

RESEARCH ARTICLE

The Effect of Intellectual Property Regulation of AI Training Data on AI Capital Accumulation

Jaehyun Kim

Director, Public Finance Strategy Institute, Republic of Korea

Corresponding Author: Jaehyun Kim (kjaihyun@daum.net)

ABSTRACT

This study quantitatively examines how the level of intellectual property protection over artificial intelligence (AI) training data affects AI capital accumulation. Korean copyright law provides no explicit exception for text and data mining; therefore, using protected works for AI training depends on ex-post interpretation of the general fair use provision. Drawing on a review of domestic and foreign regulations and disputes, the study identifies two channels through which protection operates: rising costs at the data procurement stage and binding constraints at the training stage. This distinction is consistent with 2025 U.S. court decisions that assess AI training separately from data acquisition. A dynamic general equilibrium model incorporating a Lucas span-of-control structure represents these channels as policy parameters and simulates how changes in protection affect aggregate output and knowledge capital, the model's representation of AI capital. Under the baseline calibration, tightening protection by 20 percent reduces the knowledge capital stock by 17.1 percent and output by 3.4 percent, with the training-stage channel accounting for the larger share of the overall effect. This effect arises primarily from reduced use of knowledge capital rather than substitution toward labor.

KEYWORDS

AI training data, intellectual property regulation, text and data mining, dynamic general equilibrium model, digital market competition, AI capital

Open Access

Received: July 31, 2026**Revised:** September 08, 2026**Accepted:** September 22, 2026**Published:** September 30, 2026

Funding: The author received manuscript fees for this article from Korea Institute of Intellectual Property.

Conflict of interest: No potential conflict of interest relevant to this article was reported.

© 2026 Korea Institute of Intellectual Property



This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 (<https://creativecommons.org/licenses/by-nc-nd/4.0/>) which permits use, distribution and reproduction in any medium, provided that the article is properly cited, the use is non-commercial and no modifications or adaptations are made.

원저

학습데이터에 대한 지식재산권 규제가 AI 자본 축적에 미치는 영향

김재현

공공재정전략연구소 연구소장

교신저자: 김재현 (kjaihyun@daum.net)

차례

1. 서론
2. 기존연구 검토
 - 2.1. AI 학습데이터에 대한 지식재산권 적용
 - 2.2. 데이터 접근성과 디지털 시장 경쟁
 - 2.3. 지식재산권 보호와 혁신 유인
 - 2.4. 디지털-AI 시장에서의 혁신과 경쟁
 - 2.5. 선행연구의 한계와 본 연구의 차별성
3. AI 학습데이터 관련 규제 사례 검토
 - 3.1. 국내 법제의 이중 규범 구조
 - 3.2. 주요국 TDM 규제의 비교
 - 3.3. 주요 분쟁 사례
 - 3.4. 규제 강도의 정책변수화
4. 분석모형
 - 4.1. 개요
 - 4.2. 기업의 이윤극대화
 - 4.3. 가계의 효용극대화
 - 4.4. 지식자본의 축적과 지식재산권 보호
 - 4.5. 직업선택의 문제와 시장균형
 - 4.6. 균형의 성질과 비교정확
 - 4.7. 모수의 보정(Calibration)
 - 4.8. 분석 절차
5. 분석결과
 - 5.1. 보호수준 변화의 장기효과
 - 5.2. 두 경로의 분해
 - 5.3. 규제 완화 수단에 따른 효과
 - 5.4. 이행경로 분석 결과
 - 5.5. 민감도 분석 결과
6. 결론 및 정책적 시사점
 - 6.1. 연구 결과의 요약
 - 6.2. 정책적 시사점
 - 6.3. 연구의 한계와 향후 과제

국문초록

본 연구는 AI 학습데이터에 대한 지식재산권 보호 수준이 AI 자본 축적에 미치는 영향을 정량적으로 분석한다. 현행 한국 저작권법은 텍스트·데이터마이닝에 관한 명시적 면책 규정을 두고 있지 않아, AI 학습 목적의 저작물 이용은 공정이용 조항의 사후적 해석에 의존하고 있다. 본 연구는 국내외 규제 및 분쟁 사례 검토를 통해 지식재산권 보호가 데이터 조달 단계의 비용 상승과 학습 단계의 이용 제약이라는 두 경로로 작동함을 확인하였다. 이러한 구분은 학습 행위와 데이터 취득 행위를 분리하여 평가한 2025년 미국 판례와 정합적이다. 이어 두 경로를 정책변수로 설정하고 Lucas의 span-of-control 구조를 결합한 거시 동태일반균형모형을 구축하여, 보호 수준의 변화가 지식자본 축적과 총생산에 미치는 효과를 시뮬레이션하였다. 분석 결과 기준 시나리오 대비 보호 수준을 20% 강화할 경우 지식자본 스톡은 17.1%, 총생산은 3.4% 감소하며, 이 가운데 학습 단계의 이용 제약이 데이터 조달비용 상승보다 큰 기여를 하는 것으로 나타났다. 이러한 효과는 노동으로의 대체보다 지식자본 이용 자체의 축소를 통해 주로 실현된다. 본 연구는 법제 분석의 결과를 모형의 정책변수로 직접 연결함으로써 규제 설계의 경제적 함의를 수치로 제시한다.

주제어

AI 학습데이터, 지식재산권 규제, 텍스트·데이터마이닝, 동태일반균형모형, 디지털시장 경쟁, AI 자본

1. 서론

인공지능(AI) 모델의 성능은 학습에 투입되는 데이터의 양과 질에 크게 의존한다. 이에 따라 학습데이터에 대한 접근 가능성은 AI 기업의 생산성을 결정하는 핵심 요인이자 시장 경쟁의 출발 조건으로 기능하고 있다. 문제는 이 학습데이터의 상당 부분이 저작권법과 데이터베이스제작자의 권리에 의해 보호되는 대상이라는 점이다. 결과적으로 지식재산권 보호 수준은 창작자에 대한 보상 체계를 넘어, 디지털·AI 시장의 경쟁구조를 결정하는 산업정책 변수로 그 성격이 전환되고 있다.

이러한 전환은 한국에서 특히 첨예하게 나타난다. 유럽연합은 디지털 단일시장 저작권 지침을 통해 텍스트·데이터마이닝(Text and Data Mining, 이하 ‘TDM’)에 관한 명시적 예외를 두었고, 일본은 저작권법 제30조의4를 통해 비향유 목적의 저작물 이용을 폭넓게 허용하고 있다.¹⁾²⁾

반면 현행 한국 저작권법에는 AI 학습 목적의 TDM에 관한 명시적 면책 규정이 존재하지 않는다. TDM 면책 조항을 담은 저작권법 개정안이 여러 차례 발의되었으나 입법에 이르지 못하였고, 2026년 1월 시행된 「인공지능 발전과 신뢰 기반 조성 등에 관한 기본법」(이하 ‘인공지능 기본법’)은 학습데이터 요약 공개를 포함한 투명성 확보 의무를 부과하였으나 학습 행위 자체의 적법성 문제는 다루지 않았다.³⁾

따라서 한국에서 AI 학습 목적의 저작물 이용은 저작권법 제35조의5의 포괄적 공정이용 조항에 대한 사후적 해석에 의존하고 있으며, 이 해석에 관한 국내 판례는 아직 확립되어 있지 않다. 이러한 입법 공백은 AI 기업이 부담하는 권리처리 비용과 법적 불확실성을 동시에 높이는 방향으로 작용한다.

그럼에도 이 문제에 관한 국내외 논의는 대체로 정성적 수준에 머물러 있다. 법학 문헌은 TDM 예외의 부재가 AI 산업에 부담으로 작용한다는 점을 지적하지만 그 부담의 크기를 제시하지 못하고, 경제학 문헌은 지식재산권 보호와 혁신의 관계를 다루면서도 AI 학습데이터라는 특수한 맥락과 시장 경쟁구조라는 변수를 함께 고려하지 못하였다. 이 공백은 정책 판단에 실질적인 제약을 만든다. 보호 수준을 어느 방향으로 조정할 것인지를 논의하려면 조정의 결과로 무엇이 얼마나 변하는지에 대한 정량적 근거가 필요하기 때문이다.

이러한 필요성은 최근 경쟁법 영역의 판례 변화에서 더욱 분명해졌다. 대법원은 2025년 10월 네이버의 자사우대 행위에 대한 공정거래위원회 처분을 다투는 사건에서, 시장지배적 지위 남용이 성립하려면 문제된 행위가 관련시장에서 경쟁을 제한하는 구체적 효과를 가져왔다는 점이 실증적으로 입증되어야 한다고 판시하며 원심을 파기환송하였다.⁴⁾

이 판결은 데이터·알고리즘을 매개로 한 행위의 경쟁제한성을 판단할 때 정성적 논증만으로는 충분하지 않다는 점을 명확히 하였다는 점에서, 본 연구와 같은 정량적 분석의 필요성을 뒷받침한다.

본 연구는 이러한 문제의식에서 출발하여, AI 학습데이터에 대한 지식재산권 보호 수준의 변화가 AI 자본의 축적에 미치는 영향을 정량적으로 분석하는 것을 목적으로 한다. 구체적으로 본 연구는 다음 두 가지 작업을 수행한다.

1) Directive (EU) 2019/790 of the European Parliament and of the Council of 17 April 2019 on Copyright and Related Rights in the Digital Single Market, Arts. 3-4.

2) 日本 著作権法 第30條の4.

3) 「인공지능 발전과 신뢰 기반 조성 등에 관한 기본법」 제31조. 동법은 2026년 1월 시행되었다.

4) 대법원 2025. 10. 16. 선고 2023두32709 판결.

첫째, 국내외 규제 및 분쟁 사례에 대한 검토를 통해 AI 학습데이터에 대한 지식재산권 보호가 기업에 작용하는 경로를 유형화한다. 본 연구는 이 경로가 데이터를 조달하는 단계에서 발생하는 비용 상승과, 확보한 데이터를 학습에 활용하는 단계에서 발생하는 제약이라는 두 가지로 구분될 수 있음을 보인다. 이 구분은 임의적인 것이 아니라 최근 판례에서 실제로 관찰되는 법리적 분화를 반영한 것이다.

둘째, 이렇게 유형화된 두 경로를 정책변수로 설정하고, Lucas(1978)의 span-of-control 구조를 결합한 거시 동태일반균형모형을 구축하여 보호 수준의 변화가 지식자본 축적과 총생산에 미치는 효과를 시뮬레이션한다.⁵⁾

본 연구의 차별성은 세 가지이다. 첫째, 분석 대상 측면에서 AI 학습데이터라는 특정 자산에 대한 지식재산권 보호의 경제적 효과를 직접 다루며, 혁신(지식자본 축적)과 경쟁구조(기업 수)를 모형 내에서 동시에 내생변수로 설정한다. 둘째, 분석 접근 측면에서 지식재산권 규제가 작동하는 경로를 데이터 조달 단계와 학습 단계로 분리하고, 이 구분을 2025년 이후 형성된 판례의 법리적 분화와 대응시켜 각 경로의 기여도를 정량적으로 분해한다. 셋째, 정책적 기여 측면에서 법제 분석의 결과를 모형의 정책변수로 직접 연결함으로써, 규제 설계의 구체적 선택지가 갖는 경제적 함의를 수치로 제시한다.

본 논문의 구성은 다음과 같다. 제2장에서는 선행연구를 검토하고 본 연구의 차별점을 설명한다. 제3장에서는 국내외 규제 및 분쟁 사례를 검토하고 이를 모형의 정책변수로 전환하는 논리를 제시한다. 제4장에서는 분석모형을 설정하고, 제5장에서는 정량분석 결과를 제시한다. 제6장에서는 결론과 정책적 시사점을 논의한다.

2. 기존연구 검토

2.1. AI 학습데이터에 대한 지식재산권 적용

AI 학습 과정에서의 저작물 이용이 기존 저작권 법리로 포섭될 수 있는지에 관한 논의는 비교적 이른 시기부터 축적되어 왔다. Sag(2019)는 자신이 앞서 제시한 비표현적 이용 개념을 텍스트·데이터마이닝과 기계학습에 적용하여, 이러한 이용이 전통적 저작권 침해와는 구별되는 성격을 갖는다고 보았다.⁶⁾

Lemley와 Casey(2021)는 머신러닝이 저작물로부터 사실적 정보만을 추출한다는 점에 주목하여 기존 저작권 체계와의 부정합을 지적하고, 공정이용 법리가 비표현적 목적의 복제를 허용하는 방향으로 해석되어야 한다고 주장하였으며, Samuelson(2023)은 생성형 AI의 확산 이후 학습데이터 이용이 공정이용에 해당하는지 여부가 핵심 쟁점으로 부상하였음을 정리하였다.⁷⁾⁸⁾

이러한 연구들은 AI 학습데이터를 둘러싼 법적 쟁점의 구조를 밝히는 데 기여하였으나, 대체로 법리 해석과 제도 설계의 방향을 제시하는 데 초점을 두고 있어 보호 수준의 변화가 산업에 미치는 경제적 효과를 정량적으로 다루지는 않았다.

5) Robert E. Lucas, Jr., "On the Size Distribution of Business Firms", *Bell Journal of Economics*, Vol.9 No.2(1978), pp. 508-523.

6) Matthew Sag, "The New Legal Landscape for Text Mining and Machine Learning", *Journal of the Copyright Society of the U.S.A.*, Vol.66(2019), pp. 291-367.

7) Mark A. Lemley & Bryan Casey, "Fair Learning", *Texas Law Review*, Vol.99 No.4(2021), pp. 743-785.

8) Pamela Samuelson, "Generative AI Meets Copyright", *Science*, Vol.381 No.6654(2023), pp. 158-161.

2.2. 데이터 접근성과 디지털 시장 경쟁

데이터 접근성이 디지털 시장의 경쟁구조를 결정한다는 인식은 주요국 경쟁당국의 보고서를 통해 정책 영역에 자리 잡았다. Cr  mer et al.(2019)은 유럽위원회에 제출한 보고서에서 데이터에 대한 접근성이 디지털 시장 경쟁의 핵심 요소이며 데이터 독점이 시장지배력 강화로 이어질 수 있음을 지적하였고, Furman et al.(2019)은 영국 디지털 경쟁 전문가 패널 보고서에서 플랫폼 시장의 데이터 집중 문제를 분석하고 경쟁 촉진을 위한 데이터 접근성 확보를 제안하였다.⁹⁾¹⁰⁾

OECD(2019)는 지식재산권 보호가 기술 확산을 촉진하는 한편 라이선스 거절 등 권리 행사가 경쟁을 제약할 수 있음을 분석하였으며,¹¹⁾ OECD(2025)는 생성형 AI의 가치사슬 전반에서 양질의 데이터와 연산자원 등 핵심 투입요소에 대한 접근이 경쟁의 관건이 됨을 지적하였다.¹²⁾

이들 연구는 데이터 접근성과 경쟁구조의 연관성을 제도적·정책적 관점에서 규명하였으나, 기업 분포와 시장구조의 내생적 변화를 구조적으로 분석하는 데까지는 나아가지 못하였다.

2.3. 지식재산권 보호와 혁신 유인

지식재산권 보호가 혁신 투자를 촉진한다는 명제는 실증적으로 상당한 지지를 받아 왔다. Chen과 Wang(2025)은 중국 특허법 개정을 준자연실험으로 활용하여 지식재산권 보호 강화가 기업의 연구개발 투자와 연구인력 확충을 통해 디지털 기술 혁신 성과를 유의하게 증가시킴을 보였다.¹³⁾

다만 이러한 연구는 보호 강화가 혁신에 미치는 효과를 규명하는 데 집중하여, 기업 간 격차 확대와 시장집중이라는 경쟁구조상의 결과를 함께 고려하지는 않았다.

한편 지식재산권 보호와 경쟁의 관계를 이론적으로 다룬 연구들은 그 효과가 단선적이지 않음을 지적한다. Acemoglu와 Akcigit(2012)는 기술수준이 이질적인 기업들이 존재하는 성장 모형에서, 모든 기업에 동일한 보호를 부여하는 정책보다 기업 간 기술격차에 따라 보호 수준을 차등화하는 정책이 후생 측면에서 우월할 수 있음을 보였다.¹⁴⁾

이는 지식재산권 보호의 효과가 기업 분포에 의존하며, 따라서 기업 이질성을 명시적으로 다루지 않는 분석은 불완전할 수 있음을 시사한다.

2.4. 디지털·AI 시장에서의 혁신과 경쟁

AI 투자의 기업 수준 효과에 관한 실증연구로 Babina et al.(2024)은 미국 기업의 인적자원 자료를 활용하여 AI 투자 기업이 매출·고용·기업가치 측면에서 높은 성장을 보이며 이러한 성장

9) Jacques Cr  mer et al., "Competition Policy for the Digital Era", European Commission, 2019, pp. 1-127.

10) Jason Furman et al., "Unlocking Digital Competition: Report of the Digital Competition Expert Panel", HM Treasury, 2019, pp. 1-140.

11) OECD, "Licensing of IP Rights and Competition Law", OECD Competition Committee, DAF/COMP(2019)3, 2019, pp. 1-45.

12) OECD, "Artificial Intelligence, Data and Competition", OECD Artificial Intelligence Papers, 2024, pp. 1-62; OECD, "Artificial Intelligence and Competitive Dynamics in Downstream Markets", OECD Competition Policy Roundtable Background Note No. 331, 2025, pp. 1-52.

13) Lei Chen & Jieshuang Wang, "Intellectual Property Protection, R&D Investment and Digital Technology Innovation: An Empirical Study Based on the Revision of the Patent Law", *International Review of Economics & Finance*, Vol.103(2025), Article No. 104320.

14) Daron Acemoglu & Ufuk Akcigit, "Intellectual Property Rights Policy, Competition and Innovation", *Journal of the European Economic Association*, Vol.10 No.1(2012), pp. 1-42.

이 주로 제품 혁신을 통해 실현됨을 밝혔다. 이 연구에서 특히 주목할 점은, AI를 매개로 한 성장이 대규모 기업에 집중되어 산업집중도 상승과 연관된다는 결과이다.¹⁵⁾

이는 AI 기술의 확산이 그 자체로 경쟁을 촉진하기보다 기존 격차를 확대하는 방향으로 작용할 수 있음을 시사한다. Schrepel과 Groza(2024)는 디지털 시장에서 혁신을 둘러싼 경쟁을 경쟁법 분석에 반영하기 위한 방법론을 논의하였으며, Corrigan(2025)은 기존 기술기업이 AI 공급망의 핵심 관문인 연산, 데이터, 파운데이션 모델, 유통 채널 등에 대한 통제력으로 경쟁을 저해할 수 있다는 점을 지적하고, 개방적이고 경쟁적인 AI 산업을 유지하기 위한 정책 방향을 제시하였다.¹⁶⁾¹⁷⁾

국내 연구로는 김재현·정상철(2021)이 저작권 보호 활동의 경제적 파급효과를 동태일반균형 모형을 통해 분석하여, 저작권 보호 강화가 지식자본 축적을 촉진하고 이를 통해 산업 생산과 경제 전반의 성장에 긍정적 영향을 미침을 제시하였다.¹⁸⁾ 해당 연구는 지식재산권 보호의 거시 경제적 효과를 일반균형 관점에서 다루었다는 점에서 본 연구와 방법론적 기반을 공유하나, 기업 간 이질성과 시장 경쟁구조의 변화는 분석 대상에 포함하지 않았다. 다만 해당 연구가 다루는 지식자본은 저작권 보호의 대상이 되는 창작물의 축적인 반면, 본 연구의 지식자본은 보호 대상인 학습데이터를 투입하여 형성되는 AI 자본이라는 점에서 분석 대상이 구별된다.

2.5. 선행연구의 한계와 본 연구의 차별성

선행연구의 한계는 세 가지로 정리된다. 첫째, 분석 대상의 한계이다. AI 학습데이터에 대한 지식재산권 보호의 경제적 효과를 직접 분석한 연구는 제한적이며, 법학 문헌은 정성적 논의에, 경제학 문헌은 특허 보호를 중심으로 한 분석에 머물러 있다. 둘째, 분석 범위의 한계이다. 기존 연구는 지식재산권 보호와 혁신의 관계 또는 데이터 접근성과 경쟁의 관계를 각각 다루었으나, 두 관계를 하나의 분석틀 안에서 통합적으로 고려한 연구는 드물다. 셋째, 분석 방법론의 한계이다. 기업 간 생산성 이질성과 진입·퇴출을 동시에 고려한 구조적 분석이 부족하며, 다수의 연구가 정태적 분석 또는 사례 중심 접근에 머물러 있다.

본 연구는 이러한 한계에 대응하여, AI 학습데이터에 대한 지식재산권 보호를 정책변수로 설정하고 혁신과 경쟁구조를 내생변수로 하는 거시 동태일반균형모형을 구축한다. 특히 규제가 작동하는 경로를 데이터 조달 단계와 학습 단계로 분리하여 각각을 별개의 정책변수로 대응시킴으로써, 두 경로의 상대적 기여도와 규제 설계의 구체적 선택이 갖는 경제적 함의를 정량적으로 제시하고자 한다.

3. AI 학습데이터 관련 규제 사례 검토

3.1. 국내 법제의 이중 규범 구조

AI 학습데이터를 둘러싼 국내 규제 환경은 「저작권법」과 「독점규제 및 공정거래에 관한 법

15) Tania Babina et al., "Artificial Intelligence, Firm Growth, and Product Innovation", *Journal of Financial Economics*, Vol.151(2024), Article No. 103745.

16) Thibault Schrepel & Teodora Groza, "Computing 'Innovation Competition'", Amsterdam Law & Technology Institute, 2024, pp. 1-16.

17) Jack Corrigan, "Promoting AI Innovation Through Competition: A Guide to Managing Market Power", Center for Security and Emerging Technology, 2025, pp. 1-51.

18) 김재현·정상철, "저작권 보호 활동에 따른 경제적 파급효과 분석", 「문화경제연구」, 제24권 제3호(2021), 3-22면.

를」(이하 ‘공정거래법’)이 동시에 적용되는 이중 규범 구조를 형성한다. 저작권법에서 관련되는 규정은 두 가지이다. 제35조의5는 저작물의 통상적인 이용 방법과 충돌하지 아니하고 저작자의 정당한 이익을 부당하게 해치지 아니하는 경우에 저작물을 이용할 수 있도록 하는 포괄적 공정이용 조항으로, 이용의 목적 및 성격과 저작물의 현재 또는 잠재적 시장에 미치는 영향 등을 고려요소로 규정한다. 제93조는 데이터베이스의 제작 또는 그 소재의 갱신·검증·보충에 상당한 투자를 한 자에게 해당 데이터베이스의 상당한 부분을 복제·전송할 권리를 부여하며, 개별 소재의 복제라 하더라도 반복적이거나 특정한 목적을 위하여 체계적으로 이루어져 데이터베이스제작자의 이익을 부당하게 해치는 경우 상당한 부분의 복제로 본다.¹⁹⁾

문제는 현행 저작권법에 AI 학습을 위한 텍스트·데이터마이닝 면책 규정이 존재하지 않는다는 점이다. 면책 조항을 포함한 개정안이 반복적으로 발의되었으나 입법에 이르지 못하였고, 2026년 1월 시행된 「인공지능 발전과 신뢰 기반 조성 등에 관한 기본법」은 학습데이터 요약 공개를 포함한 투명성 확보 의무를 규정하였을 뿐 학습 행위의 저작권법적 적법성은 다루지 않았다. 결과적으로 국내 AI 기업은 개별 저작물에 대한 사전 허락을 확보하거나 사후적 공정이용 판단에 의존하는 두 가지 선택지 사이에 놓여 있으며, 제35조의5의 적용에 관한 국내 판례는 아직 확립되어 있지 않다.²⁰⁾

공정거래법은 다른 방향으로 작용한다. 제5조 제1항은 시장지배적사업자의 남용행위를 금지하면서 제3호에서 다른 사업자의 사업활동을 부당하게 방해하는 행위를, 제5호에서 부당하게 경쟁사업자를 배제하기 위한 행위를 규정한다. 시장지배적 사업자가 경쟁 기업에 학습데이터 접근을 거절하거나 차별적 조건을 부과하는 경우 이들 조항의 적용이 검토될 수 있으나, 시장지배적 지위 인정을 위한 관련시장 획정이 선결 과제로 남는다. 불공정거래행위를 금지하는 제45조 제1항과 관련하여서는 공정거래위원회의 「지식재산권의 부당한 행사에 대한 심사지침」이 지식재산권의 행사가 외형상 정당해 보이더라도 그 실질이 지식재산 제도의 본질적 목적에 반하는 경우 공정거래법의 적용 대상이 될 수 있음을 명시하고 있다. 정당한 이유 없는 라이선스 제공 거절, 자사 계열사와 경쟁사 간 라이선스 조건의 차별적 설정, 데이터 접근 허락의 조건으로 자사 연산 인프라 이용을 강제하는 행위 등이 AI 학습데이터 맥락에서 적용 가능성이 높은 유형이다. 이처럼 저작권법이 데이터 생산자의 권리 보호를 위하여 접근과 이용을 제한하는 반면 공정거래법은 과도한 권리 행사를 제어한다는 점에서, 두 규범은 서로 다른 방향에서 AI 학습데이터의 이용 조건을 규율한다.²¹⁾²²⁾

3.2. 주요국 TDM 규제의 비교

AI 학습데이터에 대한 규제 강도는 국가별로 상당한 차이를 보이며, 그 차이는 주로 TDM 면책의 범위와 조건에 의해 결정된다. 주요국의 규제 체계를 데이터 접근 허용 수준에 따라 정리하면 다음과 같다.

일본은 저작권법 제30조의4를 통해 저작물에 표현된 사상 또는 감정을 스스로 향유하거나 타인에게 향유하게 하는 것을 목적으로 하지 않는 경우 필요한 한도에서 저작물을 이용할 수 있도록 규정하고 있다. 이 규정은 정보해석 목적의 이용을 상업적 목적을 포함하여 폭넓게 허용한다는 점에서 주요국 중 가장 넓은 면책 범위를 갖는 것으로 평가된다.²³⁾

19) 저작권법 제35조의5 제1항 및 제2항, 제93조 제1항 및 제2항.

20) 「인공지능 발전과 신뢰 기반 조성 등에 관한 기본법」 제31조. 동법은 2026년 1월 시행되었다.

21) 독점규제 및 공정거래에 관한 법률 제5조 제1항 제3호, 제5호.

22) 공정거래위원회, “지식재산권의 부당한 행사에 대한 심사지침”, II. 일반적 심사원칙.

23) 日本 著作権法 第30條の4.

유럽연합은 디지털 단일시장 저작권 지침 제3조를 통해 연구기관 등이 수행하는 학술연구 목적의 TDM에 대하여 강행적 예외를 규정하는 한편, 제4조를 통해 그 밖의 목적에 대해서도 예외를 인정하되 권리자가 적절한 방식으로 이용을 유보한 경우에는 적용을 배제하도록 하였다. 상업적 이용에 대한 면책이 권리자의 유보 의사에 의해 좌우된다는 점에서 일본보다 제한적인 구조이다.²⁴⁾

미국은 TDM에 관한 개별 예외 규정을 두지 않고 연방저작권법 제107조의 공정이용 법리에 의존하나, 2025년 이후 일련의 판결을 통해 학습 목적 이용에 관한 사법적 기준이 형성되고 있다. 한국은 명시적 예외 규정이 부재하다는 점에서 미국과 형식적으로 유사하나, 공정이용 조항의 적용에 관한 판례가 축적되지 않았다는 점에서 법적 불확실성의 정도가 더 크다.

3.3. 주요 분쟁 사례

3.3.1. 해외 사례

AI 학습 목적의 저작물 이용에 관한 미국 법원의 판단은 2025년을 기점으로 본격적으로 축적되기 시작했다. 캘리포니아 북부 연방지방법원은 2025년 6월 이틀 간격으로 선고된 두 건의 약식판결에서 학습 행위 자체에 대해서는 공정이용을 인정하였다. Bartz v. Anthropic 사건에서 법원은 적법하게 취득한 도서를 언어모형 학습에 이용하는 행위가 고도로 변형적인 이용에 해당한다고 보면서도, 불법적으로 취득한 복제물로 영구적인 디지털 장서를 구축하고 보유하는 행위는 학습과 구별되는 별개의 이용으로서 공정이용의 보호를 받지 못한다고 판단하였다.²⁵⁾

Kadrey v. Meta Platforms 사건에서도 법원은 학습 행위의 변형성을 인정하였으나, 원고가 시장 침해를 입증하지 못한 점을 결론의 주된 근거로 삼음으로써 향후 입증이 이루어질 경우 결론이 달라질 수 있음을 시사하였다.²⁶⁾

이들 판결이 갖는 함의는 학습 행위와 데이터 취득 행위를 법적으로 분리하여 평가하였다는 점에 있다. 학습 자체는 변형적 이용으로 폭넓게 허용되더라도, 학습에 투입될 데이터를 어떤 경로로 확보하였는지는 별도의 책임 근거가 된다. 이는 AI 기업이 부담하는 법적 위험이 데이터 조달 단계와 학습 단계에서 서로 다른 성격으로 발생함을 의미한다.

3.3.2. 국내 사례

국내에서는 AI 학습데이터를 직접 대상으로 한 판결이 아직 축적되지 않았으나, 데이터 수집 및 플랫폼의 데이터 활용에 관한 선례가 참고가 된다.

서울중앙지방법원은 경쟁 사업자가 원고의 웹사이트와 모바일 애플리케이션을 통해 접근할 수 있는 부동산 매물정보를 무단으로 수집하여 자사 애플리케이션에 게재한 사건에서, 이러한 행위가 데이터베이스제작자의 권리를 침해한다고 보아 복제·전송의 금지와 보관 중인 정보의 삭제를 명하였다.²⁷⁾ 이 판결은 상당한 투자를 통해 구축된 데이터베이스의 무단 수집에 대한 제재 근거를 확인하였다는 점에서 AI 학습 목적의 대규모 크롤링에도 유추 적용될 여지가 있다.

경쟁법 영역에서는 네이버의 자사우대 사건이 중요한 전환점을 이루었다. 공정거래위원회는 2020년 네이버가 쇼핑 검색결과 노출 알고리즘을 자사 서비스에 유리하게 조정하여 경쟁 오픈

24) Directive (EU) 2019/790, Arts. 3-4.

25) Bartz v. Anthropic PBC, No. 3:24-cv-05417 (N.D. Cal. June 23, 2025). 본 사건은 2025년 9월 당사자 간 화해로 종결되었다.

26) Kadrey v. Meta Platforms, Inc., No. 3:23-cv-03417 (N.D. Cal. June 25, 2025).

27) 서울중앙지방법원 2023. 2. 3. 선고 2020가합584719 판결.

마켓 사업자를 배제하였다고 보아 시정명령과 과징금을 부과하였고, 서울고등법원은 2022년 이를 적법하다고 판단하였다. 그러나 대법원은 2025년 10월 원심을 파기환송하면서, 시장지배적 사업자라는 사정만으로 경쟁사업자를 자사와 동등하게 대우할 의무가 도출되는 것은 아니며, 남용행위가 성립하려면 문제가 된 시장에서 경쟁제한 효과가 발생할 구체적 우려가 있었다는 점이 입증되어야 한다고 판시하였다.²⁸⁾

이 판결은 데이터와 알고리즘을 매개로 한 자사우대 행위에 대하여 경쟁제한 효과의 실증적 입증을 요구하였다는 점에서, 향후 AI 학습데이터에 대한 접근 제한 행위의 경쟁제한성을 판단할 때도 정량적 근거의 중요성이 커질 것임을 시사한다.

3.4. 규제 강도의 정책변수화

이상의 검토로부터 AI 학습데이터에 대한 지식재산권 보호가 기업에 작용하는 경로를 두 가지로 구분할 수 있다. 이 구분은 분석의 편의를 위한 임의적 설정이 아니라, 앞서 살펴본 판례에서 실제로 관찰되는 법리적 분화를 반영한 것이다.

첫째는 데이터 조달 단계에 작용하는 경로이다. 지식재산권 보호가 강화되면 AI 기업이 학습 데이터를 확보하기 위하여 부담해야 하는 라이선스료, 권리처리 비용, 법률자문 비용이 상승한다. Bartz 판결이 학습 행위와 구별하여 취득 경로의 적법성을 별도로 판단한 것, 그리고 서울중앙지방법원이 크롤링에 의한 데이터 수집을 데이터베이스제작자의 권리 침해로 인정한 것은 모두 이 경로가 독립적인 비용 요인으로 작동함을 보여준다. 본 연구는 이 경로를 지식자본 조달비용에 영향을 주는 정책변수로 설정한다.

둘째는 학습 단계에 작용하는 경로이다. 확보한 데이터라 하더라도 이를 학습에 이용할 수 있는 범위는 규제에 의해 제약된다. 비상업적 목적에 한정된 면책, 데이터 제공계약상의 학습 금지 조항, 특정 모형의 학습을 제한하는 조건 등이 이에 해당한다. 인공지능기본법에 따른 학습 데이터 공개 의무 역시 학습 과정에 추가적인 준수 부담을 발생시킨다는 점에서 이 경로에 포함된다. 본 연구는 이 경로를 데이터 사용량을 제약함으로써 지식자본의 축적 효율에 영향을 주는 정책변수로 설정한다.

두 경로의 구분은 정책적으로도 의미를 갖는다. 조달 경로의 부담이 지배적이라면 라이선스 시장의 정비와 집중관리 체계의 구축이 우선적 과제가 되는 반면, 학습 경로의 제약이 지배적이라면 TDM 면책 규정의 도입이 보다 직접적인 해법이 된다. 제5장에서는 두 경로의 상대적 기여도를 분해하여 제시한다.

나아가 3.2에서 정리한 규제 강도 순서는 시나리오 설정의 근거가 된다. 일정한 규제 수준을 기준으로 하여, 상업적 목적을 포함한 광범위한 면책을 도입하는 경우와 권리자의 유보를 허용하는 제한적 면책을 도입하는 경우를 각각 정책변수로 표현함으로써, 제도 선택이 갖는 경제적 효과를 비교할 수 있다.

4. 분석모형

본 장에서는 제3장에서 도출한 두 가지 규제 작동 경로를 정책변수로 하는 거시 동태일반균형모형을 설정한다. 모형은 Lucas(1978)의 span-of-control 구조를 결합하여 기업가 능력의 이질성이 기업규모 분포와 진입 결정으로 연결되도록 구성되며, 이를 통해 AI 학습데이터에 대

28) 대법원 2025. 10. 16. 선고 2023두32709 판결(앞의 주4). 원처분은 공정거래위원회 의결 제2021-027호이다.

한 지식재산권 보호 수준의 변화가 지식자본 축적과 총생산에 미치는 효과를 하나의 균형체계 안에서 분석한다.²⁹⁾

4.1. 개요

경제주체는 대표적 가계와 기업들로 구성된다. 가계 구성원은 경영능력 z 를 부여받는데, 능력이 일정한 임계값을 상회하는 구성원은 기업을 운영하고 나머지는 노동을 공급하며, 이 임계값은 모형 내에서 내생적으로 결정된다. 이러한 구조에 의해 기업가의 이윤과 기업의 수는 내생적으로 결정된다. 이는 노동소득, 지식자본소득, 기업가 이윤의 분배율을 정합시키는 근거가 되며, 규제의 변화가 개별 기업의 요소구성과 기업의 진입 여부 중 어느 쪽을 통해 작동하는지를 구분하여 관찰할 수 있게 한다.

생산요소는 노동과 지식자본의 두 가지이다. 여기서 지식자본은 AI 학습의 기반이 되는 데이터 및 관련 무형자산을 포괄하는 개념으로, 가계가 축적하여 기업에 공급한다.

AI 학습데이터에 대한 지식재산권 보호 수준은 두 개의 외생적 정책변수로 표현된다. 앞서 구분한 바와 같이, 데이터 조달 단계에 작용하여 기업이 지식자본을 확보하는 데 소요되는 비용을 상승시키는 부담분과, 학습 단계에 작용하여 확보한 데이터를 학습에 활용할 수 있는 범위를 제약하는 부담분을 구분하여 표현한다.

본 모형에서는 학습에 활용될 수 있는 데이터 생산기업을 별도로 상정하지 않으며, 이에 따라 학습데이터의 총량은 외생적으로 고정된다. 이러한 설정을 택한 이유는 다음과 같다. 첫째, AI 학습에 투입되는 데이터의 상당 부분은 AI 학습을 목적으로 생산된 것이 아니라 이미 존재하는 저작물이다. 웹 문서, 뉴스 기사, 이미지 등은 각기 다른 목적으로 생산되어 축적된 것이므로, 학습데이터에 대한 보호 수준의 변화가 이들의 생산량을 직접 좌우한다고 보기 어렵다. 둘째, 데이터 생산 유인의 변화가 실제 축적으로 이어지는 데에는 상당한 시간이 소요되는 반면, 이용 조건의 변화는 즉시 작동한다. 따라서 본 연구의 결과는 이미 축적된 데이터의 이용 조건이 AI 자본 축적에 미치는 영향으로 한정하여 해석되어야 하며, 보호 수준의 변화가 데이터 생산에 미치는 장기적 효과와 권리자 후생을 포함한 종합적 평가는 본 연구의 분석 범위를 벗어난다.

4.2. 기업의 이윤극대화

경영능력 z 를 보유한 기업 i 의 생산함수는 다음과 같다.

$$y_{i,t} = z_{i,t}^{1-\alpha} (n_{i,t}^\gamma A_{i,t}^{1-\gamma})^\alpha \quad (1)$$

여기서 y 는 생산, z 는 기업가 능력, n 은 고용, A 는 지식자본이다. α 는 span-of-control 모수로서 기업가 능력에 대한 수확체감의 정도를 나타내며, γ 는 노동과 지식자본으로 구성된 복합요소 내에서 노동이 차지하는 비중이다. $\alpha < 1$ 이므로 기업규모는 기업가 능력에 의해 제약되며, 이것이 균형에서 기업규모 분포가 결정되는 근거가 된다. 복합요소를 Cobb-Douglas 형태로 설정함에 따라 노동과 지식자본의 대체탄력성은 1로 고정된다. 지식자본과 노동 간 대체탄력성에 관한 신뢰할 만한 추정치가 확보되어 있지 않은 상황에서 이를 자유 모수로 도입하는 것은 식별되지 않는 모수를 추가하는 결과가 되므로, 본 연구는 대체탄력성을 1로 제약한

29) Robert E. Lucas, Jr., 앞의 논문(주5), pp. 508-523.

Cobb-Douglas 설정을 유지하였다.

다만 이 설정은 결과의 해석에 다음과 같은 함의를 갖는다. 대체탄력성이 1보다 작은 경우 지식자본의 비용 상승에 대한 노동으로의 대체가 제한되어 총요소투입 자체의 감소폭이 커지므로, 총생산의 감소는 본 연구의 추정치보다 크게 나타날 것이다. 반대로 대체탄력성이 1보다 큰 경우에는 대체가 활발하여 총생산의 감소폭이 작아지는 대신 지식자본의 감소폭은 확대된다. 따라서 본 연구가 제시하는 지식자본 축적의 감소는 그 방향에 있어서는 대체탄력성의 설정과 무관하며, 총생산에 대한 효과의 크기는 이 설정에 의존한다.

기업의 이윤극대화 문제는 다음과 같이 표현된다.

$$\max \pi_{i,t} = y_{i,t} - w_t n_{i,t} - r_t (1 + \tau p) A_{i,t} \quad (2)$$

여기서 w 는 임금, r 은 지식자본 가격, τ 는 제약계수, p 는 지식자본 조달비용에 영향을 주는 지식재산권 보호 수준이다. 기업이 실제로 부담하는 지식자본의 단위비용은 시장가격 r 에 지식재산권 보호에 따른 추가 부담이 더해진 값이며, 이를 $R_t = r_t(1 + \tau p)$ 로 정의한다. 이 추가 부담은 라이선스 협상, 권리처리, 법률자문 등 데이터 조달 과정에서 소요되는 실질적 자원비용으로 해석되며, 경제 전체의 자원제약에서 차감된다.

복합요소를 $X_{i,t} = n_{i,t}^\gamma A_{i,t}^{1-\gamma}$ 로 정의하고 비용최소화를 수행하면 복합요소의 단위비용은 다음과 같다.

$$X_{i,t} = (w_t/\gamma)^\gamma (R_t/(1-\gamma))^{(1-\gamma)} \quad (3)$$

이를 이용해 이윤극대화 문제를 풀면 기업의 공급과 이윤은 다음과 같이 도출된다.

$$y_t^*(z_t) = z_t \Phi_t, \text{ where } \Phi_t = (\alpha/X_t)^{(\alpha/(1-\alpha))} \quad (4)$$

$$\pi_t(z_t) = (1-\alpha)z_t \Phi_t \quad (5)$$

즉 기업의 공급과 이윤은 모두 경영능력 z 에 대하여 선형이며, 비례상수 Φ 는 모든 기업에 공통인 요소가격의 함수이다.

요소수요는 각각 다음과 같이 공급에 비례한다.

$$n_t(z_t) = (\gamma\alpha/w_t) Y_t^*(z_t), \quad A_t(z_t) = ((1-\gamma)\alpha/R_t) Y_t^*(z_t) \quad (6)$$

따라서 산출 대비 노동소득의 비중은 $\gamma\alpha$, 지식자본소득의 비중은 $(1-\gamma)\alpha$, 기업가 이윤의 비중은 $1-\alpha$ 로 결정된다.

4.3. 가계의 효용극대화

가계는 다음의 무한기간 효용을 극대화한다.

$$\max \sum_t \beta^t [\ln c_t - L_t^\psi / (1 + \psi)] \quad (7)$$

여기서 β 는 시간할인을, ψ 는 노동공급의 Frisch 탄력성의 역수이다. 예산제약은 다음과 같다.

$$C_t + RD_t = w_t L_t + r_t A_t^s + \int \pi(z_t) g(z_t) dz_t \quad (8)$$

가계는 노동소득, 지식자본소득, 그리고 기업을 운영하는 구성원의 이윤을 수취하며, 이를 소비 C 와 지식자본에 대한 투자 RD 에 배분한다. 노동공급에 관한 최적조건은 다음과 같다.

$$L_t^\psi = w_t / C_t \quad (9)$$

4.4. 지식자본의 축적과 지식재산권 보호

지식자본은 다음의 동태적 축적식에 따라 형성된다.

$$A_{t+1} = \phi A_t + [D(q)]^\theta RD_t \quad (10)$$

여기서 ϕ 는 지식자본의 기간 간 잔존율로서 감가상각률 $\delta = 1 - \phi$ 와 대응하며, D 는 학습에 실제로 활용 가능한 데이터의 양, q 는 데이터 접근을 제한하는 지식재산권 보호 수준이다. θ 는 데이터 접근 탄력성 계수로서 $0 < \theta < 1$ 의 값을 가지며, 데이터 투입에 대한 수확체감을 반영한다.

$D(q)$ 의 구체적 함수형태는 다음과 같이 설정한다.

$$D(q) = 1 - q, \quad q \in [0, 1) \quad (11)$$

이는 지식재산권 보호 수준이 높아질수록 학습에 활용 가능한 데이터의 비율이 선형으로 감소함을 의미하며, $D(q) < 0$ 의 조건을 충족한다. 이때 D^θ 은 D 에 대하여 증가하고 오목하므로, 학습데이터 투입의 한계생산성이 체감한다는 성질이 유지된다.

식 (10)에서 D^θ 은 소비재 1단위를 투자하였을 때 형성되는 지식자본의 양을 의미한다. 따라서 q 의 상승은 지식자본 1단위를 형성하는 데 필요한 자원의 양, 즉 지식자본의 실효 공급가격을 상승시킨다. 이 점에서 q 는 p 와 구별되는 별도의 경로로 작동한다. p 는 기업이 지불하는 임대료를 높일 뿐 지식자본이 형성되는 과정 자체에는 영향을 주지 않는다. 반면 q 는 학습에 활용 가능한 데이터를 줄여 동일한 투자로 형성되는 지식자본의 양을 감소시킨다.

한편 본 모형은 두 정책변수가 비용으로 작동하는 방식을 서로 다르게 설정하고 있다. 데이터 조달비용 p 는 그 일부가 권리자에게 귀속되는 이전으로 처리될 수 있는 반면, 학습 단계 제약 q 는 대가 없이 이용 범위만을 제한하는 형태로 설정되었다. 현실에서는 학습 목적의 이용에 대해서도 권리자가 라이선스료를 취득할 수 있다. 이 때문에 본 모형은 학습 단계 제약의 효과를 실

제보다 크게 산출하게 된다. 규제를 강화할 때는 라이선스를 통해 이용할 여지를 반영하지 않아 손실이 과대평가되고, 완화할 때는 그 대가가 사라지는 부분을 계산하지 않아 편익이 과대평가된다. 따라서 해당 두 경로의 상대적 크기는 이러한 과대평가 가능성을 고려하여 해석할 필요가 있다.

가계의 지식자본 투자에 관한 오일러 방정식은 다음과 같이 도출된다.

$$1 = \beta(C_t/C_{t+1})[D(q)]^\theta r_{t+1} \phi \quad (12)$$

정상상태에서 소비가 일정하므로 지식자본의 균형가격은 다음의 폐쇄형으로 결정된다.

$$\tilde{r} = (1 - \beta\phi)/(\beta[\tilde{D}(q)]^\theta) \quad (13)$$

식 (13)은 본 모형의 해석에 있어 중요한 성질을 보여준다. 정상상태에서 지식자본의 가격은 시간할인율, 잔존율, 그리고 학습 단계의 규제 강도만으로 결정되며 다른 내생변수에 의존하지 않는다. 이는 지식자본의 공급이 r 에 대해 완전탄력적임을 의미하며, 그 결과 모형은 순차적으로 해를 구할 수 있는 구조를 갖는다. 즉 q 가 r 을 결정하고, r 과 τp 가 R 을 결정하며, R 이 주어진 상태에서 노동시장 균형이 나머지 변수를 결정한다.

4.5. 직업선택의 문제와 시장균형

가계 구성원은 기업을 운영하여 얻는 이윤과 노동을 공급하여 얻는 소득을 비교하여 직업을 선택한다. 무차별조건은 다음과 같다.

$$\pi(z_t^*) = w_t, \quad z_t^* = w_t/[(1 - \alpha)\Phi_t] \quad (14)$$

z_t^* 는 진입 임계점으로, 이를 상회하는 능력을 보유한 구성원만이 기업을 운영한다. 경영능력의 확률분포는 최희갑(2006)의 연구를 준용해 파레토분포를 따르는 것으로 가정한다.

$$g(z_t) = \mu z_u^\mu z_t^{-(1+\mu)}, \quad z_t \geq z_u > 0, \quad \mu > 1 \quad (15)$$

여기서 μ 는 파레토분포 모수, z_u 는 경영능력의 최소값이다. 이로부터 균형에서의 기업가 비율(M)과 유효 경영능력의 총량(Z)이 다음과 같이 도출된다.

$$M_t = (z_u/z_t^*)^\mu, \quad Z_t = [\mu/(\mu - 1)]z_u^\mu (z_t^*)^{1-\mu} \quad (16)$$

기업의 산출이 z 에 선형이므로 총생산은 유효 경영능력의 총량에 비례한다.

$$Y_t = \Phi_t Z_t \quad (17)$$

시장균형 조건은 알라스 법칙에 의해 생산물시장 조건을 제외한 노동시장과 지식자본시장의 수요·공급 조건으로 구성된다. 식 (6)을 이용하면 두 조건은 다음과 같이 정리된다.

$$L_t = \gamma \alpha (Y_t / w_t) \quad (\text{노동시장}) \quad (18)$$

$$A_t^s = (1 - \gamma) \alpha (Y_t / R_t) \quad (\text{지식자본시장}) \quad (19)$$

경제 전체의 자원제약은 다음과 같다.

$$C_t + RD_t = Y_t - \tau p r_t A_t^s \quad (20)$$

우변의 마지막 항은 데이터 조달 과정에서 소요되는 권리처리 비용으로, 소비나 투자로 전환되지 않고 소멸하는 자원을 의미한다.

4.6. 균형의 성질과 비교정확

본 모형의 정상상태는 해석적으로 도출할 수 있으므로, 정책변수의 효과에 관한 두 가지 성질을 수치해에 앞서 확인할 수 있다.

첫째, 식 (19)로부터 정상상태의 지식자본-산출 비율을 도출하면 다음 식을 만족한다.

$$A / Y = (1 - \gamma) \alpha / R \quad (21)$$

이 관계와 식 (13)을 결합하면, $\tau p \geq 0$ 이므로 지식자본-산출 비율의 상한이 다음과 같이 주어진다.

$$A / Y \leq (1 - \gamma) \alpha \cdot \beta \cdot [D(q)] \theta / (1 - \beta \phi) \quad (22)$$

식 (22)는 소득분배율과 감가상각률이 주어지면 이 모형이 생성할 수 있는 지식자본-산출 비율에 상한이 존재함을 보여준다. 이 상한은 보정 과정에서 목표값의 실현 가능성을 판단하는 기준이 된다.

둘째, 식 (14)와 (16), (17)을 결합하면 총생산은 요소가격의 함수로 다음과 같이 표현된다.

$$Y_t = k \Phi_t^\mu w_t^{1-\mu}, \text{ where } k = [\mu / (\mu - 1)] z_\mu'' (1 - \alpha)^{\mu-1} \quad (23)$$

식 (3)과 (4)로부터 Φ 의 로그값이 요소가격에 대하여 갖는 탄력성은 각각 다음과 같다.

$$\partial \ln \Phi / \partial \ln R = -(1 - \gamma) \alpha / (1 - \alpha), \quad \partial \ln \Phi / \partial \ln w = -\gamma \alpha / (1 - \alpha) \quad (24)$$

이로부터 지식재산권 보호 강화의 효과에 관한 두 가지 함의가 도출된다.

먼저 지식자본의 이용에 관한 함의이다. p 또는 q 의 상승은 모두 R 을 상승시키며, 이는 Φ 를 하락시켜 총생산을 감소시킨다. 나아가 식 (21)에 의해 지식자본-산출 비율 자체도 하락하므로, 지식자본 스톡은 총생산보다 더 큰 폭으로 감소한다. 요소구성의 측면에서 보면 R 의 상승은 기업이 지식자본을 노동으로 대체하도록 유인하며, 그 결과 AI 학습기반 자본의 이용 강도가 하락한다.

다음은 기업 진입에 관한 함의이다. 식 (14)에 식 (24)를 적용하면 진입 임계점의 요소 가격탄력성은 다음과 같다.

$$\partial \ln z^* / \partial \ln R = (1-\gamma)\alpha / (1-\alpha), \quad \partial \ln z^* / \partial \ln w = 1 + \gamma\alpha / (1-\alpha) \quad (25)$$

R 상승 시 w 가 감소하는데, R 의 상승은 진입 임계점을 높이는 방향으로, w 의 하락은 진입 임계점을 낮추는 방향으로 작용하여 두 효과가 서로 반대 방향으로 작용한다. 식 (25)에서 임금에 대한 탄력성이 R 에 대한 탄력성보다 크므로 임금의 반응 크기가 순효과를 좌우한다. 다만 본 모형에서는 균형 노동투입이 정책변수에 거의 의존하지 않으므로 임금의 변화가 제한적이며, 그 결과 기업의 진입 임계점(z^*)의 변화가 매우 작게 나타난다.

4.7. 모수의 보정(Calibration)

모형의 모수는 표준적 거시경제 문헌값, 국민계정 자료에 대한 정합조건, 그리고 선행연구의 실증치를 결합하여 설정하였다.

시간할인을 β 와 노동공급의 Frisch 탄력성의 역수 ψ 는 거시 일반균형모형의 통상적 문헌값을 원용하여 각각 0.96과 1.25를 적용하였다.³⁰⁾³¹⁾

파레토분포 모수 μ 와 경영능력 최소값 z_u 는 국내 기업규모 분포에 관한 선행 추정치를 준용하여 각각 1.004와 0.005를 적용하였다. 최희갑(2006)은 한국 기업의 규모 분포가 Zipf 법칙에 근사한다는 점을 실증적으로 제시하였는데, Zipf 법칙은 형상모수가 1에 근접한 파레토분포에 해당하므로 $\mu = 1.004$ 는 이러한 실증 결과에 부합하는 값이다. 다만 μ 가 1에 근접한다는 것은 분포의 우측 꼬리가 매우 두터움을 의미하고 총계변수의 반응이 이에 민감할 수 있으므로, 제5장에서 μ 에 대한 민감도 분석을 별도로 수행한다.³²⁾

span-of-control 모수 α 와 노동비중 모수 γ 는 국민계정상 소득분배 구조에 정합하도록 설정하였다. 식 (6)에서 도출된 바와 같이 본 모형에서 산출 대비 노동소득의 비중은 $\gamma\alpha$, 지식자본 소득의 비중은 $(1-\gamma)\alpha$ 로 결정된다. 2024년 기준 피용자보수 63.9%와 재산 및 기타소득 18.6%를 목표값으로 하면 $\alpha = 0.825$, $\gamma = 0.775$ 가 도출되며, 이때 기업가 이윤의 비중은 $1-\alpha = 0.175$ 로서 잔여 항목과 정합한다.³³⁾

30) Thomas F. Cooley & Edward C. Prescott, "Economic Growth and Business Cycles", Edit. by Thomas F. Cooley, Frontiers of Business Cycle Research, Princeton University Press, 1995, pp. 1-38.

31) Raj Chetty et al., "Are Micro and Macro Labor Supply Elasticities Consistent? A Review of Evidence on the Intensive and Extensive Margins", *American Economic Review*, Vol.101 No.3(2011), pp. 471-475.

32) 최희갑, "Zipf 분포에 기초한 한국의 기업 규모 분포 분석", 『통계연구』, 제11권 제2호(2006), 73-95면.

33) 2024년 기준 피용자보수 63.9%, 재산 및 기타소득 18.6%(한국은행, "국민계정", 한국은행 경제통계시스템, <<https://ecos.bok.or.kr/#/SearchStat>>, 검색일: 2026. 7. 30.).

지식자본의 잔존율 ϕ 는 지식재산의 감가상각에 관한 실증치를 원용하여 0.9로 설정하였다. 이는 연간 감가상각률 10%에 대응한다. 데이터 접근 탄력성 계수 θ 는 AI 투자와 생산성의 관계에 관한 추정치의 범위를 준거로 0.45를 적용하였다.³⁴⁾

정책변수 p 와 q 는 현행 규제 수준을 나타내는 기준값으로 각각 0.5로 정규화하였다. 제5장의 시나리오 분석은 이 기준값으로부터의 상대적 변화를 대상으로 한다.

제약계수 τ 의 경우 식별상의 유의점이 있다. 식 (2)에서 τ 와 p 는 오직 τp 의 형태로만 모형에 진입하므로 정상상태에서 두 변수는 개별적으로 식별되지 않는다. 따라서 τ 의 값은 별도의 자료에 의해 결정되는 것이 아니라, 두 정책변수를 대칭적으로 설정한다는 원칙에 의해 부여된다. 기준값 $q = 0.5$ 는 식 (13)을 통해 지식자본의 균형가격을 1.366배로 상승시킨다. 조달 경로가 이와 동일한 크기의 비용 상승을 발생시키도록 $1 + \tau p = 1.366$ 을 요구하면 $p = 0.5$ 에서 $\tau = 0.732$ 가 도출되며, 본 연구는 이를 반올림하여 $\tau = 0.75$ 를 적용하였다. 이때 $1 + \tau p = 1.375$ 로 학습 경로의 1.366과 거의 일치하므로, 두 경로는 기준 시나리오에서 지식자본의 실효비용에 대하여 대칭적으로 기여한다.

이 설정은 $\tau p = 0.375$ 에 해당하며, 이는 권리처리 비용이 지식자본 조달비용의 27.3%를 차지하는 수준으로 해석된다. 다만 이 비중에 직접 대응하는 통계가 확보되지 않았으므로, τ 를 민감도 분석의 대상에 포함하여 결과가 이 설정에 어떻게 의존하는지를 함께 제시한다. 모형의 보정 결과는 <표1>과 같다.

<표1 모형의 보정 결과>

모수 정의	기호	값	근거
시간할인율	β	0.96	거시경제문헌 표준값
노동공급 Frisch 탄력성의 역수	ψ	1.25	거시경제문헌 표준값
파레토분포 모수	μ	1.004	기업규모 분포 실증치
경영능력 최소값	z_u	0.005	기업규모 분포 실증치
span-of-control 모수	α	0.825	소득분배율 정합
복합요소 내 노동비중	γ	0.775	소득분배율 정합
지식자본 잔존율	ϕ	0.9	지식재산 감가상각 실증치
데이터 접근 탄력성 계수	θ	0.45	AI 투자-생산성 추정치
IP 보호수준(데이터 조달비용)	p	0.5	현행 규제수준 기준값 정규화
IP 보호수준(학습단계 제약)	q	0.5	현행 규제수준 기준값 정규화
제약계수	τ	0.75	두 경로의 대칭적 설정
항목	목표값	모형 산출값	근거
노동소득 비중	0.639	0.639	2024년 국민계정
지식자본소득 비중	0.186	0.186	2024년 국민계정

4.8. 분석 절차

모형의 수치해는 다음의 절차에 따라 산출한다.

34) Bronwyn H. Hall et al., "Measuring the Returns to R&D", Edit. by Bronwyn H. Hall & Nathan Rosenberg, Handbook of the Economics of Innovation, Vol.2, North-Holland, 2010, pp. 1033-1082; Jordan Hoffmann et al., "Training Compute-Optimal Large Language Models", Advances in Neural Information Processing Systems 35 (NeurIPS 2022), 2022, p. 30021; Tamay Besiroglu et al., "Chinchilla Scaling: A replication attempt", arXiv, <<https://arxiv.org/abs/2404.10102>>, 작성일: 2024. 4. 15.

1단계에서는 정책변수와 모수값을 입력하고 식 (13)을 계산한다. 2단계에서는 R이 주어진 상태에서 임금 w 를 미지수로 하는 단일 방정식을 구성한다. 구체적으로 식 (3), (4), (14), (16), (17)을 순차적으로 적용하여 Y와 M을 w 의 함수로 표현하고, 식 (9)의 노동공급과 식 (18)의 노동수요가 일치하는 조건을 이용하여 w 를 결정한다. 이 과정에서 소비 C는 식 (20)의 자원제약과 정상상태 조건으로부터 계산된다. 3단계에서는 결정된 w 를 이용하여 나머지 내생변수의 정상상태 값을 산출한다.

이행경로는 정상상태 해를 초기값 및 종기 조건으로 설정하고, 식 (10)의 축적식과 식 (12)의 오일러 방정식을 포함한 비선형 동태 연립방정식을 구성하여 산출한다. 정책변수의 변화가 영구적 충격으로 나타나는 경우를 상정하며, 충분히 긴 기간 이후 새로운 정상상태에 수렴하도록 종기 조건을 부과한다.

5. 분석결과

5.1. 보호수준 변화의 장기효과

먼저 두 정책변수를 동일한 수준으로 놓고, 비율을 변화시켰을 때의 정상상태를 비교한다. 결과는 <표2>와 같다.

<표2 학습데이터 규제의 경제적 효과 분석 결과>

보호수준(s)	지식자본가격(r)	총산출(Y)	지식자본스톡(A)	A/Y	임금(w)	노동(L)
0.00	0.1417	0.9974 (+14.8%)	1.3069 (+115.6%)	1.3103 (+87.8%)	0.6786 (+7.4%)	0.8763 (-0.29%)
0.20	0.1566	0.9470 (+9.0%)	0.9759 (+61.0%)	1.0305 (+47.7%)	0.6516 (+3.1%)	0.8776 (-0.14%)
0.50(기준)	0.1935	0.8691	0.6063	0.6976	0.6320	0.8788
0.80	0.2923	0.7661 (-11.9%)	0.3041 (-49.9%)	0.3969 (-43.1%)	0.5563 (-12.0%)	0.8813 (+0.29%)
0.90	0.3993	0.7065 (-18.7%)	0.1961 (-67.7%)	0.2776 (-60.2%)	0.5121 (-19.0%)	0.8817 (+0.33%)

첫째, 지식재산권 보호 수준의 상승은 지식자본 스톡과 총생산을 단조적으로 감소시킨다. 기준값 $s=0.5$ 에서 $s=0.8$ 로 보호를 강화하면 지식자본 스톡은 49.85% 감소하고 총생산은 11.86% 감소한다. 반대로 $s=0.2$ 로 완화하면 지식자본 스톡은 60.96%, 총생산은 8.96% 증가한다.

둘째, 지식자본의 반응이 총생산의 반응보다 현저히 크다. $s=0.5$ 에서 $s=0.8$ 구간에서 두 변화율의 비는 약 4.2배이며, 이는 식 (21)에서 확인한 구조에 기인한다. 보호 강화는 총생산을 직접 감소시킬 뿐 아니라 지식자본-산출 비율 자체를 하락시키므로, 두 효과가 중첩되어 지식자본 스톡에 나타난다. 지식자본-산출 비율은 동일 구간에서 0.6976에서 0.3969로 43.10% 하락한다.

셋째, 보호수준을 0.5에서 0.8로 강화하면 노동 1단위당 지식자본은 49.99% 하락하는 반면 총노동투입은 0.29% 증가하는 데 그친다. 즉 지식자본의 감소가 노동의 증가로 상쇄되는 것이 아니라, 지식자본이 축소되면서 총산출도 11.86% 함께 감소한다. 규제 강화의 효과는 요소 간 대체보다 지식자본 이용 자체의 축소를 통해 실현된다.

5.2. 두 경로의 분해

앞서 구분한 두 규제 경로의 상대적 기여도를 확인하기 위하여, 기준값 0.5에서 각 정책변수를 개별적으로 0.6으로 상승시킨 경우와 동시에 상승시킨 경우를 비교하였다. 결과는 <표3>과 같다.

<표3 두 규제 경로의 분해 결과>

구분	지식자본 가격(r)	총산출(Y)	지식자본 스톡(A)	A/Y
p만 0.5 → 0.6	0.1935	-1.12%	-6.23%	-5.17%
q만 0.5 → 0.6	0.2140	-2.26%	-11.60%	-9.55%
개별 효과의 단순 합	—	-3.38%	-17.84%	-14.72%
q/p 변화율의 상대 강도	—	2.02배	1.86배	1.85배
두 변수 동시 상승	0.2140	-3.36%	-17.11%	-14.23%

학습 단계 제약(q) 경로가 데이터 조달비용(p) 경로보다 효과가 크게 나타났다. 동일한 변화에 대하여 지식자본 스톡의 감소폭은 q 경로가 11.60%, p 경로가 6.23%로 q 경로가 1.86배 크며, 총생산에서는 그 비가 2.02배로 더 커진다.

이러한 차이는 p가 지식자본의 시장가격 위에 부담을 더하여 실효비용 R을 $(1+\tau p)$ 배 높이는 데 그치는 반면, q는 식 (13)을 통해 지식자본의 균형가격 자체를 상승시키고 동시에 투자 1단위가 형성하는 지식자본의 양을 축소시키기 때문에 발생한다.

두 경로가 기준 설정에서 실효비용에 대칭적으로 기여함에도 효과의 크기가 차이를 보이는 것은, 수준에서의 기여와 한계에서의 반응이 구별되기 때문이다. 실효비용에 로그를 취하면 다음과 같이 정리된다.

$$\ln R = \ln[(1 - \beta\phi)/\beta] - \theta \ln(1 - q) + \ln(1 + \tau p) \quad (26)$$

여기서 q의 한계 효과는 $\theta/(1-q)$ 로 q가 높아질수록 커지는 반면, p의 한계 효과는 $\tau/(1+\tau p)$ 로 p가 높아질수록 작아진다. 기준값에서 두 값은 각각 0.900과 0.546으로 그 비가 1.65이다. 본 절에서 관찰된 1.86배는 이러한 한계 효과의 차이가 유한한 크기의 변화에 대하여 나타난 결과이다.

두 경로를 개별적으로 적용한 효과의 단순 합은 지식자본 스톡 기준 17.84%이며 동시 적용 시에는 17.11%로 나타나 동시 작용 시 단순 합보다 크기가 작은 것으로 분석되었다. 이는 두 정책변수가 모두 지식자본의 실효비용 R을 통해 작용하고, 지식자본-산출 비율이 R에 반비례하여 R이 높아질수록 추가 상승의 효과가 제감하기 때문이다.

5.3. 규제 완화 수단에 따른 효과

5.2에서 보호수준을 강화하는 방향의 효과를 다루었다면, 본 절에서는 완화에 따른 효과를 살펴본다. 지식재산권 규제를 완화하는 정책 수단은 작용하는 단계에 따라 두 가지로 구분된다. 하나는 학습 단계의 이용을 허용하는 면책 규정을 도입하는 것이고, 다른 하나는 데이터 조달 단계의 비용을 낮추는 것이다. 전자는 학습 단계 제약 q의 하락에, 후자는 데이터 조달비용 p의 하락에 대응한다.

학습 단계 면책의 경우 앞서 확인한 바와 같이 그 범위에 차이가 존재한다. 권리자의 유보를 조건으로 하는 제한적 면책은 이용이 허용되는 범위가 좁고, 상업적 목적을 포함하는 광범위한 면책은 그 범위가 넓다. 본 절은 이러한 차이를 반영하여 예시적으로 q 를 0.35로 낮추는 경우와 0.20으로 낮추는 경우를 각각 살펴본다. 데이터 조달비용 경로에 대해서는 두 경로의 효과를 동일한 조건에서 비교하기 위하여 p 를 같은 쪽으로 낮춘 경우를 함께 제시한다.

결과는 <표4>와 같다.

<표4 규제 완화 수단별 효과>

변화 폭	r	R	지식자본(A)	총산출(Y)	A/Y
$p\ 0.5 \rightarrow 0.35$	0.1935	0.2443	+10.89%	+1.82%	+8.91%
$p\ 0.5 \rightarrow 0.20$	0.1935	0.2226	+24.14%	+3.83%	+19.57%
$q\ 0.5 \rightarrow 0.35$	0.1720	0.2365	+15.60%	+2.73%	+12.53%
$q\ 0.5 \rightarrow 0.20$	0.1566	0.2153	+29.66%	+4.94%	+23.55%

첫째, 두 경로 모두 완화 폭이 커질수록 효과가 확대된다. 지식자본 스톡을 기준으로 p 경로는 10.89%에서 24.14%로, q 경로는 15.60%에서 29.66%로 증가한다. 총산출의 증가폭은 각각 3.83%와 4.94%로 지식자본에 비해 작는데, 이는 앞에서 확인한 구조가 완화 방향에서도 동일하게 나타나기 때문이다. 지식자본의 실효비용이 하락하면 총산출이 증가할 뿐 아니라 지식자본-산출 비율 자체가 상승하므로, 두 효과가 중첩되어 지식자본 스톡에 나타난다.

둘째, 동일한 폭의 완화에 대해서는 학습 단계 면책의 효과가 학습데이터 조달비용 절감의 효과보다 크다. 0.15만큼 완화하는 경우 지식자본 스톡의 증가율은 q 경로가 p 경로의 1.43배이며, 0.30만큼 완화하는 경우에는 1.23배이다. 두 경로가 모두 지식자본의 실효비용 R 을 통해 작용하지만, q 는 식 (13)을 통해 지식자본의 균형가격 자체를 낮추는 반면 p 는 그 위에 부과되는 비용만을 줄이기 때문이다.

셋째, 다만 q 경로의 상대적 우위는 완화 폭이 커질수록 축소된다. 완화 폭이 0.15에서 0.30으로 확대되면 두 경로의 비는 1.43배에서 1.23배로 낮아지며, 앞에서 확인한 강화 방향의 1.86배와 비교하면 그 차이가 더욱 분명하다. 이는 두 정책변수가 지식자본의 실효비용에 전달되는 방식이 서로 다르기 때문이다. q 의 한계 효과는 $\theta/(1-q)$ 로 표현되어 q 가 낮아질수록 감소하는 반면, p 의 한계 효과는 $\tau/(1+\tau p)$ 로 표현되어 p 가 낮아질수록 증가한다. 따라서 규제를 강화하는 구간에서는 학습 단계 경로의 영향력이 상대적으로 커지고, 완화하는 구간에서는 두 경로의 격차가 좁혀진다.

이러한 결과는 정책 수단의 선택에 관하여 두 가지 함의를 갖는다. 먼저 현행 수준에서 규제를 완화하고자 하는 경우, 동일한 정도의 완화라면 학습 단계 면책의 도입이 조달비용 절감보다 큰 효과를 갖는다. 다음으로 완화의 폭이 커질수록 두 수단의 효과 차이는 줄어들므로, 상당한 폭의 완화를 상정하는 경우에는 어느 한 수단에 집중하기보다 두 수단을 병행하는 접근이 유효할 수 있다.

5.4. 이행경로 분석 결과

본 절에서는 학습 단계 제약(q)을 0.5에서 0.2로 완화하는 경우 해의 동태적 경로 변화가 어떻게 나타나는지 분석했다. 이행 경로 변화는 <표5>와 같이 나타났다.

<표5 학습단계 제약 완화(q:0.5→0.2)에 따른 이행 경로 변화>

시점	지식자본(A)	A실현율	총산출(Y)	Y실현율
1	0.6063	0.0%	0.8811	27.9%
2	0.6536	26.3%	0.8901	48.9%
3	0.6885	45.7%	0.8964	63.4%
4	0.7143	60.1%	0.9007	73.6%
9	0.7652	90.2%	0.9089	93.5%
16	0.7843	99.0%	0.9118	99.4%
장기	0.7861	100.0%	0.9121	100.0%

이행경로 분석 결과 조정은 비교적 신속하게 진행되는 것으로 나타났다. 지식자본 스톡은 4기에 장기효과의 60.1%, 9기에 90%, 16기에 99%를 달성한다. 보호를 강화하는 방향의 충격에서도 조정 속도는 동일하며, 두 정책변수를 함께 0.5에서 0.6으로 상승시키는 경우도 지식자본 스톡의 반감기는 마찬가지로 4기이다.

주목할 점은 총생산의 조정이 지식자본보다 앞선다는 것이다. 이는 요소가격이 즉각 반응하여 기업의 요소구성과 노동투입이 당기에 조정되는 반면, 지식자본은 시점 간 축적을 통해 변화하기 때문이다.

이러한 시차 구조는 정책 평가에 함의를 갖는다. 규제 변화의 효과를 단기 지표만으로 판단할 경우, 총생산 측면에서는 장기효과의 절반 가까이가 2기 이내에 나타나므로 과소평가의 위험이 크지 않으나, AI 자본 축적 측면에서는 2기 시점의 실현률이 26.3%에 그쳐 과소평가가 발생한다고 볼 여지가 있다.

5.5. 민감도 분석 결과

주요 모수에 대한 민감도 분석 결과는 다음과 같다. 민감도는 데이터 접근 탄력성 모수(θ), 파레토 분포 모수(μ), 그리고 권리처리비 이전 비율(λ)에 대해 실시하였다. 권리처리비 이전 비율(λ)은 식 (20)을 다음과 같이 바꾸었을 때 나타나는 모수이다.

$$C_t + RD_t = Y_t - (1 - \lambda)\tau pr_t A_t^s \quad (27)$$

민감도 분석 결과는 <표6>과 같다.

<표6 민감도 분석>

구분	값	지식자본(A)	총산출(Y)	비고
θ	0.30	+18.90%	+3.27%	q 0.5→0.2 효과
	0.45	+29.66%	+4.94%	기준값
	0.60	+41.38%	+6.64%	결과가 민감함
μ	1.004	-17.11%	-3.36%	기업가 비율 0.096%
	1.500	-17.11%	-3.35%	기업가 비율 8.022%
	2.500	-17.10%	-3.34%	기업가 비율 14.439%

구분	값	지식자본(A)	총산출(Y)	비고
λ	0.0	-16.79%	-3.17%	전액 자원 소모(기준)
	0.5	-17.11%	-3.54%	절반 이전
	1.0	-17.41%	-3.89%	전액 이전
τ	0.25	-13.92%	-2.71%	권리처리 비용 비중 11.1%
	0.75	-17.11%	-3.36%	기준값, 권리처리 비용 비중 27.3%
	1.50	-20.01%	-3.98%	권리처리 비용 비중 42.9%
	2.00	-21.28%	-4.26%	권리처리 비용 비중 50.0%

첫째, 데이터 접근 탄력성 계수의 경우 θ 를 0.30에서 0.60으로 변화시키면 q 를 0.5에서 0.2로 낮출 경우 지식자본 스톡의 증가는 18.90%에서 41.38%의 범위를 갖는다. θ 는 식 (13)에서 q 가 지식자본 가격에 전달되는 강도를 직접 결정하는 모수이므로 이러한 민감성은 구조적인 결과로 판단된다.

둘째, 파레토분포 모수 μ 에 따라서는 결과가 크게 달라지지 않는 것으로 나타났다. μ 를 1.004에서 2.5로 변화시키면 기업가 비율은 0.096%에서 14.439%로 늘어나지만, 동일한 정책 변화에 대한 지식자본 스톡의 반응은 -17.11%에서 -17.10%로, 총생산의 반응은 -3.36%에서 -3.34%로 거의 변하지 않는 것으로 나타났다. μ 의 변화는 기업가 비율과 평균 기업규모의 수준을 크게 변화시키나 정책효과의 크기에는 영향을 주지 않는다. 이는 기업규모 분포가 경제의 수준을, 정책변수가 요소가격을 통한 반응을 각각 결정하는 구조에서 비롯된다.

셋째, 권리처리 비용의 자원 소모 가정에 대해서도 결과는 강건한 것으로 분석되었다. 4.2에서 τp 에 해당하는 부담을 실질적 자원비용으로 처리하였으나, 이 중 일부가 권리자에게 귀속되는 순수한 이전에 해당할 수 있다. 이전 비율 λ 를 0에서 1까지 변화시켜도 p 를 0.5에서 0.8로 강화하는 효과가 지식자본 스톡 기준 -16.79%에서 -17.41%, 총생산 기준 -3.17%에서 -3.89%로 나타나 방향성이 동일하고 효과의 크기 차이가 크지 않은 것으로 분석되었다. 이는 τp 가 지식자본의 실효비용을 높여 요소구성을 왜곡하는 경로가 λ 와 무관하게 작동하기 때문이다.

넷째, 제약계수 τ 에 대해서는 정책효과의 크기가 완만하게 반응한다. τ 를 0.25에서 2.00으로 변화시키면 지식자본 스톡의 감소폭은 13.92%에서 21.28%의 범위를 갖는다. τ 가 8배 증가하는 동안 효과의 크기는 1.5배 확대되는 데 그친다. 다만 두 경로의 상대적 크기는 모수 설정에 의존하며, 특정 조건에서는 우열이 역전된다. 실효비용에 대한 한계 효과를 비교하면, 학습 단계 경로가 조달 단계 경로보다 큰 효과를 갖는 조건은 다음과 같다.

$$\theta(1 + \tau p) > \tau(1 - q) \quad (28)$$

기준 설정에서는 좌변이 0.619, 우변이 0.375로 이 조건이 성립한다. 그러나 θ 가 낮고 τ 가 높은 조합에서는 부등호가 역전된다. 주어진 θ 에 대하여 역전이 발생하는 임계값 τ^* 는 <표7>과 같다. 예컨대 $\theta = 0.30$, $\tau = 1.5$ 의 조합에서는 좌변이 0.525, 우변이 0.750으로 조달 단계 경로의 효과가 더 크게 나타난다. 현재의 자료로는 두 모수 조합 중 어느 것도 배제할 이론적 근거가 없으므로, 본 연구가 제시하는 두 경로의 상대적 크기는 기준 설정에서의 결과로 한정하여 해석되어야 한다.

<표7 학습 경로와 조달 경로의 우열이 역전되는 임계 τ >

θ	임계 τ^*	τ^* 에 해당하는 권리처리 비용 비중	비고
0.30	0.857	30.0%	$\tau > 0.857$ 이면 조달 경로가 우위
0.45	1.636	45.0%	기준값 $\tau = 0.75$ 에서는 학습 경로 우위
0.60	3.000	60.0%	$\tau > 3.000$ 이어야 역전

6. 결론 및 정책적 시사점

6.1. 연구 결과의 요약

본 연구는 AI 학습데이터에 대한 지식재산권 보호 수준이 AI 자본의 축적에 미치는 영향을 거시 동태일반균형모형을 통해 분석하였다.

규제 사례 검토를 통해 본 연구는 지식재산권 보호가 AI 기업에 작용하는 경로가 데이터 조달 단계와 학습 단계로 구분된다는 점을 확인하였다. 이러한 구분은 2025년 미국 법원의 일련의 판결이 학습 행위와 데이터 취득 행위를 법적으로 분리하여 평가한 것과 정합적이며, 규제 설계의 관점에서도 서로 다른 정책 수단에 대응한다.

정량분석 결과, 지식재산권 보호 수준을 기준값에서 20% 강화할 경우 지식자본 스톡은 17.11%, 총생산은 3.36% 감소하는 것으로 나타났다. 이 효과는 두 경로를 통해 실현되었으며, 각 경로의 기여도는 지식자본 스톡 기준으로 학습 단계 제약 경로가 11.60%p, 데이터 조달비용 경로가 6.23%p로 분해되었다. 다만 두 경로의 상대적 크기는 모수 설정에 따라 달라질 수 있다.

이 효과는 요소 간 대체보다 지식자본 이용 자체의 축소를 통해 실현되었다. 보호수준을 0.5에서 0.8로 강화할 경우 노동 1단위당 지식자본은 49.99% 하락하는 반면 총노동투입은 0.29% 증가하는 데 그치는 것으로 나타났다. 이행경로 분석 결과 지식자본 스톡의 조정 반감기는 3~4기, 90% 도달 시점은 9기로 나타났으며, 총생산은 정책 변화 당기에 이미 장기효과의 27.9%를 실현하여 지식자본보다 앞서 조정되는 것으로 분석되었다.

본 연구의 결과는 김재현·정상철(2021)이 제시한 결론과 외견상 상반된다. 이들은 저작권 보호 강화가 지식자본 축적을 촉진한다고 보고한 반면, 본 연구는 학습데이터에 대한 보호 강화가 AI 자본의 축적을 위축시킨다는 결과를 제시한다. 그러나 두 결과는 서로 다른 대상을 분석한 것으로서 양립 가능하다.

김재현·정상철(2021)에서 지식자본은 저작권 보호의 대상이 되는 창작물의 축적을 의미한다. 보호 강화가 창작자의 보상을 높여 창작물의 생산을 촉진하고, 이것이 지식자본의 축적으로 이어지는 경로이다. 반면 본 연구에서 지식자본은 보호 대상인 학습데이터를 투입요소로 하여 형성되는 AI 자본이다. 보호 강화는 이 투입요소의 확보 비용과 이용 제약을 높이므로 AI 자본의 축적을 저해한다.

즉 두 연구는 지식재산권 보호가 보호 대상 자체의 축적에 미치는 효과와, 그 보호 대상을 투입물로 사용하는 다른 자산의 축적에 미치는 효과를 각각 포착한 것이다. 전자는 생산 측 유인을, 후자는 이용 측 비용을 다룬다는 점에서 상호 배타적이지 않으며, 지식재산권 제도의 양면적 성격을 서로 다른 각도에서 보여준다. 두 경로를 하나의 모형에서 동시에 다루는 것은 향후 과제로 남는다.

6.2. 정책적 시사점

본 연구의 결과는 다음과 같은 정책적 함의를 갖는다.

첫째, 정책 수단의 우선순위에 관한 것이다. 본 모형의 기준 설정에서는 학습 단계 경로의 효과가 조달 단계 경로보다 크게 나타났다. 이는 텍스트·데이터마이닝 면책 규정의 도입과 같은 학습 단계의 법적 명확화가 라이선스 시장의 정비나 집중관리 체계의 구축과 같은 조달 단계의 개선보다 큰 효과를 가질 수 있음을 시사한다.

다만 이 결과는 모수 설정에 의존한다. 앞서 확인한 바와 같이 두 경로의 우열은 $\theta(1+\tau_p) > \tau(1-q)$ 의 성립 여부에 달려 있으며, τ 가 개별적으로 식별되지 않으므로 이 조건의 성립 여부를 실증 자료로 판정할 수 없다. 따라서 본 연구의 결과는 학습 단계 규제의 완화가 유효한 정책 수단이 될 수 있음을 보이는 것이며, 두 수단 간의 우선순위를 확정하는 것은 아니다. 우선순위의 판정에는 데이터 접근 탄력성과 권리처리 비용의 크기에 관한 실증적 추정이 선행되어야 한다.

둘째, 제도 설계의 구체적 형태에 관한 것이다. 면책의 도입 여부뿐 아니라 그 설계 방식에 따라 효과의 크기가 달라진다. 면책의 범위가 넓을수록 지식자본 축적의 증가폭이 커지며, 두 설계 방식 간 차이의 크기는 데이터 접근 탄력성 θ 의 설정에 따라 8.8%p에서 20.1%p의 범위를 갖는다. 따라서 본 연구의 결과는 면책의 설계 방식이 효과의 크기에 상당한 영향을 미친다는 점을 보이는 것이며, 특정 설계가 갖는 효과의 크기를 확정하는 것은 아니다. 다만 본 모형은 권리자 측의 후생을 명시적으로 다루지 않으므로, 제도 선택에는 창작자 보상 체계에 관한 별도의 고려가 병행되어야 한다.

셋째, 정책 효과의 평가 시점에 관한 것이다. AI 자본 축적의 관점에서 정책 변화의 효과는 2기 시점에 장기효과의 26.3%만이 실현되며 완전한 조정에는 16기 내외가 소요된다. 반면 총생산 측면의 효과는 당기에 상당 부분이 나타난다. 따라서 규제 변화의 편익을 평가할 때 단기 총생산 지표만을 근거로 삼으면 AI 자본 축적에 미치는 누적적 영향을 과소평가하게 된다.

넷째, 경쟁정책과의 접점에 관한 것이다. 대법원이 자사우대 사건에서 경쟁제한 효과의 구체적인 입증을 요구한 이후, 데이터 접근 제한 행위의 경쟁제한성을 판단할 때에도 정량적 근거의 필요성이 커지고 있다. 다만 본 연구의 결과는 이 영역에서 신중한 해석을 요구한다. 본 모형에서 지식재산권 보호 강화는 기업의 진입 여부를 거의 변화시키지 않으며, 그 효과는 개별 기업의 AI 자본 이용 강도를 통해 나타났다. 이는 데이터 접근 제한의 경쟁제한성을 판단할 때 진입 봉쇄 여부만을 기준으로 삼는 접근이 효과를 놓칠 수 있음을 시사한다. 진입은 유지되더라도 기업의 기술 선택이 왜곡되는 형태의 손실이 존재할 수 있기 때문이다.

6.3. 연구의 한계와 향후 과제

본 연구는 다음과 같은 한계를 갖는다.

첫째, 기업 수가 정책변수에 거의 반응하지 않는다. 본 모형에서 기업이 비율은 균형 노동투입에 비례하여 결정되는데, 학습 단계 제약은 소비 몫과 투자 몫을 변화시키지 않으므로 노동투입에 영향을 주지 못한다. 보다 근본적으로는 본 모형이 데이터를 생산하는 기업과 이를 활용하는 AI 기업을 구분하지 않기 때문이다. 이를 보완하려면 자체 데이터를 보유한 기업과 외부 데이터에 의존하는 기업을 구분하고, 지식자본의 조달비용이 자체 보유 비율에 따라 달라지도록 설정하여야 한다. 이 경우 규제 강화가 외부 의존 기업의 진입비용을 선별적으로 높이므로 기업수와 기업규모 분포가 함께 반응하게 되며, 데이터 접근성의 격차가 시장구조에 미치는 영향을 직접 분석할 수 있다.

둘째, 데이터 생산 유인과 혁신의 구성 효과를 다루지 않았다. 본 연구는 데이터의 총량을 외생적으로 고정하였으므로, 지식재산권 보호 강화가 데이터 생산 유인을 제고하여 학습데이터의 공급을 늘리는 경로는 분석에 포함되지 않았다. 이를 보완하려면 기업가 부문을 데이터 생산과 데이터 활용으로 분리하고, 데이터의 가격과 공급량이 두 부문의 상호작용 속에서 내생적으로 결정되는 구조를 도입하여야 한다. 이 경우 보호 수준의 변화는 데이터 공급 증가라는 양의 경로와 이용 비용 상승이라는 음의 경로를 동시에 발생시키므로, 두 경로의 상대적 크기에 따라 순효과의 부호가 결정되고 최적 보호 수준의 존재 여부를 검토할 수 있게 된다. 아울러 혁신을 데이터 생산 혁신과 데이터 활용 혁신으로 구분하여 보호 수준이 혁신의 총량뿐 아니라 그 구성에 미치는 효과를 분석할 수 있다.

셋째, 규제 강도를 정책변수로 전환하는 과정과 모수의 식별에 판단이 개입한다. 제도 비교를 근거로 면책 규정의 유형별 차이를 정책변수의 완화 쪽으로 예시하였으나, 특정 제도를 특정한 수치에 대응시키기 위해서는 실증 데이터가 요구된다. 또한 제약계수와 조달비용 수준은 곱의 형태로만 모형에 진입하여 개별적으로 식별되지 않으며, 두 규제 경로의 우열은 이들 모수의 조합에 의존한다. 이를 보완하려면 두 가지 작업이 필요하다. 하나는 각국의 면책 규정과 판례의 범위를 지표화하여 규제 강도를 국가 간 비교 가능한 연속 변수로 구성하는 것이고, 다른 하나는 AI 기업의 데이터 관련 지출을 권리처리 비용과 수집·정제 비용으로 분해하여 제약계수의 크기를 직접 추정하는 것이다. 후자가 확보되면 두 경로의 우열을 실증적으로 판정할 수 있다.

넷째, 노동과 지식자본의 대체탄력성을 1로 제약하였다. 대체탄력성이 달라지면 규제 변화의 효과도 달라진다. 이를 보완하려면 복합요소를 CES 형태로 확장하고 대체탄력성을 실증적으로 추정하여야 한다. 특히 AI 자본이 노동을 보완하는 부문과 대체하는 부문이 공존할 가능성이 있으므로, 산업별 또는 직무별 대체탄력성의 차이를 반영한 분석이 요구된다. 이 경우 지식재산권 규제가 산업 간 자원 배분에 미치는 영향까지 함께 다룰 수 있을 것이다.

참고문헌

단행본(서양)

- Bronwyn H. Hall et al., “Measuring the Returns to R&D”, Edit. by Bronwyn H. Hall & Nathan Rosenberg, Handbook of the Economics of Innovation, Vol.2, North-Holland, 2010.
- Thomas F. Cooley & Edward C. Prescott, “Economic Growth and Business Cycles”, Edit. by Thomas F. Cooley, Frontiers of Business Cycle Research, Princeton University Press, 1995.

학술지(국내 및 동양)

- 김재현·정상철, “저작권 보호 활동에 따른 경제적 파급효과 분석”, 「문화경제연구」, 제24권 제3호(2021).
- 최희갑, “Zipf 분포에 기초한 한국의 기업 규모 분포 분석”, 「통계연구」, 제11권 제2호(2006).

학술지(서양)

- Daron Acemoglu & Ufuk Akcigit, “Intellectual Property Rights Policy, Competition and Innovation”, *Journal of the European Economic Association*, Vol.10 No.1(2012).
- Lei Chen & Jieshuang Wang, “Intellectual Property Protection, R&D Investment and Digital Technology Innovation: An Empirical Study Based on the Revision of the Patent Law”, *International Review of Economics & Finance*, Vol.103(2025).
- Mark A. Lemley & Bryan Casey, “Fair Learning”, *Texas Law Review*, Vol.99 No.4(2021).
- Matthew Sag, “The New Legal Landscape for Text Mining and Machine Learning”, *Journal of the Copyright Society of the U.S.A.*, Vol.66(2019).
- Pamela Samuelson, “Generative AI Meets Copyright”, *Science*, Vol.381 No.6654(2023).
- Raj Chetty et al., “Are Micro and Macro Labor Supply Elasticities Consistent? A Review of Evidence on the Intensive and Extensive Margins”, *American Economic Review*, Vol.101 No.3(2011).
- Robert E. Lucas, Jr., “On the Size Distribution of Business Firms”, *Bell Journal of Economics*, Vol.9 No.2(1978).
- Tania Babina et al., “Artificial Intelligence, Firm Growth, and Product Innovation”, *Journal of Financial Economics*, Vol.151(2024).

판례

- 대법원 2025. 10. 16. 선고 2023두32709 판결.
- 서울중앙지방법원 2023. 2. 3. 선고 2020가합584719 판결.
- Bartz v. Anthropic PBC, No. 3:24-cv-05417 (N.D. Cal. June 23, 2025).
- Kadrey v. Meta Platforms, Inc., No. 3:23-cv-03417 (N.D. Cal. June 25, 2025).

인터넷 자료

- 한국은행, “국민계정”, 한국은행 경제통계시스템, <<https://ecos.bok.or.kr/#/SearchStat>>, 검색일: 2026. 7. 30.
- Tamay Besiroglu et al., “Chinchilla Scaling: A replication attempt”, arxiv, <https://arxiv.org/abs/2404.10102>, 작성일: 2024. 4. 15.

연구보고서

- Jack Corrigan, “Promoting AI Innovation Through Competition: A Guide to Managing Market Power”, Center for Security and Emerging Technology, 2025.
- Jacques Crémer et al., “Competition Policy for the Digital Era”, European Commission, 2019.

- Jason Furman et al., “Unlocking Digital Competition: Report of the Digital Competition Expert Panel”, HM Treasury, 2019.
- OECD, “Licensing of IP Rights and Competition Law”, OECD Competition Committee, DAF/COMP(2019)3, 2019.
- OECD, “Artificial Intelligence, Data and Competition”, OECD Artificial Intelligence Papers, 2024.
- OECD, “Artificial Intelligence and Competitive Dynamics in Downstream Markets”, OECD Competition Policy Roundtable Background Note No. 331, 2025.

기타 자료

- Jordan Hoffmann et al., “Training Compute-Optimal Large Language Models”, Advances in Neural Information Processing Systems 35 (NeurIPS 2022), 2022
- Thibault Schrepel & Teodora Groza, “Computing ‘Innovation Competition’”, Amsterdam Law & Technology Institute, 2024.