유사하게 나타난다. 다음으로 (2)열은 종속변수를 디지털 콘텐츠 지출액으로 하고 디지털 콘텐츠 지출이 양(+인) 개인을 대상으로 확률효과 패널 선형 모형을 이용한 결과이다. 이 역시 프라이버시 염려 및 다른 통제변수에 대한 추정계수의 부호 및 통계적 유의성이 <표 4> (4)열의 수량 방정식의 결과와 유사한 것을 알 수 있다.

결론적으로 패널 분석을 통해 살펴보더라도, 폴링 데이터를 통한 이부모형 결과의 강건성이 검증됨을 알 수 있다. 즉, 프라이버시 역설이 디지털 콘텐츠 지출 여부의 선택에서 주로 나타나고, 통계적 유의성은 떨어지지만 프라이버시 염려가 증가할 때 디지털 콘텐츠 지출액의 결정은 감소하는 특징을 보이고 있다.

22) 패널 모형을 기본 모형(baseline model)로 하지 않은 이유는 패널데이터의 이부모형에 대한 계량분석 방법이 아직 확고하게 정립되지 않았기 때문이다.

23) 패널 프로빗 모형에서 고정효과(fixed effect) 모형을 사용하지 않는 이유는 본 자료와 같이 패널 개체의 수가 많고 시계열 관측 개체수가 작은 경우에 고정효과 모형으로 추정하면 추정계수의 한계효과 추정치가 과대 추정되는 문제가 발생하기 때문이다. 이에 대한 자세한 논의는 민인식·최필선 (2012, p.232) 및 Greene (2002, p.27)을 참조할 것. 반면 패널 선형 모형에서는 하우스만 테스트(Hausman test)를 시행한 결과 하우스만 테스트 검정통계량이 1.76으로 나타나 설명변수와 교란항 간 상관관계가 존재하지 않는다는 귀무가설을 기각할 수 없어 확률효과 모형이 적합한 것으로 나타났다.

<표 5> 디지털 콘텐츠 지출에 대한 분석 결과: 패널 모형

	(1)	(2)
	패널 프로빗 모형 (디지털 콘텐츠 지출 여부)	패널 선형 모델 (디지털 콘텐츠 지출액)
<b>프라이버시 염려</b>	<b>0.049*** (0.004)</b>	<b>-0.042 (0.060)</b>
특별/광역시 거주	0.266*** (0.048)	-0.445 (0.650)
성별	-0.115** (0.052)	1.879*** (0.672)
나이	-0.023*** (0.002)	0.033 (0.037)
결혼 여부	-0.167** (0.069)	1.228 (1.002)
직업 유무	0.071 (0.072)	1.526 (1.035)
1인 가구	-0.118 (0.138)	-1.598 (1.897)
3세대 이상 가구	0.199** (0.081)	1.563 (1.070)
소득	0.061*** (0.011)	-0.172 (0.160)
상수항	-2.447*** (0.125)	8.343*** (1.739)
관측치	17,608	941
Log Likelihood	-3204.932	
Wald Ch-sq		25.33

주: 괄호속의 수치는 추정계수의 표준오차이고, \*\*\*는  $p < 0.01$ , \*\*는  $p < 0.05$ , \*는  $p < 0.1$ 를 각각 의미함.

### 3) 강건성 검증: 처치 효과(Treatment Effect) 모형

앞선 분석에서 사용한 통제변수는 디지털 콘텐츠 지출에 영향을 미칠 개연성이 있으며, 한국미디어패널 조사에서 이용 가능한 변수를 추출한 것이다. 그런데 만일 다른 관측 불가능한 변수가 존재하고 이 변수가 프라이버시 염려 및 디지털 콘

텐츠 지출과 상관관계가 존재한다면, 프라이버시의 추정계수가 편의성(bias)을 가지는 문제가 발생한다. 이러한 문제를 우회하기 위해 본 연구에서는 최근 계량경제학에서 사용하는 처치효과(treatment effect) 모형 중 IPWRA(Inverse Probability Weighted Regression Adjustment) 방법을 이용하였다.<sup>24)</sup>

그런데 처치효과 방법이 두 집단, 즉 처치집단과 통제집단의 차이를 보여주기 때문에 본 연구에서는 프라이버시 염려의 수준에 대한 변수들을 재분류하여 평균치를 넘는 경우는 높은 그룹, 그렇지 않은 경우는 낮은 그룹의 두 가지로 구분하였다.

<표 6> 디지털 콘텐츠 지출에 대한 분석 결과: 처치효과 모형

	디지털 콘텐츠 지출 여부	디지털 콘텐츠 지출액
ATE	0.035*** (0.003)	-1.093 (0.728)

프라이버시 염려가 높은 처치집단과 낮은 통제집단의 디지털 콘텐츠 지출의 차이에 대한 처치 효과의 추정치는 <표 6>에 제시한 평균처치 효과(Average Treatment Effect; ATE)로 알 수 있다. 이에 따르면, 프라이버시 염려가 높은 그룹이 통계적으로 유의미하게 더 많이 디지털 콘텐츠 지출을 하지만, 통계적으로 유의미하지는 않지만 디지털 콘텐츠 지출액은 감소시키는 것으로 나타난다. 이에 따라 우리는 처치효과 모형의 결과 역시 풀링 데이터를 통한 이부모형 결과의 강건성이 검증됨을 알 수 있다.

## V. 요약 및 결론

본 연구에서는 온라인 구매의 측면에서 프라이버시 역설이 성립함을 실증적으로 보인 최한솔 외 (2019)의 분석에서 한 걸음 더 나아가, 디지털 콘텐츠 지출 데이터가 준연속적인(semi-continuous) 특징을 가지는 점에 기초하여 프라이버시 염려가

24) 이에 대한 보다 자세한 설명은 신일순·김현수 (2018)를 참조할 것.

디지털 콘텐츠 지출에 미치는 영향을 디지털 콘텐츠 지출 여부와 디지털 콘텐츠 지출액으로 구분하여 분석하였다. 준연속적인 자료는 0의 값과 연속형 자료의 값이 혼합되어 나타나는 형태의 자료로서, 소비자들이 디지털 콘텐츠 지출을 하지 않는 소비자와 양(+의 지출을 하는 소비자로 구분됨을 의미한다. 이 경우 Cragg (1971)의 이부모형(two-part model)을 이용하면 프라이버시 염려가 디지털 콘텐츠 지출을 하지 않던 소비자가 디지털 콘텐츠 지출을 선택하는 데에 어떠한 영향을 미치는지와 함께 이미 양(+의 디지털 콘텐츠 지출을 하는 소비자가 디지털 콘텐츠 지출액을 조정하는 데에 어떠한 영향을 미치는지를 구분하여, 두 가지 측면에서 분석하는 것이 가능하다.

본 연구는 이부모형을 이용하여 프라이버시 염려가 디지털 콘텐츠 지출 여부 및 디지털 콘텐츠 지출액에 미치는 영향을 실증적으로 추정하였으며, 패널 모형 및 처치효과 모형을 이용하여 이부모형의 강건성을 검증하였다. 이부모형, 패널 모형 및 처치효과 모형을 통한 실증분석에서는 모두 유사한 결과 도출되었는데, 프라이버시 역설은 주로 디지털 콘텐츠 지출 여부에 대한 효과에 기인하며, 이미 디지털 콘텐츠 지출을 하고 있는 사람들을 대상으로 하는 경우는 비록 통계적 유의성은 떨어지지만 프라이버시 염려가 증가할 때 디지털 콘텐츠 지출액을 줄이는 식으로 반응하는 것으로 나타났다.

본 연구의 실증결과는 다음과 같이 요약된다. 첫째, 디지털 콘텐츠 지출 여부와 지출액을 구분하지 않은 실증분석에서 프라이버시 역설이 나타난 것은 주로 참가 방정식에서의 효과, 즉 디지털 콘텐츠 지출 여부에 대한 효과에 기인한다. 다른 말로 표현하면, 프라이버시 염려가 증가함에도 불구하고 디지털 콘텐츠 지출이 증가하는 것은 디지털 콘텐츠 지출을 하지 않던 사람들이 프라이버시 염려가 증가해도 디지털 콘텐츠 지출을 선택하기 때문이다. 둘째, 이미 디지털 콘텐츠 지출을 하고 있는 사람들을 대상으로 한 수량 방정식의 결과를 살펴보면, 비록 통계적 유의성은 떨어지지만 프라이버시 염려가 증가할 때 합리적이고, 바람직하게 디지털 콘텐츠 지출액을 줄이는 식으로 반응하여 프라이버시 역설이 더 이상 성립하지 않는 것으로 나타난다.

본 연구는 자료의 한계 때문에 이러한 일이 왜 발생하는지에 대한 이유를 적극적으로 설명하는 시도는 하지 못하고 있지만, 그럼에도 불구하고 하나의 가능한 설명은 다음과 같다. 디지털 콘텐츠 지출을 안 하던 사람이 디지털 콘텐츠 지출의 편익을 느껴서 이를 시작하는 경우는 그 편익이 프라이버시 침해의 가능성을 초

과하기 때문이라고 볼 수 있으며, 이 때 참여방정식에서 프라이버시 역설 현상이 관찰된다. 이에 비해 이미 디지털 콘텐츠 지출을 하고 있는 사람은 지출액을 증가시키는 편익이 참여 여부에 대한 결정에 비해 상대적으로 크지 않기 때문에 수량 방정식에서 프라이버시 염려가 증가할 때 지출액을 줄이는 (또는 영향을 받지 않는) 반응을 보인다고 해석할 수 있다.

디지털 콘텐츠 지출 측면에서 프라이버시 역설에 대한 다음과 같은 시사점을 제공한다. 만일 이부모형을 사용하지 않고, 즉 디지털 콘텐츠 지출의 참여와 수량 조절을 구분하지 않고 분석하면 우리는 프라이버시 역설의 발생이 현저하고, 따라서 그 해결을 위해 정책적인 접근이 시급하고 필수적으로 요구된다는 식으로 결론을 내리기 쉽다. 그러나 이부모형을 이용하여 명시적으로 고려하는 경우, 프라이버시를 희생하면서 디지털 콘텐츠 지출을 하는 행동은 주로 디지털 콘텐츠 지출을 하지 않던 사람들이 디지털 콘텐츠 지출을 선택하기 때문에 발생하며, 이미 양(+의) 디지털 콘텐츠 지출을 하는 사람들은 프라이버시 염려에 대해 합리적으로 디지털 콘텐츠 지출액을 늘리지 않는 반응을 보인다는 실증분석의 결과가 도출되었다. 이러한 실증결과는 소비자들이 어느 정도 프라이버시 희생의 비용과 디지털 콘텐츠 지출의 편익을 상호 비교하여 합리적으로 자신의 행동을 결정하고 있다는 점을 시사한다. 따라서 디지털 콘텐츠 지출 및 거래에서 프라이버시 희생을 해결하기 위해 정책적인 접근방법의 당위성에 대해서는 좀 더 신중한 논의가 필요하며, 보다 상세히 타겟팅된 정책방향이 요구되는 것으로 판단된다.

## 참고문헌

- 김종기·김상희·김진성 (2018). 정보 프라이버시 역설에 대한 학술적 고찰. 『2018 NAVER Privacy White Paper』. 네이버(주).
- 민병권·김이태 (2006). 『온라인 쇼핑에서 프라이버시 염려의 원인변수와 결과변수. 『한국콘텐츠학회논문지』, 6(11), 25-37.
- 민인식·최필선 (2012). 『STATA 패널 데이터 분석』. 서울: 지필미디어.
- 민희철·정인석·신일순 (2016). 바이오기술 연구자의 상업화 활동 결정요인에 관한 실증연구. 『시장경제연구』, 45(1), 33-66.
- 신일순·김현수 (2018). 프라이버시 염려와 SNS 활동: 프라이버시 역설의 존재 및

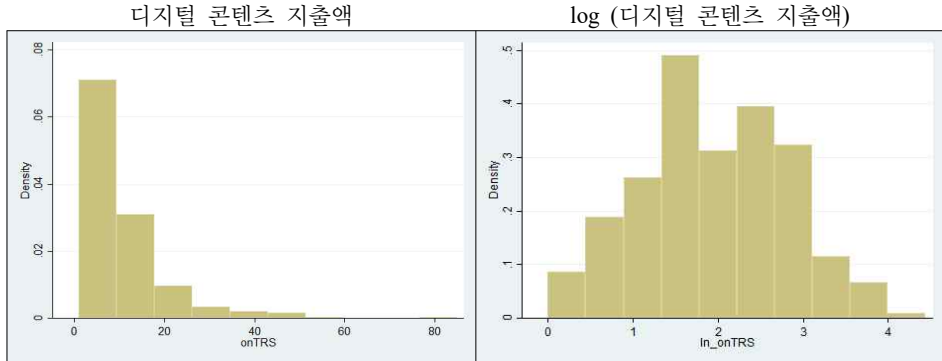
- 이유. 『정보통신정책연구』, 25(2), 33-67.
- 유일·신정신·이경근·최혁라 (2008). 프라이버시 염려 영향요인이 인터넷 이용자의 신뢰와 온라인 거래의도에 미치는 영향. 『JITAM』, 15(4), 37-59.
- 이동주·김명수 (2011). 전자상거래 이용자들의 정보 프라이버시 우려와 반응 행동에 대한 실증 연구. 『e-비즈니스연구』, 12(2), 365-383.
- 정병철 (2018). Two-part 패널모형을 이용한 미디어패널 자료 분석. 제6회 한국미디어패널 학술대회 논문집, 117-124.
- 최환솔·이명성·이혜미 (2019). 온라인 구매 맥락에서도 프라이버시 역설은 존재하는가?: 소비자의 프라이버시 염려 수준과 온라인 구매 행동에 대한 연구. 『상품학연구』, 37(5), 1-13.
- Acquisti, A., Brandimarte, L., & Loewenstein, G. (2015). Privacy and human behavior in the age of information. *Science*, 347(6221), 509-514.
- Brecht, F., Fabian, B., Kunz, S., & Mueller, S. (2011). Are You Willing to Wait Longer for Internet Privacy?. ECIS 2011 Proceedings, 236.
- Burke, W. (2009). Fitting and interpreting Cragg's tobit alternative using Stata. *The Stata Journal*, 9(4), 584-592.
- Cragg, J. G. (1971). Some Statistical Models for Limited Dependent Variables with Application to the Demand for Durable Goods. *Econometrica*, 39, 829-844.
- Duan, N., W. G. Manning, C. N. Morris, & J. P. Newhouse (1984). Choosing between the Sample-Selection Model and the Multi-Part Model. *Journal of Business and Economic Statistics*, 2, 283-289.
- Dinev, T., & Hart, P., (2005). Internet Privacy Concerns and Social Awareness as Determinants of Intention to Transact. *International Journal of Electronic Commerce*, 10(2), 7-29.
- Eastlick, M. A., Lotz, S. L., & Warrington, P. (2006). Understanding Online B-to-C Relationships: An Integrated Model of Privacy Concerns, Trust, and Commitment. *Journal of Business Research*, 59(8), 877-886.
- George, J. F. (2004). The Theory of Planned Behavior and Internet Purchasing. *Internet Research*, 14(3), 198-212.
- Greene, W. (2002). The Behavior of the Fixed Effects Estimator in Nonlinear Models. NYU Working Paper EC-02-05.  
<https://ssrn.com/abstract=1292651> (Retrieved on January 6, 2021).
- Khalifa, M., & Limayem, M. (2003). Drivers of Internet Shopping. *Communications*



- of the ACM*, 46(12), 233-239.
- Kokolakis, S. (2017). Privacy attitudes and privacy behaviour: A review of current research on the privacy paradox phenomenon. *Computers & Security*, 64, 122-134.
- Li, H., Sarathy, R., & Xu, H. (2010). Understanding Situational Online Information Disclosure as a Privacy Calculus. *Journal of Computer Information Systems*, 51(1), 62-71.
- Limayem, M., Khalifa, M., & Frini, A. (2000). What Makes Consumers Buy from Internet? A Longitudinal Study of Online Shopping. *IEEE Transactions on Systems, Man, and Cybernetics- Part A: Systems and Humans*, 30(4), 421-432.
- Suh, B., & Han, I. (2003). The Impact of Customer Trust and Perception of Security Control on the Acceptance of Electronic Commerce. *International Journal of Electronic Commerce*, 7(3), 135-161.

〈부록〉

〈부록 그림 1〉 디지털 콘텐츠 지출액(onTRS | onTRS>0) 히스토그램



〈부록 표 1〉 디지털 콘텐츠(음악) 지출에 대한 분석 결과: 이부모형

	(1)	(2)	(3)
	이부 모형-1 프로빗 모형 (디지털 콘텐츠 지출 여부)	이부 모형-2 절단된 회귀 모형 (디지털 콘텐츠 지출액)	이부 모형-2 지수 모형 (log 디지털 콘텐츠 지출액)
프라이버시 염려	<b>0.035***</b> (0.004)	<b>-0.224**</b> (0.102)	<b>-0.009*</b> (0.005)
특별/광역시 거주	0.216*** (0.044)	-1.211 (1.069)	-0.055 (0.057)
성별	-0.173*** (0.046)	1.456 (1.071)	0.066 (0.057)
나이	-0.020*** (0.002)	0.145* (0.08)	0.005 (0.004)
결혼 여부	-0.434*** (0.067)	-0.458 (1.767)	0.013 (0.098)
직업 유무	0.133* (0.073)	-0.492 (1.758)	-0.060 (0.097)
1인 가구	-0.090 (0.120)	-0.569 (2.784)	-0.037 (0.152)
3세대 이상 가구	-0.007 (0.081)	2.680 (1.804)	0.203** (0.102)
소득	0.044 (0.011)	0.184 (0.292)	0.017 (0.017)
상수항	-1.940*** (0.103)	1.082 (3.149)	1.602 (0.161)
관측치	17,608	485	485
R-sq			0.0358
Log Likelihood	-1891.2373***	-1307.462**	

주: 괄호속의 수치는 추정계수의 표준오차이고, \*\*\*는 p<0.01, \*\*는 p<0.05, \*는 p<0.1을 각각 의미함.