



# 머신러닝을 활용한 재정사업평가: 정책금융 사례를 중심으로

2021. 12

이 환 응 · 고 창 수 · 배 진 수



# 머신러닝을 활용한 재정사업평가: 정책금융 사례를 중심으로

2021. 12

이환웅 · 고창수 · 배진수



## 서 언

빅데이터의 비약적인 발전과 함께 이를 분석하기 위한 도구로서 인공지능(AI)은 다방면에서 활용되고 있다. 이 중 머신러닝은 데이터에 기반한 학습을 통해 예측성고를 향상시키는 최적의 모형을 선택하는 알고리즘으로 현재 IT분야에서 광범위하게 활용되고 있다. 최근에는 비단 IT분야뿐 아니라 사회과학분야에서도 재정정책 시행에 따른 효과를 예측하여 '누가' 가장 적절한 정책대상자가 되어야 하는지 여부를 결정할 때 머신러닝을 적극적으로 활용하는 방안에 대한 연구가 이루어지고 있다.

전통적으로 재정정책 설계관련 의사결정의 많은 부분은 전문가들의 판단 혹은 기존 통계모형에 기반한 정책효과분석 결과를 활용하여 이루어졌다. 전문가들에 의한 의사결정은 전문가들이 가지고 있는 배경지식이나 축적된 노하우를 의사결정에 반영할 수 있다는 장점이 있지만 인지능력의 한계로 수집 가능한 모든 정보를 활용하여 의사결정을 할 수 없다는 단점이 있다. 또한 선형통계모형을 활용한 효과성 분석 결과를 의사결정에 활용하는 방법 역시 통계모형의 설정에 자의적인 판단이 개입될 여지가 있다는 점, 그리고 정책효과가 클 것으로 보이는 집단을 식별하는 과정에서 다중검정의 문제가 발생할 수 있다는 점에서 한계점을 지닌다. 반면에 머신러닝 방법론은 알고리즘 자체가 데이터에 기반하여 최적 모형을 찾으며 다중검정과 관련해 과적합을 방지하는 장치가 알고리즘에 포함되어 있기 때문에 기존 통계방법론의 한계점을 보완할 수 있는 알고리즘으로 주목을 받고 있다. 바로 이러한 점에서 재정정책의 설계에 머신러닝을 접목하여 재정정책의 효과성 및 정책수혜에 따른 효과의 이질성을 추정하고 분석 결과를 정책설계에 반영하려는 시도들은 주목할 필요가 있다.

본 연구에서는 머신러닝 기법을 활용하여 재정사업을 평가하고 머신러닝이 재정사업의 효과성을 제고시키는 데 활용될 수 있는지 탐색하였다. 구체적으로

머신러닝의 활용 가능성을 탐색하기 위해 본 연구에서는 머신러닝 방법론이 정책금융의 효과를 추정하고 정책수혜자를 선정함에 있어 정책효과가 큰 집단을 식별하는 데 효과적인지 여부를 기존 연구에서 많이 사용하는 성향점수매칭 방법과 비교하여 검토하였다. 정책금융 등 다양한 재정사업의 처치효과 및 처치효과와의 이질성 추정에 머신러닝 알고리즘의 사용 가능성을 기존의 통계방법론과 비교를 통해 검토하여 머신러닝의 유용성을 확인하였고 향후 재정사업의 설계 및 성과평가를 위한 새로운 기반을 마련했다는 점이 본 연구의 핵심적인 기여이다.

본 보고서는 한국조세재정연구원의 이환웅 부연구위원, 고창수 부연구위원, 배진수 부연구위원이 집필하였다. 저자들은 본 연구에 도움을 준 익명의 논평자들과 고도화된 분석을 위해 기초자료를 조사·수집·정리해 준 박진우 연구원, 이재국 연구원, 윤다솜 연구원 그리고 최종 출판까지 행정지원을 해 준 본원의 직원들에게도 깊은 감사의 뜻을 전하고 있다. 또한 보고서 작성과정 중 원내 중간보고와 최종보고 세미나에서 가감 없는 평가와 격려를 해 준 강동익 박사, 강희우 박사, 그리고 본고 분석에 필수적이었던 신용보증기금 자료를 연구자들과 공유하고 정책금융 분석 관련 귀중한 조언들을 해준 장우현 박사께도 무한한 고마움을 전하고 있다.

끝으로 본 보고서의 내용은 저자들의 의견이며, 본 연구원의 공식 견해가 아님을 밝힌다.

2021년 12월

한국조세재정연구원  
원장 김 재 진

## 요약 및 정책적 시사점

본 연구는 머신러닝 방법론이 정책금융 수혜자를 선정하는 데 정책효과성 측면에서 개선에 도움이 될 수 있는지 탐색하였다. 머신러닝 방법론의 유용성을 검토하기 위해 신용보증기금의 정책효과를 선행연구에서 많이 사용한 성향점수매칭 방법론과 머신러닝 알고리즘인 커절 포레스트를 이용하여 분석하고 추정 결과를 비교하였다. 성향점수매칭을 활용한 추정 결과 신용보증기금의 수혜는 매출액을 7.8% 증가시키는 것으로 나타났으며 커절 포레스트를 활용한 경우 신용보증기금의 수혜는 매출액을 4.4% 증가시키는 것으로 나타났다. 두 가지 방법론에 의한 추정 결과는 질적으로 유사해 두 방법론이 상호 보완적으로 사용될 수 있음을 확인하였다. 다만 성향점수매칭은 성향점수를 로짓모형을 통해 추정할 때 추정에 사용된 변수에 따라 처치효과 크기의 크기가 많이 달라지는 현상이 관측되었다. 이는 성향점수에 의한 정책효과 추정치가 변수선택에 따라 강건하지 않음을 의미하며 커절 포레스트가 기존 통계모형의 이러한 한계점을 극복할 수 있는 좋은 대안이 될 수 있음을 확인하였다.

다음으로 본 연구에서는 성향점수매칭과 커절 포레스트를 통해 재정정책에 따른 처치효과 이질성을 분석하고 두 방법론을 비교하였다. 먼저 성향점수매칭을 사용한 정책효과 이질성은 설명변수들의 중위값을 기준으로 부집단을 구성하고 이후 각각의 집단에서 정책효과를 분석하였다. 이질성 분석 결과 사업체의 업력이 낮을수록, 매출액이 낮을수록, 그리고 영업이익이 낮을수록 정책금융이 매출액에 미치는 긍정적인 효과가 더 큰 것으로 나타났다.

한편 커절 포레스트를 사용한 이질성 분석은 주어진 독립변수의 값에 따라 조건부 평균처치효과를 계산할 수 있기 때문에 성향점수매칭과 비교하여 보다 세밀한 이질성 분석이 가능하였다. 매출액에 대한 이질적인 처치효과를

크기 순으로 나열했을 때 평균처치효과는 4.45%였지만 처치효과 기준으로 상위 10%에 속하는 사업체들의 평균적인 처치효과는 8.82%, 하위 10%에 속하는 사업체들은 0.97%로 나타나 신용보증기금의 정책효과가 매우 이질적임을 알 수 있었다. 이러한 분석 결과는 만약 생산적 중소기업 지원정책의 일환으로 본 연구의 분석 결과에 기반해 가장 높은 처치효과를 가진 그룹에게 정책금융을 제공하는 방법으로 정책금융을 타게팅하면 정책금융의 효율성을 높일 수 있음을 보여준다.

또한 커질 포레스트에 의한 이질성 분석에서는 설명변수 중 하나인 사업체의 과거 영업이익의 크기의 변화에 따라 매출액이 증대되는 효과가 매우 크게 비선형적으로 변화하는 모습을 발견하였다. 기존의 선형모형에서 발견하기 어려운 이러한 비선형적 패턴의 발견은 향후 정책금융의 처치효과가 나타나는 과정에 대한 메커니즘을 이해하는 데 도움이 될 것으로 판단된다. 또한 추정된 처치효과의 발생 원인이 설명변수 중 어떤 변수에 의한 것인지 파악하는 데 유용하기 때문에 머신러닝 알고리즘을 기반으로 한 정책수혜자 설계에 구체적인 기준을 제시하는 데에도 효과적일 것으로 보인다.

본 연구로부터 도출할 수 있는 정책적 시사점은 다음과 같다. 먼저 정책금융정책을 생산성 중소기업을 위한 정책과 공공부양적 중소기업 정책으로 이원화하고 보증심사 기준을 이에 걸맞게 보완하는 것이다. 특히 강소기업으로 성장가능성이 있으나 시장실패로 성장에 애로를 겪고 있는 생산적 중소기업을 위한 금융지원정책의 운용 시 머신러닝 방법론을 적극적으로 활용하는 방안을 검토한다면 해당 지원사업의 효과성을 향상시킬 수 있을 것으로 판단된다. 물론 정책효과가 뚜렷할 것으로 보이는 기업군에만 지원이 집중될 경우 정책금융의 본연의 목적성 달성이 저해될 가능성이 있는데 그렇기에 더욱이 정책금융 지원사업을 생산적 기업을 위한 사업과 중소기업의 복지를 위한 사업, 즉 투트랙으로 운용하는 것이 필요해 보인다.

다음으로 정부의 재정사업 효과성 평가에 머신러닝 방법론을 보완적으로 사용하는 방안을 검토할 필요가 있다. 현재 성과중심의 재정운용을 위해 정부는 주요 재정사업에 대한 평가를 실시하고 그 결과를 재정운용에 반영할 수 있도록 하고 있다. 이러한 성과중심의 재정운용이 취지에 맞게 운용되기

위해서는 무엇보다도 재정사업의 효과성 평가가 신뢰성 있게 수행되어야 한다. 머신러닝 알고리즘은 기존의 효과성 평가에 사용되는 방법론의 한계점을 보완할 수 있기 때문에 재정사업 효과성 평가에 머신러닝 알고리즘을 적극적으로 활용하는 것도 필요할 것으로 판단된다.

끝으로 본 연구의 한계점을 논의하면 본 연구는 이용 가능한 자료의 한계로 신용보증기금 지원이력 외에 기타 정책금융 지원여부를 고려하지 못하였다. 따라서 머신러닝 알고리즘을 통해 파악한 처치효과와 이질성에 기반해 정책금융의 지원대상자를 즉각적으로 선정하는 것은 어려울 것으로 판단된다. 하지만 본 연구가 정책금융의 수혜자 선정 측면에서 머신러닝의 활용 가능성을 탐색하였다는 점을 고려한다면 이러한 한계점은 본 연구의 본질적인 시사점에는 큰 영향을 미치지 않을 것으로 판단된다. 정책금융 지원이력 및 '중소기업 지원사업 통합관리시스템'의 중소기업 재정사업 지원이력 등을 연구자들이 활용할 수 있다면 금융·재정사업의 중복 수혜 이력 등을 정확히 파악할 수 있는데, 이 경우 본 연구에서 논의한 머신러닝의 방법론을 활용해 구체화된 정책설계 방안을 도출할 수 있을 것으로 판단된다. 재정사업 설계의 효과성 향상을 위해 정부의 자료공개에 대한 전향적인 자세를 기대해 본다.



## 목 차

I. 서론 .....	17
II. 정책금융의 개관 .....	22
1. 정책금융 지원의 이론적 논의 .....	22
가. 정책금융의 두 가지 접근법 .....	22
나. 시장실패로 인한 정책금융의 필요성 .....	23
다. 중소기업에 정책금융을 지원해 줘야 하는 이유 .....	25
2. 정책금융에 관한 실증연구 .....	28
가. 누가 정책금융을 지원받는가? .....	29
나. 정책금융 수혜효과의 이질성 .....	32
3. 중소기업 정책금융 지원 현황 .....	38
가. 기본체계 .....	38
나. 중소기업 정책금융 통계 .....	39
4. 지원기관별 주요사업 .....	40
가. 가용자금 공급 .....	40
나. 신용보증지원제도 .....	41
III. 정책금융의 지원효과 .....	47
1. 분석 자료 소개 및 기초통계량 .....	47
가. 한국기업데이터(KED) 자료 소개 .....	47
나. 신용보증기금 자료 소개 .....	47
다. 기초통계량 검토 .....	50
2. 분석 전략 및 해석상의 유의점 .....	53

3. 분석 결과 .....	56
가. 선행 추세 검정 .....	56
나. 성향점수매칭을 이용한 실증 분석 .....	60
다. 이질성 분석 .....	76
4. 소결 .....	104
<b>IV. 머신러닝의 개관 .....</b>	<b>106</b>
1. 머신러닝 주요 알고리즘 소개 .....	106
2. 예측문제(Prediction Problem)에 대한 선행연구 .....	111
가. 머신러닝이 사람의 결정을 개선시킬 수 있는가? .....	111
나. 머신러닝이 기존의 방법론보다 우월한가? .....	116
다. 머신러닝이 이질적인 처치효과를 잘 찾아낼 수 있는가? .....	121
<b>V. 머신러닝을 활용한 정책금융의 효과 추정 .....</b>	<b>127</b>
1. 정책효과 추정을 위한 식별가정 .....	128
2. 성향점수매칭과 커질 포레스트 .....	130
가. 성향점수매칭 .....	130
나. 커질 포레스트 .....	134
3. 분석 결과 .....	139
가. 정책효과의 추정 .....	139
나. 정책효과의 이질성 .....	152
4. 머신러닝을 활용한 기업성과 예측 .....	162
가. 데이터 .....	164
나. 머신러닝 알고리즘 및 실험방법 .....	165
다. 알고리즘별 성과 비교 .....	166

---

---

VI. 요약 및 정책적 시사점 .....	169
1. 연구 결과 요약 및 한계 .....	169
2. 정책적 시사점 .....	171
가. 정책금융 대상자 선정 .....	171
나. 재정사업 효과성 평가 .....	173
참고문헌 .....	174
부록 .....	182

---

표목차

〈표 II-1〉 중소기업 대상 정책금융의 성과분석 선행연구 요약 ..... 36

〈표 II-2〉 중소기업 금융지원 현황 ..... 39

〈표 II-3〉 금융증개지원대출 프로그램별 구성 내역 ..... 40

〈표 II-4〉 신용보증기금 주요 보증제도 ..... 42

〈표 II-5〉 최근 5년간 기본재산 및 운용배수 현황 ..... 42

〈표 II-6〉 보증공급 및 신규보증 금액 ..... 43

〈표 II-7〉 기술보증기금 주요 보증제도 ..... 43

〈표 II-8〉 최근 5년간 기본재산 및 운용배수 현황 ..... 44

〈표 II-9〉 연도별 보증지원 실적 ..... 45

〈표 II-10〉 중소벤처기업진흥공단 주요 보증제도 ..... 46

〈표 III-1〉 신용보증기금 신규 지원이력 기초통계량 1 ..... 49

〈표 III-2〉 신용보증기금 신규 지원이력 기초통계량 2 ..... 49

〈표 III-3〉 신용보증기금 신규 지원이력 기초통계량(사업체) ..... 49

〈표 III-4〉 KED 매칭 자료 기초통계량(사업체) ..... 50

〈표 III-5〉 KED 자료 기초통계량(사업체, 미지원기업) ..... 52

〈표 III-6〉 신용보증기금 지원·미지원 기업 기업크기 분포 ..... 53

〈표 III-7〉 이벤트스터디 모형 추정 결과 1(2009~2015년 표본) ..... 57

〈표 III-8〉 이벤트스터디 모형 추정 결과 2(2009~2015년 표본) ..... 58

〈표 III-9〉 실증분석을 위한 표본 구성 요약 ..... 60

〈표 III-10〉 분석자료 기초통계량(2013~2015년 표본) ..... 62

〈표 III-11〉 신용보증기금 신규지원 정책효과(2013~2015년 표본) ..... 63

〈표 III-12〉 분석자료 기초통계량(2011~2015년 표본) ..... 65

〈표 III-13〉 신용보증기금 신규지원 정책효과(2011~2015년 표본) ..... 66

---

〈표 Ⅲ-14〉 분석자료 기초통계량(2011~2016년 표본) .....	69
〈표 Ⅲ-15〉 신용보증기금 신규지원 정책효과(2011~2016년 표본) .....	70
〈표 Ⅲ-16〉 분석자료 기초통계량(2011~2017년 표본) .....	72
〈표 Ⅲ-17〉 신용보증기금 신규지원 정책효과(2011~2017년 표본) .....	73
〈표 Ⅲ-18〉 신용보증기금 신규지원 정책효과 요약 .....	74
〈표 Ⅲ-19〉 신용보증기금 신규지원 정책효과(이질성: 업력) .....	78
〈표 Ⅲ-20〉 신용보증기금 신규지원 정책효과(이질성: 매출액) .....	80
〈표 Ⅲ-21〉 신용보증기금 신규지원 정책효과(이질성: 영업이익) .....	83
〈표 Ⅲ-22〉 신용보증기금 신규지원 정책효과(이질성: 연구개발비) .....	85
〈표 Ⅲ-23〉 신용보증기금 신규지원 정책효과(이질성: 고정자산) .....	87
〈표 Ⅲ-24〉 신용보증기금 신규지원 정책효과(이질성: 업종) .....	89
〈표 Ⅲ-25〉 신용보증기금 신규지원 정책효과(이질성: 기업 규모) .....	91
〈표 Ⅲ-26〉 신용보증기금 신규지원 정책효과(이질성: 지역) .....	92
〈표 Ⅲ-27〉 신용보증기금 신규지원 정책효과(이질성: 지적재산권) .....	94
〈표 Ⅲ-28〉 신용보증기금 신규지원 정책효과(이질성: 가족주식소유 비율) .....	97
〈표 Ⅲ-29〉 신용보증기금 신규지원 정책효과(이질성: 수출기업 여부) .....	99
〈표 Ⅲ-30〉 신용보증기금 신규지원 이질성분석 요약 1 .....	101
〈표 Ⅲ-31〉 신용보증기금 신규지원 이질성분석 요약 2 .....	101
〈표 Ⅳ-1〉 머신러닝 알고리즘과 정규화 방법 .....	110
〈표 Ⅴ-1〉 다양한 처치효과의 수학적 정의 .....	129
〈표 Ⅴ-2〉 비교락성과 중첩성의 수학적 정의 .....	130
〈표 Ⅴ-3〉 커절 포레스트의 모형 적합성 검정 .....	147
〈표 Ⅴ-4〉 신용보증기금 수혜가 매출액 및 영업이익에 끼친 효과 .....	150
〈표 Ⅴ-5〉 설명변수들의 트리성장과 관련한 중요도 .....	153
〈표 Ⅴ-6〉 알고리즘별 예측 성과 비교: 직전 3개년도 실적데이터 사용 .....	167

---

## 그림목차

[그림 II-1] 중소기업금융 지원체계 .....	38
[그림 III-1] 처치군과 대조군의 매출액 추세 비교 .....	59
[그림 III-2] 신용보증기금 성향점수 분포도(2013~2015년 표본) .....	64
[그림 III-3] 신용보증기금 성향점수 분포도(2011~2015년 표본) .....	67
[그림 III-4] 처치군과 대조군의 추세 비교(2013~2015년 표본) .....	68
[그림 III-5] 처치군과 대조군의 추세 비교(2011~2015년 표본) .....	68
[그림 III-6] 신용보증기금 성향점수 분포도(2011~2016년 표본) .....	71
[그림 III-7] 신용보증기금 성향점수 분포도(2011~2017년 표본) .....	74
[그림 III-8] 2011~2015년 표본의 업력 분포 .....	77
[그림 III-9] 신용보증기금 성향점수 분포도(이질성: 업력) .....	78
[그림 III-10] 2011~2015년 표본의 매출액 분포 .....	80
[그림 III-11] 신용보증기금 성향점수 분포도(이질성: 매출액) .....	81
[그림 III-12] 2011~2015년 표본의 영업이익 분포 .....	82
[그림 III-13] 신용보증기금 성향점수 분포도(이질성: 영업이익) .....	83
[그림 III-14] 2011~2015년 표본의 연구개발비 분포 .....	85
[그림 III-15] 신용보증기금 성향점수 분포도(이질성: 연구개발비) .....	86
[그림 III-16] 2011~2015년 표본의 고정자산 분포 .....	87
[그림 III-17] 신용보증기금 성향점수 분포도(이질성: 고정자산) .....	88
[그림 III-18] 신용보증기금 성향점수 분포도(이질성: 업종) .....	90
[그림 III-19] 신용보증기금 성향점수 분포도(이질성: 기업 규모) .....	91
[그림 III-20] 신용보증기금 성향점수 분포도(이질성: 지역) .....	93
[그림 III-21] 신용보증기금 성향점수 분포도(이질성: 지적재산권) .....	95
[그림 III-22] 2011~2015년 표본의 가족주식소유 비율 분포 .....	97

---

[그림 Ⅲ-23] 신용보증기금 성향점수 분포도(이질성: 가족주식소유 비율) .....	98
[그림 Ⅲ-24] 신용보증기금 성향점수 분포도(이질성: 수출기업 여부) .....	100
[그림 V-1] 2차원 공간의 재귀적 이진분류 예시 .....	134
[그림 V-2] 결정트리의 시각화: [그림 V-1]에 상응 .....	136
[그림 V-3] 커절 포레스트의 처치효과 추정 예시 .....	138
[그림 V-4] 신용보증기금 수혜기업과 비수혜기업의 변수별 분포: IPW 가중치 적용 .....	142
[그림 V-5] 커절 포레스트의 사용하는 나무 수에 따른 분산의 증위값 .....	144
[그림 V-6] 커절 포레스트의 커절트리1 도식화: 매출액모형 .....	149
[그림 V-7] 커절 포레스트를 활용한 신용보증기금 지원의 이질적 효과 추정 ...	154
[그림 V-8] 설명변수별 처치효과의 이질성 .....	156
[그림 V-9] 예측변수별 처치효과의 이질성: 비교정태접근법 .....	160

---



---

# I. 서론

---

머신러닝이라고 불리는 방법론들은 경제, 의학, 컴퓨터 공학 등 여러 분야에서 사용되는 다양한 기법들을 포괄하는 용어이므로 명확한 정의를 내리는 것이 어렵다(Athey, 2019). 하지만 다양한 머신러닝 알고리즘의 공통점은 많은 설명변수들을 가지고 스스로 수많은 모형을 구축하고 그 다음 해당 모형들을 데이터에 대입하면서 최적의 성과를 가져다주는 모형을 찾아가는 데이터에 기반(data-driven)한 모형 선택 알고리즘이라는 점이다. 특히 빅데이터의 접근이 용이해지고 컴퓨터의 연산능력이 발전함에 따라 이러한 특성을 가진 머신러닝 방법론들은 연구자들의 많은 관심을 받고 있다. 최근에는 정책의 효과성 평가(Impact evaluation) 및 처치효과의 이질성 분석, 그리고 그에 따른 정책 수혜대상을 선정하는 데 머신러닝을 활용하는 방안에 대한 논의도 활발히 진행되고 있다(정재현 · 이환웅, 2020).

현재 재정정책 관련 의사결정은 전문가들의 판단에 의거하거나 혹은 통계 모형에 기반한 정책효과 분석을 통해 이루어지고 있다. 따라서 본 연구는 머신러닝 기법을 활용하여 재정사업을 평가하고 머신러닝이 사업의 효과성을 제고시키는 데 활용될 수 있는지 탐색하였다. 물론 전문가들에 의한 의사결정은 전문가들이 가지고 있는 배경지식이나 축적된 노하우를 의사결정에 반영할 수 있다는 장점이 있지만 인지능력의 한계로 인해 수집 가능한 모든 정보를 다 활용할 수 없다는 단점이 있다.<sup>1)</sup>

---

1) 예를 들어 Mullainathan and Obermeyer(2019)는 심장마비 의심증상이 있는 환자들에게 도관삽입술(cardiac catheterization) 수행 여부를 결정하는 과정에서 의사들은 머신러닝 알고리즘에 비해 환자들의 과거 병력에 대한 정보를 적게 반영하는 것으로 나타났으며, 이 외에도 머신러닝 방법론은 체포된 피의자들의 구속 영장을 심사하는 판사들의 의사결정이나 인력 채용과정, 승진, 재임용에 관한 의사결정을 개선시킬 수 있다는 연구 결과들이 있다(Kleinberg et al., 2018; Chalfin et al., 2016).

한편 기존의 통계모형을 통한 정책효과 분석을 의사결정에 활용하는 방법 역시 한계점을 지닌다. 연구자들은 전통적으로 선형회귀모형을 정책효과 추정에 활용하는데, 이는 설명변수가 적절하게 설정된 경우 정책효과 추정치가 통계적으로 좋은 성질들을 가지고 있기 때문이다. 하지만 모형을 설정할 만한 경제학적인 이론이 마땅치 않은 경우 연구자들은 자의적인 모형(ad-hoc model)을 설정하는 경우가 많은데 이러한 모형설정은 여러 가지 문제점들을 야기할 수 있다. 첫째로는 학자들이 휴리스틱(heuristic)에 기초해서 변수를 선정하게 되므로 학자들의 자의적인 판단이 개입된다는 문제이다. 반면 머신러닝 방법론은 알고리즘 자체가 데이터에 기반하여 최적 모형을 찾아가게 되므로 학자들의 자의적인 판단이 개입될 여지가 적다. 둘째로는 다중검정과 관련된 문제인데 학자들이 가장 좋은 모형을 찾는 과정에서 여러 가지 모형탐색을 시도하면서 다중검정의 문제가 발생할 수 있다는 것이다. 즉 여러 가지 모형을 추정하다 보면 확률적으로 우연히 유의미한 설명변수를 찾게 되는데, 이렇게 찾은 설명변수는 해당 표본에서의 과적합이 발생해서 다른 표본에서는 설명력이 없으므로 좋은 모형이 되지 못한다. 반면에 머신러닝은 모형탐색 과정에서 모형의 성과를 미리 분리시켜 둔 실험데이터를 통해 확인하고 최적의 모형을 찾으므로 모형의 성과를 담보할 수 있다. 마지막으로 선형모형은 설명변수와 종속변수의 비선형적인 관계를 잘 포착하기 어려우며 변수들 간의 상호작용을 찾아내기 어려운 문제점이 있다. 머신러닝 방법론은 매우 유연한 형태의 함수를 사용하여 최적 모형을 찾기 때문에 비선형적인 관계나 변수들과의 상호작용을 알아내기 쉽다는 장점이 있다. 따라서 이러한 머신러닝의 장점은 처치효과의 이질성을 분석하는 데 유용하게 사용될 수 있으며 또한 정책효과의 메커니즘을 알아내는 데 용이할 수 있다.

본고에서는 머신러닝 방법론이 정책금융 수혜자를 선정하는 데 정책효과성 측면에서 개선에 도움이 될 수 있는지에 관해 연구하였다. 정책금융은 머신러닝의 활용 가능성을 탐색하는 데 다음과 같은 이유에서 좋은 분석대상이 될 수 있다. 우선 정책금융은 다양한 기업들을 대상으로 하며 기업들은

다양한 특성변수들(회사유형, 업력, 기업 규모, 연구개발비, 매출액, 영업이익, 업종, 소재지 등)을 가지고 있으므로 많은 설명변수들을 통해 스스로 최적의 모형을 만드는 머신러닝 방법론이 유용할 수 있다. 또한 분석을 위한 수많은 기업들의 다양한 특성변수자료가 비교적 잘 구축되어 있어서 머신러닝 알고리즘을 수월하게 사용할 수 있다.<sup>2)</sup> 마지막으로 정책금융의 기대효과에 대해서 명확한 경제학적인 이론이 존재하지 않고 선행연구들도 각기 상이한 독립변수를 사용하고 있는 만큼 머신러닝 방법론을 통해서 알고리즘이 최적의 모형을 찾아가도록 함으로써 연구자의 자의적인 판단을 최소화할 수 있을 것으로 판단된다.

구체적으로 본 연구에서는 신용보증기금의 정책효과를 선행연구에서 많이 사용한 성향점수매칭 방법론과 머신러닝 방법론 중 하나인 커질 포레스트를 모두 이용하여 분석하고 추정 결과를 비교함으로써 머신러닝 방법론의 유용성에 대해서 연구하였다. 우선 성향점수매칭 방법론으로 정책효과를 추정한 결과 신용보증기금의 수혜는 매출액을 7.8% 증가시키는 것으로 나타났으며, 영업이익은 1,300만원 감소시키는 것으로 나타났다. 반면에 커질 포레스트로 추정한 정책효과는 매출액 4.4% 증가효과, 영업이익 300만원 감소효과<sup>3)</sup>가 있는 것으로 나타났다. 두 가지 방법론 모두 유사한 방향의 효과가 나타난 것은 두 가지 방법론 모두 데이터에서 의미 있는 해석을 이끌어 낼 수 있는 유용한 방법론이며, 서로 배척하는 관계가 아니라 상호 보완적으로 사용될 수 있음을 의미한다. 다만 성향점수매칭은 변수의 선택에 따라 정책 효과의 크기가 많이 달라지는 현상이 관측되었다. 예를 들어 성향점수를 추정하는 과정에서 변수들의 제공항을 추가하지 않는 경우 매출액에 관한 효과는 4.02%로 나타나 원래의 추정치인 7.8%와는 큰 차이를 나타냈다. 이는 성향점수에 의한 추정치가 변수 선택에 취약한 특징이 있음을 의미하며 커질 포레스트가 이러한 기존 통계모형의 취약성을 극복할 수 있는 좋은 대안

---

2) 예를 들어 본 분석에서 사용한 자료 중 하나인 한국기업데이터(KED)의 자료는 96만 9,493개의 사업체 정보를 포함하고 있으며, 기업의 다양한 특성변수 및 재무정보를 포함하고 있다.

3) 단, 해당 추정치는 통계적으로 유의하지는 않았다.

이 될 수 있음을 보여주고 있다.<sup>4)</sup>

또한 본 연구에서는 성향점수매칭 방법론과 커질 포레스트를 통해 정책효과  
의 이질성을 분석하고 두 방법론을 비교해 보았다. 성향점수매칭 방법론  
을 활용한 정책효과  
의 이질성은 독립변수들의 중위값을 기준으로 부집단  
(subgroup)으로 구성한 후 각각의 집단에서 정책의 효과를 분석해 보았으며  
업력, 매출액, 영업이익 등 11개의 변수들에 대해서 이질성을 분석해 보았  
다. 분석 결과 업력이 낮을수록, 매출액이 낮을수록, 영업이익이 낮을수록  
매출액에 미치는 긍정적 정책금융의 효과가 더 큰 것으로 나타났다.

커질 포레스트를 사용하는 경우에는 주어진 독립변수에 따라 조건부 평균  
처치효과(CATE)를 계산할 수 있었으며 성향점수매칭 방법론보다 더욱 세밀한  
이질성 분석이 가능했다. 매출액에 대한 이질적인 처치효과를 크기 순으로  
나열했을 때 평균적인 처치효과는 4.45%이나 처치효과 기준으로 상위 10%에  
속하는 사업체들의 평균적인 처치효과는 8.82%, 하위 10%에 속하는 사업체  
들은 0.97%로 나타나 신용보증기금의 정책효과가 매우 이질적인 것으로 나  
타났다. 이러한 결과는 만약 본 연구의 분석 결과에 기반하여 가장 낮은 처  
치효과를 가진 그룹들에게서 정책금융을 거두어 가장 높은 처치효과를 가진  
그룹에게 정책금융을 제공하는 방법으로 정책금융을 타게팅(targeting)한다  
면 정책금융의 효율성을 높일 수 있다는 것을 의미한다. 그리고 커질 포레  
스트에 의한 이질성 분석은 영업이익의 크기의 변화에 따라 처치효과가 매  
우 크게 비선형적으로 변화하는 모습을 발견할 수 있었다. 선형모형에서 발  
견하기 어려운 이러한 비선형적 패턴을 발견하는 것은 향후 정책금융의 처  
치효과가 나타나는 과정의 메커니즘을 이해하는 데 도움이 될 수 있을 것으  
로 판단된다.

본 보고서의 구성은 다음과 같다. 먼저 제Ⅱ장에서는 정책금융을 개관하  
였다. 정책금융에 대한 기본적인 내용 및 정책금융 지원 효과에 관한 선행

---

4) 제Ⅴ장에서 자세히 상술하겠지만 커질 포레스트는 다수의 커질트리(통계모형)의 결과값을  
평균화하기 때문에 비유하자면 성향점수매칭 방법론에서 로짓모형을 추정할 시 다수의 모형  
형태가 다른 로짓모형을 추정한 후 이를 기반으로 처치효과를 추정해서 평균값을 구하는  
것과 유사하기 때문에 커질 포레스트의 추정치는 강건성 측면에서 좋은 대안으로 판단된다.

연구를 검토함으로써 정책금융 정책설계에 머신러닝 알고리즘을 적용하는 것이 유용한지 가늠하고자 하였다. 그다음 제Ⅲ장에서는 선행연구에서 주로 사용한 성향점수매칭 방법론을 활용하여 신용보증기금의 수혜효과를 분석하였다. 제Ⅲ장의 분석과 머신러닝 알고리즘의 분석 방법을 비교 제시함으로써 머신러닝 알고리즘의 장단점을 파악하고자 함이다. 그리고 제Ⅳ장에서는 머신러닝 알고리즘을 소개하고 정책의사결정 향상을 위해 머신러닝 알고리즘을 활용한 선행연구를 소개하였다. 선행연구를 단순히 정리하지 않고 ① 머신러닝이 사람의 결정을 개선시킬 수 있는지 ② 머신러닝이 기존의 방법론보다 우월한지 ③ 머신러닝이 이질적인 처치효과를 잘 찾아내는지로 유형화하여 소개하였다.

제Ⅴ장에서 머신러닝 알고리즘을 활용하여 정책금융효과를 예측하였는데, 먼저 머신러닝 알고리즘 중에 하나인 커질 포레스트를 활용하여 처치효과를 재추정하였다. 성향점수매칭은 비교락성(Unconfoundness)과 중첩성(Overlap) 가정하에 처치효과를 편이 없이 추정할 수 있는 방법론이지만 최근 성향점수매칭을 활용한 효과 추정의 한계점에 대하여 학계에서 논의가 이루어지고 있는 만큼, 커질 포레스트를 통한 처치효과 추정은 머신러닝의 활용 가능성을 탐색하는 데 유용한 지점인 것으로 보인다. 그다음 처치효과의 이질성 분석을 통해 처치효과가 특성변수에 따라 이질적인지 살펴보았다. 기존의 선형회귀모형은 이질성 분석에서 한계가 존재하기 때문에 커질 포레스트를 통한 이질적 처치효과를 탐색함으로써 재정정책의 효과성을 높이는 정책 설계의 가능성을 탐색하고자 하였다. 마지막으로 제Ⅵ장에서는 결론 및 정책적 함의점을 논의하였다.

---

## II. 정책금융의 개관

---

### 1. 정책금융 지원의 이론적 논의

#### 가. 정책금융의 두 가지 접근법

정책금융 관련 선행연구들을 살펴보면 정책금융의 개념을 논할 때 크게 두 가지 형태로 정책금융을 구분하는 경향이 있다. 첫 번째 개념은 시장실패로 인해 시장에서 해결되지 못하는 금융 수요를 충족시켜 주는 여신을 의미하고 두 번째는 시장실패의 해결에 초점을 맞추기보다는 정부의 정책 목표를 달성하기 위해 특정부문을 우대하여 제공하는 여신을 의미한다.

손상호(2013)는 정책금융의 개념을 “시간 및 공간상 발생하는 금융시장의 다양한 시장실패로 인해 시장의 상업적 원리로 해소하지 못하는 금융수요를 자금의 가용성·만기·금융비용 측면에서 보완하는 공공적 성격의 금융”으로 정의하였다. 그러나 이와 함께 기존의 연구들을 언급하면서 정부의 정책 목표를 위해 특정부문을 우대하여 공급하는 여신이라는 개념 또한 정책금융으로 볼 수 있다고 하였다.

장우현 외(2013, pp. 46~47)는 중소기업 지원정책을 생산적 중소기업 지원정책과 공공부양적 중소기업 정책으로 명시적으로 분류하였다. 이 분류에 따르면 생산적 중소기업에 대한 지원정책은 “시장의 실패로 인해 성장에 애로를 겪고 있는 중소기업을 자립적인 강소기업 중견기업으로 발전 성장시키기 위한 정책”으로 정의되며, 이러한 생산적 중소기업 지원정책은 학술적·논리적으로 충실히 뒷받침할 수 있다고 하였다. 한편 장우현 외(2013)는 공공부양적 중소기업 지원정책을<sup>5)</sup> “시장의 실패와는 관계없이 중소기업이 직면한

---

5) 정책담당자들이나 현장에서는 복지 차원의 중소기업 정책이라고 통칭하고 있다고도 한다 (장우현 외, 2013, p. 47.).

어려움을 완화하기 위해 현재의 애로를 기준으로 중소기업을 지원하는 정책”으로 정의하였으나 학술적·정책적으로는 공식화되어 있지는 않는 편이라고 하였다.

이러한 측면에서 볼 때 시장실패를 교정하기 위한 정책금융과 정부의 정책 목표를 달성하기 위한 두 종류의 정책금융은 궁극적인 정책 목표가 다르므로 성과지표를 선정하거나 성과평과 방법론을 적용할 때도 차이가 있을 수밖에 없다(장우현 외, 2013). 따라서 본고에서도 필요한 경우 이 두 종류의 정책금융을 명시적으로 구분하여 사용할 예정이다.

## 나. 시장실패로 인한 정책금융의 필요성

시장실패로 인한 정책금융의 필요성의 이론적 근거는 Stiglitz and Weiss(1981)에서 찾아볼 수 있다. Stiglitz and Weiss(1981)는 대부자와 차입자 사이에 비대칭 정보(asymmetric information)가 존재하는 경우 시장 균형 상태에서 자본에 대한 수요가 공급을 초과할 수 있다는 것을 이론적으로 보였으며, 이와 같은 금융시장의 초과수요 상태에서 일부 차입자들이 자금을 빌리지 못하는 상황을 credit rationing이라고 하였다.

Stiglitz and Weiss(1981)에서 credit rationing이 발생할 수 있는 경제학 모형은 다양하게 언급되어 있지만 credit rationing이 발생하는 논리를 가장 간단하게 보여주는 모형은 다음과 같다. 다수의 대부자와 다수의 차입자가 존재한다고 가정하자. 차입자들은 투자 계획을 가지고 있으며 차입자들의 투자계획은 모두 동일한 기대수익을 가져다주지만 위험의 크기는 다르다고 가정하자. 구체적으로 차입자들은 자신의 프로젝트의 투자 위험도  $\theta$ 를 가지고 있으며  $\theta$ 가 높을수록 평균이 높지만 분산이 커지는 mean preserving spread의 특징을 가지고 있다. 대부자들은 차입자들의 투자 위험도  $\theta$ 를 관찰할 수 없는 비대칭정보 상황에서 처해 있다.

차입자들이 투자계획을 실행하기 위해서는 담보  $C$ 를 잡히고 대출금  $B$ 만큼을 대출받아야 하며 대출받는 경우 이자  $rB$ 만큼을 원금과 같이 상환하기로 약속한다. 만일 총투자수입과 담보의 합( $R+C$ )이 원금과 이자  $(1+r)B$ 보다 작

다면 총투자수입  $R$ 과 담보  $C$ 를 지불하고 부도를 내는 유한책임(limited liability)을 가정한다. 대부자는 대출 이자율을 정하고 차입자는 대부자가 제시하는 이자율을 택할지 결정하게 된다. 대부자는 모두 동질적이고 완전 경쟁에 놓여 있어서 자금 조달을 하는 금리는 예금자들의 예금금리 수준에서 정해진다고 가정한다.

Stiglitz and Weiss(1981)는 이러한 가정하에서 credit rationing이 발생할 수 있음을 보였는데, 이를 보인 논리는 다음과 같다. 우선 차입자의 입장에서 유한책임하에서 더욱 높은 투자 위험도  $\theta$ 를 가지고 있는 것이 더 높은 기대수익을 가져다주게 된다. 투자 실패 시에는 차입자가 모든 수입을 내어 주고 담보  $C$ 만큼만 손해를 보게 되지만, 성공하는 경우에는 원금과 이자  $(1+r)B$ 를 제하고는 모두 순수익이 되기 때문이다. 이러한 대부자의 수익은 볼록함수(convex function)이므로 더 큰 투자 위험도를 가진 대부자의 기대이익이 더 커지게 된다. 반면 대부자의 경우 차입자들이 더 큰 위험을 가진 투자 계획을 시행하는 경우 더 낮은 기대수익을 얻게 되는데, 이는 대부자의 수익구조는 차입자가 성공하면 원금과 이자  $(1+r)B$ 만큼만 받지만 차입자가 부도를 내는 경우에는 총수입과 담보가치( $R+C$ )만을 가지게 되는 오목한 함수형태를 취하고 있기 때문이다.

이러한 수익구조 아래에서 대부자가 이자율을 높이면 이자율 상승으로 인해 기대수익이 커지는 효과가 있지만 또한 점점 더 위험의 크기가 더 큰 차입자만 대출을 하는 역선택(adverse selection)에 직면하여 대부자의 기대수익이 낮아지는 효과도 동시에 발생하게 된다. 후자의 효과로 인해 이자율이 증가하는 속도보다 대부자의 기대수익의 증가가 더 낮아지게 되며 심지어 이자율이 상승하는데도 불구하고 점점 더 위험이 높은 차입자들만 대출을 원하게 되어 기대수익이 오히려 감소하는 상황이 발생할 수도 있다. 만일 이자율의 증가에도 불구하고 기대수익이 감소하기 시작하는 이자율 수준이 존재한다고 하자. 이때의 이자율을  $r^*$ 라고 하고 이때의 자금 공급을  $s(r^*)$ 의 자금 수요를  $d(r^*)$ 라고 하자. 만일 해당 이자율에서 대출수요가 공급을 초과하는 초과수요가 있다고 하더라도( $d(r^*) > s(r^*)$ ) 대부자는 이자율을 높일 유인이

없기 때문에 이자율 조정으로 인한 초과수요의 청산이 이루어지지 않고 자  
금에 대한 만성적인 초과수요가 지속된다. 이러한 시장실패 상황에서는 대  
부자가 금융시장의 공급을 모두 만족시키지 못하고 신용을 일부 차입자들에  
게만 할당(credit rationing)하는 현상이 발생할 수 있다.

Stiglitz and Weiss(1981)는 이러한 시장실패의 현상이 비대칭정보 이외의  
다른 여러 상황에서 발생할 수 있는 것을 보였는데, 예를 들어 차입자들이  
높은 이자율로 자금을 빌리게 되는 경우 상대적으로 더 위험한 투자 프로젝  
트를 택하게 되는 도덕적 해이(moral hazard)가 발생하는 경우에도 유사한  
시장실패가 발생할 수 있다. 또한 credit rationing이 발생하는 경우 차입자  
들이 더 많은 담보를 제공하고 돈을 빌리려고 하더라도 대부자들이 이들을  
더 위험한 투자자들이라고 믿을 만한 합리적인 이유가 있다면 대출을 거절  
하여 자금의 초과수요가 해소되지 않을 수도 있음을 설명하였다. 이처럼 시  
장실패로 인해 자금의 만성적인 초과수요가 발생할 수 있는 상황은 여러 가  
지 존재하므로 이러한 시장실패를 보정할 수 있는 방법으로 정부가 적절하  
게 개입할 필요성이 존재한다.

#### 다. 중소기업에 정책금융을 지원해 줘야 하는 이유

많은 정책금융은 중소기업 지원에 초점을 맞추고 있다(장우현 외, 2013).  
중소기업의 경우 정보가 부족하기 때문에 앞에서 언급한 비대칭 정보의 문  
제점이 발생하기 쉬운 구조를 가지고 있기 때문이다. 장우현 외(2013)는 중  
소기업자금융시장에서 초과수요가 더 발생하기 쉬운 이유를 다음과 같이 설명  
하였다. 첫째, 중소기업은 특정 금융기관과 관계를 맺은 기간이 짧은 점이  
고, 둘째, 중소기업은 금융기관과의 거래기록이 부족하며, 셋째, 중소기업이  
보유하고 있는 담보가 부족하기 때문이다.

하지만 중소기업이 더 큰 정보 비대칭성의 문제를 가지고 있다고 해서 정  
책금융이 반드시 중소기업에 집중되어야 하며 일반기업은 지원해 주지 말아야  
한다고 주장하기는 어렵다. 이러한 주장이 설득력을 얻기 위해서는 다음과  
같은 근거들이 추가적으로 제시되어야 한다. 첫째, 일반기업들은 정보 비대

칭성의 문제에도 불구하고 금융시장에서 배제되지 않고 충분한 자금을 조달할 수 있다는 점이며, 둘째, 동일한 금액의 정책금융을 투자하는 경우 중소기업의 (사회적) 기대수익이 일반기업의 기대수익보다 더 높아야 한다는 것이다.

Stiglitz and Weiss(1981)의 연구에서 제IV장은 대부자가 차입자들을 그룹별로 구분할 수 있는 경우로 모형을 확장하여 분석하였는데, 해당 연구의 분석은 이상에서 제시한 두 가지 근거를 제공하고 있다고 볼 수 있다. 즉 대기업의 경우에는 정보 비대칭성의 문제에도 불구하고 금융시장에서 자금을 조달할 수 있으며, 사회 전체적으로 봤을 때는 대기업에 금융을 여신하는 것보다 중소기업에 여신을 제공하였을 때 더 높은 기대수익이 발생할 수도 있음을 보이고 있다. 이 확장된 분석을 하나의 예시를 통해서 제시하고자 한다.

우선 은행은 차입자들의 투자성향을 알 수는 없지만 차입자가 대기업인지 중소기업인지는 식별 가능하다고 하자. 그리고 대기업의 경우 100을 대출하여 투자하는 경우 확실히 110을 얻는다고 가정하자. 반면 중소기업의 절반(1 유형)은 100을 대출하면 확실히 130을 얻고 다른 절반(2 유형)은 100을 대출하면 50% 확률로 0을 얻으며, 50% 확률로 200을 얻는다고 가정하자. 은행은 100을 대출하기 위해서 20의 담보를 요구한다고 하자.

대기업의 경우 10% 이하의 이자율을 제시받으면 대출을 받을 것이므로, 은행은 10%의 이자율로 대기업에 대출하는 경우 10의 이득을 반드시 얻을 수 있다. 한편 1 유형의 중소기업은 이자율이 30% 이하인 경우 대출을 받을 의사가 있고 2 유형의 중소기업은 이자율이 80%<sup>6)</sup> 이하인 경우 대출을 받을 의사가 있다. 은행이 30%의 이자율을 제시하는 경우 두 유형의 중소기업이 모두 대출을 할 것인데, 이때 은행의 기대 이윤은 2.5에 불과하다.<sup>7)</sup> 은행이 이자율을 30% 이상으로 제시하는 경우 2 유형만 대출을 할 것이므로 은행이

6) 2 유형 중소기업의 기대이윤은  $0,5 \times (-20) + 0,5 \times (200 - 100(1+r))$ 이므로, 기대이윤이 0 이하가 되는  $r$ 은 0.8이다.

7)  $0,5 \times 30 + 0,5(0,5 \times (-80) + 0,5 \times (30)) = 15 - 12,5 = 2,5$

이자율을 80%까지 올린다고 해도 기대되는 수익은 0에 불과하게 된다.<sup>8)</sup> 즉 어떠한 이자율을 택하더라도 중소기업에 대출을 해 주는 것보다 대기업에 대출을 해 주는 것이 더 높은 이윤을 가져다주게 되므로, 은행의 자금이 한정적이라면 중소기업보다는 대기업에 대출을 해 주고 자금이 남는 경우에만 중소기업에 대출을 해 주는 상황이 발생할 수 있다. 문제는 이러한 상황에서 중소기업의 (사회적) 기대수익이 115로 대기업의 기대수익 110보다 더 높은 데도 불구하고 중소기업에는 자금의 투자가 이루어지지 않는다는 점이다.

이 예시에서 나타나는 현상이 시사하는 바는 첫째로 정보 비대칭성이 심한 중소기업은 금융시장에서 자금을 대여할 수 없지만 정보 비대칭성이 상대적으로 약한 대기업은 금융시장에서 자금을 조달할 수 있다는 점이다. 두 번째로는 사회 전체적으로는 중소기업의 기대수익이 더욱 높을 수 있는 데도 불구하고 대부자들이 사적인 이익만을 추구할 때 중소기업에 자금이 공급되지 않고 오히려 (사회적) 기대수익이 낮은 대기업에 자금이 공급된다는 점이다. 이러한 현상이 실제로 발생한다면 생산적 중소기업 지원정책은 더욱 설득력을 얻게 된다(장우현 외, 2013). 단지 중소기업이 작고 보호를 받아야 되는 기업이기 때문이 아니라 중소기업이 오히려 더욱 성장성이 있고 사회적으로 더 많은 수익을 창출해 낼 수 있기 때문에 적극적인 개입이 필요하다는 것이다.

Ordober and Weiss(1981)는 이상의 예시와 같은 상황에서 정부가 중소기업에 대출을 주지 않는 것을 금지한다면 사회적으로 더 높은 기대수익을 얻을 수 있음을 보였다. 이러한 이론적인 발견은 정부가 은행으로 하여금 중소기업에 일정 수준 이상의 금융을 대출해 주도록 의무비용을 부과하는 등의 정책을 정당화해 주는 근거라고 할 수 있다.

---

8)  $0.5 \times (-80) + 0.5 \times 80 = 0$

## 2. 정책금융에 관한 실증연구

이상에서 정책금융의 필요성 특히 중소기업에 대한 정책금융 지원의 필요성을 제시하는 이론적 근거들을 살펴보았다. 정책금융의 필요성을 제시하는 이론적 근거들은 주로 정보의 비대칭성에 의한 시장실패로 금융시장에 만성적인 초과수요가 발생할 수 있으며, 이러한 초과수요를 정부가 개입하여 해결해 줘야 한다는 메시지를 제시하고 있다. 또한 중소기업과 대기업이 공존하고 있는 사회에서 중소기업이 더 높은 (사회적) 기대수익에도 불구하고 더 높은 정보 비대칭성으로 인해 금융시장에서 배제될 가능성이 있음을 보였다. 이러한 이론적 근거는 정책금융 그 자체의 필요성이나 혹은 중소기업과 같은 특정 그룹에 대한 정책금융의 필요성을 잘 설명하며, 다양한 정책금융, 특히 중소기업이나 기술혁신 기업을 지원하는 정책금융을 집행하는 것에 대한 설득력을 제공한다.

하지만 앞에서 제시된 이론들은 정책금융 현장에서 실제로 정책금융을 집행하는 현장에서 어떠한 기업을 정책금융 대상으로 선정해야 하는지에 대한 함의를 제공하기는 어려운 측면이 있다. 근본적으로 위에서 언급된 이론적 근거들은 비대칭 정보로 인해 대부자들이 경제 전체 혹은 특정 그룹에 속한 기업들의 특성을 하나하나 모두 파악할 수 없기 때문에 기업들이 속해 있는 그룹 전체에 대한 통계적인 속성을 사용하여 의사결정을 한다고 가정하고 있다. 그렇기 때문에 이론에서 제시되는 함의는 경제 전체나 특정 그룹에 대한 메시지(중소기업 지원)를 제공할 수 있으나 동일한 중소기업 범주에 속하는 두 개의 기업이 있다고 할 때 어떠한 기업에 정책금융이 제공되어야 할지에 대해서는 침묵할 수밖에 없게 된다.<sup>9)</sup> 또한 이론은 시장실패로 인해 금융시장의 자본이 공급되지 않거나 왜곡되어 사회 전체의 산출물이 낮아지는 현상을 설명할 수 있으나 산출물 이외의 경제지표(예를 들면 기업의

---

9) 물론 모형을 무한히 확장하여 중소기업을 여러 개의 집단으로 나눈 후 (사회적) 기대수익이 더 높으나 정보 비대칭성이 더 큰 집단을 식별하여 이를 지원해야 한다는 함의를 이끌어 낼 수도 있겠으나 이는 동일한 직관의 반복일 뿐이며 현실적으로 모든 기업들이 가지고 있는 이질성을 모두 이론적으로 모형화하는 것은 불가능하다고 볼 수 있을 것이다.

생존 확률, 고용 증대 등)를 개선하고자 하는 정책목표를 지닌 정책금융의 집행에 대해서는 함의를 제공하기 어렵다는 측면이 있다.

이러한 측면에서 정책금융의 효과성에 대한 실증분석은 경제학 이론이 제시하지 못하는 연구 질문들에 대해서 많은 답변을 해 줄 수가 있다. 첫째로 실증분석을 통해 관찰 가능한 기업들의 이질적인 특성들이 정책금융의 성과에 미치는 영향을 파악하기에 용이하다. 동일한 중소기업의 범주에 속하는 기업들이라고 하더라도 모든 기업들은 고유한 이질성을 가진다. 따라서 이러한 이질성들이 정책금융 성과에 미치는 영향에 대한 정형화된 사실들에 이해가 높아질수록 한정적인 재원으로 운용되는 정책금융을 효율적으로 집행하는 데 도움이 될 것이다(장우현 외, 2013). 둘째로 기업의 생존확률이나 고용의 증대 등과 같이 이론에서는 직접적으로 논의되지 않는 기업의 성과가 정책금융을 통해서 어떠한 영향을 받는지 이해하는 데 도움을 줄 수 있다.

정책금융의 지원주체 및 수혜대상 그리고 지원형태가 다양하므로 정책금융 관련 실증연구들은 한정된 지면에 모두 언급하기 힘들 정도로 많다. 본 연구는 머신러닝을 통해서 정책금융의 대상자들을 더 효율적으로 선정할 수 있는가 하는 질문을 가지고 있으므로, 이와 관련하여 두 가지 주제로 구분하여 기존의 실증분석 연구들을 소개하고자 한다. 첫 번째는 현재의 정책금융 제도에서 수혜자가 가지는 특징들에 대해서 연구한 논문들을 소개하고자 한다. 두 번째로는 기업들의 이질적인 특징에 따라 정책금융의 성과가 달라질 수 있음을 보이는 논문들을 소개하고자 한다. 본문에서 다루지 못한 연구들은 <표 II-1>에 간단하게 요약하여 관심 있는 독자들이 참고할 수 있도록 제시하였다.

## 가. 누가 정책금융을 지원받는가?

우선 현재의 정책금융 제도가 어떠한 기업들을 주로 지원하는지에 대한 정보를 실증연구들을 통해 알아보고자 한다. 장우현 외(2013)는 신보, 기보, 중진공이 지원하고 있는 기업들의 현황을 수혜지도를 통해 살펴보았다. 수혜지도는 2010년을 기준으로 볼 때 기보, 신보, 중진공 모두 30인 이하의 소

기업에 지원이 집중되었으며, 특히 업력이 짧은 7년 미만의 기업들의 경우 대부분의 지원이 30명 이하의 기업에 지원되고 있음을 보였다. 흥미로운 점은 규모가 30명 이하이긴 하지만 업력이 10년 이상인 중소기업도 많은 지원 건수를 보였다는 것이다. 저자들은 업력이 긴 기업은 시장 금융을 이용해서 자금수요를 충족시킬 여지가 있는 기업들인데도 정책금융의 대상이 되는 것은 정책금융기관이 위험이 높은 창업기업에 대한 지원을 많이 하는 만큼, 상대적으로 안정적인 10년 이상의 업력을 가진 기업을 지원함으로써 포트폴리오 구성을 안정화시키기 위한 것으로 해석하였다. 또한 지원건당 수혜금액은 상시근로자 300명 내외의 5년 미만의 기업에 집중되어 있었는데 이러한 기업들은 기존 그룹에서 분사한 그룹일 가능성이 높을 것임을 추측하며 중소기업들이 정부의 지원을 받기 위해서 규모를 늘리지 않으려고 하는 피터팬 증후군의 증거일 수도 있다고 언급했다.

최현정·하봉찬(2015)은 2007년에 중소기업진흥공단, 기술보증기금, 신용보증기금이 정책금융을 지원한 기업 16만 652개사와 지원받지 않은 기업을 대상으로 프로비트(probit) 모형을 적용하여 정책금융 지원이 어떠한 기준으로 행해지고 있는지에 대해 분석하였다. 프로비트 모형을 통해 분석한 결과, 종업원 수가 적은 소기업일수록 정책금융을 지원받을 확률이 높으며 종업원 수가 일정 수 이상이 되면 종업원의 수보다는 부채비율이나 총자산 회전율과 같은 재무비율이 높을수록 지원 확률이 높아지는 것을 확인하였다.

김상봉·김정렬(2013)은 2005~2010년에 신용보증지원을 받은 기업과 보증을 신청하였으나 받지 못한 기업의 자료를 가지고 헤크만 2단계 추정방식을 이용하여 신용보증지원의 효과를 분석하였다. 1단계 로짓분석 결과는 업력이 길거나 총자산 회전율이 큰 기업일수록 혹은 기업의 총자산 증가율과 매출액이 작을수록 보증기업이 될 확률이 높은 것으로 나타났다.

박노옥 외(2016)는 중소기업 지원사업의 효과성을 분석하기 위해 중소기업연구원 통합관리시스템의 지원이력과 한국기업데이터의 재무정보를 연계하여 분석하였으며 성향점수 추정법을 사용하였다. 2011~2014년의 데이터를 사용하였으며 기업들을 6개의 사업군으로 분류하여 분석하였다. 분석결과

창업부문의 성향점수 추정 결과를 보면 매출액이 큰 기업일수록 정책지원금을 받을 확률이 낮아지는 경향이 있었고, 부채비율이 높은 기업일수록 창업 지원을 받을 확률이 낮아지는 것으로 나타났다. 한편 영업이익률, R&D 집중도는 창업지원 확률에 유의한 양(+)의 효과가 있었다. 금융부문의 경우 50인 미만 기업들에 지원 혜택이 돌아간 것으로 나타났다. 2011년도를 제외하면 매출 규모가 작을수록 정책지원을 받을 가능성이 높아지는 것으로 나타났다. 또한 R&D 투자의 높은 집중도를 보이는 기업일 수록 정책지원의 혜택을 받았으며, 업력이 낮을수록 금융지원을 받을 확률이 높아졌다. 기술분야의 경우 매출 규모가 작고 부채비율이 낮으며, 영업 이익률이 높고 R&D 집중도가 높은 기업이 지원의 혜택을 받았다. 고용 측면에서는 고용 규모가 큰 기업일수록 정책지원을 받을 확률이 낮아졌다. 한편 수출부문의 경우 사업체 매출액 및 고용 기준 규모가 작고 R&D의 집중도가 높은 사업체들이 지원혜택을 받을 확률이 높은 것으로 나타났다. 내수부문의 경우 매출액 규모가 작으면서 업력이 짧고 고용 규모가 작을수록 정책지원을 받을 확률이 높았다. 경영부문은 전반적으로 매출액 규모는 크지만 고용 규모가 작은 기업이 정책지원을 더 많이 받는 것으로 나타났다.

Kang et al.(2008)의 경우 신용보증기금과 기술보증기금의 2001~2004년 신청기업을 대상으로 가상패널자료를 만들어 중소기업 지원정책이 기업의 생존력과 성과에 어떠한 영향을 끼치는지를 사업 규모와 산업별로 분석하였다. 분석방법으로는 OLS와 MLE를 사용하였으며, 분석 결과 이전에 지원받은 이력이 있는 기업이 상대적으로 성장 가능성이 높으나 충분한 배경 정보가 없는 기업에 비해 지원 대상으로 선정될 유인이 높은 것으로 나타났다.

이상의 연구를 종합해 보면 공통적으로 종업원 수가 적은 기업들이 일반적으로 정책금융 지원대상이 되는 것을 알 수 있다. 또한 기업의 규모가 커지면 기업의 재무적인 상황이 정책금융에 대상 여부에 영향을 주는 것을 확인할 수 있다. 다만 업력이 긴 기업들이 정책금융의 대상으로 선정되는 경향도 있음을 확인하였는데, 업력이 긴 기업들은 금융시장에서 자금수요를 충족시킬 여지가 있는 기업으로 볼 수 있는 만큼 이러한 기업들에 대한 지원

성과가 창업기업을 지원하는 것과 비교해 두드러지는지에 대한 평가는 지속적으로 이루어져야 된다고 볼 수 있을 것이다.

## 나. 정책금융 수혜효과의 이질성

이상엽 외(2014)는 중소기업청에서 2002년부터 2012년 사이에 수행한 모든 정책금융 프로그램에 대한 수혜기업 전수자료와 한국기업데이터(Korea Enterprise Data, KED)의 ‘한국기업데이터’를 이용하여 수혜기업과 비수혜기업의 재무정보를 이용해 정책금융 효과를 추정하였다. 분석을 위해 이중차분 성향점수매칭기법(Difference-in Difference Propensity Score Matching method)을 사용하였으며 중소기업청에서 수행한 전체 사업을 기업의 업력에 따라 ‘창업기 지원사업’, ‘성장기 지원사업’, ‘정체기 및 재도약기 지원사업’으로 구분하여 분석하였다. 분석 결과 창업기업 지원사업의 경우 수혜기업 집단의 자기자본순이익률(수익성), 영업이익증가율(성장성), 부채비율(안정성)이 2007년 대비 2012년에 향상되는 경향을 발견하였으나 ‘성장기 지원사업과’, ‘정체기 및 재도약기 지원사업’의 경우에는 수혜기업 집단의 성과 향상을 발견하지 못하였다

우석진·빈기범(2014)은 신용보증기금에 비해서 상대적으로 금융시장 접근성이 떨어지고 위험한 기업을 대상으로 기술평가를 하여 지원되는 기술보증이 기업 성과에 미치는 효과를 분석하였다. 한국기업데이터의 2002~2008년 재무정보와 기술보증 수혜 여부 데이터를 가지고 분위회귀모형을 사용하여 주요 재무비율의 각 분위(10%, 25%, 50%, 75%, 90%)에서 수혜기업과 비수혜기업 간 기술보증 전후의 재무비율 변화를 비교하여 이중차분모형으로 분석하였다. 분석 결과, 회계·재무적 성과가 상대적으로 낮은 기업(10% 분위)이 기술보증을 받은 경우 총자산수익률은 6.4%, 총자기자본이익률은 12~15.4% 개선되는 효과를 보여주었다. 그러나 회계·재무적 성과가 상대적으로 양호한 기업은 기술보증 지원으로 인한 수익성 개선 효과가 미미한 것으로 추정되었다. 기업의 안정성을 대표하는 부채비율을 분석한 경우에도 안정성이 상대적으로 취약한 기업에서만 기술보증의 안정성 개선 효과가 나타났다.

이기영·우석진(2015)은 공적 신용보증이 한계기업의 생존에 어떠한 영향을 미치는지 실증분석하였다. 구체적으로 2005년에 기술보증과 신용보증을 받은 기업을 3년간 추적하여 2008년까지의 생존 여부를 파악하였다. 생존분석모형을 통해 생존확률을 분석한 결과, 기술보증 수혜기업 중 한계기업의 수명은 비한계기업에 비해 37.9% 짧았다. 신용보증 수혜기업 중 한계기업의 수명은 비한계기업에 비해 31.7% 짧았다. 저자들은 이러한 결과를 바탕으로 한계기업에 대한 보증지원은 더 많은 보증과 더 긴 지원기간에도 불구하고 부도 확률이 높으며, 이는 재정적 부담의 증가로 이어질 수 있고 한계기업의 고용 유지라는 정책적 목표도 달성하기 어려울 수 있다는 시사점을 제기하였다.

장우현·우석진(2015)은 대기업과 중소기업 간의 수직관계를 고려하여 중소기업 지원정책의 효과를 분석하였다. 저자들은 중소기업에 대한 정책적 지원효과가 대기업과 협력사 간 납품단가 등의 거래조건의 조정 등으로 대기업에 흘러가는 경우를 ‘빨대효과’라고 정의하였다. 자료의 제약으로 인해 광업·제조업 분야의 중소기업과 서비스 분야의 중소기업으로 구분하여 분석하였다. 광업·제조업은 레빈슨-페트린 생산함수 추정법을 통해 사업체별 총요소생산성을 도출하였으며, 수직관계를 고려하지 않은 정책금융의 단독 효과를 성향점수 짝짓기(PSME) 추정법을 통해 추정하고, 수직관계를 고려한 빨대효과는 다중회귀분석모형을 통해 추정하였다. 서비스 산업의 중소기업 빨대효과는 이중차분법을 통해 정책자금의 효과를 협력사와 비협력사 각각에 대해 추정한 다음, 이를 다시 차분하여 계산하는 삼중차분 추정법을 통해 추정하였다. 추정 결과 광업·제조업의 경우 수직관계의 존재만으로는 중소기업의 성과에 해가 되지 않으나 수직관계에 있는 중소기업이 정책금융을 받는 경우 수직관계에 의해 그 효과가 대기업으로 흘러가는 빨대효과가 존재하여 중소기업의 생산성에 부정적인 영향을 끼치는 것을 발견하였다. 서비스업의 경우 협력사가 정책자금을 받는 경우 총자산은 커지지만 수익률은 감소하는 것을 발견하였으며, 이는 대기업과의 단가조정으로 인한 빨대효과의 결과일 수도 있다는 시사점을 제시하였다.

장우현 외(2020)는 동태적인 산업환경의 변화와 산업이 처한 상황을 고려하여 분류한 산업분류 유형을 활용하여 중소기업 정책금융 사업에 대한 분석을 하였다. 해당 연구에서는 영업이익률과 영업이익 증가율을 기준으로 산업을 우량, 성장, 축소, 침체 네 가지로 분류하였으며, 이렇게 새롭게 분류된 산업단위 내에서 정책금융의 효과를 분석하고 또한 새로운 산업분류를 이용하여 기업단위에서 정책금융 효과를 분석하였다. 산업단위 정책금융 효과의 분석에서는 산업과 시간의 효과를 배제할 수 있는 패널고정효과모형을 사용하였다. 분석 결과, 네 가지 산업 분류보다는 중소기업 여부를 포함해서 분류하는 확장된 분류 방식이 중소기업총자산영업 이익률 변화 식별에 도움을 주며, 정책금융이 1조원 증가할 때 산업의 총자산영업이익률은 1.03%p 유의하게 감소하는 것을 확인하였다. 기업단위 정책 분석은 KISData와 한국기업데이터 자료를 모두 분석하였다. KISData 분석 결과, 정책금융을 지원받은 경우 기업의 총자산영업이익률이 비지원기업에 비해 1.74%p 감소하였으며 이러한 음(-)의 효과는 우량, 성장, 축소, 침체 산업분류별로 따로 분석하여도 여전히 유효하였다. 한국기업데이터 자료를 분석한 결과, 또한 정책금융 지원이 총자산영업이익률에 음(-)의 영향을 미치는 것으로 나타났다. 추가적으로 유형자산의 변화에 대한 분석한 결과, 정책금융이 유형자산을 증가시키는 것으로 나타나 양적 투자를 증가시키는 것으로 해석할 수 있었으며, 특히 이는 축소와 침체 산업에서 두드러지게 나타났다. 재고자산에 미치는 영향을 분석할 경우 축소 및 침체 산업이 정책금융을 받을 경우 더 높은 수준의 재고자산을 보유하는 것으로 나타났으며, 이는 정책금융이 공급과잉인 시장의 상황을 더욱 심화시킬 수 있다는 시사점을 제시하였다.

장우현 외(2013)는 중소기업 정책자금의 기업의 업력에 따라 어떠한 효과를 미치는지 분석하였다. 창업 초기 기업과 업력이 상대적으로 긴 기업의 재무성과에 미친 정책자금의 효과를 이중차분 추정의 아이디어를 차용한 분위회귀분석을 통해 분석하였다(0.1, 0.25, 0.5, 0.75, 0.9 분위). 수혜기업의 재무적 성과를 살펴보기 위해서 크게 수익성 지표(ROA, ROE, OPR)와 성장성 지표(GSALES, GOP)를 사용하였으며, 단기적인 효과는 수혜 1년 후의 변화, 중기

효과는 수혜 3년 후의 변화를 통해 분석하였다. 분석 결과, 수익성 변화가 낮은 분위의 사업체에게 지원한 경우에는 업력이 높은 사업체의 수익성 개선이 창업기업에 비해서 높게 나타났지만 수익성 변화가 높은 분위의 사업체에게 지원한 경우에는 업력이 긴 사업체의 수익성 개선 정도는 창업기업에 비해 오히려 낮음을 알 수 있었다. 이는 업력이 오래된 기업 중 수익성 개선이 나뉘었던 경우는 해당 기업의 일시적 부진일 가능성이 존재하기 때문에 해당 기업들에 대한 지원은 수익성 변화를 개선시켜 줄 수 있다는 시사점을 제시한다.

Caselli et al.(2019)는 2007~2009년 이탈리아 정부의 중소기업 보증프로그램(Fondo Centrale di Garanzia - Central Guarantee Fund, FCG)이 중소기업의 수익성에 미치는 영향을 회사 규모와 업종별로 구분하여 살펴보았다. 2005~2011년 자료를 사용하였으며 글로벌 금융위기가 포함된 2007~2009년에 FCG 수혜를 처음 받은 기업을 처리집단으로, 동 기간 내에 한 번도 수혜를 받지 않은 기업을 통제집단으로 설정하여 성향점수매칭과 이중차분모형을 사용하였다. 분석 결과, FCG는 경기침체 기간에 수혜기업의 수익성을 증대시키는 역할을 하였으며 효과는 회사 규모별 업종별로 차이가 있음을 보였다. 구체적으로 회사 규모가 작을수록 종업원의 수가 적을수록 FCG의 효과가 유의미했으며, 제조업에서 그 효과가 가장 컸다.

〈표 II-1〉 중소기업 대상 정책금융의 성과분석 선행연구 요약

제목	연구목적	연구방법	연구대상	독립변수	종속변수	주요 연구내용
이준원(2019). 기술금융 중소기업과 일관성: 중소기업의 경영성과 비교분석	기술력 평가모형을 통해 선발된 기술금융 중소기업의 3년간 경영성과를 일반 중소기업과 비교하여 기술금융의 효율성 및 필요성을 검증	성향점수매칭(PSM)을 통해 기술금융 중소기업의 신원본야, 업력, 종업원 수, 자산 및 자본 규모가 유사한 중소기업업을 선별	기술금융 중소기업		성향성(매출액 증감률), 수익성(영업이익률), 안정성(부채비율)	기술금융 중소기업은 일반 중소기업에 비해 안정적인 매출액 증가율을 유지하며 이를 통해 점차 부채비율이 정상화됨
권세훈(2012). 기술신용보증기금의 특성과 성장성에 관한 실증분석	기술신용보증기금의 수익성, 성장성, 안정성에 미치는 영향을 분석	1999~2010년까지 기술보증기금의 신용보증 1회 이상 제공받은 기업에 대상으로 확률효과 패널 회귀분석을 함	기술보증기금의 신용보증 1회 이상 제공받은 중소기업	1) 최초 보증 개시연도에 해당하는 더미변수로 단기효과를 설명 2) 보증 개시 이후 모든 연도에 해당하는 더미변수로 장기효과를 설명	수익성(수익률, 영업이익률), 성장성(신성장기율, 자산 대비 매출증감률), 안정성(부채비율)	1) 기술보증 기업은 비보증 기업에 비해 수익성이 낮고 부채비율이 높지만 성장성이 높은 기업이라는 선택효과를 확인 2) 기술보증의 효과는 단기보다는 중장기적인 자원효과가 있음을 확인
신상훈 · 박정희(2010). 신용보증지원이 중소기업의 성장성에 미치는 효과에 대한 패널분석	신용보증 지원이 기업의 성장성에 어떠한 영향을 미치는지 분석	2000~2007년까지의 신용보증기금 및 한국기업데이터를 사용하여 모형을 고정효과 모형으로 사용함	신용보증기금에 의해 보증받은 기업	설명변수: 보증잔액의 존재 여부 통제변수: 유형, 업력, 기업 규모, 자본금비율, 유동비율, 제조원가비율, 영업비율, 1인당부가가치, 1인당자산, GDP, 원/달러 환율, 주가수상승률 등을 사용	총자산이익률, 매출영업이익률, 매출증감률, 부가가치증감률	보증지원은 기업의 수익성과 성장성에 부정적인 영향을 미친 것으로 나타남

〈표 II-1〉의 계속

제목	연구목적	연구방법	연구대상	독립변수	종속변수	주요 연구내용
<p>노동화(2010), 중소기업 정책자금의 미시적 성과 분석과 역할에 관한 연구</p>	<p>정책자금이 수혜기업의 경영성과 개선에 미친 영향과 민간금융시장 접근성 제고 효과를 실증분석</p>	<p>2002~2008년까지 중소기업진흥공단 정책자금 대출을 받은 중소기업을 대상으로 하였으며 정책자금 수혜 여부에 따른 지원효과를 구조 분석</p>	<p>중소기업 진흥공단의 정책자금 수혜를 받은 기업</p>	<p>실면변수: 정책자금의 규모와 대출 유형 통제변수: 업종, 지역, 업력, 고용, 정책자금 수혜시점의 부채비율, 담보능력, 연도더미</p>	<p>매출 영업이익률, 부채구성비율, 매출액증가율, 장기차입금비율, 자금차입비용비율</p>	<p>1) 정책자금은 중소기업의 수익성과 성장성 측면에서는 단기에 경영성과를 개선시키는 효과가 나타남 2) 금융접근성 측면에서는 단기에 일시적인 긍정적인 효과가 있었으나 장기효과 측면에는 한계가 존재</p>

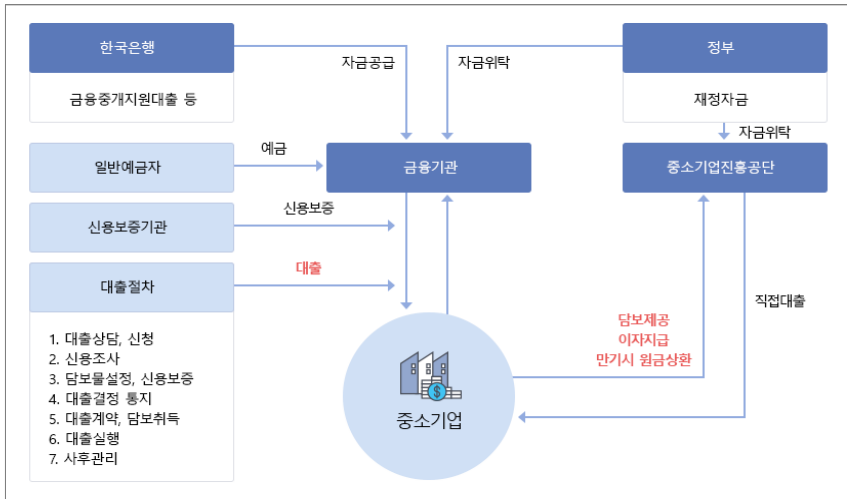
자료: 저자 작성

### 3. 중소기업 정책금융 지원 현황

#### 가. 기본체계

중소기업 정책금융은 중소기업의 신용도 혹은 담보력의 보완을 위한 제도, 이용 가능한 자금공급을 확대하기 위한 제도, 재정자금에 의한 지원제도로 구성된다. 중소기업의 가용자금 공급을 위해 한국은행은 금융중개지원대출제도와 중소기업대출비율제도 등을 통해 금융기관의 중소기업 대출 취급 촉진 및 금융지원 확대를 유도한다. 금융중개지원대출은 한국은행이 중소기업을 지원하기 위해 시중은행에 저리로 지원하는 자금이며, 중소기업대출비율제도는 시중은행들이 중소기업대출비율을 맞추지 못하는 경우 한국은행이 금융중개지원대출 한도에서 미달금액의 일부를 차감하는 제도이다.

[그림 II-1] 중소기업금융 지원체계



자료: 한국은행, 「중소기업 금융지원제도 - 기본체계」, <https://www.bok.or.kr/portal/main/contents.do?menuNo=200302>, 검색일자: 2021. 10. 21.

또한 취약한 담보력 및 신용도로 인해 자금조달에 애로를 겪는 중소기업이 금융기관으로부터 원활히 대출받을 수 있도록 신용보증기금, 기술보증기금,

지역 신용보증재단이 신용보증 제공을 중심으로 각종 금융지원을 시행하고 있다. 마지막으로 재정자금에 의한 중소기업 지원은 중소기업청을 비롯한 중앙부처에 의해 실시되며, 중소기업진흥공단은 중소기업청의 위탁을 받아 중소기업에 대한 이러한 직접대출과 은행에 위탁하는 대출을 하는 간접대출 업무를 수행하고 있다.

## 나. 중소기업 정책금융 통계

〈표 II-2〉는 중소기업의 금융지원 현황을 정책자금 및 신용보증에 따른 지원으로 구분하여 제시하였다. 정책자금은 중소벤처기업부 소관 중소기업 정책자금을 기준으로 작성되었는데, 지원금액은 2011년 3조 4,016억원에서 2019년 4조 3,580억원으로 3~4조원 내외에서 꾸준히 유지되어 왔다. 2020년에는 코로나19 확산에 따른 중소기업 지원대책에 따라 지원금액은 6조 2,900억원으로 증가하였다. 한편 신용보증 지원 현황은 신용보증기금, 기술신용보증기금, 지역 신용보증재단으로 구분하여 제시하였는데, 3개 기금에서 지원 규모가 가장 큰 기금은 신용보증기금으로 지원 규모는 2011년 38조원에 서 꾸준히 성장해 2019년 47조원에 이르렀으며, 2020년에는 코로나19 확산에

〈표 II-2〉 중소기업 금융지원 현황

(단위: 억원, 개, 조원)

구분		2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
정책자금 지원 현황 <sup>1)</sup> (억원, 개)	지원금액	34,016	36,508	48,929	44,481	39,019	45,512	46,661	44,150	43,580	62,900
	지원 업체 수	23,580	28,334	40,244	55,805	17,796	21,224	22,836	17,475	15,865	24,407
신용보증 지원 현황 <sup>2)</sup> (조원)	신용 보증기금	38	39	41	41	41	42	44	46	47	55
	기술신용 보증기금	17	18	19	19	20	21	21	22	22	25
	지역신용 보증재단	14	14	14	15	16	18	19	21	23	39

주: 1) 정책자금은 중소벤처기업부 소관 중소기업 정책자금 기준(연간 공급기준, 지방 중소기업육성자금 및 소상공인 정책자금 제외)

2) 신용보증금액은 잔액 기준

자료: e-나라지표, 「중소기업 금융지원 현황」, [https://www.index.go.kr/potal/sits/idxMain/selectPoSitsIdxMainPrint.do?idx\\_cd=2780&board\\_cd=INDEX\\_001](https://www.index.go.kr/potal/sits/idxMain/selectPoSitsIdxMainPrint.do?idx_cd=2780&board_cd=INDEX_001), 검색일자: 2021. 10. 20.

따른 경제침체 대책의 일환으로 지원 규모가 급격히 증가해 55조원에 이르렀다. 코로나19 확산에 따른 대응정책으로 지원 규모가 급격하게 증가한 기금은 지역 신용보증재단으로 2011년(14조원)과 2019년(23조원) 사이에서 완만하게 증가하였으나 2020년에는 전년도 대비 69.6% 증가한 39조원에 이르렀다.

#### 4. 지원기관별 주요사업

##### 가. 가용자금 공급

##### 1) 한국은행의 금융중개지원대출제도<sup>10)</sup>

한국은행의 금융중개지원대출에 대한 설명에 따르면 금융중개지원대출제도는 한국은행이 시중은행에 공급하는 대출의 총한도를 사전에 정하고 정해진 총한도 내에서 은행별로 중소기업의 대출실적 등의 기준을 활용해서 은행별 한도를 배정하여 저리자금을 지원하는 제도이다. 대출의 총한도와 프로그램별 한도 및 한도 유보분은 금융통화위원회에서 금융, 경제동향, 그리고 중소기업의 자금사정 등을 고려하여 필요시 수시로 조정한다.

〈표 II-3〉 금융중개지원대출 프로그램별 구성 내역

(단위: 조원, %)

구분	도입시기	한도	금리	지원 목적
무역금융지원 프로그램	1994년 3월	2.5	0.25	수출금융 지원
신성장·일자리지원 프로그램	2013년 6월	13.0	0.25	창업기업, 일자리창출기업, 소재부품장비기업 지원
중소기업대출안정화 프로그램	2017년 9월	5.5	0.25	중소기업 대출에 대한 변동성 완화, 경기 대응 등
지방중소기업지원 프로그램	1994년 3월	5.9	0.25	지방중소기업 지원

자료: 한국은행, 「금융중개지원대출제도」, <https://www.bok.or.kr/portal/main/contents.do?menuNo=200303>, 검색일자: 2021. 10. 20.

10) 금융중개지원대출제도에 대한 내용은 한국은행의 「금융중개지원대출제도」를 요약하였다 (한국은행, <https://www.bok.or.kr/portal/main/contents.do?menuNo=200303>, 검색일자: 2021. 10. 20.).

현재 금융중개지원대출은 네 개의 프로그램(무역금융지원, 신성장·일자리지원, 중소기업대출안정화, 지방중소기업지원)으로 운용되고 있다. 먼저 무역금융지원 프로그램은 물품 및 용역 수출 촉진을 위해 도입된 제도로, 수출업자와 수출용 원자재 및 완제품 제조에 필요한 자금을 대출한 은행의 취급실적을 고려해서 은행별로 한도를 배정하게 된다. 다음으로 신성장·일자리지원 프로그램은 창업 7년 이내의 창업중소기업과 일자리창출에 기여하는 중소기업 그리고 소재부품장비 제조 중소기업을 지원을 위해 도입되었으며, 은행들의 해당 기업에 대한 대출 취급 실적을 고려해서 은행별로 한도를 배정하게 된다. 한편 중소기업대출안정화 프로그램은 중소기업의 대출 변동성을 완화하고 신용경로의 원활한 작동을 통한 통화신용정책의 효율성을 제고하기 위해 도입되었고, 은행별로 한도 배정은 앞선 제도와 동일하게 은행별 중소기업에 대한 신용공급 등을 고려하여 이루어진다. 마지막으로 지방중소기업지원 프로그램은 지역 간 균형발전을 도모할 목적으로 도입되었으며, 한국은행 지역본부별로 한도를 배정하는데 배정 시 은행의 지방중소기업에 대한 대출실적과 지역별 경제사정 등을 감안한다. <표 II-3>은 금융중개지원대출 프로그램을 도입 시기, 한도, 금리, 지원 목적에 따라 요약한 것인데, 한도액은 프로그램별로 상이하나 적용되는 금리는 일정함을 알 수 있다.

## 나. 신용보증지원제도

### 1) 신용보증기금

신용보증기금은 담보 능력이 부족한 기업을 대상으로 기업의 운영과 신규 창업에 소요되는 자금의 용자에 대한 보증을 제공함으로써 금융회사로부터 대출을 받을 수 있도록 지원한다.<sup>11)</sup> 신청대상은 영리를 목적으로 사업을 영위하고 있는 중소기업이며 도박·사행성게임, 사치, 향락, 부동산 등 일부업종은 신청이 제한된다. <표 II-4>는 신용보증기금의 주요 보증제도를 정리한

11) 신용보증기금, 「신용보증기금>주요업무>신용보증>이용안내」, [https://www.kodit.co.kr/work/crdt\\_guar/crdt\\_guar/use\\_guid/user\\_outline.jsp](https://www.kodit.co.kr/work/crdt_guar/crdt_guar/use_guid/user_outline.jsp), 검색일자: 2021. 10. 30.

것으로, 일반운전자금, 유망창업기업성장지원프로그램, 구매자금, 시설자금, 수출중소기업종합지원프로그램, 지식재산보증, M&A 보증제도 등이 포함된다.

〈표 II-4〉 신용보증기금 주요 보증제도

구분	내용
일반운전자금	원재료의 구매, 생산, 판매활동 등 기업 운영에 소요되는 자금을 대한 보증지원상품
유망창업기업성장지원 프로그램	창업 초기 기업의 성장단계별로 '예비창업보증 → 신생기업보증 → 창업초기 보증 → 창업성장보증'으로 구분하여 지원하는 맞춤형 보증 프로그램
구매자금	상업어음 이용을 줄이는 대신 구매기업이 금융회사 등으로부터 대출받아 납품대금을 결제할 수 있도록 하는 금융결제수단에 대한 보증
시설자금	사업의 확장, 신설 등을 위해 사업용 공장, 고정적인 기업 설비의 취득 등에 소요되는 시설자금 지원을 위한 보증
수출중소기업종합지원 프로그램	수출을 희망하거나 수출실적을 보유한 중소기업의 수출역량 단계에 따라 지원하는 맞춤형 보증 프로그램
지식재산 보증	지식재산 창출을 위한 R&D, 기술거래 등 소요자금을 보증
M&A 보증	합병, 주식인수, 영업양수 등 M&A 소요자금을 보증

자료: 금융감독원(2021), p. 95.

신용보증기금의 최근 지원 현황을 살펴보면 2015년 총보증잔액은 49조 1,862억원에서 2019년 52조 2,161억원으로 지속적으로 증가하였는데, 이에 상응하여 기본 재산이 증가하지 않는 것으로 나타나 운용배수는 2015년 9.0 배와 비교하여 2019년 9.8배로 높아진 것으로 나타났다.

〈표 II-5〉 최근 5년간 기본재산 및 운용배수 현황

(단위: 억원, 배)

구분	총보증잔액(A)	기본재산(B)	운용배수(A/B)
2015년	491,862	54,380	9.0
2016년	495,041	50,133	9.9
2017년	508,463	51,214	9.9
2018년	505,317	50,909	9.9
2019년	522,161	53,377	9.8

자료: 신용보증기금, 『신용보증기금 연차보고서』, 각 연도.

한편 <표 II-6>에서는 연도별 보증공급 및 신규보증 금액을 정리하였는데, 2019년 기준 총보증공급은 49조 879억원으로 나타났으며, 이 중 신규보증공급은 11조 8,100억원으로 나타났다.

<표 II-6> 보증공급 및 신규보증 금액

(단위: 억원)

구분	2015년	2016년	2017년	2018년	2019년
총보증공급 <sup>1)</sup>	458,953	461,069	473,951	468,693	490,879
일반보증 신규보증공급 <sup>2)</sup>	107,790	117,514	117,788	112,540	118,100

주: 1) 총보증공급 = 일반보증 + 유동화보증 + 시장안정특별보증

2) 신규보증공급 일반보증 기준

자료: 신용보증기금, 『신용보증기금 연차보고서』, 각 연도.

## 2) 기술보증기금

기술보증기금은 담보능력이 미약한 기업이 보유하고 있는 무형의 기술을 대상으로 기술성, 시장성, 사업성 등 미래가치 위주로 평가하여 보증서를 발급함으로써 금융기관 등으로부터 자금을 지원받을 수 있도록 지원하는 제도다.<sup>12)</sup> 신청대상은 신용보증기금과 달리 신기술사업을 영위하고 있는 중소기업으로, 제조, IT, 연구 및 개발, 전문서비스업종 등이 주로 해당한다. <표 II-7>은 기술보증기금의 주요 보증제도를 정리한 것으로, 예비창업자 사전보증, 청년창업기업 우대, 수출기업 지원, 일자리창출 지원 프로그램, 소셜벤처 임팩트보증, R&D보증, 지식재산(IP) 평가보증, 우수기술사업화지원보증, 대·중소기업 상생협약보증 등이 포함된다.

<표 II-7> 기술보증기금 주요 보증제도

구분	내용
예비창업자 사전보증	창업 준비단계에서 지원가능 금액을 제시하고 창업 후 지원
청년창업기업 우대	창업 후 5년 이내로 경영주가 만 17~39세인 기술창업기업

12) 기술보증기금, 「기술보증기금>보증이용안내>보증이용절차」, <https://www.kibo.or.kr/main/work/work010101.do>, 검색일자: 2021. 10. 30.

〈표 II-7〉의 계속

구분	내용
수출기업 지원	수출중소기업의 무역금융에 대해서 기업당 70억원까지 보증지원하고, 수출실적이 연간 매출액의 10% 이상인 경우 보증비율 우대(90%)
일자리창출 지원 프로그램	‘일자리창출기업’을 대상으로 고용환경에 따른 유형별 우대지원(보증비율, 보증료, 한도 등 우대)
소셜벤처 임팩트보증	사회적 가치와 경제적 성과를 동시에 창출하는 ‘소셜벤처기업’에 대한 보증
R&D보증	R&D 단계별 맞춤형 R&D 자금을 보증
지식재산(IP) 평가보증	지식재산(IP) 가치평가 후 동 가치 이내에서 보증
우수기술사업화지원 (TECH밸리) 보증	대학, 연구기관, 대기업 연구소 재직·퇴직자가 창업한 우수 기술기업에 대한 보증
대·중소기업 상생협약보증	대기업 및 은행과 협약을 체결한 중소기업에 대한 보증

자료: 금융감독원(2021), p. 96.

기술보증기금의 최근 지원 현황을 살펴보면 2015년 총보증잔액은 20조 7,096억원에서 2019년 22조 1,122억원으로 지속적으로 증가하였는데, 기본재산은 오히려 2015년 2조 2,951억원에서 1조 6,751억원으로 감소하였다. 이에 따라 운용배수는 2015년 9.0배에서 2019년 13.2배로 가파르게 증가한 것으로 나타났다.

〈표 II-8〉 최근 5년간 기본재산 및 운용배수 현황

(단위: 억원, 배)

구분	보증잔액(A)	기본재산(B)	운용배수(A/B)
2015년	207,096	22,951	9.0
2016년	212,075	22,023	9.6
2017년	218,506	19,722	11.1
2018년	224,426	16,761	13.4
2019년	221,122	16,751	13.2

자료: 기술보증기금, 『2019 기술보증기금 연차보고서』, 2020, p. 19의 내용을 활용하여 저자 작성

한편 〈표 II-9〉에서는 연도별 보증지원 실적을 보증성격별로 정리하였는데, 2019년 기준 대부분의 보증은 기술평가보증 및 기술혁신기업에 대한 보증인

것으로 나타났으며, 청년창업기업 및 유동화회사보증의 비중은 각각 7.6%, 0.7%로 상대적으로 비중이 낮은 것으로 나타났다.

〈표 II-9〉 연도별 보증지원 실적

(단위: 억원, %)

구분	2017년		2018년		2019년	
		비중		비중		비중
보증공급	222,555	100.0	224,798	100.0	220,020	100.0
기술평가보증	215,783	97.0	219,613	97.7	215,415	97.9
기술혁신기업	213,411	95.9	217,517	96.6	213,037	96.8
지식문화산업	37,192	16.7	40,825	18.2	41,827	19.0
혁신성장산업	26,634	12.0	49,000	21.8	54,852	24.9
수출중소기업	27,734	12.5	27,200	12.1	31,467	14.3
청년창업기업	15,250	6.9	16,261	7.2	16,683	7.6
R&D보증	48,348	21.7	53,138	23.6	55,456	25.2
유동화회사보증	2,609	1.2	1,562	0.7	1,558	0.7
신규보증	52,859	-	52,465	-	48,480	-
보증잔액	218,506	-	224,426	-	221,122	-

주: 1) 중복 포함

자료: 기술보증기금, 『2019 기술보증기금 연차보고서』, 2020, p. 24.

### 3) 중소벤처기업진흥공단<sup>13)</sup>

중소벤처기업진흥공단은 정책적 지원이 필요한 기술·사업성 우수 중소기업에 장기·저리의 자금을 공급하여 중소기업의 성장 촉진을 목적으로 정책자금 지원 프로그램을 운영하고 있다. 운용 방향은 중점지원 분야인 수출, 고용 창출, 시설투자 등에 우선적인 지원을 하고, 기술 및 사업성 평가를 통해 미래의 성장 가능성이 높은 것으로 나타난 기업에 대해서는 직업 및 신용대출 위주로 지원하는 방향이다. 융자한도는 개별 기업당 중소벤처기업부 소관 정책자금의 융자잔액 기준으로 60억원(수도권을 제외한 지방소재기업은 70억원, 한도우대사업 및 기업은 100억원 이내)이다.<sup>14)</sup> 대출금리는 정책자금

13) 중소벤처기업진흥공단, 「정책자금융자>세부사업>혁신창업사업화 자금」(<https://www.kosmes.or.kr/sbc/SH/SBI/SHSBI004M0.do>, 검색일자: 2020. 10. 30.); 금융감독원(2021); 중소벤처기업부 공고 제2021-648호 「2022년도 중소기업 정책자금 융자계획 공고」를 참고하여 작성하였다.

기준금리(분기별 변동금리)에 사업별 가감금리를 적용하며, 기업별 신용위험 등급, 담보종류 등에 따라 금리를 차등 적용한다. <표 II-10>은 중소벤처기업진흥공단의 주요 보증제도를 정리한 것으로, 혁신창업사업화자금, 투융자 복합금융자금, 신시장진출지원자금, 신성장기반자금, 재도약지원자금, 긴급경영안정자금 등이 포함된다.

〈표 II-10〉 중소벤처기업진흥공단 주요 보증제도

구분	내용
혁신창업사업화자금	우수한 기술력과 사업성은 있으나 자금력이 부족한 중소·벤처기업의 창업을 활성화하고 고용창출을 도모
투융자복합금융자금	기술성과 미래 성장가치가 우수한 중소기업에 대해 융자에 투자요소를 복합한 방식의 자금지원으로 창업 활성화 및 성장단계 진입을 도모
신시장진출지원자금	중소기업이 보유한 우수 기술의 제품화·사업화 촉진 및 수출품 생산 비용을 지원하여 기술 기반 수출 중소기업을 육성
신성장기반자금	사업성과 기술성이 우수한 성장유망 중소기업의 생산성 향상, 고부가가치화 등 경쟁력 강화에 필요한 자금을 지원하여 성장동력 창출
재도약지원자금	사업전환, 구조조정, 재창업 지원을 통해 재도약과 경영정상화를 위한 사회적 기반 조성
긴급경영안정자금	경영애로 해소 등 긴급한 자금소요를 지원하여 중소기업의 안정적인 경영 기반 조성

자료: 금융감독원(2021), pp. 15~29.

14) 제조현장스마트화자금, 고성장촉진자금 등 잔액기준이 예외 적용되는 경우 최대 100억원 이내에서 지원한다.

---

### Ⅲ. 정책금융의 지원효과

---

#### 1. 분석 자료 소개 및 기초통계량

##### 가. 한국기업데이터(KED) 자료 소개

본고에서는 개별 기업 단위의 정보는 한국기업데이터(KED) 자료를 이용하여 구축하였다. 한국기업데이터 자료는 기업의 기본적 특성, 재무정보, 지적재산권 보유 현황, 주주 현황 등 다양한 정보를 포함하고 있는 기업 단위 자료이다. 이 자료는 대기업, 중기업, 소기업, 소상공인 등 다양한 기업 크기 및 산업에 속하는 기업들을 망라하고 있는 자료이며, 따라서 적절한 통제그룹 선정에 위해 최대한 많은 관측치를 확보하는 것이 바람직한 본고의 분석 방향에 부합하는 자료이다. 해당 자료의 매출액 등 기업의 연간 재무 상황을 반영하는 변수 등은 연간 단위로 제공되어 패널 형태의 자료를 구축하는 것이 가능하지만, 기업의 일반적인 특성과 관련된 변수 등 많은 변수들은 데이터가 업데이트된 마지막 시점의 정보에 한정하여 제공된다는 특징이 존재한다.

##### 나. 신용보증기금 자료 소개<sup>15)</sup>

본고에서 이용한 개별 기업 단위의 정책금융 수혜 이력 자료는 2014년 및 2015년의 신용보증기금 신규지원 이력 자료이다.<sup>16)</sup> 사업자등록번호 확인이 불가능한 기업을 제외하고 2014년은 4만 3,798건, 2015년은 5만 1,710건의 지

---

15) 본고의 분석에 필수적이었던 신용보증기금 자료를 정리, 제공해 주신 한국조세재정연구원 장우현 박사님께 감사드린다.

16) 이후 본고에서는 신용보증기금과 정책금융이라는 단어를 혼용하여 사용한다.

원이력 자료가 포함되어 있으며, 자료에 포함된 기업의 숫자는 2014년 3만 4,081개소, 2015년 2만 7,584개소이다. 지원이력 단위의 기초통계량을 <표 Ⅲ-1>과 <표 Ⅲ-2>에, 지원받은 기업 단위의 기초통계량을 <표 Ⅲ-3>에 제시하였다.

<표 Ⅲ-1>과 <표 Ⅲ-2>는 신규 지원이력을 기준으로 기초통계량을 정리하였다. 신규지원 이력의 대부분은 대출에 대한 보증으로 나타났으며, 제2 금융보증 유형까지 고려하면 절대적인 비중을 차지한다. 용도별로 살펴보면 운전자금과 시설자금 중 시설자금이 10%를 약간 상회하는 정도로 무시할 수 없는 비중을 차지하나, 운전자금 용도의 비중이 절대적인 것으로 나타났다. 평균적인 보증 기간은 2년에 조금 미치지 못하며, 평균적인 보증 비율은 90% 수준으로 상당히 높은 모습을 보인다. 지원이력 단위의 평균 보증금액은 2014년 2억 7천만원 수준에서 2015년 3억 3천만원 수준으로 다소 상승하였다.

<표 Ⅲ-3>으로부터 신규 지원이력을 보유한 사업체의 특성에 대해 살펴보면, 제조업에 비해 서비스업이 더 많은 비중을 차지하는 것을 확인할 수 있다.<sup>17)</sup> 또한 신규 지원이력을 가진 사업체의 60% 이상이 창업기업으로 분류되었다. 사업을 일정 기간 이상 영위하고 있는 기업과 비교했을 때 창업 기업의 영업활동 및 성과는 다소 상이한 모습을 보일 수 있으므로, 이를 고려하여 지원받은 기업의 성과를 해석할 필요가 있다. 예를 들어 초기 시장 진입을 위해 다소의 적자를 감수하더라도 적극적인 영업활동을 펼치는 것이 창업기업의 입장에서는 합리적인 의사결정일 수 있다.

17) 한국표준산업분류 기준 제조업 대분류 코드를 이용하여 제조업 기업을 정의하였으며, 서비스업을 정의하는 과정에서는 통계청 서비스업조사에서 활용되는 서비스업 포함 대분류 코드를 이용하였다. 해당 코드는 E(수도, 하수 및 폐기물 처리, 원료재생업), G(도매 및 소매업), H(운수 및 창고업), I(숙박 및 음식점업), J(정보통신업), K(금융 및 보험업), L(부동산업), M(전문, 과학 및 기술서비스업), N(사업시설 관리, 사업지원 및 임대서비스업), O(공공 행정, 국방 및 사회보장 행정), P(교육서비스업), Q(보건업 및 사회복지 서비스업), R(예술, 스포츠 및 여가 관련 서비스업), S(협회 및 단체, 수리 및 기타 개인 서비스업), T(가구 내 고용활동 및 달리 분류되지 않은 자가소비 생산활동, U(국제 및 외국기관)이다.

〈표 III-1〉 신용보증기금 신규 지원이력 기초통계량 1

(단위: 건수)

변수명	세부유형	2014년	2015년
취급종류	납세보증	2	-
	대출보증	38,235	46,394
	상거래담보보증	1,282	1,478
	어음보증	708	288
	이행보증	796	604
	제2금융보증	2,645	2,830
	지급보증(제2금융)	7	11
자금용도	지급보증의보증	123	105
	시설자금	4,478	6,222
	운전자금	39,320	45,488
총 신규보증 수		43,798	51,710

자료: 신용보증기금 내부자료

〈표 III-2〉 신용보증기금 신규 지원이력 기초통계량 2

(단위: 지원이력 건수)

변수	2014년		2015년	
	평균	표준편차	평균	표준편차
보증기간(일)	677.09	748.76	697.46	793.08
보증비율	89.92	6.11	90.86	5.51
신규보증금액(백만원)	272.08	592.16	334.69	694.42
총 신규보증 수	43,798		51,170	

자료: 신용보증기금 내부자료

〈표 III-3〉 신용보증기금 신규 지원이력 기초통계량(사업체)

변수	2014년		2015년	
	평균	표준편차	평균	표준편차
제조업 비중	0.37	0.48	0.38	0.49
서비스업 비중	0.57	0.49	0.56	0.50
수출기업 여부	0.09	0.29	0.10	0.30
이노비즈기업 여부	0.04	0.18	0.08	0.26
벤처기업 여부	0.02	0.14	0.04	0.20
창업기업 여부	0.60	0.49	0.64	0.48
사업체 수	34,081		27,584	

주: 계산 단위는 사업체  
 자료: 신용보증기금 내부자료

#### 다. 기초통계량 검토

신용보증기금 신규 지원이력 자료에 포함된 사업자등록번호를 이용하여 한국기업데이터 자료의 기업개요 자료와 매칭한 후, 매칭이 성공한 기업의 특성에 대한 보다 다양한 기초통계량을 검토한다. 특히 한국기업데이터 자료에 담긴 정보를 통해 사업체에 대한 기초통계량을 보다 풍부하게 검토하는 것도 중요하지만, 매칭에 실패한 기업들의 특성이 성공한 기업과 비교했을 때 크게 이질적인 경우 분석 결과를 일반화하는 것에 대한 적절성에 영향을 미칠 가능성도 있으므로 앞서 소개한 기초통계량과 특성을 비교 검토하는 것 또한 중요하다. <표 Ⅲ-4>는 매칭에 성공한 기업에 대한 기초통계량을 2014년, 2015년 지원이력 보유 기업으로 분리하여 제시하였다. 기초통계량을 제시하기에 앞서 2014년 지원이력 보유기업 중 한국기업데이터 자료에 매칭된 기업은 2만 9,103개소이며, 2015년 지원이력 보유기업 중 매칭된 기업은 2만 4,390개소이다. 그리고 두 연도 모두에 대해 지원이력을 보유한 기업은 3,683개소이다. 마지막으로 한국기업데이터 자료에 포함된 기업 중 두 연도 모두에 지원이력을 보유하고 있지 않는 기업은 91만 9,685개소이다. 앞서 언급했듯 한국기업데이터 「기업개요 자료」 정보 중 상당수는 기업 별로 자료가 수집된 최신 시점을 기준으로 한 자료이므로, 이하에 제시된 기초통계량의 변수들이 2014년 또는 2015년 시점의 실제 수치와 다소 다를 수 있음에 유의할 필요가 있다. 그러나 대분류 단위의 업종, 소재지 등의 특성은 시간에 따라 민감하게 변동하는 변수가 아닐 것으로 생각되므로, 여전히 비교 자료로서의 의미는 가질 수 있다고 판단된다.

<표 Ⅲ-4> KED 매칭 자료 기초통계량(사업체)

변수	2014년 지원기업		2015년 지원기업	
	평균	표준편차	평균	표준편차
제조업 비중	0.38	0.49	0.39	0.49
서비스업 비중	0.48	0.50	0.47	0.50
수도권 소재 여부	0.49	0.50	0.50	0.50
벤처기업 여부	0.04	0.19	0.05	0.22

〈표 III-4〉의 계속

변수	2014년 지원기업		2015년 지원기업	
	평균	표준편차	평균	표준편차
창업 후 기간(연) <sup>1)</sup>	6.76	6.39	6.79	6.53
사업체 수	29,103		24,390	

주: 1) 창업 후 기간은 2014년 지원기업은 2014년 기준, 2015년 지원기업은 2015년 기준으로 계산되었으며, 창업연도가 결측치인 경우 또는 지원연도보다 이후 연도로 창업연도가 기록된 기업의 경우 계산에서 제외하였음

자료: 신용보증기금 내부자료; 한국기업데이터 자료

지원기업의 제조업 비중은 앞서 지원이력 자료로부터 계산된 기초통계량과 거의 유사한 수준이다. 그러나 서비스업 비중은 10% 정도 하락한 모습을 보이는데, 이는 한국기업데이터 자료에 사업자등록번호가 등록되어 있지 않은 이유로 인해 매칭에 실패한 사업체 중 서비스업 분류에 속하는 사업체가 비교적 많음을 의미한다. 그럼에도 불구하고 서비스업 종사 기업은 전체 매칭기업 중 여전히 과반에 가까운 큰 비중을 차지하고 있다. 기업의 위치는 두 연도 모두 수도권에 소재한 기업이 대략 절반 정도를 차지한다. 벤처기업의 경우 지원이력 자료에서 살펴본 기초통계량과 비교하였을 때 매칭에 성공한 기업의 벤처기업 비율이 다소 높은 것을 확인할 수 있다. 창업 후 기간은 두 연도 모두에서 6.8년 정도이며, 지원이력 자료에서 창업기업으로 분류된 기업이 60%를 상회하는 비중을 보이는 것을 고려하면 다소 높은 수치라고 생각할 수 있다. 이는 창업기업에 대한 신용보증기업의 지원 프로그램이 창업 전이거나 갓 창업한 기업뿐 아니라 창업 후 일정 기간이 경과한 기업에 대한 지원 유형을 포함하고 있기 때문일 수 있다. 예를 들어 신용보증기금의 유망창업기업성장지원프로그램의 경우 예비창업보증(창업 전 6개월), 신생기업보증(창업 후 1년 내), 창업초기보증(창업 후 1~3년), 창업성장보증(창업 후 3~7년)으로 지원대상 기업의 업력이 비교적 다양하게 분포되어 있다.<sup>18)</sup> 한국기업데이터 자료에 매칭하는 과정에서 서비스업의 비중 등 일부

18) (주) 중소기업경영혁신연구원, 「신용보증기금 설명자료」, [http://smbc.re.kr/bbs/board.php?bo\\_table=policyfund02&wr\\_id=39&listsca=true](http://smbc.re.kr/bbs/board.php?bo_table=policyfund02&wr_id=39&listsca=true), 검색일자: 2021. 7. 17.

특성에서 차이가 발생하는 것으로 보이나, 대부분의 기업이 매칭되었고 여전히 매칭에 성공한 그룹 내부에서도 다양한 기업 특성이 관찰된다는 점을 고려하면 분석 진행에는 무리가 없을 것으로 판단한다.

〈표 Ⅲ-5〉에서는 한국기업데이터 자료에 포함되어 있고 2014, 2015년 두 연도 모두에 대해 지원이력이 매칭되지 않은 기업에 대한 기초통계량을 제시하였다. 제조업의 비중이 매칭기업에 비해 다소 낮으며, 서비스업 비중과 수도권 소재기업의 비중은 대체로 유사한 편이다. 창업 후 경과한 연도는 지원기업과 비교해서 다소 긴 편이다.

〈표 Ⅲ-5〉 KED 자료 기초통계량(사업체, 미지원기업)

변수	평균	표준편차
제조업 비중	0.30	0.46
서비스업 비중	0.51	0.50
수도권 소재 여부	0.51	0.50
벤처기업 여부	0.03	0.18
창업 후 기간(연) <sup>1)</sup>	11.86	8.43
사업체 수	919,685	

주: 1. 계산 단위는 사업체

1) 창업 후 기간은 2015년 기준으로 계산되었으며, 창업연도가 결측치인 경우 또는 지원연도보다 이후 연도로 창업연도가 기록된 기업의 경우 계산에서 제외하였음

자료: 신용보증기금 내부자료; 한국기업데이터 자료

다음으로 신용보증기금 지원기업과 미지원기업의 기업 크기에 따른 분포를 〈표 Ⅲ-6〉에 제시하였다. 기업의 규모는 기업의 특성변수를 구성하는 중요한 요소이나, 앞서 언급한 대로 한국기업데이터 자료에 수록된 기업 규모 관련 정보는 신용보증기금 지원이 이루어진 2014 또는 2015년의 정보와 상이할 수 있다. 특히 업종 등의 변수와 달리 기업 크기는 연도에 따라 활발하게 변할 수 있는 변수이기 때문에 유의할 필요가 있다. 기업 규모 정보의 제공시점은 한국기업데이터 「기업개요 자료」에 등록된 96만 9,493개의 사업체 중 2019년 시점 정보를 제공한 기업이 33만 3,643개, 2018년 제공 기업이 12만 7,780개로 이를 합하면 47.59%로 과반에 가까운 비율을 차지한다.

이를 포함하여 2015년 이후 정보를 제공한 기업을 더하면 67.64%가량의 비율을 보이므로, 상당한 숫자의 관측치는 지원이 이루어진 시점으로부터 일정한 기간이 지난 이후의 기업 규모를 나타낸다. 이 점에는 유의할 필요가 있으나 대략적으로나마 지원기업과 미지원기업의 기업 규모 분포를 비교하려는 목적에는 <표 Ⅲ-6>에서 제공하는 정보가 유용할 것으로 판단된다.

<표 Ⅲ-6> 신용보증기금 지원·미지원 기업 기업크기 분포

변수명	세부유형	2014년 지원	2015년 지원	미지원
기업 규모	대기업	10	6	2,141
	중기업	4,853	4,907	77,133
	소기업	11,469	8,755	496,715
	한시성중소기업	9	9	160
	중견기업	17	12	1,657
	소상공인	10,390	9,134	299,293
	보호대상중견기업	51	52	2,857
	판단대상제외	64	137	3,190
사업체 수		26,863	23,012	883,146

주: 해당 자료의 관측 연도는 사업체에 따라 상이함  
 자료: 신용보증기금 내부자료; 한국기업데이터 자료

<표 Ⅲ-6>은 한국기업데이터 자료에서 기업 규모를 확인할 수 없는 기업은 제외하고 작성된 표이다. 기업 규모 정보가 존재하는 기업 중 신용보증기금 지원이력이 연결된 기업의 경우 중기업, 소기업, 소상공인으로 분류된 기업이 절대다수를 차지하는 것을 확인할 수 있으며, 이는 정책금융의 취지를 고려하면 예상할 수 있는 분포이다.

## 2. 분석 전략 및 해석상의 유의점

본 절에서는 2014~2015년 신용보증기금 정책금융 지원의 효과를 추정하기 위한 전략을 소개한다. 추정을 위한 데이터를 구축하는 과정에서는 공통적으로 다음과 같은 처리 과정을 수행하였다.

(1) 추정에 요구되는 변수들이 추정 대상 기간 동안 관측된 기업이어야 한다. (2) <표 Ⅲ-6>에서 확인할 수 있듯이 대부분의 지원 대상 기업은 중기업, 소기업, 소상공인 분류에 속하므로 잠재적인 대조군도 해당 분류에 속하는 기업으로 제한하였다. (3) 장우현 외(2013)를 참고하여 2014년과 2015년 각 연도에 대해 지원이력이 있는 기업이 존재하지 않는 산업소분류에 속한 기업들을 표본에서 제외하였다. (4) 재무정보가 동일 연도에 2회 이상 관측된 기업 및 1900년 이전 설립기업의 경우 분석을 위한 표본 구축 과정에서 제외하였다. (5) 업력, 매출액, 연구개발비가 음수로 나타나는 기업을 제외하였다. (6) 분석 기간 중 연도별 매출액 상위 1%, 하위 1% 표본을 제외하였다. (7) 기업 유형이 공공기관으로 기록된 기업을 제외하였다. (8) 재무정보의 결산일이 각 연도 말일(12월 31일)이 아닌 기업을 제외하였다.

분석 전략에 대해 언급하기에 앞서, 본고 및 본 장의 목적 및 실증분석 결과 해석의 유의점에 대해 먼저 밝혀 두고자 한다. 본 장의 주안점은 실증분석을 통해 정책금융 지원의 인과관계를 새롭게 추정하려는 목적보다는 실증분석에서 연구자들이 채택할 수 있는 접근법을 이용한 추정을 시도하는 과정에 대해 논의하고, 이에 따른 추정 결과를 본고의 후반부에서 이루어질 머신러닝 방법론을 적용하여 추정한 결과와 비교 검토함으로써 잠재적인 머신러닝의 활용 방안에 대해 탐구하고자 하는 것이다. 따라서 본 장의 실증분석 결과는 비록 통계분석이 매우 엄밀하게 수행되었지만 자료의 한계점을 종합적으로 고려했을 때 정책금융 지원의 인과관계로 해석하기보다는 실증분석 방법론 수행의 예로서 제한적으로 해석할 필요가 있다.<sup>19)</sup>

정책금융 지원의 동적 효과 분석을 위해서는 지원이 이루어진 시점 이후의

19) 본 장에서 이용하는 분석 자료는 정책금융 제공 기관 측면에서 한계점이 존재한다. 정책금융을 제공하는 기관은 신용보증기금 이외에도 기술보증기금과 중소기업진흥공단이 존재하는데, 본 연구에서는 신용보증기금 수혜정보만을 활용하였다는 한계점이 있다. 따라서 분석의 대조군에 속하는 기업들 중 신용보증기금이 아닌 타 기관의 정책금융 지원을 받은 기업이 존재할 수 있다. 만약 신용보증기금을 지원받지 못한 사업체들이 다른 정책금융 지원을 신용보증기금 수혜사업체와 비교하여 많이 받은 경우가 발생한다면 본 연구의 추정 결과는 긍정적 효과는 과소평가, 부정적 효과는 과대평가하게 된다. 단 신용보증기금이 전체 정책금융에서 차지하는 비중이 다른 기금과 비교해서 크기 때문에 이러한 한계점은 추정편의에 큰 영향을 끼치지지는 못할 것으로 판단된다.

자료를 풍부하게 포함하는 것이 바람직하다. 그러나 작성 시점에서는 2016년 이후 신용보증기금 지원 자료의 정보가 존재하지 않는 상황이므로 분석 기간이 2016년 이후로 확장될수록 대조군과 처치군의 구분이 어려워지는 문제가 발생할 수 있다. 또한 자료가 포함하는 기간이 길어질수록 자동적으로 해당 기간 동안 생존한 기업만 표본에 포함된다는 점에도 유의할 필요가 있다. 이를 고려하더라도 여러 기간의 효과를 추정하여 비교해 보는 분석은 의미가 있으므로, 본 장에서는 지원연도(2014년) 이후 1, 2, 3년에 대한 효과 추정을 시도하였다.

지원 시점 이전 자료의 기간을 설정하는 것 또한 중요하다. 2014년 지원 이력을 분석하는 경우를 가정하면 효과 분석을 위해 2014년 이전 시점의 자료가 최소 1기간 이상 존재해야 하는 것은 필수적인 조건이다. 이에 더해 정책금융 지원 여부는 무작위 선택으로 보기 어렵기 때문에, 지원기업(처치군)과 대조군의 특성과 차별적인 추세를 통제할 필요가 있다. 일반적으로 고려해볼 수 있는 방법은 지원 이전의 일정 기간을 설정하고, 해당 기간의 매출액을 비롯한 주요 변수들의 특성을 처치군과 대조군 간 유사하게 매칭시키는 전략이다. 지원 시점 이전 기간이 길어질수록 보다 정교한 매칭을 시도할 수 있다는 장점이 존재하는 반면, 해당 기간의 관측치가 존재하지 않는 신규 창업기업들이 표본 구축 과정에서 자동적으로 분석 자료로부터 탈락하는 문제가 심각해진다는 단점이 동시에 존재한다. 본 연구에서는 이를 고려하여 지원 직전 연도인 2013년만을 포함한 자료, 그리고 2년을 추가하여 2011년까지의 자료를 활용한 분석을 시도하고 처치군 및 대조군 통제 정도를 검토하였다.

주된 실증분석은 장우현 외(2013) 등 많은 연구에서 널리 이용되는 성향점수매칭 방법론을 기반으로 수행하였다. 성향점수매칭은 본 연구의 분석 대상 사건(event)인 신용보증기금 지원 여부를 예측하는 확률모형을 통해, 자료에 포함된 모든 기업을 대상으로 정책금융 지원 예측 확률(성향점수, propensity score)을 계산하고, 이를 바탕으로 지원이력이 없는 대조군에 속하지만 지원을 실제로 받은 처치그룹에 속하는 기업과 유사한 특성을 가지는

기업들을 식별하는 작업이다. 성향점수매칭의 구체적인 과정은 후술하였다.

이 과정을 거쳐 최대한 유사한 특성을 가지는 대조군과 처치군 그룹을 설정하고, 두 그룹의 성과를 비교함으로써 정책금융 지원 여부라는 변수와 본질적으로 동반하여 존재하는 선택편향(Selection Biases)의 문제를 비롯하여 올바른 식별을 위협하는 문제들을 극복하여 일치추정량을 구할 수 있을 것으로 기대할 수 있다.

본 장에서는 성향점수매칭 방법론을 이용한 분석 결과를 제시하기에 앞서 선행회귀분석 방법론을 기반으로 처치군과 대조군 사이에 구분되는 추세가 존재하는지의 여부를 점검하고, 추세가 존재한다는 사실을 확인한 후 성향점수매칭을 이용하여 정책금융 지원의 성과를 검토한 결과를 제시한다.

### 3. 분석 결과

#### 가. 선행 추세 검정

성향점수매칭 방식을 이용하기에 앞서, 이벤트 스터디 형태를 취한 회귀 분석을 통해 실제 정책금융 지원이 이루어진 기업들인 처치군과 지원을 받지 않은 기업들인 대조군 사이에 통계적으로 상이한 선행 추세가 존재하는지 여부를 검토하였다. 2014년 신용보증기금 지원이라는 사건을 기준으로 할 때, 선행 추세가 존재한다는 것의 의미는 지원받은 기업이 지원받지 않은 기업에 비해 2013년 및 이전 기간에 매출액, 영업이익 등 기업 경영 관련 주요 변수들의 시계열 추세가 다른 모습을 보였다는 의미이다. 이러한 추세의 존재는 회귀분석을 통한 효과 식별에 상당한 위협이 된다. 설명을 위해 처치군에 속한 기업들의 경영이 2013년까지 지속적으로 개선되는 추세였다고 가정한다면, 이를 고려하지 않고 신용보증기금 지원의 정책효과를 분석할 경우 기존 상승세의 연장으로 인해 양(+, positive)의 정책효과가 추정될 수 있다. 그러나 실제 이는 신용보증기금 지원의 효과가 아닌 기존에 경영이 개선되는 추세였던 기업이 지원 대상으로 선정된 결과일 수 있다.

이러한 추세가 실제 존재하는지 검토하기 위해 식 (1)을 이용한 회귀분석

을 수행하였다. 회귀분석은 2015년 신용보증기금 지원기업을 표본에서 제거하고 2014년 신용보증기금 지원 대상 기업을 처치군으로 설정하여 수행되었으며, 분석 기간은 2014년을 기준으로 5년 이전 시점인 2009년부터 2015년까지의 기간으로 설정하였다.

$$y_{i,t} = \beta_1 + \sum_t \beta_t \text{treat}_i \times I_t + \alpha_i + \eta_g \text{igroup}_g \times \gamma_t + \epsilon_{i,t} \quad \text{식 (1)}$$

종속변수는 본 분석의 관심변수인 매출액, 영업이익, 당기순이익, 연구개발비이다.  $\text{treat}_i \times I_t$ 항은 신용보증기금 지원 직전 해인 2013년을 제외하고, 2014년 신용보증기금 지원 여부를 나타내는 더미변수와 각 연도를 나타내는 더미변수를 교차시킨 항들을 의미한다. 추세가 잘 통제되었을 경우 2009년부터 2012년 기간에 대한 계수  $\beta_i$ 는 0에 가까우며 유의하지 않게 추정되어야 바람직하다고 할 수 있다.

$\eta_g \text{igroup}_g \times \gamma_t$ 항은 신용보증기금 지원 여부가 아닌 다른 요소로 인해 발생하는 추세를 통제하는 항이다. 본 절에서는 2009년 기준 매출액을 5분위로 나누어 각 분위에 속하는 기업들을 범주화하여 개별적인 추세를 통제할 추정 결과를 <표 III-7>에, 그리고 2012~2013년 매출액을 각각 10분위로 나누어 총 100개의 그룹을 생성한 후 추세를 통제할 결과를 <표 III-8>에 수록하였다. [그림 III-1]은 처치군과 대조군의 매출액 변화 추세를 보여준다.

<표 III-7> 이벤트스터디 모형 추정 결과 1(2009~2015년 표본)

(단위: 백만원)

구분	매출액	영업이익	당기순이익	연구개발비
처치×2009년	-1259.908*** (68.702)	-85.710*** (8.348)	-85.795*** (9.271)	2.054 (2.546)
처치×2010년	-1120.515*** (57.520)	-90.947*** (8.009)	-89.900*** (9.116)	0.742 (2.692)
처치×2011년	-916.785*** (49.351)	-69.507*** (7.524)	-68.906*** (9.935)	1.196 (2.359)
처치×2012년	-587.160*** (35.421)	-45.274*** (6.614)	-52.603*** (8.215)	-0.819 (1.852)

〈표 III-7〉의 계속

(단위: 백만원)

구분	매출액	영업이익	당기순이익	연구개발비
처치×2013년				
처치×2014년	166,952*** (38,942)	-36,105*** (9,408)	-58,231*** (14,237)	9,577*** (2,070)
처치×2015년	245,269*** (56,301)	-36,035*** (11,263)	-65,830*** (13,510)	11,779*** (2,668)
관측치	610,165			

주: 1. 2009년 매출액 5분위 범주변수와 연도더미의 교차항이 통제변수로 포함되었음  
 2. 표준오차는 기업 단위 클러스터를 고려하여 계산  
 3. 통계적 유의성은 \*\*\* 1%, \*\* 5%, \* 10%와 같음  
 자료: 신용보증기금 내부자료; 한국기업데이터 자료

〈표 III-8〉 이벤트스터디 모형 추정 결과 2(2009~2015년 표본)

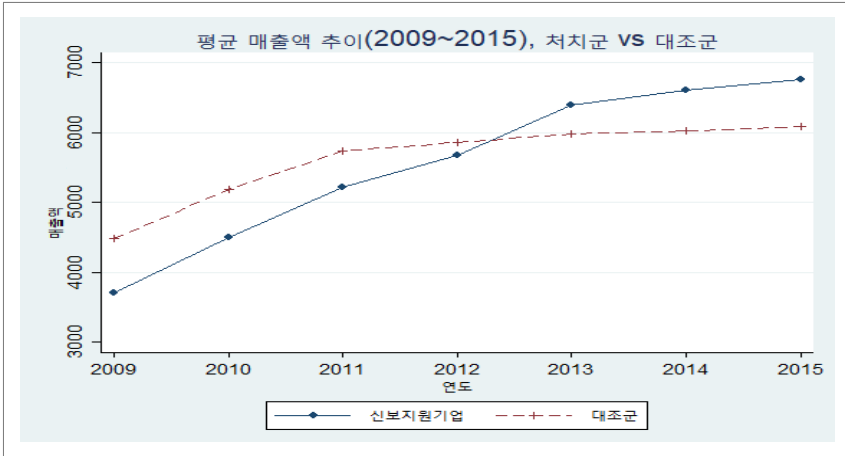
(단위: 백만원)

구분	매출액	영업이익	당기순이익	연구개발비
처치×2009년	-840,845*** (53,008)	-55,021*** (8,347)	-59,322*** (9,879)	6,818*** (2,606)
처치×2010년	-755,822*** (47,813)	-62,298*** (8,037)	-61,929*** (9,633)	4,776* (2,763)
처치×2011년	-564,198*** (43,691)	-38,319*** (7,614)	-40,263*** (10,249)	4,090 (2,527)
처치×2012년	-174,047*** (27,616)	-4,417 (6,575)	-16,370* (8,673)	2,164 (1,898)
처치×2013년				
처치×2014년	199,548*** (39,367)	-25,491*** (9,569)	-46,837*** (14,945)	8,945*** (2,136)
처치×2015년	277,834*** (56,843)	-22,860*** (11,385)	-53,521** (14,646)	11,379*** (2,744)
관측치	610,165			

주: 1. 2012년 매출액 10분위 범주변수, 2013년 매출액 10분위 범주변수와 연도더미의 교차항이 통제변수로 포함되었음  
 2. 표준오차는 기업 단위 클러스터를 고려하여 계산  
 3. 통계적 유의성은 \*\*\* 1%, \*\* 5%, \* 10%와 같음  
 자료: 신용보증기금 내부자료; 한국기업데이터 자료

[그림 III-1] 처치군과 대조군의 매출액 추세 비교

(단위: 백만원)



자료: 신용보증기금 내부자료; 한국기업데이터 자료

〈표 III-7〉과 〈표 III-8〉에서 확인할 수 있는 추정치는 질적으로 유사하며, 매출액, 영업이익, 당기순이익의 경우 두 집단 간 별도의 선행 추세가 강하게 존재하는 것으로 보인다. 특히 2012~2013년 두 연도의 매출액 10분위 변수를 교차시켜 그룹을 생성하여 추세를 별도로 통제하였을 경우에도 처치군과 대조군 간 별도의 추세가 존재하는 것으로 보임에 따라 여전히 선행 추세가 완전히 통제되지 않은 것으로 판단된다. 다만 연구개발비의 경우는 다소 불명확한 결과를 보여준다.

신보지원기업과 미지원기업을 단순히 분리하여 연도별 평균 매출액의 추세를 살펴본 [그림 III-1]에서도 신보 지원기업의 평균 매출액이 2009년에는 대조군의 매출액을 크게 하회하였으나, 지원기업의 가파른 상승폭으로 인해 2013년 시점에는 오히려 대조군의 평균보다 더 큰 매출 규모를 보이는 상이한 추세를 확인할 수 있다.

[그림 III-1]과 회귀분석 결과를 통해 선행 추세를 검토한 결과는 선행 추세가 존재한다는 점을 분석에 고려할 필요가 있음을 시사한다. 이를 고려하지 않은 추정 방법을 이용할 경우 편향된 추정치를 도출할 가능성이 매우 높을 것으로 보이며, 이후 분석에서는 이를 최대한 분석에 반영하고자 성향점수

에 기반한 매칭 방법을 활용하여 적절한 대조군을 선정하는 방식으로 실증 분석을 수행하고자 한다.

### 나. 성향점수매칭을 이용한 실증 분석

본 분석에서는 성향점수매칭 방법을 활용하여 신용보증기금 지원이 매출액, 영업이익, 연구개발비에 미친 영향을 분석한다. 영업이익과 당기순이익에 대한 분석 결과는 변수의 정의를 고려하면 대단히 유사할 가능성이 높으므로, 실제 기업의 영업활동과 보다 밀접하게 연관되어 있을 것으로 생각되는 영업이익을 분석 대상으로 선정하였다.

앞서 언급했듯 자료 구축 대상 연도를 길게 설정할수록 분석 측면의 장점이 존재하지만 동시에 많은 기업이 표본 구축 과정에서 탈락하며, 이 탈락이 무작위가 아니라는 단점이 존재한다. 성향점수 추정에 이용될 지원시점 이전 과거 변수들의 포함 범위를 적절한 수준에서 결정하기 위해 먼저 2013년 단일 연도만을 이용한 성향점수 추정, 그리고 2011년부터 2013년까지의 연도를 이용한 성향점수 추정을 수행하고 그 결과를 제시하여 비교한다. 그 이후 성향점수매칭을 이용하여 비교적 적절하게 대조군이 선정된다고 판단할 수 있는 2011~2013년 기간 포함 자료를 이용하여 2015, 2016, 2017년 변수들에 대한 분석 결과를 제시한다. 따라서 본 분석에서 제시할 표본은 총 네 가지 유형이며, <표 III-9>에 요약·정리하였다. 처치연도는 분석 대상 정책금융이 실행된 연도이며, 기간은 처치연도 이후 효과 분석 시점까지의 연도 차이이다.

<표 III-9> 실증분석을 위한 표본 구성 요약

표본유형	처치연도	기간	처치 이전 연도	분석 대상 연도
1	2014년	1	2013년	2015년
2	2014년	1	2011~2013년	2015년
3	2014년	2	2011~2013년	2016년
4	2014년	3	2011~2013년	2017년

자료: 저자 작성

성향점수 추정 과정은 다음과 같은 과정을 통해 이루어진다. 우선 앞서 언급한 자료 구축 과정에서의 표본 제거 기준에 더해 매출액, 영업이익, 당기순이익, 업력, 고정자산, 연구개발비 상위 1%, 하위 1%에 속하는 기업체들의 변수 값을 상위 1%, 하위 1%에 해당하는 수치로 대체하였다. 다음으로 신용보증기금 2014년도 지원 여부를 종속변수로, 이전 기간의 매출액과 제공항, 영업이익과 제공항, 산업중분류 기준의 2013년 매출액 평균치와 제공항, 산업대분류, 업력과 제공항, 2013년 연구개발비와 제공항, 기업 유형을 주식회사, 개인, 그 외로 구분한 변수, 수도권 소재 여부, 2013년 고정자산과 제공항을 독립변수로 설정한 로짓모형을 추정하였다. 로짓모형을 추정하면 기업별 성향점수의 예측값을 얻을 수 있고, 성향점수가 0.05 아래인 경우 극단값으로 처리하여 표본에서 제거하였다. 추가로 관측 첫 연도의 매출액 값이 0으로 기록된 관측치가 상당수 존재하여, 이를 표본에서 제거하였다.<sup>20)</sup>

표본에 포함된 기업마다 구한 성향점수의 예측값을 바탕으로 가장 가까운 매칭(Nearest Neighbor Matching) 방식을 통해 각 기업마다 성향점수 측면에서 가장 유사한 대조군 기업을 짝짓는 방식으로 매칭이 이루어진다. 본 분석을 위한 표본은 대체로 신용보증기금을 지원받은 기업에 비해 지원받지 않은 기업의 표본 수가 월등히 많으며, 따라서 지원기업과 유사한 성향점수를 가진 미지원기업을 탐색하는 것이 그 반대방향과 비교했을 때 매칭의 질이 높을 것으로 예상할 수 있다. 성향점수매칭의 질에 크게 의존하는 본 분석의 방법론을 고려하여 이후 진행될 분석에서는 처치군에서의 처치효과(Average Treatment Effects on the Treated)를 추정하였다. 원칙적으로 성향점수가 가장 가까운 대조군 기업을 짝짓는 과정을 거치며, 복수의 기업이 존재할 경우 가중치를 부여함으로써 처리한다. 처치군 기업  $i$ 에 짝지어진 대조군 기업을  $i^c$ 라고 하면, 처치효과는 다음과 같이 간단히 계산된다.

$$\frac{1}{N_T} \sum_{i \in T} (Y_i - Y_{i^c}) \quad \text{식 (2)}$$

20) 2011~2015년 표본을 기준으로 8만 5,027개의 기업 중 1,130개 기업의 매출액이 0으로 기록되었다.

$Y$ 는 분석하고자 하는 종속변수이며,  $N_i$ 는 처치군에 속하는 기업의 개수이다.  $i$ 와  $i'$ 를 연결하는 과정에서 성향점수가 이용되며, 이 매칭의 질, 즉 유사한 기업 선정이 추정치의 신뢰성을 결정하는 중요한 요소이다.

### 1) 2013~2015년 자료 분석 결과

2013년부터 2015년까지 관측치가 확보된 기업 중 2015년에 신용보증기금 지원이력이 있는 기업을 제거한 뒤 앞서 언급한 자료 정리 과정을 거쳐 최종적으로 구축된 자료는 11만 5,927개의 기업을 포함한다. 이 중 2014년 신용보증기금 지원이력이 존재하는 기업은 1만 792개로 전체 관측치의 약 9.31%를 차지한다. 추정 결과를 제시하기에 앞서 표본에 포함된 기업들의 재무정보를 비롯한 주요 정보들에 대한 기초통계량을 <표 Ⅲ-10>에 제시하였다.

<표 Ⅲ-10> 분석자료 기초통계량(2013~2015년 표본)

변수	2013년		2014년		2015년	
	평균	표준편차	평균	표준편차	평균	표준편차
매출액	3255.56	3762.70	3463.94	3985.75	3582.42	4114.72
영업이익	159.96	230.77	165.40	331.93	173.05	387.67
연구개발비	17.83	76.13	24.63	125.94	27.97	151.42
	평균			표준편차		
제조업 비중	0.47			0.50		
서비스업 비중	0.44			0.50		
수도권 소재 여부	0.54			0.50		
창업 후 기간(연)	6.74			5.48		
사업체 수	115,927(신보지원: 10,792)					

주: 1. 매출액·영업이익·연구개발비의 금액 단위는 백만원이며, 명목 금액임  
 2. 제조업 비중·서비스업 비중·수도권 소재 여부·창업 후 기간(연) 단위는 사업체임  
 자료: 신용보증기금 내부자료; 한국기업데이터 자료

앞 절에서 소개한 기초통계량과 비교하였을 때 서비스업에 속한 사업체의 비중이 소폭 감소하였고, <표 Ⅲ-10>이 지원기업과 미지원기업을 모두 포함하고 있다는 것을 고려하면 제조업 기업 비중은 표본 구축 과정에서 다소 상

승한 것으로 보인다. 3년간의 경영 관련 지표들은 대체적으로 조금씩 개선되고 있는 것으로 보이나, 물가상승의 영향이 혼재되어 있다는 것을 감안할 필요가 있다.

분석에 이용된 종속변수인 2015년 결산 기준 매출액(로그), 영업이익, 연구개발비에 대한 추정 결과를 <표 III-11>에 제시하였다. 로짓모형을 이용하여 성향점수를 추정하는 과정에서 처치변수인 신용보증기금 지원이력 보유 더미변수에 대한 예측값이 0.001 이하로 추정된 이상치를 제거하였다.

<표 III-11> 신용보증기금 신규지원 정책효과(2013~2015년 표본)

(단위: 백만원)

구분	(로그)매출액		영업이익		연구개발비	
	계수	P값	계수	P값	계수	P값
ATT	0.101*** (0.012)	0.000	-19.593*** (5.391)	0.000	-1.058 (1.941)	0.586
표본 수	108,593					

주: 통계적 유의성은 \*\*\* 1%, \*\* 5%, \* 10%와 같음  
 자료: 신용보증기금 내부자료; 한국기업데이터 자료

추정 결과 지원이 이루어진 1년 후 시점의 매출액에 대해서는 유의한 양(+)의 효과가 관측되었다. 종속변수가 로그 형태를 취하고 있으므로 약 0.1의 추정치는 지원기업의 매출액이 미지원 대조군에 비교해서 10% 정도 높음을 의미한다. 반면 영업이익에 대해서는 유의한 음(-)의 효과가 관측되었으며, 약 20백만원 정도 낮아지는 효과가 관측되었다. 영업이익의 2015년 기준 평균이 109백만원 수준임을 고려하면 음(-)의 효과의 크기는 상당히 부정적인 것으로 나타났다. 연구개발비에 대한 추정치는 부정적으로 나타났으나 통계적으로 유의하지 않았다.

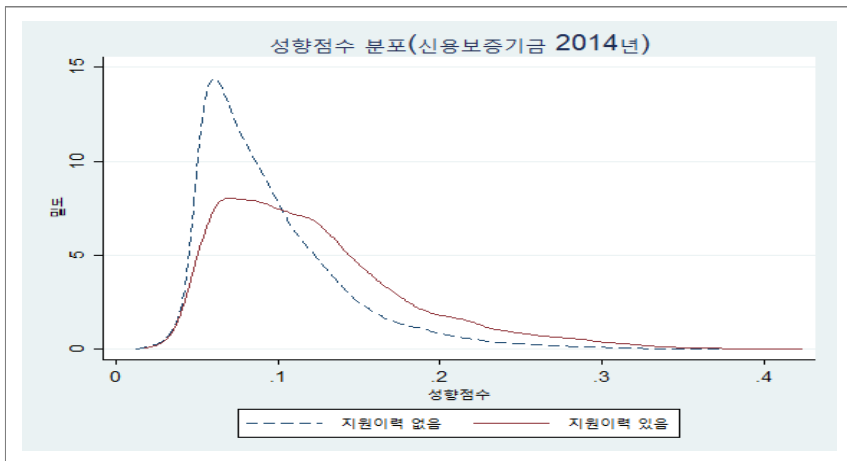
이러한 결과가 추정된 원인을 정확하게 파악하기는 어려우나 다음과 같은 가설을 고려해 볼 수 있다. 우선 창업기업이 정책금융을 지원받음으로써 단기적인 성과를 덜 고려하고 보다 장기간의 시계에서 영업활동을 계획할 수 있고, 이 경우 단기간의 부진한 영업이익 성과가 반드시 영업활동의 실패를 의미한다고 단정하기는 어렵다. 반대로 지원받은 기업이 미지원기업과 사전

적으로 유사한 특성을 지닌 기업일지라도 지원받은 이후 행태가 변화하여 낮은 수익성이 예상되는 영업활동에 상대적으로 더 종사할 가능성 또한 존재한다. 또한 매출액에 비해 영업이익 측면에서 다소 부진한 결과가 추정된 것은 장우현 외(2013)에서 도출된 추정 결과와 일치하는 측면이 있다.

〈표 Ⅲ-11〉의 분석 결과를 해석하는 과정에서 이 추정 결과는 지원 직후 1년에 한정된 분석 결과임을 유의해야 한다. 제시된 추정 결과가 여러 기간 지속성을 유지하는지 여부에 대해서는 추가적인 분석이 필요하다. 또한 2013년부터의 관측값이 존재하는 기업만 분석에 포함되었다는 점도 결과 해석에 영향을 미칠 수 있는 중요한 요소이다. 2013년에 실적이 포착되지 않았을 수 있는 창업기업도 정책금융의 중요한 대상일 수 있으나, 효과 분석을 위해서는 이전 기간의 관측치를 사용하는 것이 필수적이므로 제시된 결과를 추정하는 과정에서 이러한 기업들을 포함시키는 것은 불가능하다.

성향점수매칭 과정에서 지원이력을 보유한 기업과 그렇지 않은 기업 사이에 성향점수를 일정 수준 이상 공유해야 매칭을 통한 신뢰성 있는 추정치 획득을 기대할 수 있다. 이를 확인하고자 [그림 Ⅲ-2]에서 성향점수의 예측값 분포를 제시하였다.

[그림 Ⅲ-2] 신용보증기금 성향점수 분포도(2013~2015년 표본)



자료: 신용보증기금 내부자료; 한국기업데이터 자료

지원이력이 있는 기업과 대조군의 성향점수 분포는 상당히 유사하게 나타났다. 지원이력이 있는 기업이 대조군에 비해 다소 우측으로 치우친 모습이 나타나지만, 치우침의 정도는 매우 낮은 수준으로 두 그룹 간 성향점수 수준의 특성을 상당한 정도로 공유하고 있는 것으로 보인다.

## 2) 2011~2015년 자료 활용

이 표본에서는 지원 이전 기간을 보다 풍부하게 포함하는 자료를 구축하여 분석을 수행하였다. 2011년부터 2015년까지 관측치가 확보된 기업 중 2015년에 신용보증기금 지원이력이 있는 기업을 제거한 뒤, 앞서 언급한 자료 정리 과정을 거쳐 최종적으로 구축된 자료는 8만 3,897개의 기업을 포함한다. 이 중 2014년 신용보증기금 지원이력이 존재하는 기업은 8,496개로 전체 관측치의 약 10.13%를 차지한다. 추정 결과를 제시하기에 앞서 표본에 포함된 기업들의 재무정보를 비롯한 주요 정보들에 대한 기초통계량을 <표 III-12>에 제시하였다.

<표 III-12> 분석자료 기초통계량(2011~2015년 표본)

변수	2013년		2014년		2015년	
	평균	표준편차	평균	표준편차	평균	표준편차
매출액	3735.04	4040.54	3792.81	4189.07	3859.41	4301.86
영업이익	187.26	248.19	184.61	329.15	186.86	393.43
연구개발비	23.19	89.81	29.68	134.95	32.62	147.05
	평균			표준편차		
제조업 비중	0.50			0.50		
서비스업 비중	0.43			0.50		
수도권 소재 여부	0.53			0.50		
창업 후 기간(연)	9.70			5.90		
사업체 수	83,897(신보지원: 8,496)					

주: 1. 매출액·영업이익·연구개발비의 금액 단위는 백만원이며, 명목 금액임  
 2. 제조업 비중·서비스업 비중·수도권 소재 여부·창업 후 기간(연) 단위는 사업체임  
 자료: 신용보증기금 내부자료; 한국기업데이터 자료

〈표 Ⅲ-10〉과 비교했을 때 영업활동의 양적 지표가 소폭 증가했음을 알 수 있고, 2011년부터 관측된 기업으로 표본을 한정함에 따라 업력이 증가했음을 확인할 수 있다. 상대적으로 장기간 생존한 기업에만 한정된 표본으로 생각할 수 있으므로, 영업활동의 양적 지표가 큰 것은 자연스럽다고 판단된다.

분석에 이용된 종속변수인 2015년 결산 기준 매출액(로그), 영업이익, 연구개발비에 대한 추정 결과를 〈표 Ⅲ-13〉에 제시하였다. 로짓모형을 이용하여 성향점수를 추정하는 과정에서, 처치변수인 신용보증기금 지원이력 보유 더미변수에 대한 예측값이 0.001 이하로 추정된 이상치를 제거하였다.

〈표 Ⅲ-13〉 신용보증기금 신규지원 정책효과(2011~2015년 표본)

구분	(로그)매출액		영업이익		연구개발비	
	계수	P값	계수	P값	계수	P값
ATT	0.078*** (0.013)	0.000	-13.405** (5.678)	0.018	-2.480 (2.771)	0.371
표본 수	83,367					

주: 1. 금액 단위는 백만원  
 2. 통계적 유의성은 \*\*\* 1%, \*\* 5%, \* 10%와 같음  
 자료: 신용보증기금 내부자료; 한국기업데이터 자료

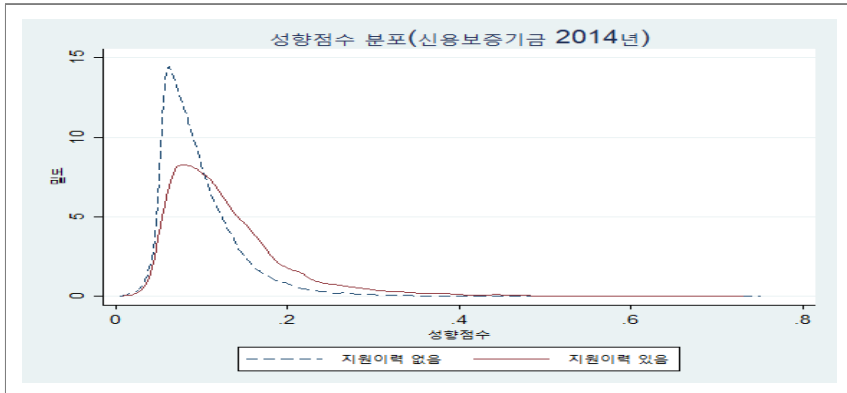
추정 결과, 지원이 이루어진 1년 후 시점의 매출액에 대해서는 유의한 양(+)의 효과가 관측되었다. 종속변수가 로그 형태를 취하고 있으므로 약 0.08의 추정치는 지원기업의 매출액이 미지원 대조군에 비교해서 8% 정도 높음을 의미한다. 반면 영업이익에 대해서는 유의한 음(-)의 효과가 관측되었으며, 약 13백만원 정도 낮아지는 효과가 관측되었다. 연구개발비에 대한 추정치는 통계적으로 유의하지 않았다. 2013년 기준 지표에 한정하여 성향점수 추정 과정에서 이용한 〈표 Ⅲ-11〉의 추정 결과와 비교하면 보다 장기간의 과거 변수를 매칭 과정에 이용하였고, 보다 적절하게 대조군 선정이 이루어졌을 것으로 기대할 수 있다. 그러나 표본 구축 과정에서 3년 이상 생존 기업들로 표본이 보다 한정됨으로써 발생하는 이질성이 혼재되어 있을 가능성도 존재한다.

〈표 Ⅲ-11〉과 동일하게 〈표 Ⅲ-13〉의 분석 결과를 해석하는 과정에서 이

추정 결과는 지원 직후 1년에 한정된 분석 결과임을 유의해야 한다. 또한 창업기업의 성격을 가진 기업들이 분석에 포함되지 않았을 경향성이 <표 Ⅲ-11>에 비해 높다는 점 또한 해석에 고려할 필요가 있다.

[그림 Ⅲ-3]에서는 [그림 Ⅲ-2]와 같이 성향점수의 예측값 분포를 제시하였다.

[그림 Ⅲ-3] 신용보증기금 성향점수 분포도(2011~2015년 표본)



자료: 신용보증기금 내부자료; 한국기업데이터 자료

지원이력이 있는 기업과 대조군의 성향점수 분포는 상당히 유사하게 나타났다. 지원이력이 있는 기업이 대조군에 비해 다소 우측으로 치우친 모습이 나타나지만, 치우침의 정도는 매우 낮은 수준으로 두 그룹 간 성향점수 기준의 특성을 상당한 정도로 공유하고 있는 것으로 보인다.

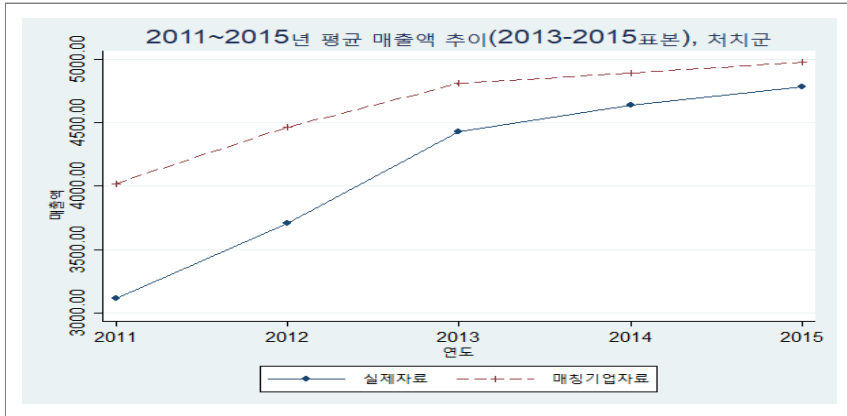
가) 표본 선택: 2013~2015년 자료와 2011~2015년 표본 비교

앞서 제시한 두 개의 분석 결과는 비교적 유사하다. 그러나 분석 기간을 다르게 설정함으로써 발생하는 표본 포함 기업의 특성 차이와 잠재적인 대조군 매칭의 질 문제가 혼재되어 있으므로, 두 표본 중 어떤 표본이 신용보증기금 지원의 효과를 올바르게 추정하는 목적에 부합하는지 판단하여 이후 분석을 진행할 필요가 있다. 본고에서는 성향점수매칭의 근본적인 목적을 고려하여 과거 추세를 보다 유사하게 추종하는 대조군을 선정하는 과정이 원활하게 이루어지는 것에 주안점을 두고 표본을 선정하고자 한다.

이를 위해 2013~2015년 추정 자료와 2011~2015년 추정 자료를 구축하는 과정에서 표본에 포함된 지원기업과 각 지원기업마다 성장점수를 기준으로 매칭된 대조군 기업에 대해 2011~2015년 기간의 평균 매출액 추이를 비교하여 두 표본 중 본고의 선호 표본을 선정하고자 한다. [그림 Ⅲ-4]는 2013~2015년 표본, [그림 Ⅲ-5]는 2011~2015년 표본으로부터 계산된 추세를 보여준다.

[그림 Ⅲ-4] 처치군과 대조군의 추세 비교(2013~2015년 표본)

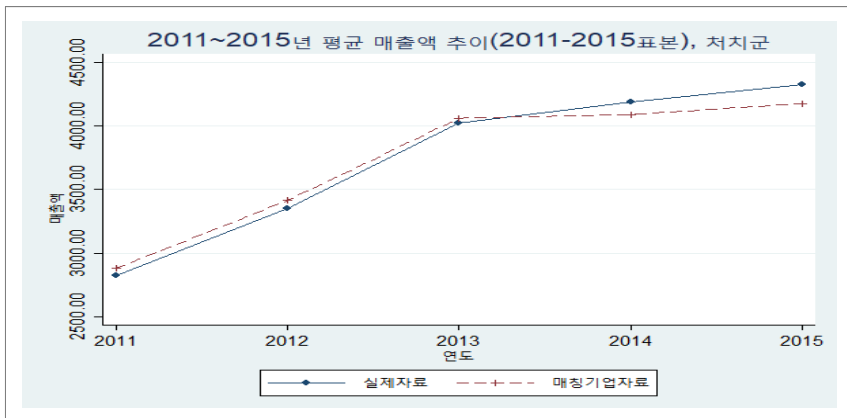
(단위: 백만원)



자료: 신용보증기금 내부자료; 한국기업데이터 자료

[그림 Ⅲ-5] 처치군과 대조군의 추세 비교(2011~2015년 표본)

(단위: 백만원)



자료: 신용보증기금 내부자료; 한국기업데이터 자료

두 그림을 비교하면 2011~2015년 표본이 2014년 지원 이전 3년간의 추세를 2013~2015년 표본에 비해 유사하게 통제하고 있는 것을 확인할 수 있다. 이는 2013~2015년 표본이 2011, 2012년의 표본값을 매칭 과정에서 이용하지 않았다는 점에서 예상된 결과라고 생각할 수 있다. 단 2013~2015년 표본에 대한 그래프에 선행 추세 통제의 정도를 간략하게 살펴보는 것 이외의 큰 의미를 부여하는 것에는 주의할 필요가 있다. 그 이유는 2013~2015년 표본의 추정 과정에서 포함되었던 많은 기업들이 2011~2012년의 매출액 정보를 가지고 있지 않으므로 [그림 Ⅲ-5]를 작성하는 과정에서 자동적으로 탈락하였기 때문이다.

검토 결과 본 분석에서는 선행 추세를 상대적으로 적절하게 통제하는 것으로 판단되는 2011~2015년 표본을 기본 표본으로 설정하고, 이후 제시될 2016, 2017년에 대한 효과 분석 및 이질성 분석에서 이 표본을 이용하였다.

### 3) 2011~2016년 자료 활용 - 2014년 지원의 2년 후 효과

본 분석은 2014년 신용보증기금 지원 대상 기업의 2년 시차의 성과를 분석한다. 2011년부터 2016년까지 관측치가 확보된 기업 중 2015년에 신용보증기금 지원이력이 있는 기업을 제거한 뒤 앞서 언급한 자료 정리 과정을 거쳐 최종적으로 구축된 자료는 7만 4,933개의 기업을 포함한다. 이 중 2014년 신용보증기금 지원이력이 존재하는 기업은 7,654개로 전체 관측치의 약 10.21%를 차지한다. 추정 결과를 제시하기에 앞서 표본에 포함된 기업들의 재무정보를 비롯한 주요 정보들에 대한 기초통계량을 <표 Ⅲ-14>에 제시하였다.

<표 Ⅲ-14> 분석자료 기초통계량(2011~2016년 표본)

변수	2013년		2014년		2015년	
	평균	표준편차	평균	표준편차	평균	표준편차
매출액	3810.55	4006.45	3870.29	4109.24	3953.89	4199.62
영업이익	191.29	227.65	189.59	292.13	194.10	328.02
연구개발비	23.63	90.01	30.53	136.66	33.98	150.36

〈표 III-14〉의 계속

변수	2013년		2014년		2015년	
	평균	표준편차	평균	표준편차	평균	표준편차
	평균			표준편차		
제조업 비중	0.50				0.50	
서비스업 비중	0.43				0.49	
수도권 소재 여부	0.52				0.50	
창업 후 기간(연)	9.85				6.02	
사업체 수	74,933(신보지원: 7,654)					

주: 1. 매출액·영업이익·연구개발비의 금액 단위는 백만원이며, 명목 금액임  
 2. 제조업 비중·서비스업 비중·수도권 소재 여부·창업 후 기간(연) 단위는 사업체임  
 자료: 신용보증기금 내부자료; 한국기업데이터 자료

〈표 III-12〉와 비교했을 때 소폭이지만 영업활동의 양적 지표 및 업력이 증가했음을 알 수 있다. 2016년 관측된 기업이라는 조건이 추가되었기 때문에 상대적으로 장기간 생존한 기업에만 한정할 표본으로 생각할 수 있으므로, 영업활동의 양적 지표가 큰 것은 자연스럽다고 판단된다.

분석에 이용된 종속변수인 2015년 결산 기준 매출액(로그), 영업이익, 연구개발비에 대한 추정 결과를 〈표 III-15〉에 제시하였다. 로짓모형을 이용하여 성향점수를 추정하는 과정에서 처치변수인 신용보증기금 지원이력 보유 더미변수에 대한 예측값이 0.001 이하로 추정된 이상치를 제거하였다.

〈표 III-15〉 신용보증기금 신규지원 정책효과(2011~2016년 표본)

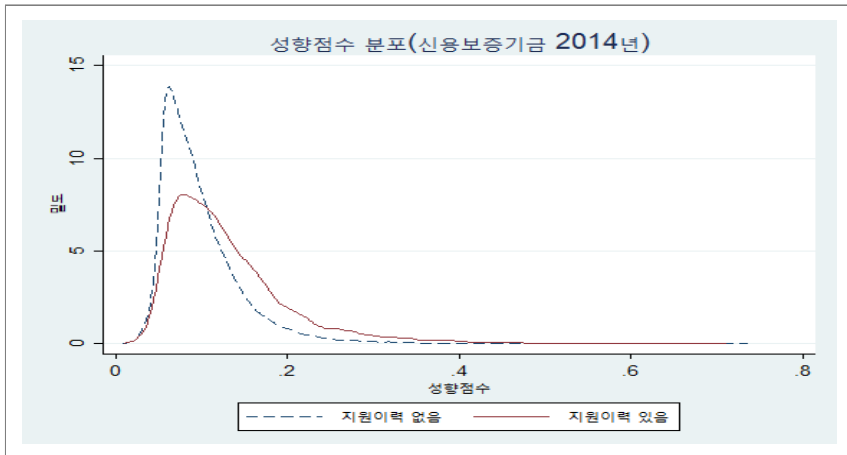
구분	(로그)매출액		영업이익		연구개발비	
	계수	P값	계수	P값	계수	P값
ATT	0.103*** (0.013)	0.000	-3.120 (5.497)	0.570	-1.433 (2.107)	0.497
표본 수	74,461					

주: 1. 금액 단위는 백만원  
 2. 통계적 유의성은 \*\*\* 1%, \*\* 5%, \* 10%와 같음  
 자료: 신용보증기금 내부자료; 한국기업데이터 자료

추정 결과, 지원이 이루어진 2년 후 시점의 매출액에 대해서 유의한 양(+)의 효과가 관측되었다. 종속변수가 로그 형태를 취하고 있으므로 약 0.1의 추정치는 지원기업의 매출액이 미지원 대조군에 비교해서 10% 정도 높음을 의미한다. 지원 후 1년 시점의 추정치가 8%가량이었던 것에 비하면 점추정치 기준 매출액 증가효과는 소폭 증가하였다. 반면 영업이익 및 연구개발비에 대한 효과는 통계적으로 유의하지 않았다. 그러나 앞서 언급했듯이 (1) 지원 이후 2년 시점인 2016년 말일 기준 생존한 기업만이 추정 과정에 포함되었다는 점, (2) '대조군'에 속하는 2014~2015년 신용보증기금 미지원 기업이 2016년에 신용보증기금을 지원받거나 혹은 다른 정책금융 지원을 받았을 수도 있다는 점 등 분석의 한계점을 종합적으로 고려하여 본 추정치를 제한적으로 해석할 필요가 있다.

[그림 Ⅲ-6]에서는 [그림 Ⅲ-3]과 같이 성향점수의 예측값 분포를 제시하였다.

[그림 Ⅲ-6] 신용보증기금 성향점수 분포도(2011~2016년 표본)



자료: 신용보증기금 내부자료; 한국기업데이터 자료

지원이력이 있는 기업과 대조군의 성향점수 분포는 상당히 유사하게 나타났다. 지원이력이 있는 기업이 대조군에 비해 다소 우측으로 치우친 모습이

나타나지만, 치우침의 정도는 매우 낮은 수준으로 두 그룹 간 성향점수 기준의 특성을 상당한 정도로 공유하고 있는 것으로 보인다.

#### 4) 2011~2017년 자료 활용 - 2014년 지원의 3년 후 효과

본 분석은 2014년 신용보증기금 지원 대상 기업의 3년 시차의 성과를 분석한다. 2011년부터 2017년까지 관측치가 확보된 기업 중 2015년에 신용보증기금 지원이력이 있는 기업을 제거한 뒤 앞서 언급한 자료 정리 과정을 거쳐 최종적으로 구축된 자료는 6만 6,452개의 기업을 포함한다. 이 중 2014년 신용보증기금 지원이력이 존재하는 기업은 6,870개로, 전체 관측치의 약 10.34%를 차지한다. 추정 결과를 제시하기에 앞서 표본에 포함된 기업들의 재무정보를 비롯한 주요 정보들에 대한 기초통계량을 <표 Ⅲ-16>에 제시하였다.

<표 Ⅲ-16> 분석자료 기초통계량(2011~2017년 표본)

변수	2013년		2014년		2015년	
	평균	표준편차	평균	표준편차	평균	표준편차
매출액	3969.23	4148.15	4032.22	4230.23	4129.45	4305.84
영업이익	197.92	256.63	196.85	300.77	202.39	337.05
연구개발비	26.37	96.21	33.37	142.57	36.95	156.47
	평균			표준편차		
제조업 비중	0.51			0.50		
서비스업 비중	0.41			0.49		
수도권 소재 여부	0.52			0.50		
창업 후 기간(연)	10.06			6.21		
사업체 수	66,452(신보지원: 6,870)					

주: 1. 매출액·영업이익·연구개발비의 금액 단위는 백만원이며, 명목 금액임  
 2. 제조업 비중·서비스업 비중·수도권 소재 여부·창업 후 기간(연) 단위는 사업체임  
 자료: 신용보증기금 내부자료; 한국기업데이터 자료

<표 Ⅲ-12> 및 <표 Ⅲ-14>와 비교했을 때 소폭이지만 영업활동의 양적 지표 및 업력이 증가했음을 알 수 있다. 2011~2015년 표본과 비교했을 때 2016년 및 2017년에 관측된 기업, 즉 2014년 신용보증기금 지원 시점 이후 3년 이상

더 생존한 기업이라는 조건이 추가되었기 때문에 상대적으로 장기간 생존한 기업에만 한정된 표본으로 생각할 수 있으므로, 영업활동 및 업력이 평균적으로 증가하는 모습은 자연스럽다고 판단된다.

분석에 이용된 종속변수인 2015년 결산 기준 매출액(로그), 영업이익, 연구개발비에 대한 추정 결과를 <표 III-17>에 제시하였다. 로짓모형을 이용하여 성향점수를 추정하는 과정에서 처치변수인 신용보증기금 지원이력 보유 더미변수에 대한 예측값이 0.001 이하로 추정된 이상치를 제거하였다.

<표 III-17> 신용보증기금 신규지원 정책효과(2011~2017년 표본)

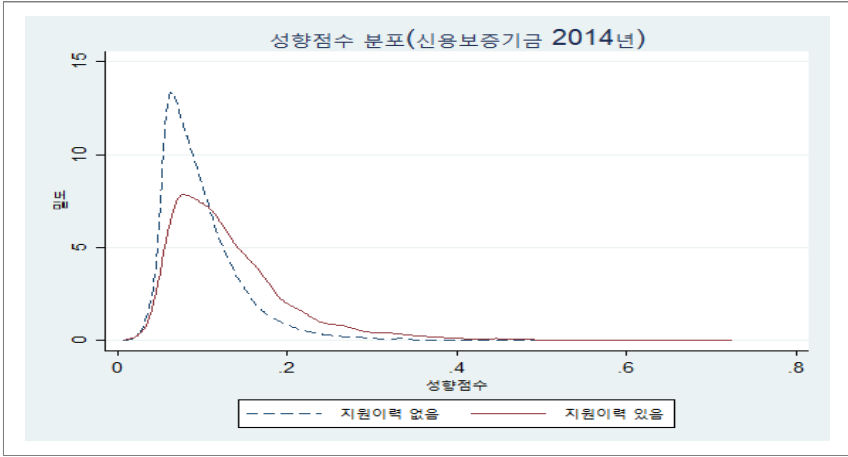
구분	(로그)매출액		영업이익		연구개발비	
	계수	P값	계수	P값	계수	P값
ATT	0.059*** (0.015)	0.000	-22.482*** (6.879)	0.001	-1.489 (2.569)	0.562
표본 수	66,054					

주: 1. 금액 단위는 백만원  
 2. 통계적 유의성은 \*\*\* 1%, \*\* 5%, \* 10%와 같음  
 자료: 신용보증기금 내부자료; 한국기업데이터 자료

추정 결과, 지원이 이루어진 3년 후 시점의 매출액에 대해서 유의한 양(+)의 효과가 관측되었으나 효과의 크기는 6% 수준으로 다소 감소하였다. 또한 영업이익에 대해 유의한 음(-)의 효과가 관측되었으며, 크기는 -22백만원 가량이다. 그러나 이러한 부정적인 효과가 나타난 이유가 근본적으로 정책 금융 지원 효과가 장기적으로 유지되지 않아서인지, 표본을 지원 이후 3년 이상 생존기업으로 국한해서인지 정확히 파악하기는 어렵다. 그러나 일정 기간 이상 생존하는 기업은 그렇지 않은 기업에 비해 영업활동을 보다 잘 수행하는 기업으로 생각할 수 있으며, 이러한 기업이 정책금융을 수혜받았을 경우 생존에 실패한 기업과 비교해서 오히려 이를 잘 활용하는 것이 직관적일 수 있다. 이에 따르면 1년 수혜의 효과가 장기간 지속되지 않을 수도 있다는 해석이 가능할 것으로 생각된다. 물론 앞서 지속적으로 언급한 표본 구축의 한계점을 고려하며 해석할 필요가 있다.

[그림 III-기에서는 성향점수의 예측값 분포를 제시하였다.

[그림 III-7] 신용보증기금 성향점수 분포도(2011~2017년 표본)



자료: 신용보증기금 내부자료; 한국기업데이터 자료

지원이력이 있는 기업과 대조군의 성향점수 분포는 상당히 유사하게 나타났다. 지원이력이 있는 기업이 대조군에 비해 다소 우측으로 치우친 모습이 나타나지만, 치우침의 정도는 매우 낮은 수준으로 두 그룹 간 성향점수 기준의 특성을 상당한 정도로 공유하고 있는 것으로 보인다.

### 5) 결과 종합

〈표 III-18〉 신용보증기금 신규지원 정책효과 요약

표본 유형	(로그)매출액		영업이익		연구개발비	
	계수	P값	계수	P값	계수	P값
1	0.101*** (0.012)	0.000	-19.593*** (5,391)	0.000	-1.058 (1,941)	0.586
2	0.078*** (0.013)	0.000	-13.405** (5,678)	0.018	-2.480 (2,771)	0.371
3	0.103*** (0.013)	0.000	-3.120 (5,497)	0.570	-1.433 (2,107)	0.497
4	0.059*** (0.015)	0.000	-22.482*** (6,879)	0.001	-1.489 (2,569)	0.562

주: 1. 금액 단위는 백만원  
 2. 통계적 유의성은 \*\*\* 1%, \*\* 5%, \* 10%와 같음  
 자료: 신용보증기금 내부자료; 한국기업데이터 자료

〈표 Ⅲ-18〉은 앞서 추정한 신용보증기금 지원의 효과를 요약한 표이다. 앞서 언급했듯 선행 추세를 비교적 잘 통제하는 것으로 보이는 2, 3, 4를 본 분석의 주된 결과로 해석하였다. 대체적으로 매출액에는 긍정적인 효과, 영업이익에는 부정적인 효과가 나타난 것을 확인할 수 있으며, 정책금융 지원 시점 이후 1, 2, 3년에 대해 효과를 살펴보았을 때 질적인 측면에서 뚜렷한 차이를 발견할 수 없었다. 연구개발비에 대해서는 일관되게 통계적으로 유의한 효과가 파악되지 않았다.

앞서 언급했듯 이 결과를 인과관계로 해석하기에는 분석 과정에서 유의해야 할 점들이 존재한다. ① 지원 이전 3년간, 그리고 표본에 따라 지원 이후 2~3년간 생존한 기업만이 추정 과정에 포함되었기 때문에 표본 구축 단계에서 편이가 발생할 수 있으며 ② ‘대조군’에 속하는 2014~2015년 신용보증기금 미지원 기업이 2013년 이전, 2016년 이후에 신용보증기금을 지원받거나 혹은 전 기간에 걸쳐 다른 기관의 정책금융 지원을 받았을 수도 있다는 점을 고려하면 처치군과 대조군이 명확하게 분리되지 않을 수 있다는 점을 비롯한 분석의 한계점을 종합적으로 고려하여 본 추정치를 제한적으로 해석할 필요가 있다.

추가적인 강건성 검정을 위해 2015년 기준 GDP디플레이터를 이용해 변수들을 실질변수로 전환하여 분석한 결과와 100만원 이하 연구개발비 수치를 0으로 처리하고 로그 변환 연구개발비를 이용한 결과를 본 장의 부록에서 제시하였다.<sup>21)</sup> 분석 표본은 표본 2와 동일하다. 실질변수 변환 결과 매출액과 연구개발비에 대해서는 매우 유사한 결과가 추정되었으나, 영업이익은 음(-) 추정치는 동일하지만 계수의 절대적 크기가 다소 줄어들고 통계적 유의성이 하락하였다. 세 종속변수에 대한 변화의 경향이 일치하지 않으므로 영업이익 계수의 변화를 물가의 영향으로 온전히 단정하기에는 어렵지만, 영업이익 계수에 대한 해석은 다소 주의할 필요가 있음을 시사한다. 로그 변환 연구개발비를 이용한 분석 결과는 연구개발비의 감소가 10% 수준에서

21) 로그 연구개발비를 종속변수로 취급한 분석 과정에서 이용되는 성향점수는 로그로 변환되지 않은 연구개발비를 이용하여 추정된 성향점수이다.

유의하게 나타났으며, 감소의 크기는 약 12% 수준으로 추정되었다. 2015년 기준 연구개발비의 평균은 약 3,261만원이고 3,261만원의 12%는 대략 390만원 수준이다. 표본 2의 점추정치에 따른 연구개발비의 감소분은 248만원으로, 표준오차를 감안하면 질적으로 크게 다르지 않은 추정치로 판단된다.

#### 다. 이질성 분석

본 절에서는 정책금융 지원 효과가 지원기업의 다양한 특성에 따라 이질적일 수 있다는 점을 고려하여, 가용 자료를 바탕으로 이질성 분석을 수행하였다. 먼저 정책금융 지원의 효과가 이질적으로 나타날 수 있을 것으로 여겨지는 특성들을 다양하게 선정하였다. 다음으로 특성에 따라 중위값 혹은 0을 비롯한 특정한 값을 기준으로 설정하고, 표본을 해당 값보다 큰 표본과 작은 표본의 하위 표본으로 분리하였다. 분리된 하위 표본에 대해 앞서 제시된 성향점수매칭 추정방법을 적용하여 분석을 수행한 후 추정 효과를 비교하였다.

표본을 세밀하게 나누어 분석할수록 다각적인 이질성 분석을 수행할 수 있는 장점이 있으나, 나뉜 표본마다 적절한 수준의 처치군과 대조군 기업들이 포함되어 있어야 하는 제약 또한 동시에 고려해야 한다. 처치군 기업의 숫자가 많지 않은 상황을 고려하여 본 분석에서는 두 개의 하위 표본을 구축하여 이질성 분석을 수행하였다.

또한 한 개의 변수에 대해 기준을 정하고, 두 개의 표본을 분리하여 추정하는 방식으로는 정확한 이질성의 원인을 알기 어렵다는 한계가 존재한다. 예를 들어 매출액이 높은 기업 그룹과 낮은 기업 그룹은 기준으로 이용된 매출액이라는 수치 이외에도 다양한 특성이 크게 상이할 것임을 짐작할 수 있다. 따라서 본 이질성 분석을 통해 이질적 효과가 나타나더라도 그 효과의 원인을 짐작하기는 제한적이다. 그러나 본 분석의 목적이 이질적 효과에 대한 정확한 결론을 내리는 것이 아닌 이후 제시될 머신러닝 방법론을 이용한 이질적 효과 탐구를 위한 기초 작업인 것을 고려하여 이질성 분석을 진행하였다.

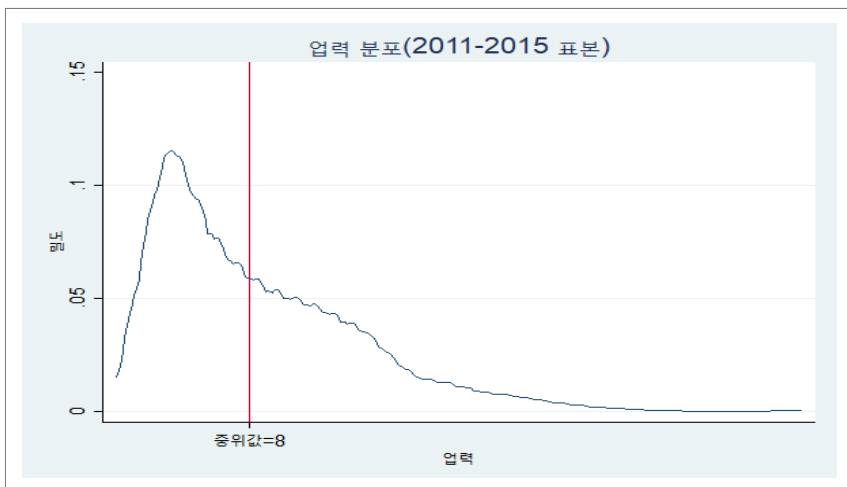
이질성 분석은 선행 추세를 보다 잘 통제하는 것으로 판단되는 2011~2015년 표본을 바탕으로 수행하였으며, 각 특성에 따른 분석에서는 추정 결과를 간략하게 보고하고 마지막으로 추정 결과들을 종합하여 공통된 함의에 대해 간략히 서술하였다.

### 1) 이질성 분석 1: 기업의 존속 연도

기업이 신생 기업인지, 일정 기간 이상 영업을 지속해 온 기업인지에 따라 생산성, 네트워크 등이 상이할 것이고 궁극적으로 지원받은 정책금융의 활용 방식 및 효과성 측면에서 차이가 나타날 수 있다. 이에 표본에 포함된 기업의 2014년 기준 업력 분포를 고려하여 중위값을 기준으로 이질성 분석을 수행하였다. 중위값은 8이며, 업력이 8년 이상인 기업은 4만 6,877개 기업이고 그중 2014년도 지원기업은 3,950개이다. 8년 미만 기업은 3만 6,585개이며, 그중 지원기업은 4,544개이다.

업력 분포는 [그림 III-8]에 제시하였고, 추정 결과는 <표 III-19>에 제시하였다.

[그림 III-8] 2011~2015년 표본의 업력 분포



자료: 신용보증기금 내부자료; 한국기업데이터 자료

〈표 III-19〉 신용보증기금 신규지원 정책효과(이질성: 업력)

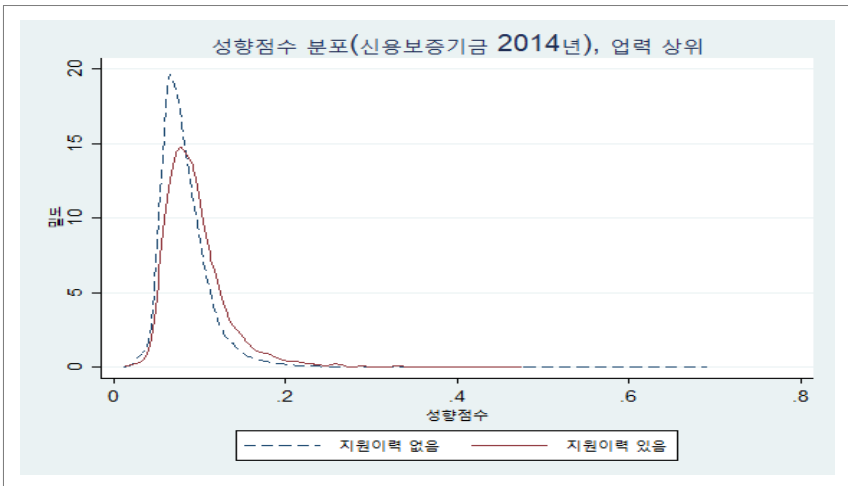
구분	표본 크기	(로그)매출액		영업이익		연구개발비	
		ATT	P	ATT	P	ATT	P
정책효과 (중위값 이상)	46,808	0.083*** (0.019)	0.000	-10.144 (9.258)	0.273	-0.538 (2.652)	0.839
정책효과 (중위값 미만)	36,537	0.106*** (0.017)	0.000	-3.139 (8.634)	0.716	-1.561 (2.092)	0.455

주: 1. 금액 단위는 백만원  
 2. 통계적 유의성은 \*\*\* 1%, \*\* 5%, \* 10%와 같음  
 자료: 신용보증기금 내부자료; 한국기업데이터 자료

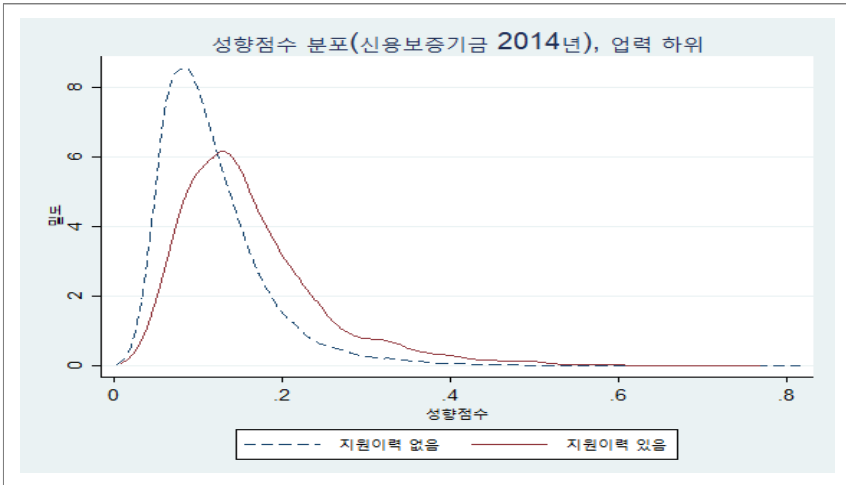
표준오차를 고려하면 업력에 따른 이질성이 뚜렷하게 관측된다고 판단하기는 어려운 것으로 보인다. 영업이익에는 음(-)의 효과가 나타났으나 두 그룹 모두에서 통계적으로 유의하지 않았는데, 표본을 분리하며 추정의 안정성이 떨어졌을 가능성이 있다.

다음 [그림 III-9]는 이질성 분석에 이용된 두 그룹의 성향점수 추정 분포를 나타내며, 그룹별로 유사한 특성을 공유하는 처치군과 대조군이 상당수 존재하는 것을 확인할 수 있다.

[그림 III-9] 신용보증기금 성향점수 분포도(이질성: 업력)



[그림 Ⅲ-9]의 계속



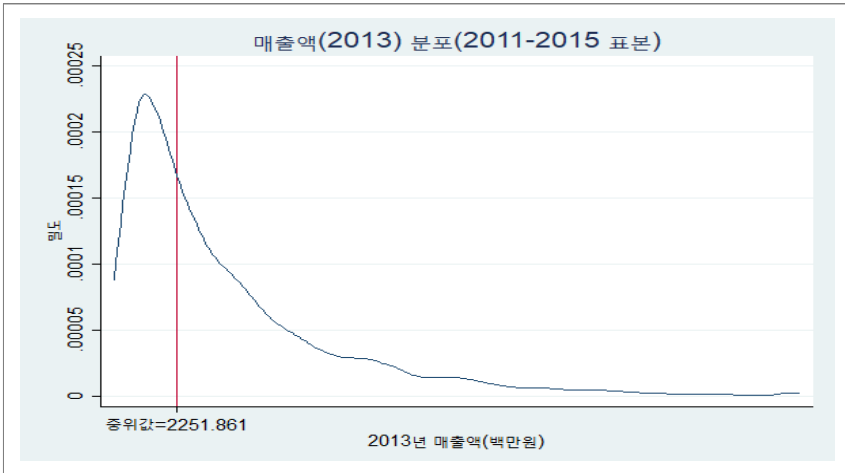
자료: 신용보증기금 내부자료; 한국기업데이터 자료

## 2) 이질성 분석 2: 매출액 규모

기업의 매출액은 기업의 크기, 산업, 영업활동의 정도 등 다양한 차원의 정보가 집약된 변수이다. 따라서 매출액의 차이는 정책금융의 효과성 측면에서 나타날 수 있는 이질성과 밀접하게 관련되어 있을 것으로 예상된다. 표본에 포함된 기업의 2013년 기준 매출액의 분포를 고려하여 중위값을 기준으로 이질성 분석을 수행하였다. 중위값은 약 2,252백만원이며, 2013년 매출액이 중위값 이상인 기업은 4만 1,949개 기업이고, 그중 2014년도 지원기업은 4,831개이다. 중위값 미만 기업은 4만 1,948개이며, 그중 지원기업은 3,665개이다.

업력 분포는 [그림 Ⅲ-10]에 제시하였고, 추정 결과는 <표 Ⅲ-20>에 제시하였다.

[그림 III-10] 2011~2015년 표본의 매출액 분포



자료: 신용보증기금 내부자료; 한국기업데이터 자료

<표 III-20> 신용보증기금 신규지원 정책효과(이질성: 매출액)

구분	표본 크기	(로그)매출액		영업이익		연구개발비	
		ATT	P	ATT	P	ATT	P
정책효과 (중위값 이상)	41,796	0.037*** (0.012)	0.002	-17.122* (9.127)	0.061	-1.626 (2.688)	0.545
정책효과 (중위값 미만)	41,534	0.069*** (0.014)	0.000	-4.726 (3.778)	0.211	-0.654 (1.144)	0.567

주: 1. 금액 단위는 백만원  
2. 통계적 유의성은 \*\*\* 1%, \*\* 5%, \* 10%와 같음

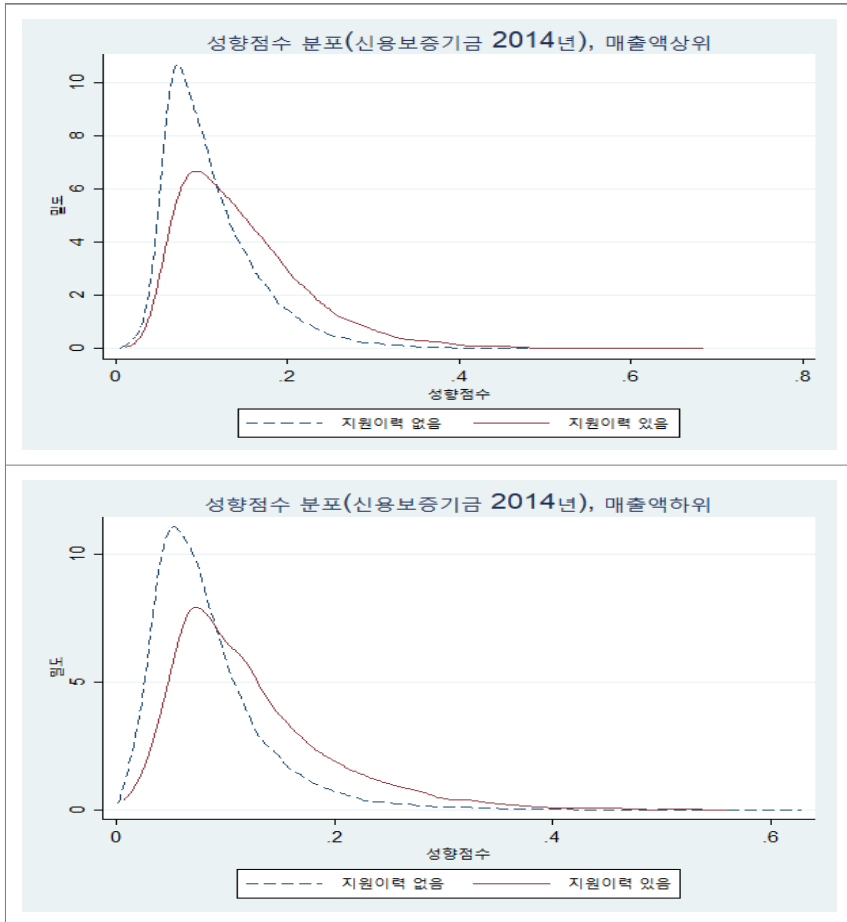
자료: 신용보증기금 내부자료; 한국기업데이터 자료

표준오차를 고려하면 단정하기 어려우나, 매출액이 중위값 미만인 그룹이 다소 양호한 방향의 효과가 추정되는 경향성이 있는 것을 확인할 수 있다. 매출액이 낮은 기업은 상대적으로 영업활동이 부진한 기업일 수도 있지만 기업의 규모, 업력 등 여러 측면과 매출액의 규모 사이에 상관관계가 존재할 것이므로 원인을 단정하기는 어렵다. 예를 들어 기업의 크기가 작아서 의사결정이 빠른 기업일수록 정책금융을 효율적으로 활용할 가능성도 생각해 볼 수 있을 것이나, 이 경우 매출액에 따른 이질성이라기보다는 기업의

크기에 따른 이질성이 될 것이므로 정확한 해석을 시도하기에는 주의할 필요가 있다.

[그림 Ⅲ-11]은 이질성 분석에 이용된 두 그룹의 성향점수 추정 분포를 나타내며, 그룹별로 유사한 특성을 공유하는 처치군과 대조군이 상당수 존재하는 것을 확인할 수 있다.

[그림 Ⅲ-11] 신용보증기금 성향점수 분포도(이질성: 매출액)



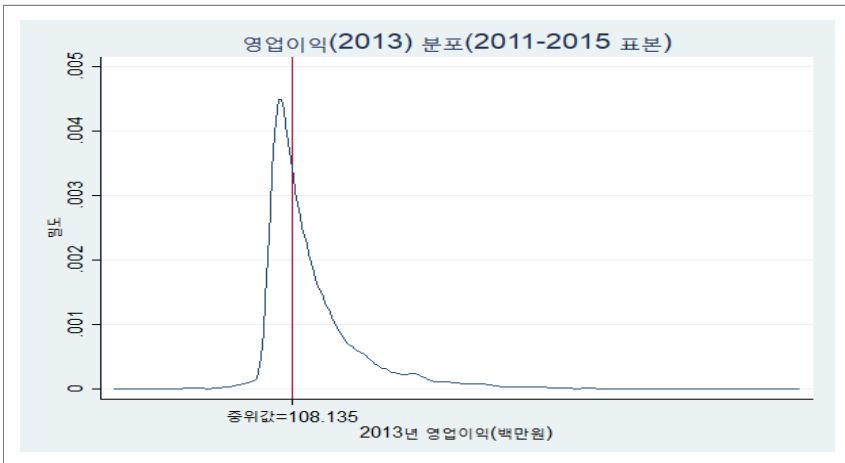
자료: 신용보증기금 내부자료; 한국기업데이터 자료

### 3) 이질성 분석 3: 영업이익 규모

기업의 영업이익은 매출액과 유사하게 기업의 크기, 산업, 영업활동의 정도, 생산성 등 다양한 차원의 정보가 집약된 변수이다. 자연히 영업이익의 차이는 정책금융의 효과성 측면에서 나타날 수 있는 이질성과 밀접하게 관련되어 있을 것으로 예상된다. 표본에 포함된 기업의 2013년 기준 영업이익의 분포를 고려하여 중위값을 기준으로 이질성 분석을 수행하였다.<sup>22)</sup> 중위값은 약 108.1백만원이며, 2013년 매출액이 중위값 이상인 기업은 4만 1,949개 기업이고, 그중 2014년도 지원기업은 4,649개이다. 중위값 미만 기업은 4만 1,948개이며, 그중 지원기업은 3,847개이다.

영업이익 분포는 [그림 III-12]에 제시하였고, 추정 결과는 <표 III-21>에 제시하였다.

[그림 III-12] 2011~2015년 표본의 영업이익 분포



자료: 신용보증기금 내부자료; 한국기업데이터 자료

22) 후술할 머신러닝을 이용한 분석 결과와의 비교를 위한 참고 목적으로, 2011년과 2012년 기준 영업이익의 분포에 따른 이질성 분석 또한 수행하였고, 추정 결과는 본 장의 부록에 수록하였다.

〈표 III-21〉 신용보증기금 신규지원 정책효과(이질성: 영업이익)

구분	표본 크기	(로그)매출액		영업이익		연구개발비	
		ATT	P	ATT	P	ATT	P
정책효과 (중위값 이상)	41,744	0.040*** (0.015)	0.008	-20.024** (9.771)	0.040	-5.249 (4.755)	0.270
정책효과 (중위값 미만)	41,161	0.115*** (0.017)	0.000	4.828 (5.211)	0.354	-1.159 (2.019)	0.566

주: 1. 금액 단위는 백만원

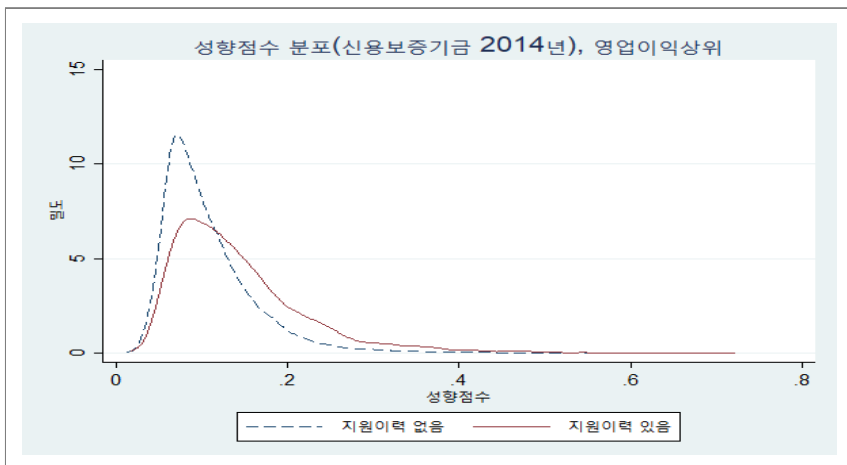
2. 통계적 유의성은 \*\*\* 1%, \*\* 5%, \* 10%와 같음

자료: 신용보증기금 내부자료; 한국기업데이터 자료

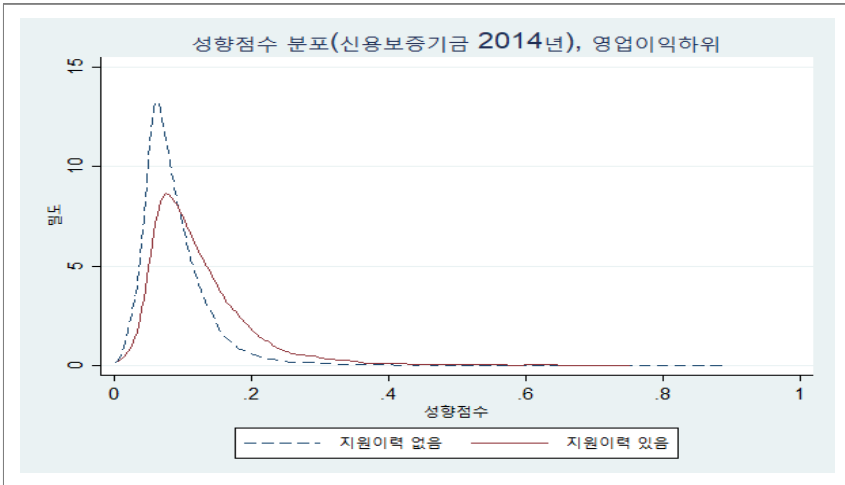
매출액에 대한 이질성 분석에서 추정된 결과와 유사하게 영업이익이 중위값 미만인 그룹으로부터 다소 양호한 방향의 효과가 추정되는 경향성이 존재한다. 영업이익의 절대적인 규모를 기준으로 그룹을 나누었으므로 매출액이 낮은 기업은 영업이익도 낮은 경향을 보일 것이므로, 이러한 경향성이 나타나는 원인도 유사할 가능성이 크다.

[그림 III-13]은 이질성 분석에 이용된 두 그룹의 성향점수 추정 분포를 나타내며, 그룹별로 유사한 특성을 공유하는 처치군과 대조군이 상당수 존재하는 것을 확인할 수 있다.

[그림 III-13] 신용보증기금 성향점수 분포도(이질성: 영업이익)



[그림 III-13] 의 계속



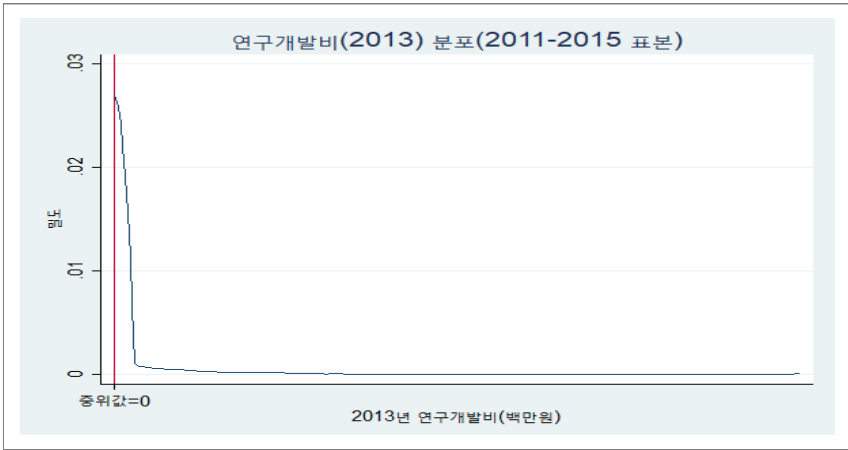
자료: 신용보증기금 내부자료; 한국기업데이터 자료

#### 4) 이질성 분석 4: 연구개발비 규모

기업의 연구개발비는 기업 경영의 방향, 산업의 특성, 인적구성의 특성 등과 관련된 변수로, 연구개발비 규모에 따라 정책금융의 효과성 측면에서 이질성이 발생할 가능성이 있다. [그림 III-14]에 제시된 연구개발비 분포에 따르면 연구개발비가 0인 기업이 대다수이므로, 연구개발비가 0인 기업과 0보다 큰 기업으로 분리하여 이질성 분석을 수행하였다. 2013년 연구개발비가 0보다 큰 기업은 1만 5,045개 기업이고, 그중 2014년도 지원기업은 1,081개이다. 0인 기업은 6만 8,852개이며, 그중 지원기업은 7,415개이다.

추정 결과는 <표 III-22>에 제시하였다. 연구개발비가 0인 그룹은 성향점수에 대한 로짓모형 추정 과정에서 연구개발비가 변수로 이용되지 않았다.

[그림 III-14] 2011~2015년 표본의 연구개발비 분포



자료: 신용보증기금 내부자료; 한국기업데이터 자료

<표 III-22> 신용보증기금 신규지원 정책효과(이질성: 연구개발비)

구분	표본 크기	(로그)매출액		영업이익		연구개발비	
		ATT	P	ATT	P	ATT	P
정책효과 (0 초과)	14,872	0.052* (0.029)	0.069	-38.108 (23,614)	0.107	10,078 (9,729)	0.300
정책효과 (0)	68,463	0.090*** (0.014)	0.000	-9,931* (5,721)	0.083	-2,485** (1,226)	0.043

주: 1. 금액 단위는 백만원

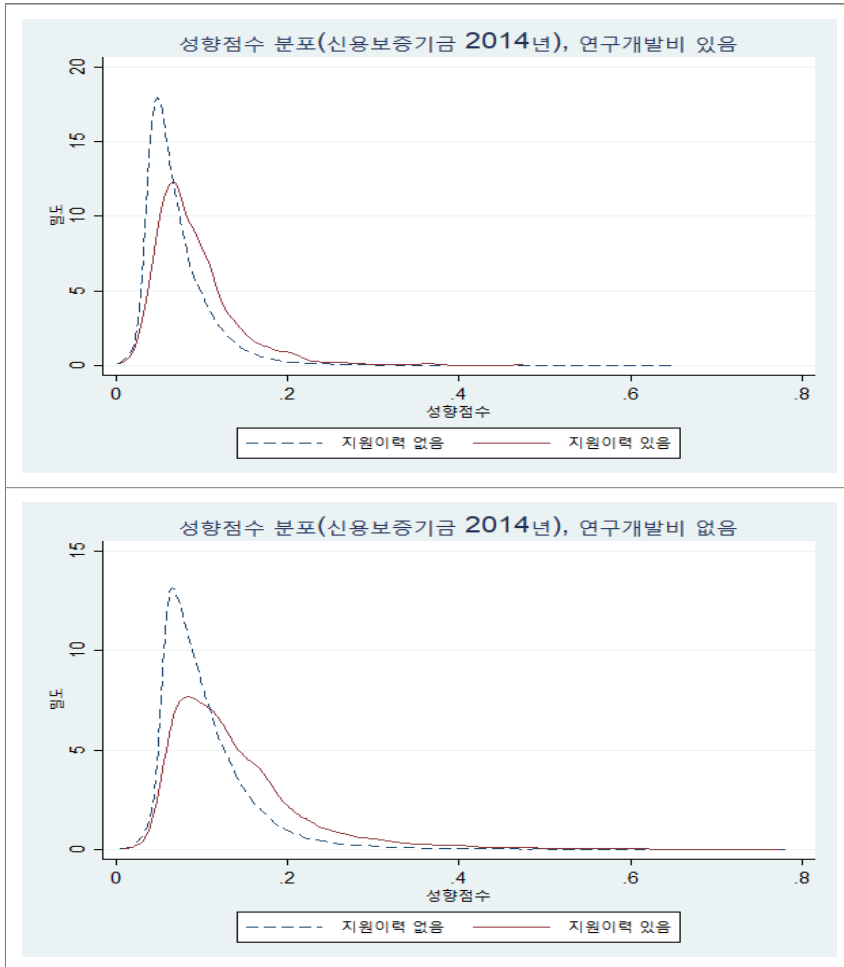
2. 통계적 유의성은 \*\*\* 1%, \*\* 5%, \* 10%와 같음

자료: 신용보증기금 내부자료; 한국기업데이터 자료

분석 결과에 따르면 2013년 연구개발비 지출이 없던 그룹의 효과가 매출액과 영업이익 측면에서는 상대적으로 양호한 편이라고 할 수 있으나, 연구개발비 지출 측면에서는 오히려 유의한 음(-)의 효과가 추정되었다. 매출액과 영업이익에 대한 결과는 앞서 소개한 이질성 분석 결과 매출액이 낮은 기업, 영업이익이 낮은 기업에서 상대적으로 긍정적인 효과가 추정되었던 것과 유사한 이유에서 나타났을 수 있다. 연구개발비에 대한 결과는 방향이 상이하며, 연구개발비가 0인 그룹이 공유하는 업종, 업력 등의 특성과 연관되어 나타날 가능성이 있다.

다음 [그림 Ⅲ-15]는 이질성 분석에 이용된 두 그룹의 성향점수 추정 분포를 나타내며, 그룹별로 유사한 특성을 공유하는 처치군과 대조군이 상당수 존재하는 것을 확인할 수 있다.

[그림 Ⅲ-15] 신용보증기금 성향점수 분포도(이질성: 연구개발비)



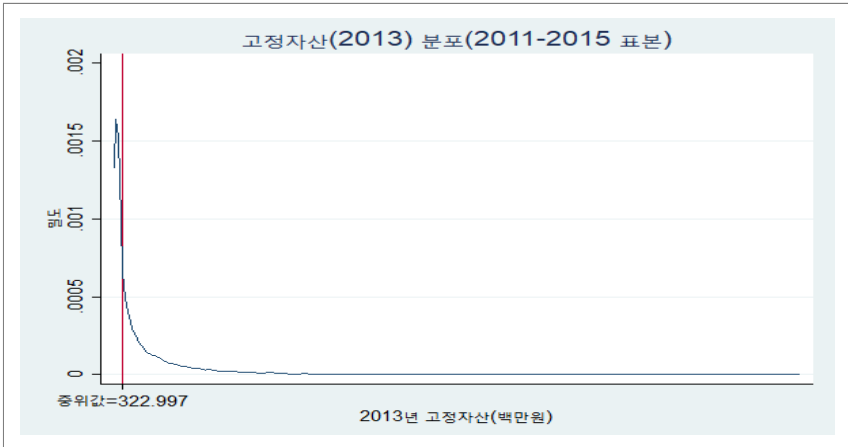
자료: 신용보증기금 내부자료; 한국기업데이터 자료

### 5) 이질성 분석 5: 고정자산 규모

기업의 자산 규모는 과거의 영업활동 내력, 산업의 특징, 업력 등의 다양한 요소를 반영한다. [그림 Ⅲ-16]에 제시된 고정자산 분포에 따르면 고정자산의 중위값은 322.997백만원이며, 중위값을 기준으로 그룹을 분리하여 이질성 분석을 수행하였다. 2013년 고정자산이 중위값 이상인 기업은 4만 1,949개 기업이고, 그중 2014년도 지원기업은 3,840개이다. 중위값 미만인 기업은 4만 1,948개이며, 그중 지원기업은 4,656개이다.

추정 결과는 <표 Ⅲ-23>에 제시하였다.

[그림 Ⅲ-16] 2011~2015년 표본의 고정자산 분포



자료: 신용보증기금 내부자료; 한국기업데이터 자료

<표 Ⅲ-23> 신용보증기금 신규지원 정책효과(이질성: 고정자산)

구분	표본 크기	(로그)매출액		영업이익		연구개발비	
		ATT	P	ATT	P	ATT	P
정책효과 (중위값 이상)	41,743	0.044*** (0.016)	0.007	-40,784*** (11,403)	0.000	-3,971 (3,743)	0.289
정책효과 (중위값 미만)	41,587	0.093*** (0.017)	0.000	6,773 (4,511)	0.133	-0,329 (1,316)	0.802

주: 1. 금액 단위는 백만원

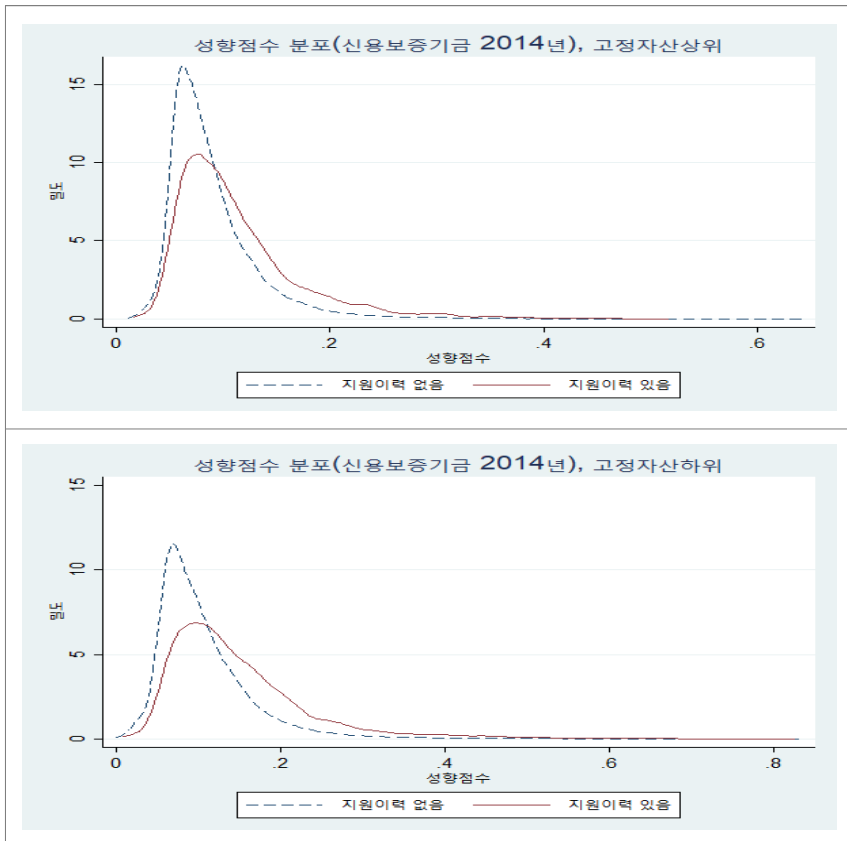
2. 통계적 유의성은 \*\*\* 1%, \*\* 5%, \* 10%와 같음

자료: 신용보증기금 내부자료; 한국기업데이터 자료

분석 결과에 따르면 고정효과 규모가 낮은 그룹의 효과가 매출액과 영업이익, 연구개발비 측면에서는 모두 양호한 편이다. 고정자산 규모와 경영 상태 사이의 기대되는 양(+)의 관계를 고려하면 이 결과는 앞서 소개한 이질성 분석 결과 매출액이 낮은 기업, 영업이익이 낮은 기업에서 상대적으로 긍정적인 효과가 추정되었던 것과 유사한 이유에서 나타났을 수 있다.

다음 [그림 Ⅲ-17]은 이질성 분석에 이용된 두 그룹의 성향점수 추정 분포를 나타내며, 그룹별로 유사한 특성을 공유하는 처치군과 대조군이 상당수 존재하는 것을 확인할 수 있다.

[그림 Ⅲ-17] 신용보증기금 성향점수 분포도(이질성: 고정자산)



자료: 신용보증기금 내부자료; 한국기업데이터 자료

6) 이질성 분석 6: 업종

기업의 업종에 따라 영업의 행태, 경쟁 정도, 운영자금의 상대적인 중요성 등 다양한 특징이 상이할 것임에 따라 이질적인 효과가 나타날 가능성이 존재한다. 산업대분류를 기준으로 제조업과 서비스업으로 표본을 분리하여 이질성 분석을 시도하였다. 제조업으로 분류되는 기업은 4만 2,026개, 서비스업으로 분류되는 기업은 3만 6,296개가 존재하였고, 그중 2014년 신용보증기금 지원을 받은 기업은 각각 4,147개와 3,863개이다.

추정 결과는 <표 Ⅲ-24>에 제시하였다.

<표 Ⅲ-24> 신용보증기금 신규지원 정책효과(이질성: 업종)

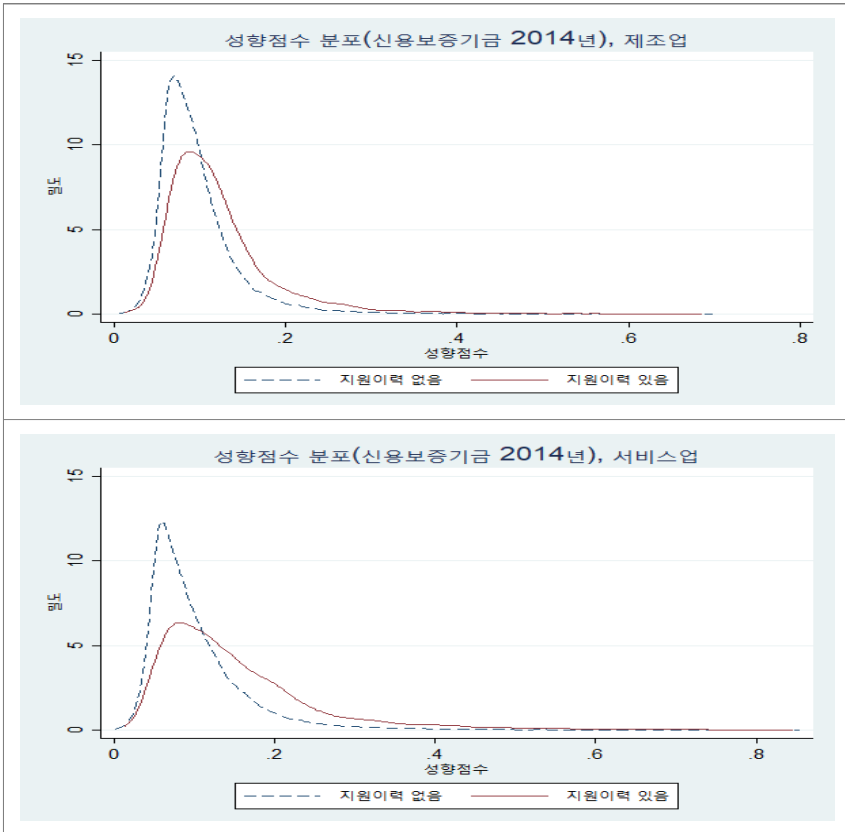
구분	표본 크기	(로그)매출액		영업이익		연구개발비	
		ATT	P	ATT	P	ATT	P
정책효과 (제조업)	41,704	0.059*** (0.018)	0.001	-30.592*** (9,040)	0.001	-3.039 (2,706)	0.261
정책효과 (서비스업)	36,210	0.085*** (0.019)	0.000	-0.800 (6,841)	0.907	0.061 (2,189)	0.978

주: 1. 금액 단위는 백만원  
 2. 통계적 유의성은 \*\*\* 1%, \*\* 5%, \* 10%와 같음  
 자료: 신용보증기금 내부자료; 한국기업데이터 자료

분석 결과에 따르면 서비스업 그룹의 효과가 매출액과 영업이익 측면에서 상대적으로 양호한 모습을 보이며, 표준오차를 고려하면 매출액에 비해 영업이익에서 보다 분명히 대비되는 효과가 나타난다. 이러한 차이는 업종의 특성으로부터 나타난 효과일 가능성도 있고, 업종에 관계된 기업 규모, 업력 등 업종 특성 이외의 다른 측면에서 기인한 이질성일 가능성도 있다.

[그림 Ⅲ-18]은 이질성 분석에 이용된 두 그룹의 성향점수 추정 분포를 나타내며, 그룹별로 유사한 특성을 공유하는 처치군과 대조군이 상당수 존재하는 것을 확인할 수 있다.

[그림 III-18] 신용보증기금 성향점수 분포도(이질성: 업종)



자료: 신용보증기금 내부자료; 한국기업데이터 자료

### 7) 이질성 분석 7: 기업 규모

매출액, 영업이익, 연구개발비의 규모 및 업력과도 밀접하게 관련되어 있으나, 기업 규모 또한 중요한 특성이며 의사결정의 문제 및 다양한 정부 정책 측면에서의 차이 등 잠재적으로 이질적인 효과를 발생시킬 수 있다. 기업 규모에 따른 이질성은 중기업 및 소기업, 그리고 소상공인의 두 그룹으로 분리하여 이질성 분석을 시도하였다. 중기업 및 소기업으로 분류되는 기업은 4만 4,355개이며, 지원기업은 4,648개이다. 서비스업으로 분류되는 기업은 3만 9,542개이며, 그중 지원기업은 3,848개이다.

추정 결과는 <표 Ⅲ-25>에 제시하였다.

<표 Ⅲ-25> 신용보증기금 신규지원 정책효과(이질성: 기업 규모)

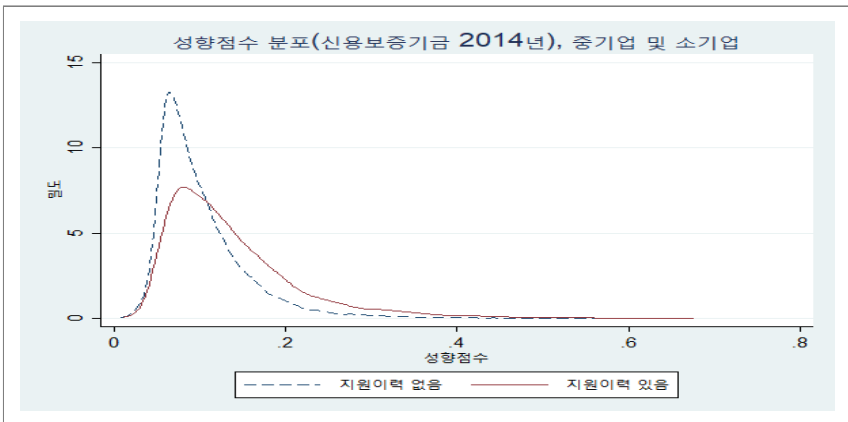
구분	표본 크기	(로그)매출액		영업이익		연구개발비	
		ATT	P	ATT	P	ATT	P
정책효과 (중·소기업)	44,029	0.045*** (0.015)	0.003	-18.532* (9.737)	0.057	-1.292 (3.724)	0.729
정책효과 (소상공인)	39,078	0.117*** (0.016)	0.000	2.269 (4.273)	0.595	-0.704 (0.960)	0.463

주: 1. 금액 단위는 백만원  
 2. 통계적 유의성은 \*\*\* 1%, \*\* 5%, \* 10%와 같음  
 자료: 신용보증기금 내부자료; 한국기업데이터 자료

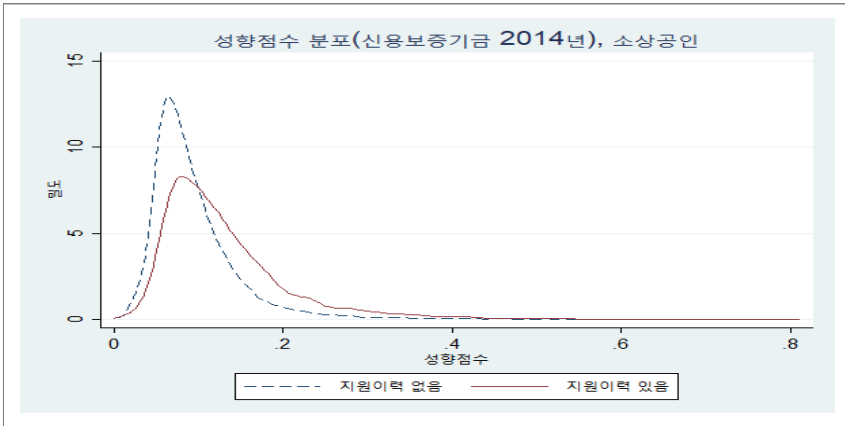
분석 결과에 따르면 소상공인 그룹의 효과가 매출액과 영업이익 측면에서 상대적으로 양호한 모습을 보인다. 기업 규모는 매출액의 규모, 영업이익의 규모, 업력, 고정자산의 규모 등 앞서 분석한 다양한 특성들과 밀접하게 관련되어 있을 것이며, 기업 규모에 따른 이질적인 효과도 이러한 특성들과 종합하여 고려할 필요가 있다.

[그림 Ⅲ-19]는 이질성 분석에 이용된 두 그룹의 성향점수 추정 분포를 나타내며, 그룹별로 유사한 특성을 공유하는 처치군과 대조군이 상당수 존재하는 것을 확인할 수 있다.

[그림 Ⅲ-19] 신용보증기금 성향점수 분포도(이질성: 기업 규모)



[그림 III-19]의 계속



자료: 신용보증기금 내부자료; 한국기업데이터 자료

#### 8) 이질성 분석 8: 수도권 여부

기업의 위치는 토지의 점유 비용과 밀접하게 관련되어 있으므로 일반적인 재무 상태와도 관련성이 존재할 수 있으나, 업종의 밀집도 및 노동시장의 차이, 집적으로 인한 생산성 등 다양한 측면과도 또한 관련되어 있다. 기업의 위치에 따른 이질성은 수도권과 비수도권의 두 그룹으로 분리하여 이질성 분석을 시도하였다. 수도권 소재 기업은 4만 4,198개이며, 지원기업은 4,656개이다. 비수도권 소재 기업은 3만 9,699개이며, 그중 지원기업은 3,840개이다. 그러나 자료의 한계로 인해 소재지 정보가 기업이 자료를 업데이트한 최신 시점에 한정하여 이용할 수 있는 관계로, 정보가 2014년 지원 당시와 일치하지 않는 단점이 존재한다는 점에는 유의할 필요가 있다.

추정 결과는 <표 III-26>에 제시하였다.

<표 III-26> 신용보증기금 신규지원 정책효과(이질성: 지역)

구분	표본 크기	(로그)매출액		영업이익		연구개발비	
		ATT	P	ATT	P	ATT	P
정책효과 (수도권)	43,636	0.076*** (0,018)	0,000	-17,544** (8,225)	0.033	-5,382 (3,319)	0.105

〈표 III-26〉의 계속

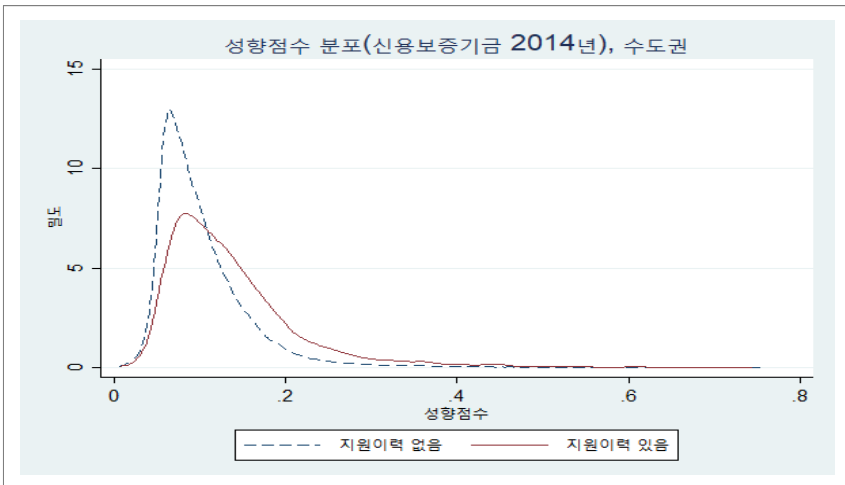
구분	표본 크기	(로그)매출액		영업이익		연구개발비	
		ATT	P	ATT	P	ATT	P
정책효과 (비수도권)	39,375	0.079*** (0.018)	0.000	-15.366* (8.288)	0.064	1.010 (1.939)	0.602

주: 1. 금액 단위는 백만원  
 2. 통계적 유의성은 \*\*\* 1%, \*\* 5%, \* 10%와 같음  
 자료: 신용보증기금 내부자료; 한국기업데이터 자료

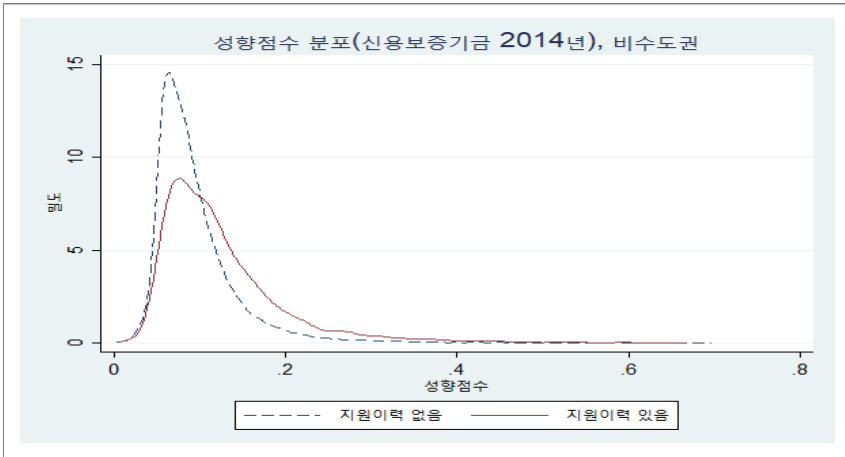
분석 결과에 따르면 지역에 따른 이질성은 거의 존재하지 않는 것으로 보이며, 매출액과 영업이익에 대한 추정치는 질적 및 양적 측면에서 매우 유사한 수준이다. 연구개발비는 수도권의 경우 음(-)의 효과가 존재하는 것으로 여길 수 있으나, 비수도권은 통계적으로 효과가 존재하지 않는다고 볼 수 있는 차이가 존재한다. 그러나 표준오차를 종합하여 고려하면 분명하게 이질성이 드러난다고 판단하기에는 어려움이 있다.

〔그림 III-20〕은 이질성 분석에 이용된 두 그룹의 성향점수 추정 분포를 나타내며, 그룹별로 유사한 특성을 공유하는 처치군과 대조군이 상당수 존재하는 것을 확인할 수 있다.

〔그림 III-20〕 신용보증기금 성향점수 분포도(이질성: 지역)



[그림 III-20]의 계속



자료: 신용보증기금 내부자료; 한국기업데이터 자료

### 9) 이질성 분석 9: 지적재산권 보유 여부

지적재산권 보유 숫자는 기업의 기술력과 밀접하게 연관되어 있을 수 있고, 기술력의 차이는 자금 지원 이후의 성과 차이로 연결될 수 있다. 분석을 위해 2010년 기준 지적재산권 보유 현황을 바탕으로 지적재산권 보유 기업과 미보유 기업의 두 그룹으로 분리하여 이질성 분석을 시도하였다. 지적재산권을 보유한 기업은 6,823개이며, 지원기업은 503개이다. 미보유 기업은 7만 7,074개이며, 그중 지원기업은 7,993개이다.

추정 결과는 <표 III-27>에 제시하였다.

<표 III-27> 신용보증기금 신규지원 정책효과(이질성: 지적재산권)

구분	표본 크기	(로그)매출액		영업이익		연구개발비	
		ATT	P	ATT	P	ATT	P
정책효과 (보유)	6,771	0.023 (0.039)	0.564	-54.982* (33,367)	0.099	15.913 (13,268)	0.230
정책효과 (미보유)	75,686	0.081*** (0.013)	0.000	-8.696 (5,533)	0.116	-5.427* (2,833)	0.055

주: 1. 금액 단위는 백만원

2. 통계적 유의성은 \*\*\* 1%, \*\* 5%, \* 10%와 같음

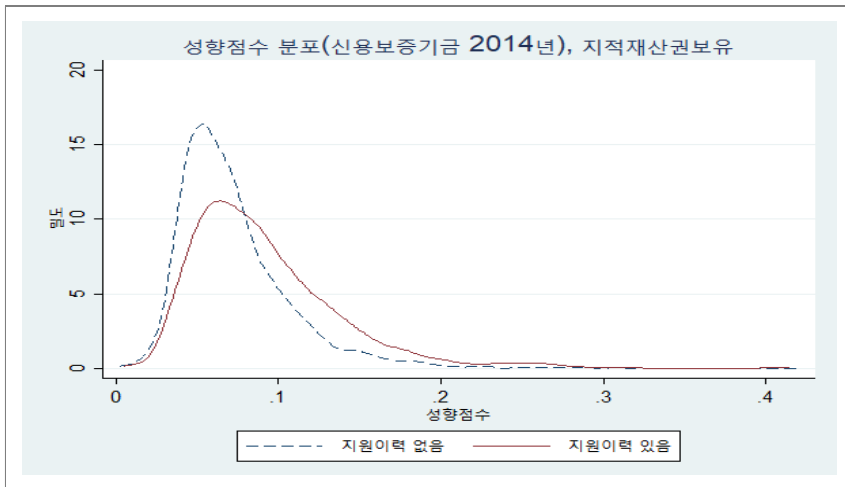
자료: 보증기금 내부자료; 한국기업데이터 자료

분석 결과에 따르면 지적재산권을 보유하고 있지 않았던 그룹에서 매출액에 대한 긍정적인 효과가 강하게 나타났으며, 영업이익에 대한 부정적인 효과는 상대적으로 거의 나타나지 않았다. 반면 연구개발비에 대한 음(-)의 효과는 더 강하게 나타났다. 이 결과는 앞서 소개된 이질성 분석 결과에서 경영상 지표가 낮게 나타난 그룹이 효과성 측면에서는 상대적으로 양호한 모습을 보인 것과 유사한 결과로 생각할 수 있으며, 관련 변수들이 모두 상관관계가 존재할 것으로 짐작되는 변수들인 것을 고려하면 일관된 결과로 볼 수 있다.

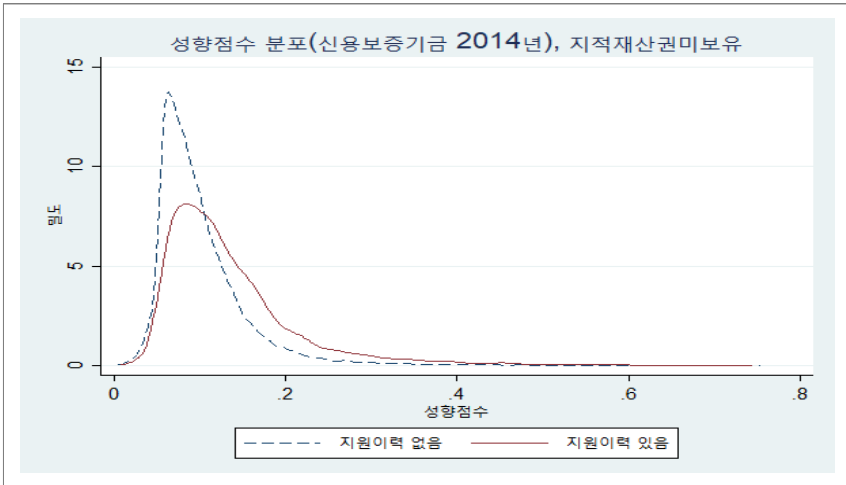
그러나 앞서 제시된 결과들과 다르게 지적재산권을 보유한 기업의 수가 적으며 따라서 성향점수매칭의 신뢰성이 다소 낮을 수 있다는 점에는 유념할 필요가 있다.

[그림 Ⅲ-21]은 이질성 분석에 이용된 두 그룹의 성향점수 추정 분포를 나타내며, 그룹별로 유사한 특성을 공유하는 처치군과 대조군이 상당수 존재하는 것을 확인할 수 있다.

[그림 Ⅲ-21] 신용보증기금 성향점수 분포도(이질성: 지적재산권)



[그림 III-21]의 계속



자료: 신용보증기금 내부자료; 한국기업데이터 자료

#### 10) 이질성 분석 10: 가족주식소유

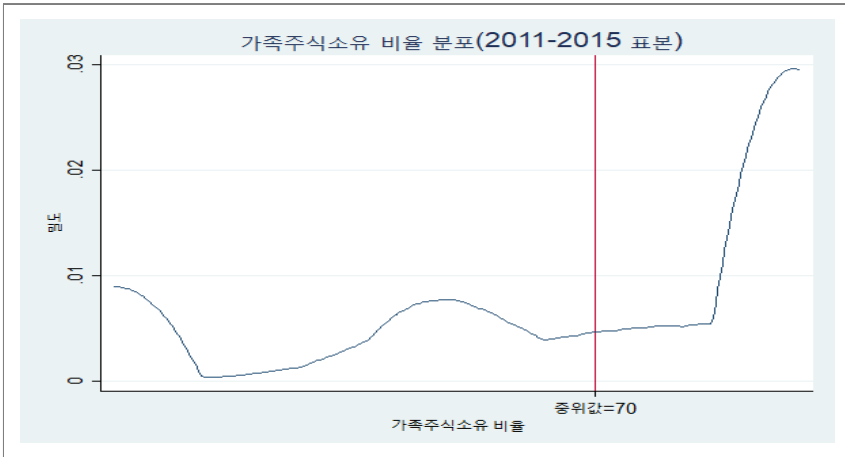
가족주식소유에 대한 정보는 한국기업데이터 주주 현황 자료를 바탕으로 구축하였다. 주주현황 자료는 주주별 경영실권자 관계에 대한 정보를 포함하며, 이는 본인, 가족, 친척, 관계회사, 관계회사임원, 타인, 기타로 정의되어 있다. 이 중 본인, 가족, 친척으로 경영실권자와의 관계를 정의한 경우 해당 주주를 가족주식소유자로 정의하고, 가족주식소유자로 분류된 주주들의 전체 주식 수 대비 소유 비율을 기업 단위로 합산하여 기업별 가족주식소유 비율을 구하였다. 비율이 100을 약간 상회하는 관측치가 소수 존재하였으나 이 경우 합산 오류로 판단하여 100으로 값을 수정하였다.

데이터의 제약으로 인해 자료의 기준 연도는 대부분 2019년이라는 한계점이 있으나, 본 연구의 분석 기간과 일정 수준의 경향성을 가지는 것으로 가정하였고 이는 해석 과정에서 유의할 필요가 있다. 가족주식소유 비율의 분포는 [그림 III-22]에 제시하였으며, 중위값인 70을 기준으로 두 그룹으로 분리한 후 이질성 분석을 수행하였다.

가족주식소유 비율 변수가 유효한 기업 중 중위값 이상인 기업은 3만 888개

이며, 지원기업은 3,974개이다. 중위값 미만 기업은 2만 9,485개이며, 그중 지원기업은 2,343개이다.

[그림 III-22] 2011~2015년 표본의 가족주식소유 비율 분포



자료: 신용보증기금 내부자료; 한국기업데이터 자료

추정 결과는 <표 III-28>에 제시하였다.

<표 III-28> 신용보증기금 신규지원 정책효과(이질성: 가족주식소유 비율)

구분	표본 크기	(로그)매출액		영업이익		연구개발비	
		ATT	P	ATT	P	ATT	P
정책효과 (비율상위)	30,396	0.060*** (0.018)	0.001	-33.464*** (11.163)	0.003	1,894 (2,449)	0.439
정책효과 (비율하위)	29,380	0.056** (0.024)	0.021	-19.462* (11.688)	0.096	-4,344 (4,457)	0.330

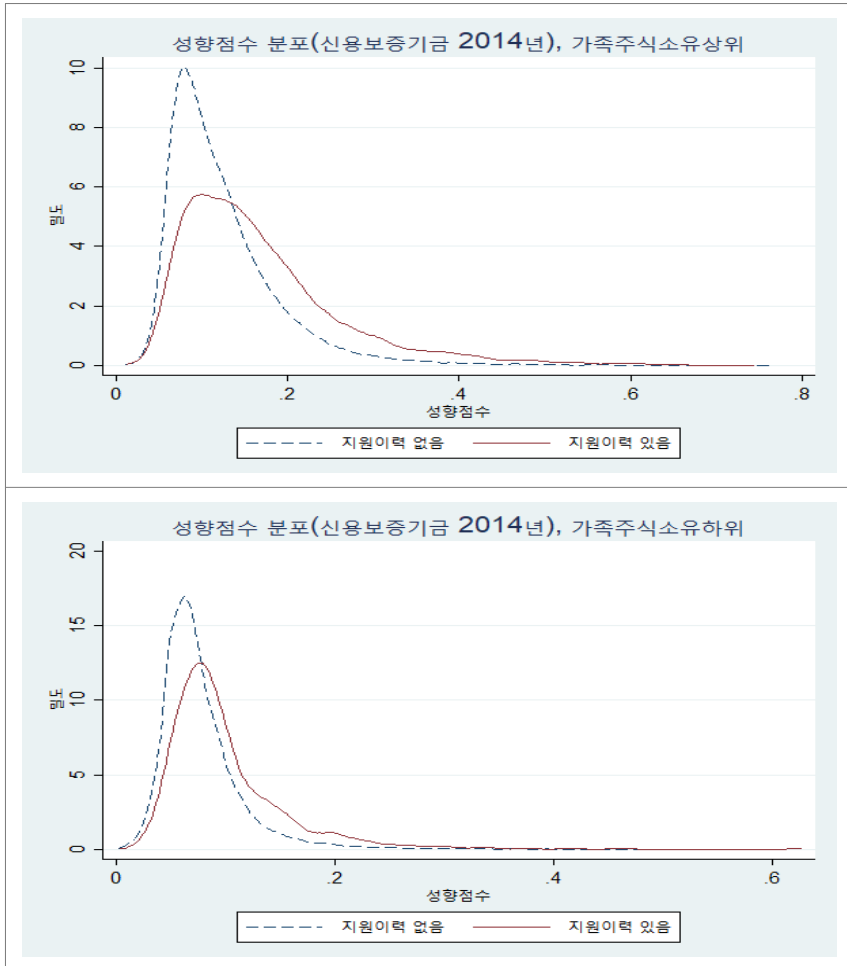
주: 1. 금액 단위는 백만원  
2. 통계적 유의성은 \*\*\* 1%, \*\* 5%, \* 10%와 같음

자료: 신용보증기금 내부자료; 한국기업데이터 자료

추정 결과 매출액에는 양(+)의 효과, 영업이익에는 음(-)의 효과가 통계적으로 유의하게 관찰되었으나 가족주식소유 비율에 따른 이질성은 뚜렷하게 나타나지 않았다. 연구개발비에 대한 효과는 유의하지 않았다.

[그림 Ⅲ-23]은 이질성 분석에 이용된 두 그룹의 성향점수 추정 분포를 나타내며, 그룹별로 유사한 특성을 공유하는 처치군과 대조군이 상당수 존재하는 것을 확인할 수 있다.

[그림 Ⅲ-23] 신용보증기금 성향점수 분포도(이질성: 가족주식소유 비율)



자료: 신용보증기금 내부자료; 한국기업데이터 자료

### 11) 이질성 분석 11: 수출기업 여부

수출기업에 대한 정보는 한국기업데이터 「제품별 매출 현황 자료」를 바탕으로 구축하였다. 이 자료는 제품별 매출액 및 수출금액에 대한 정보를 연도별로 포함하고 있으며, 수출금액이 음수(-)로 기록되었거나 결측값인 관측치를 제거한 후 기업별로 2013년 이전 기간에 수출 실적을 기록한 제품을 보유한 기업을 수출기업으로 정의하였다.

수출실적을 보유한 기업은 1,027개 기업이며 255개 기업이 2014년 신용보증기금 지원기업이다. 미보유기업은 4만 2,882개 기업이며, 4,670개 기업이 지원기업이다. 이질성 분석은 실적 보유기업과 미보유기업으로 표본을 분리하여 수행하였으며 추정 결과는 <표 III-29>에 제시하였다.

<표 III-29> 신용보증기금 신규지원 정책효과(이질성: 수출기업 여부)

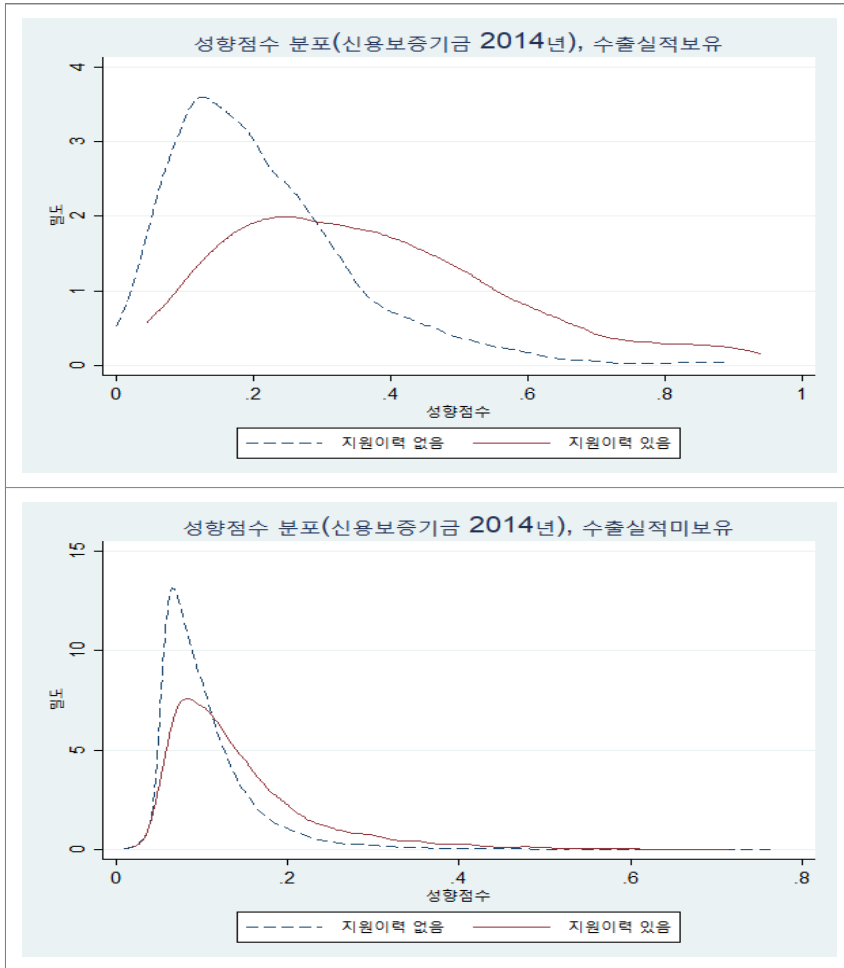
구분	표본 크기	(로그)매출액		영업이익		연구개발비	
		ATT	P	ATT	P	ATT	P
정책효과 (실적보유)	1,010	0.048 (0.062)	0.436	-0.475 (43,562)	0.991	-5,536 (8,014)	0.490
정책효과 (실적 미보유)	42,267	0.063*** (0.016)	0.000	-22,754** (9,604)	0.018	-6,991** (2,981)	0.019

주: 1. 금액 단위는 백만원  
 2. 통계적 유의성은 \*\*\* 1%, \*\* 5%, \* 10%와 같음  
 자료: 신용보증기금 내부자료; 한국기업데이터 자료

추정 결과 점추정치 기준으로는 영업이익을 제외하면 유사한 수준으로 결과가 추정되었으며, 영업이익 또한 표준오차를 고려하면 이질적인 결과로 단정짓기 어렵다. 또한 수출 실적을 보유한 기업의 숫자가 적은 관계로 성향점수가 충분히 가까운 기업들을 적절하게 탐색하지 못하였을 가능성이 높으므로 결과를 제한적으로 해석할 필요가 있다.

[그림 III-24]는 이질성 분석에 이용된 두 그룹의 성향점수 추정 분포를 나타내며, 그룹별로 유사한 특성을 공유하는 처치군과 대조군이 상당수 존재하는 것을 확인할 수 있다.

[그림 III-24] 신용보증기금 성향점수 분포도(이질성: 수출기업 여부)



자료: 신용보증기금 내부자료; 한국기업데이터 자료

## 12) 이질성 소결

〈표 III-30〉과 〈표 III-31〉에서 앞서 보고된 이질성 분석 결과를 종합하여 제시하였다.

〈표 III-30〉 신용보증기금 신규지원 이질성분석 요약 1

이질성	표본 수	(로그)매출액		영업이익		연구개발비	
		ATT	P	ATT	P	ATT	P
업력 (중위값 이상)	46,808	0.083*** (0.019)	0.000	-10,144 (9,258)	0.273	-0.538 (2,652)	0.839
업력 (중위값 미만)	36,537	0.106*** (0.017)	0.000	-3,139 (8,634)	0.716	-1.561 (2,092)	0.455
매출액 (중위값 이상)	41,796	0.037*** (0.012)	0.002	-17,122* (9,127)	0.061	-1,626 (2,688)	0.545
매출액 (중위값 미만)	41,534	0.069*** (0.014)	0.000	-4,726 (3,778)	0.211	-0.654 (1,144)	0.567
영업이익 (중위값 이상)	41,744	0.040*** (0.015)	0.008	-20,024** (9,771)	0.040	-5,249 (4,755)	0.270
영업이익 (중위값 미만)	41,161	0.115*** (0.017)	0.000	4,828 (5,211)	0.354	-1,159 (2,019)	0.566
연구개발비 (0 초과)	14,872	0.052* (0.029)	0.069	-38,108 (23,614)	0.107	10,078 (9,729)	0.300
연구개발비 (0)	68,463	0.090*** (0.014)	0.000	-9,931* (5,721)	0.083	-2,485** (1,226)	0.043
고정자산 (중위값 이상)	41,743	0.044*** (0.016)	0.007	-40,784*** (11,403)	0.000	-3,971 (3,743)	0.289
고정자산 (중위값 미만)	41,587	0.093*** (0.017)	0.000	6,773 (4,511)	0.133	-0,329 (1,316)	0.802

주: 1. 금액 단위는 백만원  
 2. 통계적 유의성은 \*\*\* 1%, \*\* 5%, \* 10%와 같음  
 자료: 신용보증기금 내부자료; 한국기업데이터 자료

〈표 III-31〉 신용보증기금 신규지원 이질성분석 요약 2

이질성	표본 수	(로그)매출액		영업이익		연구개발비	
		ATT	P	ATT	P	ATT	P
업종 (제조업)	41,704	0.059*** (0.018)	0.001	-30,592*** (9,040)	0.001	-3,039 (2,706)	0.261
업종 (서비스업)	36,210	0.085*** (0.019)	0.000	-0,800 (6,841)	0.907	0,061 (2,189)	0.978
기업 규모 (중·소기업)	44,029	0.045*** (0.015)	0.003	-18,532* (9,737)	0.057	-1,292 (3,724)	0.729
기업 규모 (소상공인)	39,078	0.117*** (0.016)	0.000	2,269 (4,273)	0.595	-0,704 (0,960)	0.463

〈표 III-31〉의 계속

이질성	표본 수	(로그)매출액		영업이익		연구개발비	
		ATT	P	ATT	P	ATT	P
지역 (수도권)	43,636	0.076*** (0.018)	0.000	-17.544** (8.225)	0.033	-5.382 (3.319)	0.105
지역 (비수도권)	39,375	0.079*** (0.018)	0.000	-15.366* (8.288)	0.064	1.010 (1.939)	0.602
지적재산권 (보유)	6,771	0.023 (0.039)	0.564	-54.982* (33.367)	0.099	15.913 (13.268)	0.230
지적재산권 (미보유)	75,686	0.081*** (0.013)	0.000	-8.696 (5.533)	0.116	-5.427* (2.833)	0.055
가족주식 (중위 이상)	30,396	0.060*** (0.018)	0.001	-33.464*** (11.163)	0.003	1.894 (2.449)	0.439
가족주식 (중위 미만)	29,380	0.056** (0.024)	0.021	-19.462* (11.688)	0.096	-4.344 (4.457)	0.330
수출기업 (실적 보유)	1,010	0.048 (0.062)	0.436	-0.475 (43.562)	0.991	-5.536 (8.014)	0.490
수출기업 (실적 미보유)	42,267	0.063*** (0.016)	0.000	-22.754** (9.604)	0.018	-6.991** (2.981)	0.019

주: 1. 금액 단위는 백만원  
 2. 통계적 유의성은 \*\*\* 1%, \*\* 5%, \* 10%와 같음  
 자료: 신용보증기금 내부자료; 한국기업데이터 자료

〈표 III-30〉은 업력 및 기업의 재무상태 관련 변수들에 대한 이질성 분석 결과를 담고 있다. 표준오차를 종합적으로 고려했을 때 일부 제한적으로 해석할 필요가 있으나 업력이 낮을수록, 매출액이 낮을수록, 영업이익이 낮을수록, 연구개발비가 0인 그룹에서, 그리고 고정자산이 작은 그룹에서 매출액 증가에 대한 양(+)의 효과가 크고 영업이익에 미치는 음(-)의 효과는 상대적으로 긍정적인 방향을 띠는 경향이 존재한다. 상대적으로 신생 기업 혹은 재무적으로 상황이 좋지 않은 기업에 대해 정책금융의 효과가 양호하게 나타나는 측면이 있는 것을 의미한다. 2013년 기준 매출액과 영업이익은 0.5 이상의 상관계수를 보이므로 두 변수에서 기인한 이질성은 사실상 유사한 이유에서 기인했을 가능성이 높을 것으로 예상된다. 매출액, 영업이익과 고정자산의 상관계수도 각각 0.38, 0.44 정도로 높은 편이다. 업력과 연구개발비는 매출액, 영업이익, 고정자산과의 상관계수가 최대 0.2 수준으로, 매출액

등의 변수와 동일한 이유로 이질성이 발생했다고 판단하기에는 상대적으로 어렵다.

〈표 Ⅲ-31〉은 업종, 기업 규모, 기업 위치, 지적재산권 보유 여부, 가족주식소유 비율, 수출기업 여부에 대한 이질성 분석 결과를 정리하였다. 업종의 경우 제조업에 비해 서비스업이 영업이익 측면에서 상대적으로 양호한 추정 결과를 보여 준다. 이질성이 나타나는 이유에 대해서는 업종과 앞서 검토한 변수들과의 뚜렷한 관계를 예상하기 쉽지 않으므로 단정짓기 어렵다. 기업 규모에 대해서는 소상공인 그룹이 상대적으로 양호한 결과를 보이며, 〈표 Ⅲ-30〉과 연계하여 검토해 보면 기업 규모와 매출액 등의 경영실적 변수와의 관련성이 높을 것으로 보인다. 가령 2013년 매출액과 소상공인에 0을 할당하고 중기업, 소기업에 1을 할당한 더미변수의 상관관계는 0.44 수준이다.

지적재산권 보유 여부는 연구개발비와의 상관성이 높을 수 있다. 그러나 표본에서의 연구개발비는 2011년부터의 자료이고 지적재산권 보유 여부는 2010년 기준인 것을 고려하면 표본으로부터 관계성을 밝히기는 어려우나, 지적재산권을 보유하지 않은 기업이 연구개발비에서는 더욱 부정적, 매출액과 영업이익 측면에서는 상대적으로 긍정적이라는 결과는 〈표 Ⅲ-30〉의 연구개발비 대상 이질성 분석 결과와 궤를 같이 하며 이질성이 발생한 근본 원인이 유사할 것이라는 예상을 어느 정도 뒷받침한다. 지역, 가족주식소유 비율, 수출기업 여부에 대한 이질성은 상대적으로 그 경향성이 뚜렷하게 드러났다고 보기 어렵다.

제시된 이질성 분석의 한계는 정확한 메커니즘에 대한 식별이 누락되어 있다는 점에 더해, 표본이 포함하는 기업의 수가 적어짐에 따라 처치군과 대조군이 적절하게 식별되고 있는지에 대한 우려가 존재할 수 있다. 이질성이 통계적으로 유의하게 발생했는지, 그 원인은 무엇인지를 심도 있게 검토하는 것은 중요하며, 이질성 분석 결과가 신뢰성 있게 추정되었는지의 여부도 매우 중요하다. 그러나 앞서 언급했듯이 제시된 본 장의 이질성 분석의 주된 목적은 신용보증기금 지원의 이질적 효과에 대한 정확한 결론을 내리는 것이 아닌 이후 본고에서 소개할 머신러닝 방법론을 이용한 이질적 효과

분석을 위한 기초 작업을 수행하는 것이다. 따라서 제시된 이질성 분석을 최종적인 분석 결과로 받아들이기보다는 본고의 중간 단계로 이해하는 것이 적절하다.

#### 4. 소결

본 장에서는 2014~2015년 신용보증기금 지원 이력을 바탕으로 성향점수 매칭 방법론을 활용하여 정책금융 신규지원이 매출액, 영업이익, 그리고 연구개발비에 미친 영향에 대한 실증분석을 시도하였다. 그리고 매출액, 영업이익, 업력, 기업의 위치 등 기업의 다양한 특성에 대한 이질적인 정책효과의 존재 및 경향성에 대한 검토를 수행하였다.

지원기업과 대조군의 지원 시점 이전의 추세를 적절하게 통제하기 위한 목적으로 지원 시점의 3년 이전인 2011년부터 지원 이후 1년 시점인 2015년까지의 표본을 주된 표본으로 설정하여 분석을 수행하였으며, 해당 표본을 이용한 성향점수매칭을 수행한 결과 처치군과 대조군의 지원 시점 이전 3년간의 추세는 적절하게 통제된 것으로 판단된다.

분석 결과 매출액에는 대체로 양(+)의 효과가, 영업이익과 연구개발비에 대해서는 유의하지 않거나 음(-)의 효과가 나타났다. 이러한 결과는 선행연구에서도 보고된 바 있으며, 정책금융을 이용한 기업의 행태 변화로 인한 장기적 안목의 영업활동 증가 또는 반대로 수익성이 낮은 영업활동 증가 등 다양한 원인을 통해 나타날 수 있다. 그러나 정확한 원인에 대해 본 연구에서 분석하는 것은 한계가 있었다.

이질성 분석 결과 신생 기업, 재무적 상황이 좋지 않은 기업, 서비스업, 기업 규모가 작은 기업, 지적재산권을 보유하지 않은 기업으로부터 추정된 정책효과가 그렇지 않은 기업에 비해 상대적으로 긍정적인 방향으로 추정되었다. 종합하면 상대적으로 영업활동이 안정적인 궤도에 오르지 않은 기업들에 지원하였을 때의 효과가 긍정적인 방향을 나타낸다고 생각된다. 그러나 앞서 언급한 특성들은 모두 상관관계가 강하게 존재할 것으로 예상됨에 따라 이질성을 유발하는 원인을 정확하게 특정하기는 어렵고, 매칭의 질을

최대한 유지하기 위해 표본을 세밀하게 분리하지 않고 두 개의 집단으로 나누어서 분석하였다는 한계점이 존재함을 언급하고자 한다.

그러나 본 장의 주된 목적은 정책금융에 대한 인과관계 및 이질성 분석 결과를 새롭게 제시하는 것이 아닌, 본고의 후반부에서 소개될 머신러닝 방법론을 이용한 분석 결과와의 비교를 위한 기초 작업이기에 이상으로 본 장을 갈무리하고자 한다.

---

## IV. 머신러닝의 개관

---

대규모 데이터셋을 구하기가 점점 더 쉬워지고 컴퓨터의 연산능력이 더욱 발전하면서 많은 예측변수들을 사용해서 결과를 예측하려는 시도들도 더 많은 관심을 끌고 있다(Giannone et al., 2021). 예를 들어 GDP 예측에 관심이 있는 연구자들은 섹터별 성장률, 섹터별 고용률, 이자율, 날씨, 인구이동 등 수많은 변수들을 이용해 GDP 예측을 시도해 볼 수 있을 것이다. 그 뿐만 아니라 수만 개의 유전 정보를 통해서 질병의 발생 가능성을 예측해 볼 수도 있고 사진에 있는 픽셀의 정보를 모아서 사진이 개인지 고양이인지를 예측해 볼 수도 있다.

머신러닝이라고 불리는 기법들은 이처럼 많은 변수들을 이용해 결과를 예측하는 데 광범위하게 사용되고 있다. 사실 머신러닝이라는 용어는 위에서 보았듯이 경제, 의학, 컴퓨터 공학 등 매우 다양한 형태의 데이터를 다루는 여러 분야에서 사용되는 매우 다양한 기법들을 포괄하는 용어이므로 명확한 정의를 내리기가 어렵고(Athey, 2019), 다양한 머신러닝의 방법론과 적용 사례를 모두 소개하기는 불가능에 가까울 것이다. 따라서 본 장 제1절에서는 현실에서 자주 쓰이는 머신러닝 기법들에 대하여 간단히 소개하고 제2절에서는 실제 적용 사례에서 머신러닝이 사람의 판단보다 더 나은 의사결정을 할 수 있는지, 기존의 계량 방법론보다 예측성고를 향상시킬 수 있는지에 관한 연구 논문들을 소개하고자 한다.

### 1. 머신러닝 주요 알고리즘 소개

우선 머신러닝의 기법들을 소개하기 전에 머신러닝이 기존의 전통적인 방법론들에 비해 어떠한 특징을 가지고 있는지에 대해서 간단히 짚고 넘어가

고자 한다. 최소자승법(Ordinary Least Square, OLS)을 이용한 선형회귀 분석과 같은 전통적인 방법론들은 대부분 통계적으로 좋은 성질을 가지고 있다. 예를 들어 OLS 선형회귀모형은 계수 추정에 있어서 불편성과 일치성을 가지고 있으므로 관측치의 수가(설명변수의 수에 비해) 충분히 많고 참 모형이 선형이라면 OLS의 예측성과는 다른 어떤 방법론보다 뛰어날 가능성이 크다.

하지만 다루는 데이터가 매우 많은 설명변수를 가지는 경우에는 설명변수의 개수가 관측치의 수에 비해서도 상당히 큰 경우가 많다. 또한 설명변수들 간에 복잡한 상호작용이 존재하는 경우에는 선형모형이 이를 잘 포착해 내기 어려울 수도 있다. 예를 들어 수만 개의 유전정보가 건강에 미치는 영향을 예측하고자 할 때 설명변수의 수는 관측치의 숫자에 대비해서 상당히 많을 것이며, 또한 한 유전정보가 다른 유전 정보와 어떠한 상호작용을 통해 질병을 유발하게 될지에 대해서는 사전적으로 매우 짐작하기 어려운 측면이 존재한다.

이처럼 잠재적인 설명변수가 매우 많은 경우 (혹은 변수들의 복잡한 상호작용을 기대하여 이를 모두 터미변수로 설정하는 경우) OLS 선형회귀는 모형을 과대적합(over-fitting)하는 경향을 가지게 된다. 모형이 과대적합되었다는 것은 참 모형에서 오차(error)로 발생한 부분까지도 모형 안으로 끌어오게 되는 경향을 의미하는데, 이러한 과대적합이 발생하는 경우 새로운 표본에서 기존 표본과는 다른 오차가 실현되면 추정된 모형이 새로운 표본을 잘 예측하지 못할 수 있다. 이러한 상황을 다르게 표현하자면 추정량의 편향(bias)은 작지만 분산(variance)이 큰 경우이다. 설명변수가 많아서 OLS 추정량이 모형을 과대적합하는 경우 설명변수의 수를 줄이면 모형의 bias는 다소 증가할 수 있지만 모형이 단순화되면서 어떤 표본을 가져와도 모형의 예측성과가 크게 변동하지 않는데(즉 낮은 변동성), 이를 가리켜 편향·분산 트레이드 오프(bias-variance trade-off)라고 말한다.

머신러닝 기법들은 편향·분산 트레이드 오프를 사용하여 특별히 예측의 성과를 높이기 위해 개발되었다고 할 수 있다(Kleinberg et al., 2015). 물론 기

존에도 편향·분산 트레이드 오프를 이용하여 최선의 모형을 찾으려고 하는 시도는 있어 왔다.<sup>23)</sup> 하지만 머신러닝이 전통적인 방법론들과 다른 것은 최적의 모형을 찾는 과정에서 사람의 휴리스틱에 의존하는 기준을 따라 모형을 선택하는 것이 아니라 실제로 수만 가지 모형을 만들어 본 다음, 그 모형들을 샘플데이터에 대입해서 모형들의 예측성과를 파악하여 성과가 가장 좋은 모형을 스스로 선택해 준다는 것이다. 이러한 점에서 머신러닝은 본질적으로는 데이터에 기반(data-driven)하여 모형 선택을 해주는 알고리즘을 말한다고 볼 수 있다(Athey, 2019).

머신러닝이 어떻게 데이터에 기반하여 스스로 모형 선택을 하는지 머신러닝 관련 문헌에서 자주 사용되는 Lasso와 결정트리(decision tree)의 사례를 간단하게 소개하고자 한다. Lasso는 쉽게 말하면 OLS에서 설명력이 떨어지는 변수를 완전히 배제시키는(변수의 계수를 0으로 추정하는) 알고리즘이라고 할 수 있다. 구체적으로 Lasso 추정량은 다음과 같이 정의되는데, 계수의 추정량이 잔차의 제곱합을 줄이는데 많은 기여를 하지 못할 경우 우측의 패널티 항만을 키우게 되므로 계수를 0으로 추정하게 된다.

$$\beta^{lasso} = \arg \min_{\beta} \sum_{i=1}^n (y_i - x'_i \beta)^2 + \lambda \sum_{j=1}^p |\beta_j| \quad \text{식 (3)}$$

이때 Lasso 알고리즘은 다음과 같은 과정을 거쳐서 스스로 모형 선택을 한다. 우선 임의의 람다( $\lambda$ )를 정한 다음 훈련 표본(training set)을 통해 모형을 추정한다. 모형을 추정한 다음 검증 표본(validation set)에서 모형의 예측성과를 측정한다. 그리고 또 다른 람다( $\lambda$ )를 선택한 다음 동일한 과정을 거친다. 이러한 과정을 여러 번 거친 다음 가장 좋은 예측성과를 가져다 주는 람다( $\lambda$ )를 정하고 계수를 모두 추정하면 모형 선택이 마무리된다.<sup>24)</sup> 모형 선택은 데이터의 특징과 머신러닝 방법론의 특징에 의해서 이루어지기

23) AIC, BIC, Adjusted R<sup>2</sup>와 같은 기준을 가지고 완전탐색(Exhaustive Search), 전진선택(Forward Selection), 후진제거(Backward elimination), 스텝와이즈(stepwise selection) 등을 통해 모형 선택을 할 수 있다.

24) 이러한 과정을 empirical tuning이라고 하기도 한다(Mullainathan and Spiess, 2017).

때문에 연구자의 의도가 모형 선택에 개입될 여지는 적다고 할 수 있으며, 연구자의 직관이나 경제이론이 잘 발견하지 못하는 모형의 패턴을 발견할 가능성은 더 크다고 할 수 있다(Sansone, 2019).

한편 결정트리의 경우에는 Lasso와는 다르게 비모수적인 방법으로 통해 모형 선택을 하는 머신러닝 방법이다. 결정트리는 한 개 이상의 조건문들을 이용해 설명변수들을 조건에 따라 분류하여 데이터를 파티션으로 구분한 후 각각의 파티션에 대해 하나의 예측값을 부여하는 모형을 설정한다. 일반적으로 결정트리는 탐욕적(greedy) 알고리즘을 통해서 모형을 구성한다. 조건문이 하나도 없는 트리의 깊이가 0인 뿌리노드(root node)에서 시작하여 여기서 데이터의 종속변수를 가장 잘 분할하는 조건문을 설정하여 깊이가 1인 자식노드를 형성한다. 이후 각각의 자식노드에서 또 다시 종속변수를 가장 잘 분할하는 조건문을 설정하여 깊이가 2인 자식노드를 형성한다. 이렇게 분할이 이루어지다 최종노드(terminal node, 잎이라고도 함)에 이르게 되면 하나의 트리모형이 완성된다.

결정트리는 가장 적절한 노드의 깊이를 모색하면서 스스로 최적의 모형을 찾아낸다. 우선 임의의 노드의 깊이를 정하고 종속변수를 가장 잘 분할하는 조건문들을 훈련표본에서 찾아낸다. 그리고 검정표본에서 예측성과를 측정한다. 다음 또 다른 노드의 깊이를 정하여 동일한 과정을 거친다. 이러한 과정을 여러 번 거치면서 가장 좋은 예측성과를 가져다주는 노드의 깊이를 정하게 되고 해당 노드의 깊이에서 트리모형의 선택이 완료된다.

Lasso와 결정트리는 상당히 다른 접근법을 취하지만 공통적으로 모형의 복잡도를 조절해 주는 정규화(Regularization) 방법을 가지고 있으며-Lasso의 람다( $\lambda$ ), 결정트리의 노드 깊이가 이에 해당 - 정규화 수준은 데이터에 기반하여 예측성과를 최적화<sup>25)</sup>하는 방향으로 스스로 정해진다. 이러한 특징은 대부분의 머신러닝 기법들을 규정하는 특징이라고 할 수 있다(Mullainathan and Spiess, 2017). <표 IV-1>은 자주 쓰이는 머신러닝 방법론들의 특징과

25) 예측성과를 평가하는 방법은 여러 가지가 있으며 자세한 사항은 정재현·이환웅(2020, p. 23)을 참고할 수 있다.

정규화 방법을 소개하고 있다. 여기서 랜덤포레스트는 데이터의 부분집합으로부터 여러 개의 결정트리들을 만들어 트리들의 평균을 내는 앙상블 형태의 방법이다. 최근접 이웃 탐색 방법은 설명변수로 구성된 공간에서 어떠한 점을 잡았을 때 그 점에서 가장 근접한 관측치가 가지는 결과변수의 평균을 해당 점에 대한 예측치로 선정하는 방법이다. 이웃의 개수를 적게 택할수록 개별 데이터에 더 적합한(fit) 모형을 얻을 수 있으나 새로운 데이터에서의 오차는 더 심해질 수 있으므로 이웃의 개수를 적절히 조절하여 최선의 모형을 얻을 수 있다. 마지막으로 서포트 벡터 머신(Support vector machine)은 일반적으로 분류를 위한 머신러닝 방법론으로서 두 개의 클래스를 분리하는 분리초평면을 찾는 것이 기본 아이디어이며, 최적의 분리초평면은 양쪽의 클래스로부터 가장 가까운 데이터 포인트들과 분리초평면과의 거리를 최대화 시키는 방법으로 찾게 된다.<sup>26)</sup> 만일 두 개의 클래스를 완전하게 분리하는 분리초평면이 존재하지 않는 경우에는 오분류를 허용하는 방법으로 분리초평면을 구성하게 되는데, 오분류를 허용하는 정도에 따라서 과적합, 과소적합을 조절하여 최적의 모형을 얻을 수 있다.

〈표 IV-1〉 머신러닝 알고리즘과 정규화 방법

머신러닝 방법	설명	정규화 방법
선형모형	선형모형에 페널티 항을 추가	Lasso $\lambda \sum_{j=1}^p  \beta_j $ , Ridge $\lambda \sum_{j=1}^p \beta_j^2$ Elastic Net $\lambda_1 \sum_{j=1}^p  \beta_j  + \lambda_2 \sum_{j=1}^p \beta_j^2$
결정트리	조건문들을 통해 데이터를 분류	깊이, 노드 및 잎의 개수, 잎의 크기 등
랜덤포레스트	여러 개의 트리를 앙상블	트리 개수, 개별 트리의 복잡성
최근접 이웃 탐색	표본공간을 가장 근접한 이웃의 특징에 따라 분류	이웃의 개수
Support Vector Machine	초평면을 통해 서로 다른 분류에 속한 데이터를 나눔	커널의 $\gamma$ 값, C값

자료: Mullainathan and Spiess(2017), p. 93.을 참고하여 저자 작성

26) 여기서 분리초평면과 가장 가까운 데이터 포인트들을 서포팅 벡터라고 지칭한다.

머신러닝을 사용할 때는 머신러닝 기법마다 특징이 다양하고 성질이 다르다는 것을 유념할 필요가 있다. Lasso와 같은 모형은 변수를 줄여주는 특징을 가지고 있기 때문에 변수의 개수가 많은 고차원(high-dimensional) 자료에서 유용하다. 하지만 Lasso는 변수들의 복잡한 상호작용을 스스로 찾아주는 특징을 가지고 있지 않다. 만일 Lasso를 통해서 변수들의 복잡한 비선형적인 관계를 찾고자 한다면 연구자가 임의로 가능한 많은 상호작용들을 변수로 생성한 다음 Lasso를 통해서 변수를 선택하는 과정을 추가적으로 거쳐야 될 것이다. 반면 결정트리에서는 사전적으로는 찾기 어려울 수도 있는 변수들의 복잡한 상호작용들을 찾아내는 데 유용할 수 있다. 하지만 결정트리는 비모수적인 접근을 하므로 고차원 자료에서 차원 축소 목적으로는 유용하지 않을 수 있다. 이런 측면에서 하나의 절대적으로 최선인 머신러닝 방법은 존재하지 않으며(Abadie and Kasy, 2019), 연구의 목적에 따라 적절한 머신러닝 방법을 사용해야 할 것이다.

## 2. 예측문제(Prediction Problem)에 대한 선행연구

본 절에서는 머신러닝을 사용한 연구들에 대한 간략한 문헌 정리를 해 보고자 한다. 우선 첫 번째로 머신러닝이 사람의 의사결정보다 더 나은 의사결정을 내릴 수 있는지에 대해 살펴보기 위해 사람의 의사결정과 머신러닝 알고리즘의 성과를 비교하는 논문을 먼저 살펴보고자 한다. 그리고 두 번째로 머신러닝과 기존의 전통적 계량 모형의 예측성과를 비교하는 논문들을 살펴보고자 한다. 세 번째로는 머신러닝이 전통적인 계량방법론(conventional methodology)들에 비해 얼마나 처치효과의 이질성(heterogeneity effect, 변수들 간의 교차항을 포착하는 것과 연관)을 잘 파악하는지에 관한 논문을 살펴보고자 한다.

### 가. 머신러닝이 사람의 결정을 개선시킬 수 있는가?

Kleinberg et al.(2018)은 머신러닝이 체포된 피의자들의 구속영장을 심사

하는 판사들의 결정을 더 개선시킬 수 있는지에 대하여 연구하였다. 체포된 피의자를 구속할지 말지 결정하는 것은 법적으로 오직 해당 피의자가 도망을 치거나 증거를 인멸할 확률 혹은 재범의 위험성 등을 고려해서 이루어져야 하기 때문에 구속 여부를 판단하는 것은 예측의 문제라고 할 수 있다. 그렇기 때문에 구속 여부에 관한 결정은 예측에 강한 머신러닝을 사용하는 것이 적합한 분야라고 할 수 있다. Kleinberg et al.(2018)은 2008년부터 2013년까지 뉴욕주에서 이루어진 146만여 건의 법원심문 자료를 Gradient boosted decision trees(Friedman et al., 2001)를 통해서 분석하였다. 예측변수로는 심문에서 얻어진 정보들(예를 들어 현재 혐의, 과거 범죄경력 등)만을 이용하였으며 인종, 성별 등과 같은 정보는 사용하지 않았다.

분석 결과 머신러닝 알고리즘을 이용하여 구속 여부를 결정한다면 현재와 동일한 구속률을 유지하면서 도주나 재범확률을 24.7% 낮추거나 혹은 구속률을 41.9% 낮추면서도 도주나 재범확률을 늘리지 않을 수 있는 것으로 나타났다. 이러한 결과가 나타난 몇 가지 이유는 다음과 같다. 첫째, 판사들의 결정은 고위험 피의자들을 생각보다 많이 석방시키는 것으로 나타났다. 예를 들어 상위 1% 위험군의 경우 석방되었을 때 56.3%가 재판에 출석하지 않으며 62.7%는 다시 체포되는 것으로 나타났는데, 판사들은 상위 1% 위험군을 48.5%나 석방시키는 것으로 나타났다. 둘째, 구속을 더 엄격하게 하는 판사들은 고위험군만을 더 많이 구속하는 것이 아니라 모든 피의자들을 더 많이 구속하는 것으로 나타났다. 머신러닝 알고리즘이 판사들로 하여금 고위험군을 더 많이 구속하고 저위험군을 덜 구속하는 방향으로 도울 수 있다면 판사들의 의사결정을 개선시킬 수 있을 것이다.

저자들은 또 머신러닝과 기존에 많이 사용되는 로짓회귀(logistic regression) 알고리즘과의 성과를 비교해 보았다. 머신러닝 알고리즘이 상위 1%의 고위험군이라고 지목한 피의자들 중 오직 30.6%만이 로짓회귀 알고리즘을 통해 고위험군으로 지목되었다. 머신러닝이 지목한 고위험군의 도주 및 재범 확률은 56.3%로 예측되었지만 로짓회귀가 예측한 확률은 46.3%였다. 이러한 결과는 머신러닝이 기존의 방법론보다 더 나은 예측성과를 가질 수 있다는

예시를 보여준다고 할 수 있을 것이다.

저자들은 머신러닝이 구축 결정을 개선시킬 수 있다는 결과가 판사들이 의사결정을 잘못하고 있다는 뜻으로 단정 짓기는 어렵다고 언급했다. 판사들은 머신러닝이 고려하지 않은 요소들을 고려하여 판결을 내릴 수도 있다. 예를 들어 피의자가 가족의 생계를 책임져야 하는 상황에 있는 경우 판사들은 이들을 더 많이 석방시킬 수도 있을 것이다. 하지만 만일 판사들이 그 당시의 기분이나 피의자들의 외모 등에 영향을 받아서 일관성 없는 결정을 하는 경우에는 머신러닝이 판사들의 의사결정을 개선시킬 수 있을 것이다.

Mullainathan and Obermeyer(2019)는 심장마비 의심증상이 있는 환자들에 대해서 도관삽입(cardiac catheterization) 테스트 여부를 선택하는 의사들의 결정을 머신러닝이 개선시킬 수 있는지에 대해 연구하였다. 테스트 결정에 영향을 미치는 중요한 요소 중 하나는 환자가 실제로 관상동맥 폐쇄 등 질환이 있는지에 대한 예측이므로, 머신러닝의 사용이 유용할 수 있다. 다만 관상동맥 폐쇄가 의심되더라도 만일 환자가 다른 요인으로 치료의 유용성이 떨어지는 경우에는 테스트에 들어가는 시간과 비용 그리고 고통을 지불하지 않는 것이 나은 결정일 수도 있다. 그래서 저자들은 테스트의 질이 보정된 수명(Quality adjusted life years)으로 측정한 비용-효율 비율을 기준으로 삼아 머신러닝 알고리즘과 응급실 의사들의 결정을 비교 분석해 보았다.

저자들은 2010년부터 2015년까지 뉴욕주에서 이루어진 146만 여건의 응급실 방문 기록을 Lasso와 Gradient boosted decision trees(Friedman et al., 2001)를 통해서 분석하였다. 예측변수로는 과거 의료기록, 의사들의 노트, 혈액 검사 결과 등 고차원 변수들을 사용하였다. 머신러닝으로 분석한 결과, 테스트의 평균 비용-효율 비율은 8만 6,683달러로(1년의 질이 보정된 수명을 늘리기 위해서 드는 비용), 일반적으로 수용되는 수준인 것으로 나타났다. 그러나 머신러닝 알고리즘이 저위험군으로 예측하는 그룹들의 비용-효율 비율은 매우 높아서 수행된 테스트의 49%는 비용-효율 비율이 15만달러가 넘는 것으로 나타났다(over-test). 또한 의사들은 매우 고위험군의 환자들에 대해서는 충분한 수의 테스트를 하지 않는 것처럼 보이는데, 머신러닝 알고

리즘에서 예측한 가장 고위험 환자들을 테스트하지 않는 경우에는 30일 내에 치명적인 심장 문제가 발생한 빈도가 9.16%나 되었으며 이 중 3.82%는 사망에 이르렀다. 이러한 점을 볼 때 머신러닝은 저위험군의 테스트를 줄이거나 고위험군의 테스트를 늘리는 데 도움이 될 것으로 생각할 수 있다.

저자들은 의사들의 실제 결정 기준을 추정해 보기 위해 의사들의 테스트 결정과 가장 유사한 결정 결과를 가져오는 모형을 구현해 보았다. 이 모형은 머신러닝에 비해서 겉으로 보이는 증상들(예를 들어 가슴 통증)에 많은 가중치를 두었고, 보이지 않는 정보들(예를 들면 과거 병력)에 낮은 가중치를 두는 모형으로 잘 근사되는 것을 확인했다. 이러한 결과는 만일 의사들이 머신러닝보다 결정을 잘 못하는 것이 사실이라면 위와 같은 편향 때문임을 의미한다.

Chalfin et al.(2016)은 머신러닝을 방법론을 이용하여 인력채용이나 승진 상황에서 머신러닝이 사람들의 판단보다 더 나은 결정을 할 수도 있음을 보였다. 저자들은 우선 1991년에서 1998년 사이에 필라델피아 경찰청에 의해 고용된 1,949명의 경찰들에 관한 자료 분석을 통해 머신러닝이 위험한 경찰관(과도한 무력사용, 총기 사용 등)들을 고용 시점에 잘 예측할 수 있는지 분석하였다. 예측변수로는 지원자들의 사회 경제적 특징, 참전용사 여부, 결혼상태, 그리고 과거 직장에서의 해고 경력 등을 사용하였다. 분석 결과 머신러닝이 예측한 가장 위험한 10% 경찰관들을 평균 위험의 경찰관들로 대체했을 때는 총기 사용은 4.81%, 무력 남용은 약 3%, 언어폭력은 약 5% 정도 감소할 것으로 예측되었다. 하지만 지원자들의 입사순위를 기준으로 가장 위험한 하위 10%를 교체했을 때는 오히려 총기 사용이 1.92%, 무력 남용이 약 0.5%, 언어폭력은 약 2.5% 증가할 것으로 예측되었다. 이는 머신러닝이 위험한 경찰관을 예측하는 데 평가점수보다 더 유용할 수 있음을 보여준다.

또한 저자들은 ‘Measures of Effective Teaching’ 프로그램에서 사용된 데이터를 통해서 어떠한 교사들을 재임용하는 것이 효과적인지 머신러닝과 교장들의 평가를 기준으로 비교해 보았다. 분석에 사용된 데이터들은 2009년과 2010년 사이에 측정된 교사들의 사회인구학적 특성, 설문, 수업 평가, 학생

들의 시험 점수 등이 사용되었다. 분석 결과 머신러닝이 예측한 하위 10%의 선생님들을 평균 수준의 선생님이로 교체하는 것이 교장들이 평가한 하위 10%의 선생님들을 평균 수준의 선생님이로 교체하는 것보다 학생들의 수학과 언어 성적에 더 큰 향상을 가져다 줄 것으로 예측되었다. 이러한 결과는 다양한 특징을 가지고 있는 근로자들이 일하는 환경에서, 평가자들의 평가보다는 머신러닝의 평가가 근로자들의 성과를 더욱 잘 평가할 가능성을 보여준다고 할 수 있다.

이상의 논문들은 머신러닝이 사람의 판단보다 나올 수 있다는 점을 보였지만 머신러닝이 사람의 판단을 개선시킬 수 있다는 연구만 있는 것은 아니다. McKenzie and Sansone(2019)는 신규 사업자들의 3년 후 사업성과를 예측하는 데 전문평가자들의 평가와 전통적 계량 모형, 그리고 여러 머신러닝 방법론들의 예측성과를 비교해 보았다. 저자들은 나이지리아에서 시행되는 사업계획 공모전인 Youwin 프로그램의 지원자들 중 입상하여 창업자금을 지원받은 기업들과 입상하지 못한 기업들 중 일부 자료를 분석하였다. 총 2,506개의 기업들의 성과를 분석하였으며 공모전 지원 과정에서 얻어진 다양한 변수들(예를 들어 자기 효능감에 대한 리커트 척도, 레이븐 지능검사 지표 등)을 포함하여 566개의 예측변수를 이용하여 모형을 분석하였다.

저자들은 평가자들의 사업계획 평가점수, 기존 학술연구에서 사용된 단순 회귀모형, 그리고 Lasso와 SVM, Boosting 같은 머신러닝 모형으로 신규 사업자들의 3년 후 생존확률과 총고용, 그리고 월별 이윤을 예측해 보았다. 분석한 결과 평가자들의 사업계획 평가점수는 기업들의 성과를 예측하는 데 유의미하지 않은 것으로 나타났다. 학술모형에 근거한 선형회귀분석의 경우 사업가의 성별, 나이, 사업섹터가 성과에 유의미한 영향을 주는 것으로 나타났다. 그러나 모형 전체의 예측력은 평가자들의 평가에 비해 큰 개선이 없었다. 마찬가지로 Lasso와 SVM, Boosting과 같은 머신러닝 알고리즘의 경우에도 평가자들의 점수나 선형회귀모형에 비해 예측성과의 개선이 없는 것으로 나타났다.

저자들은 이처럼 머신러닝이 평가자들의 판단이나 단순회귀모형에 비해

나은 성과를 내지 못한 것에 대해 여러 각도로 이유를 찾으려고 시도해 보았으나 설명력 있는 원인을 제시하지는 못하였다. 이에 저자들은 기업들의 성과를 예측하는 것은 본질적으로 매우 어려운 문제일 수 있으며 해당 데이터에서 얻을 수 있는 최대한의 예측성과를 머신러닝 방법론과 기존의 방법론이 모두 달성하고 있기 때문에 두 방법론 사이의 성과차이가 없을 수 있다고 평가하였다.

### 나. 머신러닝이 기존의 방법론보다 우월한가?

머신러닝 기법들과 기존의 계량 방법론의 예측성과를 비교한 연구들은 최근 그 수가 많아지고 있다. 본고에서는 저명한 저널에 출판이 되었거나 출판 예정인 몇몇 연구들을 소개하고자 한다. 머신러닝과 기존의 방법론을 비교한 논문들 중에는 머신러닝이 기존연구에 비해 우수한 성과를 내었다는 연구도 있으나 머신러닝이 기존의 방법론에 비해 큰 개선이 없다는 연구도 존재한다.

Sansone(2019)는 미국 고등학생들의 학업중단을 예측하는 데에 머신러닝 방법이 기존의 계량 모형보다 더욱 효과적일 수 있다는 것을 보였다. 그는 2009년 고등학교 종단연구 자료를 이용해 2만 1,440명의 학생들에 대한 다양한 변수들(인종, 성별, 성적, 가족 특성, 출석률, 징계 경험 여부, 과학 수업 수강 여부, 수강 학점 수)을 사용하여 학생들의 학업중단을 예측하는 여러 모형의 성과를 비교해 보았다. 먼저 전통적으로 자주 사용하는 계량 방법론인 로짓, OLS, 프로빗 등을 사용하여 GPA와 수강학점 만을 예측변수로 사용하였을 때는 학업중단에 대한 정확도(Accuracy)는 90% 정도였으며 재현율(Recall)은 15%를 크게 초과하지 않았다. 반면에 자주 사용되는 머신러닝 기법인 SVM(support vector machine), Boosting, Lasso, Logit-Lasso 등을 사용하고 9학년에 얻어진 다양한 변수들을 함께 사용했을 때는 정확도를 90% 수준에서 유지하면서 재현율을 23% 수준까지 향상시킬 수 있는 것으로 나타났다. 저자는 미국의 2009년도 9학년 학생들 중 48만 3,270명이 학업중지를 하였으므로, 학업중지 예측에 대해 재현율이 8%p 향상된 것은 약 4,830

명의 추가적인 학업중지 위험군을 포착할 수 있는 개선이라고 주장하였다.

Sansone(2019)는 이 결과를 바탕으로 머신러닝 기법이 전통적인 학업중단 조기경보 시스템보다 학업중단 위험에 처해 있는 학생들을 찾아내는 데 더 큰 도움이 될 수 있다고 주장했다. 전통적인 학업중단 조기경보 시스템은 낮은 성적이라든가 적은 이수학점 등을 지표로 이용해서 학업중단 위험에 처해있는 학생들을 파악하지만, 머신러닝은 획득 가능한 모든 변수들(예를 들어 잦은 전학 횟수, 친구들보다 빠른 출생연도 등)을 고려하여 학업중단 위험에 처해 있는 학생들을 신속하게 파악할 수 있는 장점이 있다는 것이다.

Gentzkow et al.(2019)은 1873년부터 2016년까지 미국의 국회 연설문 분석을 통해 공화당과 민주당의 당파 차이의 변화를 분석하였다. 분석에 사용된 연설문은 두 개의 단어로 구성된 구문을 기준으로 데이터화하였으며, 중복되지 않는 50만 8,352개의 구문들과 7,732명의 연설자들을 포함하여 총 2억 8천만 구문이 데이터셋에 포함되었다. 만일 공화당 의원들이 사용하는 구문들과 민주당 의원들이 사용하는 구문들이 현저하게 차이가 난다면 이는 당파의 차이가 더욱 심해진 것이라고 볼 수 있을 것이다. 저자는 전통적인 방법론인 최우추정법(Maximum likelihood estimation)과 머신러닝 방법론인 leave-out 추정치와 Lasso 추정치를 사용하여 당파 차이의 변화를 추정하였다.

저자들은 최우추정법이 해당 데이터에는 좋은 추정치를 내지 못할 것으로 예상하였다. 이러한 예상에 대한 근거는 최우추정치는 일치성을 가지고 있기 때문에 관측치의 숫자가 많은 경우에 좋은 추정량을 가져다 줄 수 있지만, 해당 연구에서 사용된 데이터의 경우 비중복 구문의 숫자가 매우 많아서 각 구문이 관측되는 횟수는 많지 않기 때문이다. 이처럼 각 구문의 관측치가 많지 않은 경우에는 유한표본의 한계로 인해서 최우추정법의 성과가 좋지 않을 수 있다. 반면 머신러닝 방법론은 유한 표본에서 좋은 성과를 내는 경향을 가지고 있기 때문에 해당 데이터에서도 좋은 성과를 낼 수 있을 것으로 기대하였다.

분석 결과 저자들의 예측대로 최우추정법은 당파 차이 정도를 잘 추정하지 못하는 것으로 나타났다. 정확한 예측치는 당파의 구별 없이 전체 사용

된 구문에서 임의로 추출된 데이터를 가지고 당파 차이를 추정했을 때에는 당파의 차이가 없는 것으로 추정해야 한다. 하지만 최우추정치는 당파 구별이 없는 임의 데이터에서도 당파를 구별하여 추정한 추정치와 거의 같은 결과를 나타냈다. 이는 전체 사용한 구문의 수가 많아질수록 의원들이 서로 다른 단어를 사용할 가능성이 크므로 최우추정치가 당파 차이가 커진 것으로 착각을 하는 문제가 발생하기 때문이다. 반면 leave-out 추정치와 Lasso 추정치는 당파의 차이가 1990년대까지는 상대적으로 안정적이다가 1990년 이후로 급격하게 늘어난 것으로 추정하였다. 이 둘의 추정치는 당파의 구별 없이 전체 구문에서 임의로 추출된 데이터를 사용하는 경우 당파의 차이가 없다고 올바르게 추정하였다. 이러한 결과를 바탕으로 볼 때 머신러닝은 유향표본의 문제가 심해질 수 있는 상황에서 전통적인 방법론보다 더 적절한 대안이 될 수 있다는 것을 보여준다.

Fuster et al.(2020)은 머신러닝 방법론이 모기지 시장에서 차입자들의 부도율을 예측하는 데 어떠한 영향을 주게 될지 분석하였다. 분석을 위해 2009년부터 2013년까지 발생한 937만개의 모기지 대출 자료를 사용하였으며 부도 예측에 사용된 변수들은 신용점수, 대출금액, 직업, 보증 여부, 대출 목적, 대출 기간 등 14개의 변수를 사용하였다. 저자들은 세 가지 모형을 설정하였다. 첫 번째는 일반적으로 사용되는 로짓모형을 사용하였다. 두 번째는 로짓모형에 비선형성을 포착할 수 있는 수많은 더미변수들(예를 들면 신용 점수에 20개 구간을 나타내는 더미변수)을 포함시킨 비선형 로짓모형이었다. 세 번째는 머신러닝 방법론에 자주 사용되는 랜덤 포레스트 모형이었다.

분석 결과 랜덤 포레스트 모형은 두 가지의 로짓모형보다 더 좋은 성과를 나타냈다. 랜덤 포레스트는 평균 정확도(precision)에서 비선형 로짓보다 5.1% 개선된 성과를 보여주었으며,  $R^2$ 에서도 14.3%의 향상을 보여주었다. 특히 표본에서 추출한 100개의 부스트랩 샘플에서 예측성과를 비교한 결과, 랜덤 포레스트가 로짓모형들보다 매번 더 좋은 성과를 내는 것을 확인할 수 있었다. 이는 비선형 모형이나 변수들 간의 복잡한 상호관계를 잘 포착해 낼 수 있는 머신러닝 방법론이 연구자가 임의로 더미변수들을 설정하여 모형에 추

가한 비선형 로짓모형보다 나은 성과를 가져다주는 예시라고 볼 수 있을 것이다.

이상의 연구들을 살펴보면 머신러닝은 ① 연구자의 눈에 잘 띄지 않는 변수들이 예측 결과의 개선을 가져올 수 있는 경우 ② 관측치의 개수가 부족하여 유한 샘플의 문제가 심각한 경우 ③ 연구자들이 쉽게 파악하기 어려운 비선형성이나 변수들 간의 복잡한 상호작용이 있는 경우 기존의 계량 모형에 비해 개선된 예측성고를 낼 수 있다고 생각해 볼 수 있다. 하지만 머신러닝이 모든 경우에 예측성고를 향상시킬 수 있는 것은 아니다. 이하에서 살펴볼 논문들은 머신러닝이 기존의 방법론에 비해 예측성고의 큰 개선을 가져오지 못했다고 주장한다.

Beattie et al.(2018)는 대학교 신입생들 중 성취도가 예상보다 심하게 떨어지는 그룹(Divers)과 예상보다 크게 올라가는 그룹(Thrivers)을 예측하는 연구를 진행하였다. 기존의 연구들은 과거의 성적이 대학교 성적을 예측하는데 가장 중요한 예측변수인 것을 보였지만, 과거의 성적이 유사한 학생들 사이에도 대학교의 학업성취도에는 굉장히 큰 편차가 있었다. 저자들은 학생들의 비학업적 특징(non-academic characteristic)이 이와 같이 과거 성적으로 설명되지 않는 성취의 편차를 설명할 수 있는지 연구하였다.

분석을 위해 저자들은 2015년 가을학기에 토론토대학교(University of Toronto)에서 경제학입문을 수강한 6천명가량의 학생들의 설문 응답과 입학 시점의 행정자료를 사용하였다. 설문자료에는 일을 미루는 정도에 대한 자가 점검, 노력을 위해 인내하는 정도, 관심을 일관성 있게 유지하는 정도 등이 포함되어 있었으며 위험에 대한 선호, 시간선호 등 또한 포함되어 있었다. 입학 행정자료에는 성별, 국적, 학교 등록 형태, 성적, 고등학교 수강 과목 및 성적 등이 포함되어 있었다.

분석에서는 예상 밖의 위험그룹(Divers)과 성취그룹(Thrivers)을 구별하기 위해 과거 성적이 예측하는 성적과 실제 성적 간의 차이를 통해 잔차를 구축했으며 이 잔차의 상위 10%p를 성취그룹, 하위 10%p를 위험그룹으로 정의하였다. 그리고 비학업적 특징들이 이들 그룹을 예측하는 데 설명력이 있

는지 분석하였다. 분석 결과, 시험을 위해 베타치기를 하는 성향, 공부하는 시간, 돈을 벌기 위해 일하는 시간, 자신이 기대하는 성적, 설문조사를 기입한 시기(설문조사 초기 시점에 입력하는 사람 vs 설문조사 종료 시점에 입력하는 사람), 성실성(conscientiousness), 참을성에 대한 지표가 위험그룹과 성취그룹을 예측하는 데 설명력이 좋을 것으로 확인했다. 저자들은 LARS (Least-angle regression) 알고리즘을 통해 모든 변수를 집어 넣고 알고리즘이 선택한 모형과 이상의 7개 변수의 평균만을 가지고 단순 예측한 모형의 성과를 비교했는데, LARS 알고리즘이 선택한 모형의 성과는 7개 변수의 평균을 가지고 예측한 모형에 비해 큰 예측성과의 향상을 가져오지 못했다.

Orlov et al.(2021)은 경제학 과목에서 수강철회를 하거나 낮은 성적을 받은 것으로 예상되는 학생을 수학 성적으로 예측할 수 있는지에 대한 연구를 진행하였다. 이를 위해서 2019년 코넬대학교에서 미시경제학입문을 수강한 학생들 711명과 중급미시경제학을 수강한 207명의 학생들에게 수업 초기에 수학 시험을 보게 하였으며, 이 수학 성적을 이용하여 하위 25%(B- 이하 성적) 그룹을 예측할 수 있는지 분석하였다. 미시경제학입문에서는 수학 성적을 단일변수로 사용했을 때 수학 점수 72점 미만 학생을 하위그룹으로 예측하는 경우 True positive(실제 하위그룹이면서 예측도 하위그룹으로 한 경우)가 54%, False positive(실제 하위그룹이 아닌데 예측을 하위그룹으로 한 경우)가 20%였다. 중급미시경제학에서는 수학 성적이 66점 미만 학생을 하위그룹으로 예측하는 경우 True positive가 52%, False positive가 12%였다.

저자들은 수학 성적 이외에 다른 정보들이 예측력을 높일 수 있는지 확인하기 위해 성별, 학년, 수학시험에서 각 문항당 정답 여부 등을 변수로 추가한 후 lasso를 사용하여 예측력을 비교해 보았다. 미시경제학입문의 경우에는 55% 수준의 True positive를 유지할 때 False positive는 14% 정도로, 수학 성적 단일변수 모형보다는 약간의 개선이 있었다. 그러나 중급미시경제학 과목에서는 lasso 모형이 예측력에 아무런 개선을 가져다주지 못하였다.

Bjerre-Nielsen et al.(2021)은 머신러닝과 전통적인 계량 방법론을 직접적으로 비교한 것은 아니지만, 사생활을 침해할 수 있는 빅데이터의 사용보다는

설명력이 있는 잘 알려진 변수들을 몇 가지만 사용하는 것이 더 높은 예측력을 가져다줄 수 있음을 보였다. 저자들은 대학교 학생들의 성적을 예측하는 데 입학행정정보(성별, 나이, 과거 성적 등)와 빅데이터(출석횟수, 강의 중 스마트폰 사용시간, 주류소비량, 성격 유형 등)의 예측력을 비교하였다. 분석 결과, 빅데이터의 설명력보다는 과거의 성적이 포함된 입학행정정보의 성적 예측력이 더 좋았다.

이상의 연구와 같이 머신러닝이 기존의 방법론의 예측력을 개선시키지 못하는 경우는 다음과 같은 특징이 있다고 할 수 있을 것이다. ① 설명변수의 수가 많기는 하지만 설명변수의 구조가 원천적으로 고차원적인 경우가 아니라 단지 유사한 의미의 정보를 여러 번 측정하는 것에 불과한 경우(예를 들어 Orlov et al.(2021)에서와 같이 수학시험의 각 문항당 정답 여부를 각각의 변수로 사용하는 경우) ② 그리고 연구자들이 손으로 직접 고르거나 기존 연구에서 잘 알려진 설명변수들이 이미 좋은 예측력을 가진 경우에는 머신러닝의 개선 효과는 기대하기 어려울 수 있다.

#### 다. 머신러닝이 실질적인 처치효과를 잘 찾아낼 수 있는가?

이상의 논의는 예측문제(prediction problem)에서 머신러닝 방법론이 사람들의 결정보다 나은 의사결정을 할 수 있는지 혹은 기존의 전통적인 계량 방법론에 비해서 우수한 예측력을 보일 수 있는지에 대해서 논하였으며 특정한 환경에서는 머신러닝의 예측성도가 기존의 방법론에 비해 더욱 우수할 수 있음을 언급하였다.

하지만 정책 결정자의 입장에서는 특성변수들을 통해 결과변수를 잘 예측하는 모형만을 가지고는 효과적인 의사결정을 하기 어려울 수 있다. 예를 들어 한정된 자원을 가지고 정책금융 지원 정책을 집행해야 하는 정책 결정자가 있다고 하자. 그리고 기업의 특성변수들로 결과변수인 기업의 부도율을 예측하는 모형을 만들어 부도가 날 확률이 높은 기업의 특성들을 발견했다고 하자. 그렇다면 정책 결정자는 부도가 날 확률이 높은 특성을 가진 기업들에 정책금융을 지원해야 하는가? 이에 답변하기 위해서는 해당 특성을

가진 기업들이 정책금융을 받았을 때 정책금융을 받지 않았을 때보다 확실히 부도율이 낮아진다는 증거를 확보해야 한다. 즉 정책금융의 대상자를 선정하는 의사결정과 같은 문제에서는 어느 기업이 부도가 날 가능성이 높을까 하는 예측의 문제에 대한 답이 아니라, 정책금융이라는 처치효과에 대한 인과추론의 문제에 대한 답이 있어야 효과적인 의사결정을 할 수 있다는 것이다.<sup>27)</sup>

인과추론은 하나의 관측 대상이 처치를 받았을 때의 실제적 결과와 처치를 받지 않았을 때의 잠재적 결과를 비교해서 처치효과를 찾아내는 문제이므로 본질적으로 반사실적(counterfactual) 상황에 대해 가정해야 하는 어려운 문제이다.<sup>28)</sup> 이러한 인과추론을 위해서 많이 사용되는 방법 중 하나는 무작위 통제 실험(Randomized Controlled Trials)으로, 무작위로 처치를 받은 실험군과 처치를 받지 않는 대조군을 선택하여 그 두 그룹의 차이를 비교하는 것이다. 이러한 무작위 통제 실험이 이루어지는 경우 개별 관측 대상 수준의 처치효과를 발견하기는 어렵더라도 평균적인 수준에서의 처치효과를 찾아낼 수는 있는데, 이를 평균처치효과(Average treatment effect)라고 한다.

평균처치효과를 아는 경우에 정책 결정자는 정책이 유효한 효과가 있을지에 대한 정보를 확보한 것과 같다. 하지만 평균처치효과는 어떤 대상을 정책의 대상으로 선정해야 하는지에 대한 답을 제시하지는 못한다. 정책의 대상자를 효율적으로 선정하기 위해서는 해당 정책으로 인해 가장 많은 효과를 누릴 수 있는 집단의 특성을 파악하는 것이 중요한데, 정책의 처치효과가 집단별로 다르게 나타나는 것을 가리켜 정책의 이질적 효과(Heterogenous effects)라고 한다.

기존의 선형모형에서는 정책의 이질적인 효과를 분석하기 위해 데이터를 집단별로 나누어서 분석하거나 집단의 특성변수의 교차항을 추가하여 분석

---

27) 반면 대조적으로 제IV장 제2절 '가. 머신러닝이 사립의 결정을 개선시킬 수 있는가?'에서 언급된 구속 여부를 결정하는 상황에서는 구속의 처치효과에 대해서는 크게 고려할 필요가 없는데, 이는 구속의 처치효과는 사전적으로 명확하기 때문이다. 이 경우에는 체포된 피의자들의 도주나 재범의 확률을 예측하는 것이 더 중요하기 때문이다.

28) 이러한 인과적 추론에 있어서의 근본 문제(Fundamental Problem for Causal Inference)에 대해서는 정재현·이환웅(2020)의 '제4장 가'를 참고할 것을 추천한다.

하는 방법을 사용할 수 있다. 하지만 이러한 접근 방법에는 연구자들이 임의적으로 집단을 선정해야 하는 문제가 존재하는데, 이는 이질적 처치효과를 발견하기 어렵게 만드는 요인이 된다. 연구자들은 사전적으로 이질적 처치 효과가 어떻게 발생할 것인가에 대해서 알지 못하기 때문에 여러 가지 선형모형을 추정해 보면서 이질적인 처치효과를 발견하려고 시도할 유인이 존재한다. 하지만 여러 가지 모형을 시도하는 경우 다중검정(multiple testing)의 문제에 빠지기 쉽다. 이는 여러 번 모형을 추정하는 경우에 유의한 결과가 하나라도 나올 확률이 자연스럽게 올라가게 되지만 연구자가 이를 감안하지 않아 실제로는 유의하지 않은 결과를 유의하다고 잘못 해석하게 되는 상황을 의미한다.<sup>29)</sup>

머신러닝 방법은 본질적으로 데이터가 스스로 모형을 선택하도록 짜인 알고리즘이기 때문에 연구자의 임의적인 판단을 최소화한 채 이질적 처치효과를 찾아내는 데 유용하게 사용될 수 있다. 처치효과를 찾아내는 것은 예측의 문제와는 다른 목적을 가지고 있기 때문에 앞에서 언급한 머신러닝 방법론들과는 다른 형태의 접근이 필요한데, 가장 대표적인 방법론은 Wager and Athey(2018)가 제안한 커질 포레스트(Causal forest)가 있다. 간단히 소개하자면 커질 포레스트는 처치효과가 가장 극대화되는 집단을 찾아내는 방식으로 모형을 만들어 가는 방법으로, 결과의 예측 오차를 최소화하는 방식으로 모형을 만들어 가는 랜덤 포레스트와는 다른 특징을 가진다.<sup>30)</sup> 본 절에서는 커질 포레스트를 중심으로 머신러닝을 통한 이질적 처치효과를 분석한 선행연구들을 소개하고자 한다.

Knittel and Stolper(2021)는 Home Energy Report(HER) 프로그램이 전력 소비에 미친 이질적 처치효과를 커질 포레스트를 통해 분석해 보았다. HER 프로그램에 참여한 70만가구의 전력 소비량을 분석하였으며 가구의 특징이 처치효과에 어떠한 영향을 미쳤는지 분석하였다. 전통적인 이중차분법 방법론을 통해 구한 평균적인 처치효과는 9.4kWh의 전력소비 감축이 있는 것으

29) 이렇게 얻어진 결과를 가리켜 허구적 발견(Spurious finding)이라고 하기도 한다.

30) 커질 포레스트에 대한 자세한 이론적 소개는 정재현 · 이환웅(2020) '제4장 다'를 참고

로 확인되었으나, 이질적인 처치효과는 40kWh 전력소비 감축에서 10kWh 전력 소비 증가까지 다양하게 분포하는 것으로 확인되었다. 이처럼 처치효과가 이질적이라는 것은 해당 프로그램을 시행할 때 대상을 잘 선택하는 경우 더욱더 프로그램의 효과를 극대화할 수 있다는 것을 의미한다.

구체적으로 처치효과에 많은 영향을 주는 가구의 특징은 프로그램 참가 전 전력의 기초사용량과 주택의 시장가치인 것으로 나타났다. 다만 이러한 특징들이 처치효과에 미치는 영향은 매우 비선형적으로 나타나서 만일 일반적인 회귀모형에서 이질적인 처치효과를 분석하였다면 발견하기 어려울 수도 있는 패턴인 것으로 나타났다. 예를 들어 기초사용량이 처치효과에 미치는 영향은 기초사용량이 높을수록 HER 프로그램의 처치효과가 크게 나타나다가, 기초사용량이 2천kWh가 넘어가는 경우 기초사용량이 높을수록 처치효과가 낮아지는 것을 확인하였다. 주택의 가치 또한 처치효과에 많은 영향을 주었는데 주택의 가치가 낮은 가구의 경우 매우 큰 폭의 전력 감축을 이룬 경우가 많았으나, 주택의 가치가 10만달러 이상이 되는 경우에는 전력 소비 감축의 효과가 매우 제한적이었으며 23kWh 이상을 감축한 경우는 전혀 발견되지 않았다. 저자들은 이러한 커질 포레스트를 통해 발견한 이질적인 처치효과를 바탕으로 정책의 대상을 선정하고 정책을 재단한다면 더욱 효과적인 정책 집행이 가능할 것으로 주장하였다.

Davis and Heller(2017)는 시카고에서 수행된 여름 청소년 고용 프로그램(Summer Youth Employment Programs, SYEPs)의 이질적인 처치효과를 커질 포레스트를 통해 분석하였다. 우선 전통적인 방법론을 통해 처치효과를 분석하였을 때 폭력범죄율은 크게 감소하는 경향이 있었으나 고용이나 학업에 대한 성과지표는 개선되지 않는 것으로 나타났다. 이뿐만 아니라 전체적인 체포 횟수는 줄어들지 않았으며 재산 관련 범죄는 오히려 증가한 것으로 확인되었다. 이는 기존의 이론들이 고용 증대 프로그램은 고용 상황을 개선시키고 범죄의 기회비용을 증가시켜 범죄를 줄이는 효과가 있을 것이라고 예상한 것과는 부합하지 않는 결과였으며, 이러한 결과에 대한 메커니즘을 좀 더 깊게 이해할 필요성이 존재하였다.

저자들이 커절 포레스트를 통해 이질적인 처치효과를 연구해 본 결과 기존의 연구 방법론에서는 얻을 수 없었던 두 가지 추가적인 결론을 얻을 수 있었다. 첫째는 청소년 고용 프로그램에 대해서 큰 이질적인 처치효과가 있었다는 것이다. 평균적으로는 SYEP는 고용 상황을 개선하는 효과가 거의 없었지만 상대적으로 젊고 학교에 더 잘 나오며 히스패닉, 여성 집단의 경우에는 해당 프로그램을 통한 고용 상황이 크게 개선된 것으로 확인되었다. 저자들은 이러한 결과를 기초로 커절 포레스트에서 고용의 효과가 극대화되는 집단을 정책대상으로 타게팅하는 경우 고용 효과는 네 배나 증가할 수 있음을 보였다.

두 번째는 청소년 고용 프로그램에 대한 기존의 이론을 뒷받침할 만한 근거가 약하며 오히려 고용 프로그램의 효과에 대한 새로운 메커니즘의 연구가 필요하다는 것이다. 저자들이 이질적인 처치효과를 분석하였을 때 폭력 범죄의 감소는 특정 집단에 국한된 현상이 아니었다는 것을 확인하였다. 반면 고용 상황이 훨씬 개선된 집단의 경우에는 재산 관련 범죄가 오히려 증가하는 현상을 발견하였다. 이러한 사실은 고용 프로그램이 고용 상황이 좋지 않던 사람들의 고용 가능성을 높여 범죄의 기회비용을 증가시키므로 범죄가 줄어든다는 기존의 기회비용 이론을 뒷받침하지 않는 결과였다. 저자들은 이러한 결과를 바탕으로 기존의 이론과는 대안적인 메커니즘을 주장하였다. 고용 프로그램은 일시적으로라도 사회생활을 해 보는 기회를 제공하므로 고용주의 비판을 수용하고 받아들이는 과정을 통해서 사회적 갈등 상황을 잘 해결하는 능력을 기르는 데 도움을 줄 수 있다는 것이다. 반면 고용 상황이 많이 개선된 집단의 경우에는 더욱 부유한 일터에 근무하거나 부유한 주변 사람들을 많이 만나게 되어 재산 관련 범죄를 저지를 기회가 많아질 수 있다는 것이다. 이처럼 커절 포레스트를 통해 이질적인 처치효과를 살펴보는 것은 정책효과의 메커니즘을 연구하는 데에도 도움이 될 수 있다.

Knaus et al.(2020)은 스위스의 취업 탐색 프로그램(Job Search Program, JSP)의 이질적인 처치효과를 Lasso를 응용한 방법론을 통해 분석하였다. 분석에 사용된 데이터는 2003년도 스위스의 지방 고용청에서 얻은 자료로서,

개별 실업자의 정보를 포함하고 있으며 실업자 중 취업탐색프로그램에 참여한 사람들과 참여하지 않은 사람들의 취업일수를 분석하여 취업탐색프로그램의 효과를 분석하였다. 이질적인 처치효과를 분석하기 위해 저자들은 실업자들의 특성(취업이력, 사회인구적 특성 등)과 담당 사회복지사의 평가 등과 같은 설명변수의 4차까지의 다항식과 변수들의 이차교차항, 그리고 로그 등을 포함하여 총 1,268개의 변수들을 구성하였으며, Lasso 알고리즘을 통해서 이 변수들 중에서 설명력이 떨어지는 변수들을 탈락시키는 방법으로 모형을 설정하였다. 모형을 설정한 후 개별 관측치에 대한 처치효과를 계산하였으며 이러한 과정을 30번 반복하여 평균적인 개별 처치효과를 계산하였고 평균 개별 처치효과에 대한 분산의 계산은 부스트랩 방법을 이용하여 수행하였다.

분석 결과 저자들은 취업탐색프로그램에 참여한 참가자들의 취업일수는 참여일로부터 6개월 이내의 경우 참가하지 않는 실업자들보다 줄어드는 부정적인 효과가 있는 것을 발견하였다. 이는 실업자들이 취업탐색 프로그램에 참여하는 경우 초기에 적극적인 구직활동을 줄이는 경향이 있기 때문이다. 이러한 단기의 부정적 효과는 참가자들의 특성에 따라서 이질적으로 나타났다는데, 교육을 적게 받고 자격증이 별로 없는 사람들은 이러한 부정적인 효과가 작게 나타났다. 또한 외국인들의 경우 부정적인 효과가 작은 것으로 나타났으며 미숙련 노동자일수록 다른 노동자들에 비해 긍정적인 프로그램 참여효과가 있는 것으로 나타났다. 만일 머신러닝 알고리즘이 제시하는 최선의 프로그램 참여자를 선택한다면 단기에서 나타나는 부정적 효과는 60% 감소하는 것으로 예상 되었다. 한편 장기적으로는 취업탐색프로그램 참여여부는 취업일수에 큰 영향을 주지 않는 것으로 나타났다.

---

## V. 머신러닝을 활용한 정책금융의 효과 추정

---

제Ⅲ장에서는 준실험적 방법 중에 하나인 성향점수매칭을 활용하여 신용보증기금 수혜 집단의 평균처치효과(Average Treatment Effect on the Treated, ATT)를 추정하였다. 본 장에서는 머신러닝 알고리즘 중에 하나인 커질 포레스트를 활용하여 ATT, ATE 등 처치효과를 재추정하며 처치효과가 특성변수에 따라 이질적인지 살펴보았다. 성향점수매칭은 비교락성(Unconfoundness)과 중첩성(Overlap) 가정하에 처치효과를 편이 없이 추정할 수 있는 방법론 중 하나이다. 하지만 최근 성향점수매칭을 활용한 효과 추정의 한계점에 대하여 학계에서 논의가 이루어지고 있는 만큼 커질 포레스트를 통한 처치효과 추정은 기계학습 알고리즘의 활용 가능성을 탐색하는 데 유용한 지점인 것으로 판단된다.

한편 기존의 선형모형을 통한 이질성 분석은 처치효과와 특성변수 간의 교차항을 추가하여 해당 계수가 통계적으로 유의한지 살펴보거나 특정 변수에 따라 그룹별로 표본을 재구성한 후에 추정된 처치효과가 그룹별로 이질적인지 살펴보았다. 하지만 처치효과의 영향을 끼칠 수 있는 설명변수의 수가 많을 시 전통적인 방법론은 연구자가 몇 개의 변수를 취사선택해서 살펴보기 때문에 연구자의 자의적인 판단이 요구된다. 또한 설명변수와 처치효과간의 교차항을 추가한 이질성 분석은 선형적인 이질성만을 파악할 수 있다는 한계점이 존재한다. 즉 기존의 선형회귀모형은 이질성 분석에서 한계가 존재하는데 이에 따라 최근에 많은 연구자들은 머신러닝의 이질적 효과의 탐색 가능성에도 주목하고 있다. 따라서 본 연구에서는 커질 포레스트를 통한 ATE 효과 추정과 이질적 처치효과를 탐색함으로써 재정정책의 효과성을 높이는 정책 설계의 가능성에 대해서 추가적인 탐색을 진행하였다.<sup>31)</sup>

---

31) 머신러닝 알고리즘이 조세재정정책 효과성 개선에 어떻게 활용될지에 대한 일반적인 논

## 1. 정책효과 추정을 위한 식별가정

본 연구에서 비교 대상인 성향점수매칭과 컷절 포레스트 방법론 모두 비교락성 가정하에 편이 없는 추정이 가능하다. 따라서 두 방법론의 비교에 앞서 먼저 해당 방법론들의 식별을 위한 가정들과 추정하고자 하는 파라미터에 대해서 논의하였다. 한편 식별가정 및 추정하고자 하는 파라미터들은 루빈의 인과모형(Rubin Causal Mode, RCM)을 활용하여 표현하는 것이 효과적이기 때문에 본 연구에서는 식별가정 및 추정이 필요한 파라미터에 대한 논의에 앞서 루빈의 인과모형의 기본적인 세팅을 먼저 소개한다.

루빈의 인과모형은 1974년 루빈에 의해 소개되었는데 먼저 표본은  $N$ 개의 관측치로 구성되고 각각의 관측치  $i$ 에 대해 가상성과(Potential Outcomes)는  $Y_i(0)$ 와  $Y_i(1)$ 로 표현한다. 여기서  $Y_i(0)$ 는 처치를 받지 못한 관측치  $i$ 의 결과값인데 본 연구에서는 신용보증을 지원받지 못한 경우의 매출액이며,  $Y_i(1)$ 은 신용보증을 지원받았을 경우의 매출액이다. 한편 처치 여부를 나타내는 확률변수는  $W_i$ 로, 관측치  $i$ 가 처치(신용보증기금 지원)를 받으면 1, 받지 못한 경우 0의 값을 가진다. 그리고  $N_c$ 는 처치를 받지 못한 기업체의 수(통제그룹의 수),  $N_t$ 는 처치를 받은 기업체의 수이다. 따라서 실현된 성과(realized outcome)는 다음과 같이 표기할 수 있다.

$$Y_i^{obs} = \begin{cases} Y_i(0) & \text{if } W_i = 0, \\ Y_i(1) & \text{if } W_i = 1 \end{cases} \quad \text{식 (4)}$$

한편 처치 여부를 제외한 결과값에 영향을 끼치는 관측치들의 예측변수<sup>32)</sup>(Predictors)는  $X_i$ 로 표기하며 예측변수의 특정값에 따라(예를 들어  $X_i = x$ ) 관측치들을 그룹화하는 것이 가능하다.

제Ⅲ장에서는 신용보증지원에 따른 수혜집단의 처치효과를 추정하였는데,

---

의는 정재현·이환웅(2020)의 pp. 64~69를 참조하라.

32) 머신러닝 연구들에서는 예측변수 혹은 특성변수로 지칭하지만 계량경제학에서는 설명변수로 지칭하기 때문에 본 연구에서는 예측변수와 설명변수를 혼용해서 사용하였다.

처치효과들을 루빈의 수식에 따라 표기하면 <표 V-1>과 같다. 여기서 평균 처치효과는 관측치가 표본에서 처치를 받았는지 여부와 상관없이 관측치의 처치 여부에 따른 가상성과의 차이를 평균한 값이다. 수혜집단의 처치효과(ATT)의 경우 처치를 받은 관측치들의 가상성과의 차이를 평균한 값이며, 마지막으로 조건부 평균처치효과(Conditional Average Treatment Effect, CATE)는 예측변수 값이  $X_i = x$ 인 관측치들의 가상성과의 차이의 평균값으로 설명 변수 값이  $x$ 인 집단의 처치효과를 의미한다.<sup>33)</sup>

<표 V-1> 다양한 처치효과의 수학적 정의

구분	표기
평균처치효과(ATE)	$\tau = E[Y_i(1) - Y_i(0)]$
수혜집단의 처치효과(ATT)	$\tau = E[Y_i(1) - Y_i(0)   W_i = 1]$
조건부 평균처치효과(CATE)	$\tau(x) = E[Y_i(1) - Y_i(0)   X_i = x]$

자료: Rubin(1974); Imbens(2014), pp. 2~6을 참조하여 저자 작성

다음은 관측 데이터를 활용하는 준실험적 방법의 편의 없는 추정을 위한 식별가정을 RCM에 따라 표기하였는데 <표 V-2>에서 제시하였다. 먼저 비교락성은 일련의 설명변수들( $X_i$ )이 통제된 상태에서 처치그룹 여부는 해당 관측치의 가상 성과와는 독립적임을 의미한다. 예를 들어 신용보증을 지원 받는 기업들이 지원받지 않는 기업들과 비교하여 매출액이 성장하는 기업이라 가정한다면 지원을 받지 않는 경우에도 매출액은 상승하였을 가능성이 높기 때문에 신용보증을 지원받은 기업의 매출액  $Y_i(1)$ 과 신용보증을 지원받지 못한 기업의 매출액  $Y_i(0)$ 은 다른 값을 가질 가능성이 높다. 이 경우 관측데이터에서 신용보증기금을 수혜한 사업체와 지원받지 못한 사업체의 비교에 의한 처치효과의 추정은 정책효과를 과대추정하게 된다. 반면에 비교락성 가정은 사업체들의 과거 매출액 정보를 통제된 상태에서는 지원받지 않았을 때의 결과값이 처치그룹 여부와 관계없이 독립적임을 의미한다.

33) 이후로는 평균처치효과는 ATE, 수혜집단의 처치효과는 ATT, 조건부 평균처치효과는 CATE로 표기하였다.

그다음 중첩성은 설명변수의 값이  $x$ 인 관측치들이 처치를 받을 확률이 0과 1 사이에 위치해야 함을 의미한다. 만약 특정 값을 가지는 관측치( $X_i = x$ )들이 처치그룹에 편중되어 있는 경우 처치효과를 계산할 때 해당 그룹(A그룹)에는 통제관측치가 없음에도 불구하고 모형에 대한 가정에 의해 통제관측치의 값이 추정되고 A그룹에서의 처치효과는 처치집단과 계산된 통제집단의 비교를 통해 추정된다. 이럴 경우 모형에 대한 가정에 따라(예를 들어, 종속변수와 설명변수 간의 선형관계를 가정) 추정치가 민감하게 변하는 문제점이 발생할 수 있는데, 중첩성 가정은 이러한 문제점을 완화하는 장치로 볼 수 있다.

〈표 V-2〉 비교락성과 중첩성의 수학적 정의

구분	표기
비교락성 <sup>1)</sup>	$Y_i(1), Y_i(0) \perp W_i   X_i$
중첩성	$0 < \Pr(W_i = 1   X_i = x) < 1 \forall x$

주: 1) 비교락성 가정은 조건부 독립 가정(Conditional Independence Assumption)으로 명칭하기도 함  
 자료: Rubin(1974); Imbens(2014), pp. 2~6을 참조하여 저자 작성

## 2. 성향점수매칭과 커절 포레스트

### 가. 성향점수매칭<sup>34)</sup>

제Ⅲ장에서는 성향점수매칭을 통해 신용보증기금 지원효과를 추정하였다. 주지하다시피 성향점수매칭은 비교락성과 중첩성 가정하에 ATE와 ATT를 편의 없이 추정 가능하다. 물론 해당 가정하에서 회귀분석 및 다양한 매칭 방법론들도 편의 없는 추정이 가능한데, 정책금융의 효과를 연구한 선행연구들은 대체로 성향점수매칭을 사용하여 정책효과를 추정하였기에 제Ⅲ장에서는 선행연구의 연구 결과와 용이한 비교를 위해 성향점수매칭을 활용하였다. 하지만 해당 가정하에 성향점수매칭을 활용한 효과 추정의 한계점에 대하여

34) 성향점수매칭에 대한 논의는 Angrist and Pischke(2008); Imbens(2014); Gulen et al. (2020)의 논의를 참조하여 정리하였다.

다양한 분야의 연구자들 사이에서 논의되고 있는데(Imbens, 2014; Shipman et al., 2017; King and Nielsen, 2019; Gulen et al., 2020), 이에 따라 본 연구에서는 성향점수매칭의 한계점을 보완하는 측면에서 머신러닝의 활용 가능성을 탐색하였다.

성향점수매칭의 한계점 그리고 대안으로의 머신러닝의 활용 가능성을 논의 하기 앞서 성향점수를 활용한 매칭방법론들이 어떻게 편이 없는 추정이 가능한지 간략하게 논의하면 다음과 같다. 먼저 앞서 언급한 비교락성 조건은  $X_i$ 를 조건부로 성립하게 되는데, 이 경우 처치확률인  $e(x) = \Pr(W_i = 1 | X_i = x)$ 를 조건부로도 성립하게 된다.<sup>35)</sup> 따라서 앞서 논의한 모집단에서의 ATT는 다음과 같이 표현 가능하다.

$$E\{Y_{1i} - Y_{0i} | W_i = 1\} = E\{E\{Y_i | e(X_i), W_i = 1\} - E\{Y_i | e(X_i), W_i = 0\} | W_i = 1\} \quad \text{식 (5)}$$

즉 ATT는 처치그룹과 동일한 처치확률을 지닌 통제그룹과 처치그룹의 결과 값의 차이의 평균을 통해서 추정할 수 있음을 의미하는데, 표본에서 계산 시 조건부 기대확률의 차이는 처치그룹과 동일한 처치확률을 지닌 통제 집단과 처치집단의 평균값의 차이를 먼저 계산하고 바깥쪽의 기댓값은 해당 차이값들의 가중평균으로 대체하면 추정이 가능함을 의미한다(Heckman et al., 1998).

물론 실제 표본을 활용하여 추정 시 이와 같은 추정 방법은 다양한 어려움이 수반하게 되는데 Robins et al.(1992)은 성향점수를 활용하여 ATE과 ATT를 다음과 같이 제시하였다.

$$E\{Y_i(1) - Y_i(0)\} = E\left[\frac{Y_i D_i}{e(X_i)} - \frac{Y_i(1 - D_i)}{1 - e(X_i)}\right] = E\left[\frac{(D_i - e(X_i)) Y_i}{e(X_i)(1 - e(X_i))}\right] \quad \text{식(6)}$$

$$E\{Y_i(1) - Y_i(0) | W_i = 1\} = E\left[\frac{(W_i - e(X_i)) Y_i}{(1 - e(X_i)) P(W_i = 1)}\right] \quad \text{식(7)}$$

35) 관련해서 수학적 증명은 Angrist and Pischke(2008)의 p. 80을 참조하라.

즉 처치 정도의 내생성에 의한 편이 가능성을 성향점수를 가중치로 활용하여 조정하는 방법인데, 직관적으로 처치 관측치 중에서 처치확률이 낮은 관측치에 높은 가중치를 부여하고 처치확률이 높은 관측치에 낮은 관측치를 부여함으로써 임의배정에 의한 처치집단과 통제집단의 결과값 비교와 최대한 유사하게끔 가중치를 조정해 주는 것이다.

이와 같이 성향점수매칭은 연구자들이 처치에 영향을 주는 모든 설명변수를 고려할 필요 없이 차원 축소에 따른 성향점수만 고려하게끔 하고 성향점수를 가중치로 활용할 경우 처치효과를 용이하게 추정할 수 있기 때문에 많은 연구에서 활발하게 사용되고 있다. 하지만 성향점수매칭이 다른 매칭이나 회귀분석과 비교하여 여러 측면에서 우월하지 않음을 많은 연구들이 지적하고 있다. 먼저 추정량의 점근적 효율성 측면에서 성향점수매칭은 모든 설명변수를 활용하여 처치그룹과 통제그룹을 매칭하는 추정량과 비교하여 효율적이지 않는 것으로 나타났다(Hahn, 1998). 이는 처치 여부에 크게 관여하지 않는 변수들 중에서 결과값의 변이를 일부 설명하게 된다면 해당 변수들을 추가로 통제함으로써 추정량의 표준오차를 줄일 수 있기 때문이다. 즉 점증적으로 표본이 증가하는 상황에서는 성향점수매칭 방법론이 효율적이지 않음을 의미한다. 그럼에도 불구하고 성향점수매칭의 활용의 정당성은 Angrist and Hahn(2004)에서 찾을 수 있다. 해당 연구에서는 유한한 표본(Finite Sample)에서는 추정량의 효율성 측면에서 성향점수매칭이 다른 추정량과 비교해서 효율적일 수 있음을 주장하였다. 즉 유한한 표본에서는 표본이 점증적으로 증가하는 경우와 다르게 결과값의 변이를 설명하는 데 큰 영향을 끼치지 못하는 수많은 변수들을 제외하고 성향점수를 활용하는 것이 효율적임을 의미한다. 이는 유한한 표본에서 해당 변수들을 모두 통제하는 경우 해당 변수들과 결과값 간의 영향을 추가로 추정해야 하기 때문에 통계적 부담으로 작용하게 되고 결과적으로 추정량의 효율성을 하락시킬 수 있다는 이야기이다.

반면에 Imbens(2014)는 성향점수매칭이 기타 매칭방법들에 비해 강건하지 못함을 주장하였고 회귀분석과 블록(Block)을 결합하는 방식을 사용할 것을

제안하였다. 또한 King and Nielsen(2019)은 연구자들이 실제로 성향점수매칭을 활용할 때 발생할 수 있는 문제점을 크게 두 가지 측면에서 지적하였다. 첫 번째는 성향점수매칭이 처치집단과 통제집단 간의 설명변수 측면에서의 불균형을 심화시킨다는 점이다. 성향점수매칭은 특히 ATT를 추정할 때 성향점수가 처치집단과 가장 유사한 통제집단을 제외한 나머지 관측치들을 제거하는데, 이 경우 설명변수 측면에서는 두 집단 간의 불균형이 오히려 심해질 수 있으며 결과적으로 추정량의 모형 의존도와 비효율성이 증가할 수 있음을 주장하였다. 두 번째로 처치 여부를 모형화할 때 모형의 형태, 즉 로지스틱 모형에 사용할 변수들의 선택에 따라 처치효과의 추정치가 민감하게 변화할 수 있다는 점이다. 이 경우 연구자들이 성향점수 추정 시 본인이 선호하는 결과를 제공해 주는 모형을 선택할 여지를 제공하기 때문에 성향점수매칭 사용에 주의해야 한다.

성향점수매칭의 한계점을 보완하는 방법론으로 본 연구에서는 커절 포레스트의 활용 가능성을 탐색하였는데, Gulen et al.(2020)의 주장처럼 성향점수매칭의 한계점을 보완하는 데 유용할 것으로 판단된다. King and Nielsen (2019)이 지적한 처치·통제 집단 간의 불균형 문제와 관련해서 성향점수매칭은 성향점수가 유사한 처치관측치와 통제관측치를 비교하지만 커절 포레스트는 설명변수 측면에서 유사한 처치 및 통제 관측치를 비교하기 때문에 커절 포레스트를 활용 시 두 집단의 불균형 문제를 완화할 수 있다. 또한 처치 여부를 모형화할 때 연구자들의 자의적인 판단 문제도 커절 포레스트를 활용하면 최소화될 것으로 보인다. 만약 연구자들이 처치 여부에 영향을 끼치는 요소들에 대한 사전지식이 없거나 혹은 영향을 끼치는 요소들과 처치 여부와의 관계를 구체적으로 파악하기 어렵다면 모형 선택에 대한 연구자들의 자의적인 판단이 필요하다. 이 경우 성향점수매칭은 처치 여부를 결정하는 모형 형태에 따라 처치효과의 추정치가 민감하게 변화하기 때문에 연구자의 판단에 따라 추정치가 변화할 수 있다는 한계점이 있다. 반면에 커절 포레스트 뒤에서 자세히 논의하겠지만 처치효과의 추정 과정에서 트리를 성장시키는 데 가장 중요한 변수들만을 알고리즘이 선택하기 때문에 연

구지들의 자의성이 개입될 여지를 최소화할 수 있다는 장점을 가진다. 특히 행정데이터 등의 빅데이터의 사용이 급격하게 증가해서 처치효과 추정에도 대량의 변수가 제공되는 경우가 증가하고 있다는 점을 고려한다면 커절 포레스트 활용의 중요성은 더욱 증가할 것으로 판단된다.

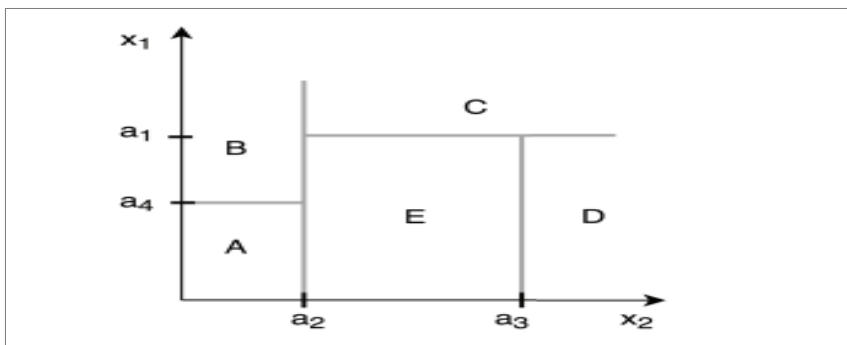
## 나. 커절 포레스트

Athey and Imbens(2016), Wager and Athey(2018), Athey et al.(2019)은 Brieman et al.(1984)에서 처음 소개된 CART(Classification and Regression Trees) 알고리즘을 수정·보완하여 조건부 평균처치효과(CATE)를 추정할 수 있는 알고리즘인 커절트리와 커절 포레스트를 제안하였다. 커절트리와 커절 포레스트의 특성은 CART와의 비교를 통해서 명확하게 드러나기 때문에 본 고에서는 CART의 기본적인 개념과 특성에 대해서 먼저 논의하였다.

### 1) CART의 개념 및 특성

CART는 분류와 회귀문제에 널리 사용되는 머신러닝 알고리즘 중에 하나이며 예측 대상인  $y$ 가 0,1의 값을 가지는 경우 분류트리(Classification Tree)를

[그림 V-1] 2차원 공간의 재귀적 이진분류 예시



자료: Schonlau and Zou(2020), p. 4, Figure 1.

사용하며, 예측변수가 연속형 변수이면 회귀트리(Regression Tree)를 사용한다(정재현 · 이환웅, 2020). 해당 알고리즘은 설명변수로 구성된 다차원 공간  $X$ 을 이진분류(Binary Classification)를 활용하여  $M$ 개의 상호배타적인 공간( $R_1, \dots, R_M$ )으로 나누게 된다.

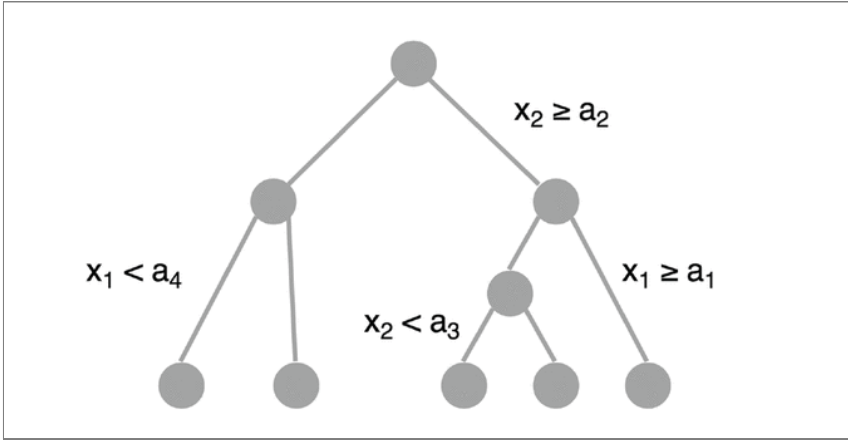
[그림 V-1]은 2차원 공간을 재귀적 이진분류를 활용하여 5개의 상호배타적인 공간으로 나누는 과정을 보여주며, [그림 V-2]는 [그림 V-1]에 상응하는 결정트리를 시각화하였다. [그림 V-1]에 따르면 첫 번째 분류는  $X_2 \geq a_2$ 에서 발생하여 두 개의 공간 {A, B}, {C, D, E}로 구분되며, 해당 공간들은 다시 한 번 이진분류에 의해 분화된다. 여기서 왼쪽 공간({A,B})은  $X_1 \geq a_1$ 에 의해 다시 {A}와 {B}로 구분되며 오른쪽 공간({C, D, E})은  $X_1 \geq a_1$  기준으로 {C}와 {D, E}로 분류된다. 마지막으로 {D, E} 공간은  $X_2 \geq a_3$ 에 의해 {D}와 {E}로 분류되고 트리는 성장을 멈춘다.

$$\min_{j,s} = \left[ \sum_{x_i \in R_1(j,s)} (y_i - \bar{y}_1(j,s))^2 + \sum_{x_i \in R_2(j,s)} (y_i - \bar{y}_2(j,s))^2 \right] \quad \text{식 (8)}$$

여기서,  $R_1(j,s) = \{X|X_j \leq s\}$ ,  $R_2(j,s) = \{X|X_j > s\}$

이와 같이 설명변수공간이 분류되면  $X_i = x$ 인 기업의 결과값은  $x$ 가 속한 공간  $R_m$ 의 평균값을 통해서 예측하게 된다. CART는 설명변수로 이루어진 다차원 공간을 어떤 변수를 활용하여 이진분류할지, 그리고 해당 변수의 어떤 값을 기준으로 분류할지를 특정변수와 해당 변수의 기준값의 수많은 조합 중에서 평균제곱오차(Mean Squared Error, MSE)를 최소화해 주는 변수와 기준값을 선택하게 된다. 즉 CART는 변수  $j$ 와 기준값  $s$ 를 식 (8)에서 제시한 목적함수를 최소화하는 조합에서 선택하게 된다.

[그림 V-2] 결정트리의 시각화: [그림 V-1]에 상응



자료: Schonlau and Zou(2020), p. 5, Figure 2.

## 2) 커절트리와 커절 포레스트의 특성

Athey and Imbens(2016)는 CART를 기반으로 처치효과를 추정할 수 있는 커절트리(Causal Tree)를 제안하였는데, 커절트리는 다음 두 가지 측면에서 CART와 차이점이 있다. 먼저 CART와 커절트리는 이진분류에 사용할 변수 및 기준값을 정할 때 판단 기준이 되는 목적함수에 차이점이 존재한다. CART의 목적함수는 앞서 식 (8)에서 확인하였는데, 실제값과 예측값의 차이의 제곱항을 최소화하는 분할조합(변수 및 해당변수의 기준값)을 찾는 데 있다. 하지만 커절트리는 실제 처치효과를 관측할 수 없기 때문에 실제 처치효과와 예측된 처치효과 간 차이의 제곱항을 최소화할 수 없다. 따라서 Athey and Imbens(2016)는 다음과 같이 중단노드의 처치효과를 추정하였다.

$$\hat{\tau}(x) = \frac{1}{i: \{W_i = 1, x_i \in L_i\}} \sum_{\{W_i = 1, x_i \in L_i\}} y_i - \frac{1}{\{i: W_i = 0, x_i \in L_i\}} \sum_{\{W_i = 0, x_i \in L_i\}} y_i$$

식 (9)

즉 커절트리의 마지막 노드에 포함된 처치그룹과 통제그룹의 결과값 차이를 통해서 처치효과를 계산하는데, Athey and Imbens(2016)는 추정된 처치

효과의 분산을 최대화하는 것이 결정트리에서의 MSE를 최소화하는 것과 유사함을 보였다.<sup>36)</sup> 따라서 커절트리는 설명공간을 분류할 때 해당 분류로 인해 처치효과의 분산이 증가하는 쪽으로 트리를 성장시키는 것이다.

커절트리가 기존의 CART와 또 다른 차이점은 트리를 훈련하는 과정에서 어니스티 접근법(Honesty Approach)을 도입하였다는 점이다. 즉 알고리즘을 훈련시키는 데이터를 임의로 훈련데이터( $S^{tr}$ )와 추정데이터( $S^{est}$ )로 분류해서 훈련데이터는 개별트리를 성장시키는 데 활용하고 추정데이터는 처치효과를 예측하는 데 사용된다. 즉 모형은 훈련데이터에서 적합화되며 처치효과는 추정데이터에서 추정되기 때문에 어니스티 접근법은 CART와 비교하여 모형의 과대적합을 방지해 주며 편의를 최소화하게 된다. 커절트리를 제시한 이후 Wager and Athey(2018)는 다수의 커절트리를 결합한 커절 포레스트를 제시하였으며 해당 알고리즘에 의해 추정된 조건부 평균 처치효과(CATE)는 비교락성과 중첩성 가정하에서 일치추정량이며 점증적으로 정규분포를 따름을 증명하였다.

$$[\hat{\tau}(x) - \tau(x)] / \sqrt{\text{Var}[\hat{\tau}(x)]} \Rightarrow N(0,1)$$

식 (10)

where,  $\tau(x) = E[Y_i(1) - Y_i(0) | X_i = x]$

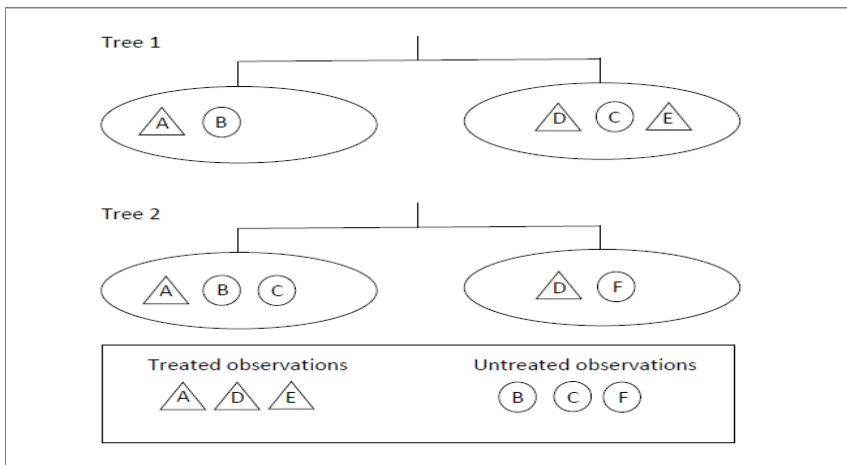
한편 커절 포레스트는 다수의 커절트리를 결합하여 구성하는데, 다수의 커절트리 생성 시 원표본으로부터 복원추출을 통한 무작위 표본을 구성한 후 트리를 성장시키게 된다. 따라서 이러한 무작위성에 기인하여 커절트리들은 조금씩 다른 특성을 가지며 커절 포레스트는 커절트리에서 사용하는 어니스티 접근법에 의해서 증가된 변동성을 다수의 커절트리의 결합으로 최소화하게 된다. 커절 포레스트는 B개의 트리를 성장시키는데, 커절 포레스트에 의한  $x$ 의 처치효과  $\hat{\tau}(x)$ 는 각각의 커절트리에 의한 추정된 처치효과의 평균값으로 추정된다( $\hat{\tau}(x) = \sum_{b=1}^B \hat{\tau}_b(x)$ ).

36) 자세한 내용은 Athey and Imbens(2016)의 pp. 7356~7357을 참조하라.

[그림 V-3]은 커절 포레스트를 통해서 처치효과가 추정되는 과정을 제시하였다. 여기서 커절 포레스트는 두 개의 커절트리를 결합하였으며(트리 1, 트리 2) 관측치 A, D, E는 처치집단, B, C, F는 통제집단이다. 여기서 주목할 점은 트리 1의 경우 처치효과를 추정하는 추정데이터 ( $S^{est}$ )에 관측치 C가 트리 2에서는 관측치 E가 제외되었다는 점이다. 해당 데이터들은 나무를 성장시키는 훈련데이터로 사용되었기 때문인데 이에 따라 추정데이터에서 제외된 것이다. 이렇듯 처치효과 추정 시 트리마다 사용하는 관측치가 다른 것은 어니스티 접근법의 특성으로 모형을 구축하는 데 사용한 데이터와 처치효과를 추정하는 데 사용하는 데이터가 분류되었음을 보여준다.

[그림 V-3]에서 A의 처치효과 추정과정을 살펴보면 먼저 트리 1에서는 A와 B의 결과값의 비교를 통해서 추정되며, 트리 2에서는 A의 결과값과 B·C의 결과값의 차이를 통해서 추정되어 트리마다 A의 처치효과가 다르게 추정될 수 있다. 즉 커절 포레스트는 훈련데이터와 처치데이터를 임의로 배정하기 때문에 각 커절트리의 최종노드에 포함되는 사업체들의 변이 역시 발생하며, 이에 따라 커절트리별 처치효과에도 변이가 발생하게 되는 것이다.

[그림 V-3] 커절 포레스트의 처치효과 추정 예시



자료: Gulen et al.(2020), p. 57, Figure 2를 참조하여 작성

마지막으로 커절 포레스트를 통해서 추정 가능한 처치효과를 살펴보면 앞서 언급하였듯이 커절 포레스트는 설명변수가 특정값을 가지는 사업체의 조건부 평균 처치효과(CATE)를 비교락성과 중첩성 가정하에서 편의 없이 추정 가능하다. 구체적으로  $\hat{\tau}(x)$ 은 개별 커절트리의 최종노드들 중에서  $X=x$ 가 포함된 최종노드에서 추정된 처치효과들의 평균을 통해서 추정된다. 한편 CATE가 추정되면 해당 추정치들을 활용하여 ATE와 ATT 등의 계산이 가능하다. ATE와 ATT는 CATE의 단순평균을 통해서 간편하게 추정될 수 있으나, 본 연구에서는 처치효과들을 계산할 때 AIPW(Augmented Inverse Propensity Weighted) 추정량을 사용하였다. 만약 단순 가중평균과 같이 IPW 추정량을 사용하는 경우 유한표본에서 예측된 성향점수가 0 혹은 1에 근접할 때 IPW의 성과가 좋지 못하며, 처치 여부에 대한 모형과 결과값에 대한 모형 설계에 오류가 있는 경우 AIPW의 성과가 IPW와 비교해서 우월하기 때문이다(Glynn and Quinn, 2010). 한편 본 연구에서는 ATE, ATT 외에도 ATO(Average Treatment effect with an Overlap correction)를 추가로 추정하였는데 ATO는 ATE 추정과 유사하지만 중첩성이 높거나 혹은 처치집단과 통제집단의 비율이 유사한 노드에 높은 가중치를 부여하는 방식으로 추정된다.

### 3. 분석 결과

#### 가. 정책효과의 추정

##### 1) 분석데이터 및 가정 검토

커절 포레스트를 사용하여 신용보증기금 수혜 효과를 추정하기 위해 본 분석에서는 제Ⅲ장의 2011~2015년 표본을 사용하였다. 즉 종속변수는 2015년 매출액 및 영업이익이며, 처치변수는 특정 사업체가 2014년에 신용보증기금을 수혜했는지 여부, 그리고 통제그룹으로는 2014~2015년에 신용보증기금을 수혜하지 않는 사업체로 한정하였다. 2011~2015년 표본을 사용한 이유는 [그림 Ⅲ-5]에서 확인한 것처럼 2011~2013년 매출액 및 영업이익을 통제한 상태에서는 신용보증기금 수혜 여부와 가상성과가 독립적이라는 비교락성 가정이

충족될 가능성이 높기 때문이다.

한편 제Ⅲ장에서 처치효과 분석 시 사용한 지역정보는 사업체의 수도권 위치 여부였는데, 본 분석에서는 사업체의 자세한 지리적 정보를 알고리즘이 충분히 활용하기 위해 시·군·구 수준의 사업체 지리정보를 구축하여 수도권 여부 변수 대신에 사용하였다.<sup>37)</sup> 그리고 영업이익과 순이익의 경우 서로간의 상관관계가 매우 높기 때문에 본 분석에서는 영업이익만을 예측변수로 활용하였다.<sup>38)</sup> 2011~2015년 표본의 기초통계량은 <표 Ⅲ-12>에서 확인할 수 있는데, 총 사업체 수는 8만 3,897개를 포함하고 있다. 하지만 본 분석에서는 시·군·구 수준의 사업체 지리적 정보를 구축하는 과정에서 지리적 정보가 부정확한 관측치 1,574개를 분석표본에서 제외해서 최종 분석표본의 관측치는 8만 2,323개로 구성되었다.

비교락성의 성립 여부와 관련해서 본 분석에서는 처치그룹과 통제그룹이 과거 영업이익과 매출액을 통제할 경우 결과값의 동일한 추세를 지니는 것을 확인한 후 비교락성이 성립함을 가정하였다. 한편 Imbens(2014)는 교육 훈련프로그램의 참여하는 구직자들이 참가 직전 연도의 임금이 하락하는 ‘Ashenfelter Dip’과 같은 상황에서 비교락성을 검정하는 방법을 제안하였다. 최근 사업체 단위에서 조세·재정 사업의 효과성 평가가 활발히 진행되고 있는데, 이환웅(2021)은 정책의 수혜 사업체가 비수혜 사업체와 비교해서 정책 수혜 이전부터 결과 값이 하락 혹은 상승하는 경우가 빈번하게 발생함을 지적하였다. 따라서 비록 본 분석에서는 Imbens(2014)의 방법론을 차용해서 비교락성을 검정하지는 않았지만 활용도가 높은 것으로 판단하였기 때문에 ‘Ashenfelter Dip’과 같은 상황에서의 비교락성 검정방법을 간략하게 소개한다.

먼저 과거 한 시점 전의 결과변수  $Y_{i(-1)}$ 부터  $T-1$  시점 전의 결과변수  $Y_{i-(T-1)}$ 와 그 외의 설명변수  $Z_i$ 를 통제하였을 때 비교락성이 성립함을 가정하였다.

37) 지리적 정보를 세세하게 사용하였음에도 불구하고 커질 포레스트를 활용하여 추정된 ATE는 사업체 지역적 정보를 수도권 여부 혹은 시·군·구와 관련 없이 유사하게 추정되었다.

38) 과거의 순이익 정보를 추가로 사용한 커질 포레스트의 처치효과 추정치는 본 분석의 추정치와 유사하게 추정되었다.

$$Y_i(1), Y_i(0) \perp W_i | Y_{i,-1}, \dots, Y_{i,-(T-1)}, Z_i \quad \text{식 (11)}$$

Imbens(2014)에 따르면 결과변수의 한계분포에 대하여 정상성(Stationarity)와 교환가능성(Exchangeability)을 추가로 가정하면<sup>39)</sup> 다음과 같은 비교락성 조건을 만족하게 된다.

$$Y_{i,(-1)} \perp W_i | Y_{i,(-2)}, \dots, Y_{i,-(T-1)}, Z_i \quad \text{식 (12)}$$

즉 과거 한 시점 전의 결과변수  $Y_{i(-1)}$ 는 두 시점 전의 결과변수  $Y_{i(-2)}$  및 기타 설명변수를 통제하는 경우 처치 여부와 독립적임을 의미한다. 따라서 식 (11)의 비교락성이 성립하는지에 대한 검정은 과거 한 시점 전의 결과변수를 종속변수로 처치 여부 및 기타 설명변수들을 활용해서 처치효과를 추정 한 후 추정계수가 통계적으로 0과 유의하게 다른지 검토하면 된다. 만약 처치효과가 0에 가깝게 추정되면 식 (12)가 성립함을 의미하기 때문에 해당 방법론을 활용해서 비교락성을 검정할 수 있게 된다. 물론 한계분포에 대한 정상성 및 교환 가능성을 추가로 검정하는 부분이 연구자들에게 요구된다는 한계점을 지니지만 확률분포 가정에 대한 믿음이 연구자들에게 있을 경우 Imbens(2014)가 제시한 방법론도 유용할 것으로 판단된다.

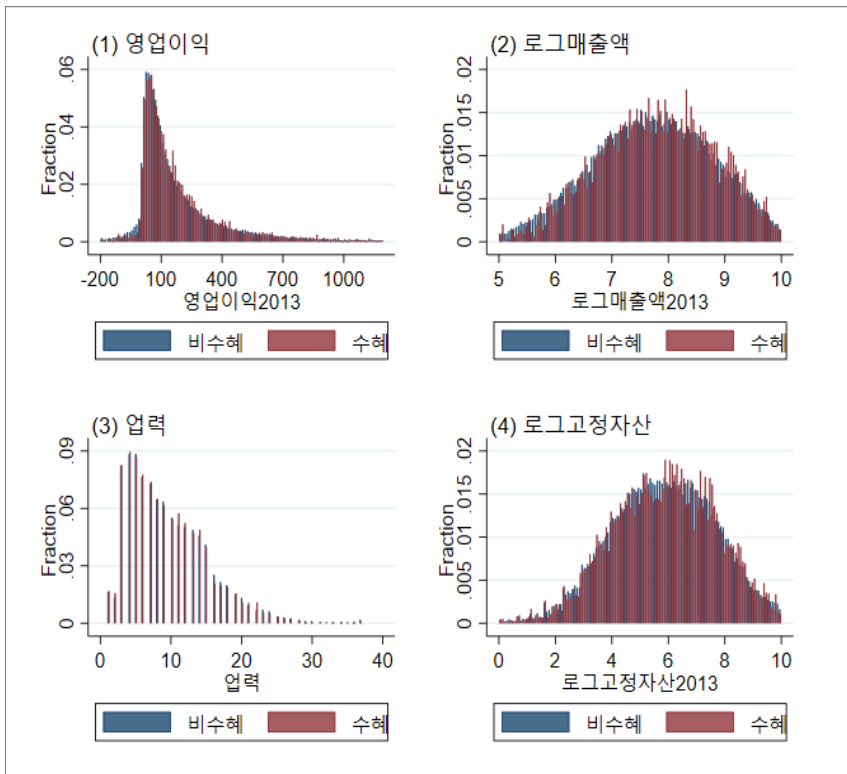
한편 커질 포레스트를 통해서 처치효과를 편이 추정하기 위해서는 추가적으로 중첩성 조건이 요구된다. 중첩성 가정의 성립 여부와 관련해 제 Ⅲ장의 [그림 Ⅲ-3]은 처치그룹과 통제그룹의 성향점수 분포를 제시하였는데, 처치그룹의 경우 통제그룹과 비교하여 성향점수 분포가 다소 우측으로 치우친 모습이지만 치우침의 정도는 매우 낮은 수준으로 중첩성 가정을 만족하고 있는 것으로 보인다.

물론 커질 포레스트에서 추정한 성향점수와 성향점수매칭에서 추정한 성향점수는 동일한 변수를 활용함에도 불구하고 차이가 발생한다. 이는 커질 포레스트는 설명변수와 처치 여부의 비선형적인 관계를 고려하기 때문인데,

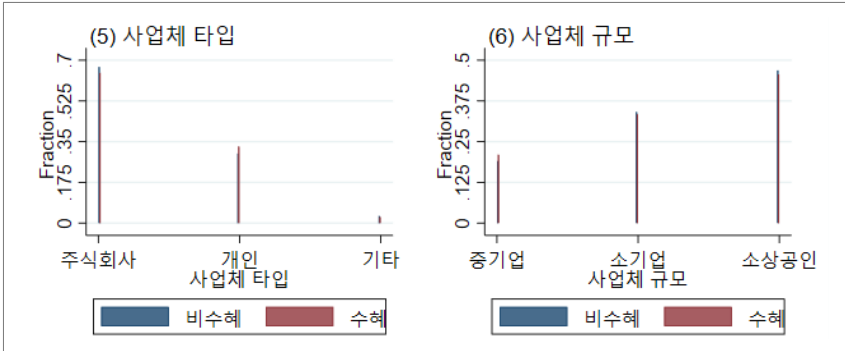
39) 정상성과 교환가능성에 대한 추가가정은 Imbens(2014)의 p. 32를 참조하라.

이에 따라 두 방법론이 계산한 성향점수에 차이가 발생하게 된다. 따라서 커질 포레스트를 사용할 때 중첩성 여부가 성립하는지를 추가적으로 검증하였는데, [그림 V-4]은 중요 설명변수들 측면에서 수혜집단과 비수혜집단의 분포가 유사한지 보여주고 있다. 해당 그림에서 확인할 수 있듯이 IPW 가중치를 적용한 후 수혜집단과 비수혜집단은 각각의 설명변수에 대하여 유사한 분포를 가지고 있음을 확인할 수 있었다. 결론적으로 본 분석에서 사용하는 2011~2015년 표본은 과거의 영업이익 및 매출액 정보를 추가로 통제할 경우 비교락성이 성립될 가능성이 높아지는 것으로 판단되면 중첩성도 만족시키는 것으로 판단된다.

[그림 V-4] 신용보증기금 수혜기업과 비수혜기업의 변수별 분포: IPW 가중치 적용



[그림 V-4]의 계속



자료: 「기업정보데이터」; 한국기업데이터 자료를 활용하여 저자 작성

## 2) 커절 포레스트 하이퍼 파라미터 튜닝

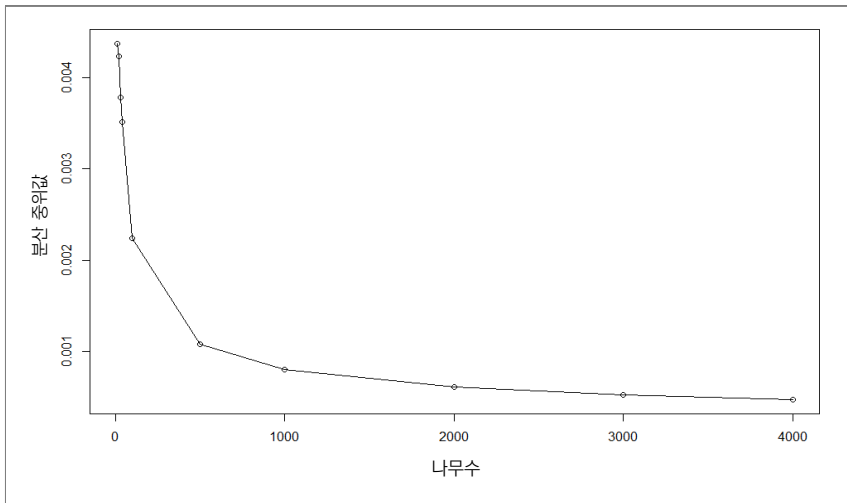
본 분석에서는 r의 `grf` 패키지를 사용하여 처치효과 및 처치효과 이질성을 분석하였다. 해당 패키지는 연구자들이 설정해야 하는 하이퍼 파라미터들이 존재하는데, 추정 결과를 제시하기에 앞서 먼저 해당 값을 어떻게 설정하였는지를 논의하였다.

커절 포레스트는 어니스티 접근법과 어댑티브 접근법을 사용할 수 있는데, 처치효과 추정의 편의를 최소화하기 위해 본 분석에서는 Wager and Athey(2018)을 참조하여 어니스티 접근법에 기반한 커절 포레스트를 사용하였다. 어니스티 접근법을 사용함에 따라 모형 훈련에 사용한 총 관측치 5만 7,382개 중에서 50%는 훈련데이터로, 50%는 추정데이터로 사용하였다.<sup>40)</sup> 즉 모형의 훈련에 사용하는 관측치는 2만 8,691개이며 처치효과를 추정하는 데 사용한 관측치 역시 2만 8,691개이다. 앞서 간략하게 논의하였듯이 어니스티 접근법은 모형 훈련에 적은 데이터를 사용하기 때문에 모형의 성과가 좋지 않을 가능성이 존재한다. 하지만 본 분석에서는 훈련데이터에 2만 8,691개가 배정되기 때문에 충분한 데이터가 훈련데이터로 활용되고 있는 것으로 판단하였다.<sup>41)</sup>

40) 전체 데이터 8만 2,323개 중 30%는 테스트 데이터로, 70%는 전통적인 개념의 훈련데이터로 활용하였다. 해당 훈련데이터를 다시 50:50으로 배분하여 50%는 훈련데이터로, 50%는 추정데이터로 사용하였다.

한편 어니스티 접근법을 사용함에 따라 증가하는 추정치의 변동성은 다수의 트리를 성장시킴으로써 완화할 수 있다. 따라서 연구자들은 몇 개의 트리를 성장시켜야 하는지 설정이 필요하다. 물론 다수의 트리를 성장시키면 추정치의 변동성을 줄일 수 있다는 장점이 있으나, 그와 동시에 알고리즘을 혼련시키는 데 소요되는 시간이 비례적으로 증가하게 된다. 따라서 이러한 상충관계에서 최적의 트리 수를 설정하기 위해 커질 포레스트에서 사용한 트리 수에 따라 추정치의 분산(중위값)이 어떻게 변화하는지를 [그림 V-5]에서 제시하였다. 당연하게도 트리 수가 증가함에 따라 추정치의 분산이 감소함을 확인할 수 있는데, 특히 트리 수가 10개에서 500개로 증가하는 구간에서는 분산의 중위값이 급격하게 감소하는 것을 확인할 수 있었다. 하지만 1천개 이후에서는 트리 수의 증가에 따른 분산의 감소효과가 약하게 나타났다. 따라서 본 분석에서는 1천개의 트리를 사용하여 처치효과를 추정하였다.

[그림 V-5] 커질 포레스트의 사용하는 나무 수에 따른 분산의 중위값



자료: 「기업정보데이터」; 한국기업데이터 자료를 활용하여 저자 계산

41) 훈련데이터의 70%의 관측치를 배정한 커질 포레스트와 본 분석에서 사용한 50%를 배정한 커질 포레스트의 모형의 적합성을 검정한 결과, 두 알고리즘 모두 유사한 것으로 나타났으며 ATE 추정치도 유사하게 추정되었다.

커팅 포레스트의 경우에도 랜덤포레스트와 동일하게 알고리즘이 이진분류법을 활용하여 나무를 성장시킬 때 모든 변수를 고려하지 않고 그중에 몇 개의 변수만을 임의적으로 선택해서 이진분류 시 활용하게 된다. 일반적으로 임의로 선택하는 변수의 개수는 다음의 공식을 활용하여 결정된다.

이진분류에 사용하는 변수 수 =  $\min(\sqrt{p} + 20, p)$

본 분석에서 설명변수로 사용하는 변수는 총 14개인데, 이상의 공식에 따르면 14개 변수를 모두 사용해야 한다. 따라서 해당 옵션에 대해서는 14개의 변수로 설정하였다.

마지막으로 grf 패키지에서는 최종노드에 포함되는 최소 처치관측치 수를 설정할 수 있는데, 본 분석에서는 10개로 설정하였다. 이에 따라 특정 노드에서 이진분류법을 사용해 생성되는 두 개의 노드에서 10개 이하의 처치 관측치가 포함될 경우 알고리즘은 더 이상 트리를 성장시키지 않고 해당 노드가 최종노드가 되게 된다. 최소 요구치를 10개로 설정한 이유는 처치관측치를 너무 작게 설정하면 처치효과를 추정데이터( $S^{est}$ )를 활용해서 추정할 때 해당 노드에 처치관측치가 없는 경우가 발생하며, 이에 따라 해당 트리에서는 처치효과가 추정되지 않고 제거되기 때문이다. 이렇게 특정 트리가 포레스트에서 제거되는 것을 방지하기 위해 최종노드에 포함되는 최소 처치집단의 수를 10개로 설정하였다.

### 3) 모형의 적합도 검증

grf 패키지는 커팅 포레스트가 처치효과 및 처치효과 이질성을 잘 추정하였는지 일종의 적합도(goodness of fit)를 가늠할 수 있는 테스트를 연구자들에게 제공한다. 구체적으로 Chemozhukov et al.(2018)의 최량선형예측(Best Linear Predictor) 방법론을 활용하여 CATE가 커팅 포레스트의 OOB(Out-of-Bag) 추정량  $\hat{\tau}^{-1}(X_i)$ 의 선형함수로 얼마나 잘 설명되는지 검증하는 것이다.<sup>42)</sup> 구체

적으로 결과값  $Y_i$ 에서  $X_i$ 에 의한 조건부 기댓값을 차분한 값을 종속변수로, 그리고 이하의 두 개 변수를 설명변수로 회귀분석을 실시한다.

$$C_i = \bar{\tau}(W_i - \hat{e}^{(-i)}(X_i)),$$

$$D_i = (\hat{\tau}^{(-i)}(X_i) - \bar{\tau})(W_i - \hat{e}^{(-i)}(X_i))$$

여기서  $C_i$ (mean, forest, prediction)의 추정계수는 평균처치효과가 얼마나 잘 추정되었는지를 검정하며, 1의 값을 가질 경우 커절 포레스트가 평균처치효과를 잘 추정한 것으로 판단할 수 있다. 그리고  $D_i$ (differential, forest, prediction)의 추정계수는 커절 포레스트가 처치효과의 이질성을 얼마나 잘 추정하였는지를 검정하는 것으로 1의 값을 가지면 처치효과의 이질성이 잘 추정되었음을 의미한다. 물론 추정계수가 1보다 작은 경우에도 통계적으로 유의하게 0보다 클 경우  $\hat{\tau}^{(-i)}(X_i)$ 와 CATE 간에 유의한 관계가 있음을 의미한다(Athey and Wager, 2019).

커절 포레스트에 의한 추정 결과를 제시하기 앞서 먼저 추정된 커절 포레스트의 적합성을 검토하였는데 검정 결과는 <표 V-6>에서 제시하였다. 2015년 매출액을 종속변수로 사용한 경우 평균처치효과에 대한 추정계수는 0.988로, 1과 유사한 것으로 나타나 평균처치효과를 잘 추정하는 것으로 판단된다. 한편 처치효과 이질성 여부의 추정계수는 0.481로 1과 유사하지 않은 것으로 나타나 처치효과의 이질성을 커절 포레스트를 통해서 정확하게 추정하지 못한 것으로 나타났다. 하지만 통계적으로 유의하게 0보다 큰 것으로 나타났기 때문에 커절 포레스트를 통해 추정된 이질성과 CATE 간의 어느 정도 유의한 관계가 있음을 알 수 있다. 반면에 영업이익을 종속변수로 사용한 경우 추정계수는 매출액을 사용했을 때와 유사하게 추정되었으나 추정치의 표준오차가 크게 추정되어 통계적 유의성을 띠지 못하였다. 이질성을 검정하는 계수 역시 통계적으로 유의하지 않는 것으로 나타나 영업이익을 종속변수로 사용한 커절 포레스트는 처치효과 및 처치효과의 이질성을 적절하게

42) 해당 테스트의 자세한 내용은 Athey and Wager(2019)의 p. 3을 참조하라.

추정하지 못한 것으로 보인다.

〈표 V-3〉 커절 포레스트의 모형 적합성 검정

	추정계수	표준오차	t-value	p-value
<b>종속변수: 15년 매출액</b>				
mean,forest,prediction	0,988***	0,138	7,172	0,000
differential,forest,prediction	0,481*	0,294	1,634	0,051
<b>종속변수: 15년 영업이익</b>				
mean,forest,prediction	1,056	1,723	0,613	0,270
differential,forest,prediction	0,422	0,456	0,925	0,178

주: 통계적 유의성은 \*\*\* 1%, \*\* 5%, \* 10%와 같음  
 자료: 「기업정보데이터」; 한국기업데이터 자료를 이용하여 저자 계산

커절 포레스트를 통해 추정된 처치효과가 이질적인지 살펴보는 또 다른 방법은 커절 포레스트를 통해서 추정된 CATE를 크기 순으로 나열한 후 중위값을 기준으로 두 그룹으로 나누어 평균처치효과를 계산한 후에 두 집단 간의 처치효과의 차이가 통계적으로 유의한지 일종의 t-test를 실시하는 것이다(Athey and Wager, 2019). 2015년 매출액을 종속변수로 사용한 커절 포레스트는 두 집단 간의 처치효과 차이를 2.44%로 계산하였으며 10% 수준에서 통계적으로 유의하게 0과 다른 것으로 나타났다. 하지만 2015년 영업이익을 사용한 경우에는 두 집단 간 처치효과의 차이가 2,264만원으로 나타났으나, 〈표 V-3〉의 결과와 동일하게 표준오차가 크게 추정되어 통계적으로 유의하지 않았다.

모형검정 결과를 종합하자면 커절 포레스트는 2015년 매출액을 종속변수로 사용하는 경우에는 평균처치효과와 처치효과의 이질성을 어느 정도 모형을 통해서 추정 가능한 것으로 판단된다. 반면에 영업이익의 경우 처치효과의 이질성을 커절 포레스트를 통해서 추정하기 어려운 것으로 보인다. 따라서 본 분석에서 처치효과 추정 결과는 매출액과 영업이익을 종속변수로 사용한 결과 모두를 제시하였으나 처치효과의 이질성 분석은 2015년 매출액을 종속변수로 사용한 경우에만 추정 결과를 제시하였다.

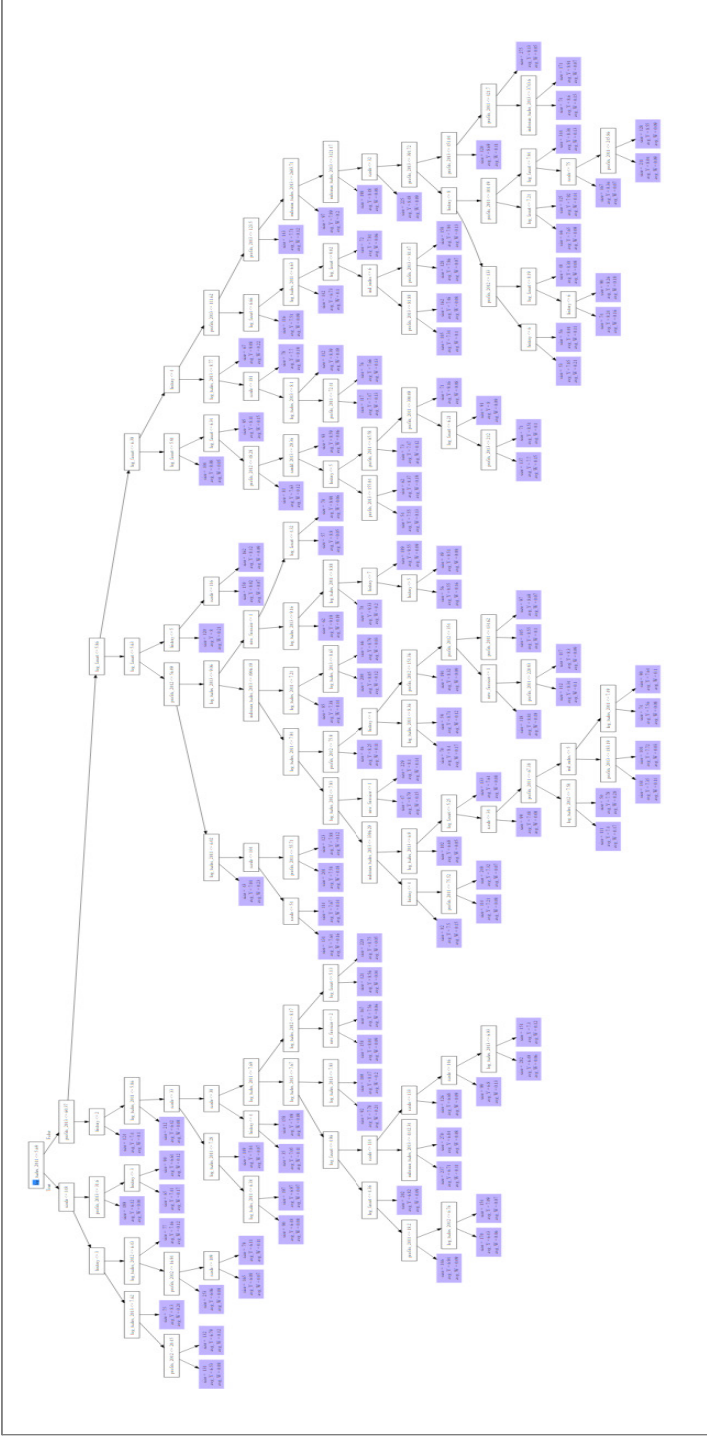
#### 4) 정책금융 지원효과

그림 V-6은 매출액을 종속변수로 한 커질 포레스트의 첫 번째 트리를 도식화하였다. 첫 번째 이진분류는 2011년 로그매출액이 5.69 이하를 기준으로 참이면 왼쪽 노드로, 거짓이면 오른쪽 노드로 분류하였음을 알 수 있다. 그다음 두 개의 노드에서 첫 번째 노드는 지역코드가 148 이하 여부에 따라 트리를 성장시키고 두 번째 노드에서는 2013년 영업이익이 6,037만원 이하 여부를 기준으로 트리를 성장시켰다. 참고로 최종 노드는 보라색 박스로 색칠이 되어 있는데, 노드에 포함되어 있는 데이터의 수, 해당 노드의 평균 매출액(결과값), 처치그룹의 비중 등이 제시되어 있다.

〈표 V-6〉은 커질 포레스트를 기반으로 추정된 신용보증기금 수혜 여부가 매출액 및 영업이익에 끼친 효과를 보여주는데, 매출액과 영업이익을 종속변수로 사용한 경우로 구분하여 ATE, ATT, ATC, ATO를 제시하였다. 먼저 매출액 효과를 살펴보면 신용보증기금 수혜는 평균적으로 2015년 매출액을 4.45% 증가시키는 것으로 추정되었다. 한편 실제로 신용보증기금을 수혜 하였던 사업체들을 대상으로 추정된 처치효과인 ATT의 경우에도 4.46%로 추정되었으며 1% 수준에서 통계적으로 유의하였다. 통제집단이 신용보증기금을 수혜하였을 때의 효과를 계산한 ATC의 경우에도 1% 유의수준에서 통계적으로 유의한 4.44%로 추정되었다. ATC와 ATT가 매우 유사하게 추정된 것은 표본을 구성하는 처치그룹과 통제그룹이 설명변수 측면에서 유사한 사업체로 이루어져 있음을 의미한다. 이는 ATO 추정치를 확인해 보면 더욱 명확해지는데, ATO는 중첩성을 보다 잘 만족시키기 위해 성향점수가 0이나 1에 가까운 관측치에는 적은 가중치를, 0.5에 근접한 관측치에는 높은 가중치를 부여해 CATE를 평균하기 때문이다. ATO 추정치 역시 ATE와 매우 유사하게 추정되었는데, 이는 분석에 사용한 표본의 처치집단과 통제집단의 성향점수 분포가 유사하고 중첩성 가정을 만족하고 있기 때문으로 판단된다.<sup>43)</sup>

43) 2011~2015년 표본의 구축 시에 Crump et al.(2009)의 조언을 고려하여 성향점수가  $0.05 < e(x_i) < 0.9$ 에 포함되는 관측치만 남긴 것도 처치그룹과 통제그룹 간의 균형을 잡는 데 도움을 준 것으로 보인다.

[그림 V-6] 커질 포레스트의 커질트리1 도식화: 매출액모형



자료: 「기업정보데이터」; 한국기업데이터 자료를 활용하여 저자 작성

한편 동일한 표본을 활용하여 성향점수매칭으로 추정된 추정치는 <표 III-13>에서 확인할 수 있듯이 7.8%로 컷절 포레스트에 기반한 추정치보다 크게 추정되었다. 흥미로운 점은 성향점수매칭을 활용하여 추정된 ATE는 13.48%로 ATT와 큰 차이를 보였다는 점이다. 이러한 차이는 King and Nielsen(2019)이 지적한 것처럼 성향점수에 기반해 제일 근접한 관측치를 제외한 다른 관측치들을 제거하는 방식이 처치그룹과 통제그룹 간의 불균형을 심화시킨 데서 기인한 것으로 보인다.

〈표 V-4〉 신용보증기금 수혜가 매출액 및 영업이익에 끼친 효과

처치효과	추정치	표준오차	관측치
2015년 로그매출액			
ATE	0.0445***	0.0063	57,382
ATT	0.0446***	0.0062	57,382
ATC	0.0444***	0.0064	57,382
ATO	0.0444***	0.0062	57,382
2015년 영업이익(백만원)			
ATE	-1.827	4.793	57,382
ATT	-3.222	4.537	57,382
ATC	-1.672	4.853	57,382
ATO	-3.156	4.579	57,382

주: 통계적 유의성은 \*\*\* 1%, \*\* 5%, \* 10%와 같음  
 자료: 「기업정보데이터」; 한국기업데이터 자료를 활용하여 저자 계산

King and Nielsen(2019)은 또한 성향점수매칭의 한계점으로 처치 여부를 모형화할 때 사용하는 로짓모형의 형태에 따라 처치효과와 추정치가 민감하게 변화함을 지적하였다. 제Ⅲ장에서 성향점수매칭을 활용하여 처치효과를 추정할 때 본 연구에서는 Dehejia and Wahba(1999)의 조언에 따라 처치 여부를 로짓모형으로 모형화할 때 매출액과 영업이익 등 연속형 변수들의 제곱항을 추가하였다. 흥미로운 점은 제곱항을 제외하고 로짓모형을 추정할 경우 ATT 추정치는 4.02%로, 본 연구에서 선호하는 성향점수매칭 추정치인 7.8%와 큰 차이가 있다는 점이다. 즉 연속형 변수의 제곱항 추가 여부에 따라

처치효과가 민감하게 변화하였음을 보여준다. 이 부분에서 커절 포레스트는 기존의 방법론을 보완할 수 있을 것으로 판단된다. 왜냐하면 각각의 커절트리 는 성향점수 추정 시 설명변수와 처치 여부 간의 비선형적인 관계를 유연하게 고려할 뿐 아니라 커절 포레스트는 다수의 커절트리(본 분석에서는 1천개)의 결과값을 평균화하기 때문이다. 즉 성향점수매치에서 로짓모형을 추정 시 1천개의 모형형태가 다른 로짓모형을 추정한 후, 이를 기반으로 처치효과를 추정해서 처치효과와 평균값을 구한 것과 유사하기 때문에 추정치의 강건성 측면에서 커절 포레스트는 성향점수모형의 한계점을 보완할 수 있는 것으로 보인다.

앞선 분석 결과에 따르면 신용보증기금 수혜가 해당 업체의 매출액을 증가시키는 것을 확인할 수 있었는데, 매출 증가가 실제로 사업체들의 수익성 개선으로 이루어졌는지 확인하기 위해서는 영업이익 등의 지표를 추가로 살펴볼 필요가 있다. 커절 포레스트를 활용한 추정 결과, 신용보증기금은 평균적으로 2015년 영업이익을 183만원 감소시키는 것으로 나타났으나, 통계적으로 유의하지 않았다. 반면 처치확률의 중첩성을 고려한 ATO는  $-3.16$ 으로 추정되어 ATE와 비교하여 부정적인 영향이 좀 더 크게 추정되었으나 ATE와 마찬가지로 통계적으로 유의하지 않았다. 결론적으로 신용보증기금이 영업이익에 부정적인 영향을 끼쳤다는 증거를 찾기는 어려웠다. <표 III-13>의 성향점수매치를 기반한 추정치가 통계적으로 유의한  $-13.41$ 로 추정되었음을 고려해 본다면, 이는 큰 차이라 할 수 있다. 물론 커절 포레스트의 적합도 검정 결과 영업이익을 종속변수로 사용할 경우 커절 포레스트의 성과가 좋지 못했던 것을 감안한다면 해당 추정치를 기반으로 강한 결론을 이야기하기는 어려울 것으로 판단된다. 따라서 2015년 영업이익을 종속변수로 한 커절 포레스트는 검정력을 높이기 위해 하이퍼 파라미터들의 값을 다양하게 설정하고 알고리즘을 훈련시키는 과정이 추가적으로 필요할 것으로 보인다.

커절 포레스트를 이용한 처치효과 분석 결과는 머신러닝의 활용 영역이 비단 예측 문제에 국한되는 것이 아니라 프로그램 효과 평가에서도 활용할 수 있음을 보여준다. 조세·재정 사업의 효과성 평가뿐만 아니라 법정평가

인 조세특례심층평가, 재정사업심층평가, 고용영향평가 등의 경우에 최근 사업체 단위에서 준수실적 방법론을 활용한 효과성 평가가 활발하게 이루어지는 만큼 본 연구에서 탐색한 커절 포레스트를 보완적으로 사용하는 것도 가능할 것으로 판단된다. 물론 성향점수매칭을 사용해도 만약 연구자가 중첩성 확보를 위한 선행연구들이 제안하고 있는 가이드라인을 준수하고(Dehejia and Wahba, 1999; Crump et al., 2009; Imbens, 2014), 식별 가정을 면밀히 검토하며, 로짓모형에 기반하여 처치 여부를 모형화할 때 다양한 모형형태를 실험해서 강건성을 검증한다면 성향점수매칭을 통해서도 신뢰성 있는 추정치를 구축할 수 있을 것으로 판단된다. 다만 연구자들 입장에서 특히 로짓모형의 강건성 검증은 손쉽게 이루어지기 어려운 반면에, 커절 포레스트는 앞서 언급한 것처럼 개별 트리마다 처치 여부를 모형화하고 처치효과를 추정하기 때문에 모형에 의존적이지 않는 추정치를 비교적 손쉽게 제공한다는 장점이 명확히 존재한다. 따라서 처치효과의 추정 시 커절 포레스트의 유용성이 큰 것으로 판단된다.

마지막으로 커절 포레스트 분석 결과를 신용보증기금을 분석한 선행연구의 결과와 비교하면 장우현 외(2013)는 정책금융 지원이 개별 사업체의 성과에 일관되게 부정적인 방향으로 영향을 미치는 것으로 추정하였다. 본 연구에서는 신용보증기금을 지원받은 기업의 경우 단기적으로 매출액이 증가하는 것으로 나타났으나 영업이익은 통계적으로 유의하지 못한 부정적인 효과가 추정되어 정책금융 지원이 기업의 수익성을 개선시키지 못한다는 선행연구의 결과는 질적으로 유지되는 것으로 나타났다.

#### 나. 정책효과의 이질성

‘grf 통계패키지’는 특성변수(혹은 설명변수)들이 트리를 성장시키는 데 얼마나 많이 사용되었는지를 비중으로 전환하여 연구자들에게 제공해 준다. 예를 들어 두 개의 트리를 성장시키는 데 커절 포레스트가 20번의 이진분류를 활용하여 트리를 성장시켰고 이 중 전년도 영업이익이 10번 사용되었다면 해당 변수의 중요도는 0.5가 되는 것이다. 설명변수들이 트리성장과 관련

하여 얼마만큼 중요한지는 <표 V-5>에서 제시하였다. 매출액을 종속변수로 모형을 훈련시킨 경우 2013년 영업이익이 트리 성장에서 차지하는 비중이 가장 높았으며 비중은 26.7%인 것으로 나타났다. 그다음 2011년 영업이익과 2011년 로그매출액이 각각 12.7%, 11.6%를 차지하고 있어서 처치효과와 이질성 추정에 과거 매출액과 영업이익이 중요한 역할을 하였음을 파악할 수 있다. 영업이익과 매출액을 제외하고 처치효과와 이질성 추정 시 중요했던 변수는 비록 비중은 작지만 산업중분류별 매출액(6.1%), 산업대분류(4.8%), 고정자산(5.4%), 시·군·구코드(4.6%) 등이 있다.

<표 V-5> 설명변수들의 트리성장과 관련한 중요도

변수명	변수레이블	종속변수	
		매출액	영업이익
X1	2013년 로그매출액	0.065	0.224
X2	2012년 로그매출액	0.040	0.140
X3	2011년 로그매출액	0.116	0.063
X4	2013년 영업이익	0.267	0.204
X5	2012년 영업이익	0.080	0.093
X6	2011년 영업이익	0.127	0.098
X7	산업중분류별 매출액	0.061	0.027
X8	r&d	0.034	0.040
X9	업력	0.034	0.012
X10	산업대분류	0.048	0.010
X11	사업체 형태	0.004	0.000
X12	고정자산	0.054	0.072
X13	지역코드	0.046	0.012
X14	사업체 규모	0.024	0.005

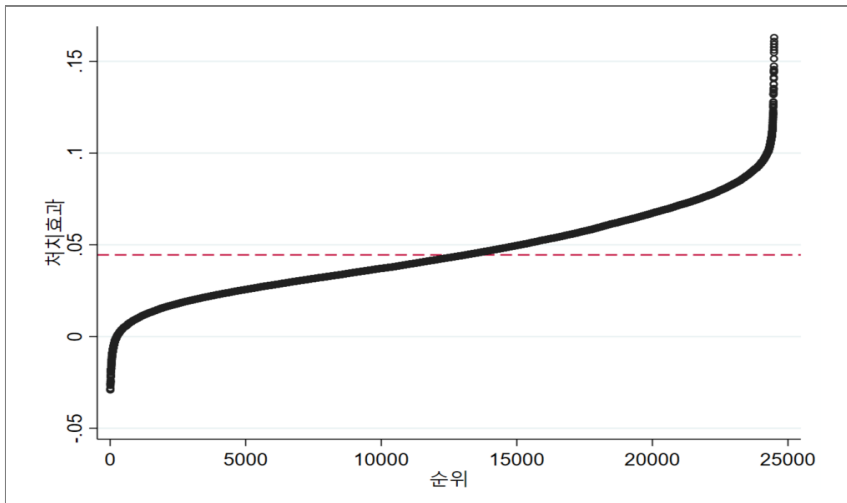
주: 1. 연도가 표시되지 않은 변수들의 경우 2013년 정보  
 자료: 「기업정보데이터」; 한국기업데이터 자료를 활용하여 저자 작성

한편 영업이익을 예측변수로 사용한 알고리즘의 경우 2013년 로그매출액과 2013년 영업이익이 트리를 성장시키는 데 가장 중요한 변수로 활용되었으며, 2012년 로그매출액, 2011년 영업이익, 2012년 영업이익 순으로 중요하다.

였다. 하지만 <표 V-3>의 모형검정에서 확인할 수 있었듯이 매출액을 예측 변수로 활용한 커절 포레스트만이 처치효과의 이질성을 통계적으로 유의하게 추정하는 것으로 나타났다. 따라서 본 분석에서는 매출액을 종속변수로 사용한 커절 포레스트에 한정하여 처치효과의 이질성을 살펴보았다.

[그림 V-7]은 2만 4,484개의 관측치를 포함한 테스트데이터를 앞서 추정된 커절 포레스트를 활용하여 관측치별 처치효과를 추정한 후 해당 처치효과를 오름차순으로 정리하여 제시한 것이다. 참고로 앞서 추정된 평균처치효과는 4.45%로 점선으로 표시되어 있는데 해당 선을 기준으로 왼쪽 끝부분에서는 부(-)의 처치효과가 발생하였고, 오른쪽 끝에는 10%가 넘는 처치효과가 발생하였음을 확인할 수 있다. 즉 평균처치효과와 비교하여 관측치의 특성에 따라 처치효과가 크게 추정이 되고 어떤 경우에는 작게 추정되었음을 알 수 있다.

[그림 V-7] 커절 포레스트를 활용한 신용보증기금 지원의 이질적 효과 추정



자료: 「기업정보데이터」; 한국기업데이터 자료를 사용하여 저자 계산

다음으로 관측치별 어떤 특성이 처치효과의 이질성과 관련 있는지 살펴보기 위해 추정된 처치효과를 설명변수별로 검토하였다. 이때 모든 변수를 검

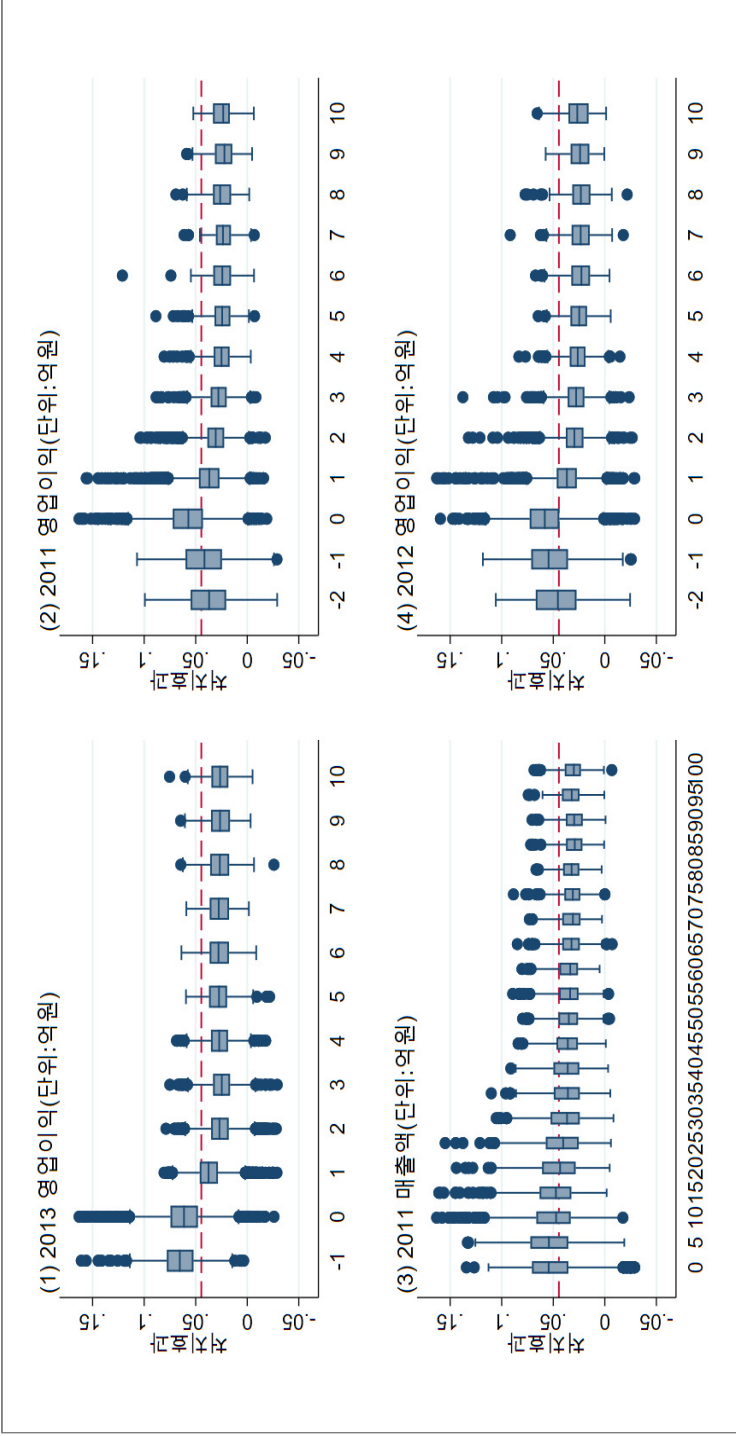
토하지 않고 트리를 성장시키는 데 중요 변수였던 2013년 매출액, 2011년 매출액, 2013~2011년 영업이익, 산업중분류 매출액, 산업대분류, 지역코드를 중심으로 설명변수의 값에 따라 추정된 처치효과가 어떻게 변화하는지를 살펴 보았다. 설명변수별 처치효과를 상자그림(box plot)을 활용하여 [그림 V-8]에서 제시하였다.

2013년 영업이익을 기준으로 처치효과를 살펴보면 영업이익이 ‘-1억~0억원’ 구간, 즉 해당 연도에 소규모 적자를 기록하였던 사업체들의 처치효과가 큰 것으로 나타났다. 또한 동일 구간에 포함되어 있는 사업체 중에서 처치효과 의 크기가 하위 25%에 해당하는 사업체들의 처치효과 역시 평균처치효과인 4.45%보다 큰 것으로 확인되어 해당 구간에 포함된 사업체들의 처치효과는 전반적으로 평균처치효과보다 크게 나타났다. 영업이익 ‘0억~1억원’ 구간에 속한 사업체들도 처치효과가 평균처치효과와 비교하여 큰 것으로 나타났는데 앞서와 동일하게 해당 구간의 처치효과가 하위 25%에 해당하는 사업체들의 처치효과도 평균처치효과보다 큰 것으로 나타났다. 하지만 이후 영업이익이 증가함에 따라 처치효과는 중위값을 기준으로 평균처치효과보다 작아지고 이후 처치효과는 일정한 값으로 추정되었다.

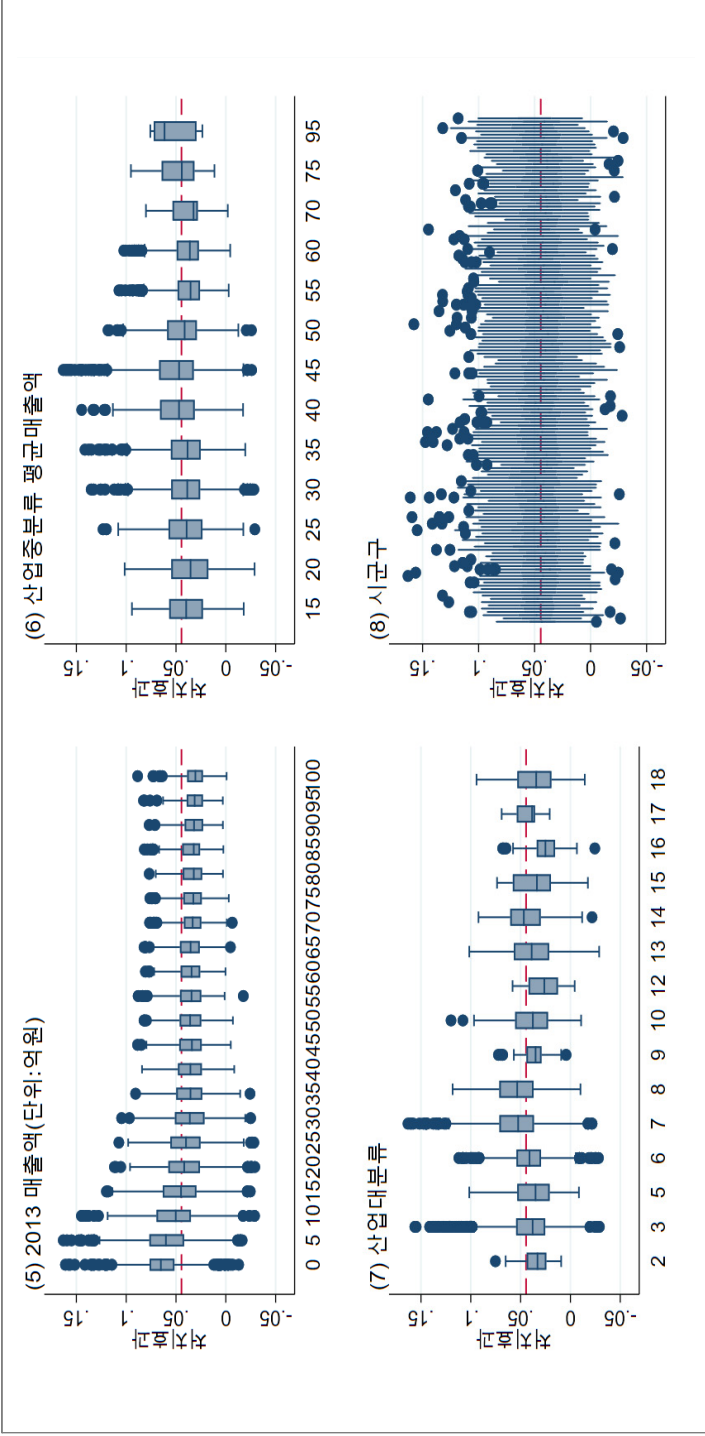
2011년 영업이익을 기준으로 사업체들의 처치효과를 살펴보면 영업이익이 ‘-2억~-1억원’, ‘-1억~0억원’, ‘0억~1억원’ 구간으로 이동함에 따라 처치효과는 증가하였고, 이후 구간에서 처치효과는 중위값을 기준으로 서서히 감소하여 일정한 값으로 추정되었다. 이어서 2012년 영업이익을 기준으로 처치효과를 살펴보면 2011년 영업이익별 처치효과에서 관측되었던 패턴과 유사한 패턴을 보이는 것을 확인할 수 있었다.

한편, 과거 매출액을 기준으로 처치효과를 살펴보면 2013년 매출액을 기준으로 매출액이 ‘0~5억원’, ‘5억~10억원’ 구간에 포함된 소규모 사업체는 해당 구간의 하위 25%에 해당하는 처치효과를 기준으로 평균처치효과보다 큰 것으로 나타났다. 하지만 매출액이 증가함에 따라 중위값을 기준으로 처치효과는 꾸준히 감소하여 ‘35억~40억원’ 구간부터 처치효과는 지속적으로 작은 값으로 추정되었다. 다만 2011년 매출액은 처치효과의 이질성을 설명하는 데

[그림 V-8] 설명변수별 처치효과와의 이질성



[그림 V-8]의 계속



자료: 「기업정보데이터」; 한국기업데이터 자료를 활용하여 저자 계산

큰 영향이 없는 것으로 나타났다. 마지막으로 산업중분류별 평균매출액, 산업대분류, 시·군·구에 따른 처치효과의 이질성을 살펴보면 해당 변수들과 처치효과는 큰 영향이 없는 것으로 판단된다. 단 산업대분류에서 7번(도매 및 소매업)과 8번(운수 및 창고업)에 속하는 사업체들의 처치효과는 중위값을 기준으로 다른 산업에 속한 사업체들과 비교해서 처치효과가 조금 큰 것으로 추정되었다.

[그림 V-8]의 분석에서는 설명변수별로 처치효과가 어떻게 추정되었는지를 시각화하여 제시하였다. 하지만 설명변수 간의 상관관계가 존재하기 때문에 상자그림 분석은 어떤 변수가 처치효과의 이질성에 중요한지 파악하기 어렵다. 예를 들어 영업이익과 업력의 상관관계가 매우 높다고 가정하였을 때 특정 영업이익 구간에서 처치효과가 큰 것으로 나타난 것이 영업이익에 의한 것인지 아니면 영업이익과 상관관계가 큰 업력의 차이에 의한 것인지 상자그림 분석을 통해서 구별하는 것은 불가능하다.

따라서 본 분석에서는 Gullen et al.(2020)이 제안한 방법을 차용하여 추가적인 분석을 실시하였다. 해당 방법은 일종의 비교정태접근법이라 할 수 있다. 왜냐하면 처치효과의 이질성에 특정변수의 역할을 살펴보기 위해서 다른 변수들의 값을 고정하기 때문이다. 구체적으로 본 분석에서는 2013년 영업이익에 따른 처치효과의 이질성을 살펴보기 위해 상자그림에서 사용한 동일한 데이터를 사용하되, 2013년 영업이익 외에 다른 변수들의 값은 해당 변수들의 중위값으로 대체하였다. 이렇게 새롭게 구축된 데이터를 활용하여 처치효과를 다시 계산하고 계산된 처치효과를 2013년 영업이익에 따라 도식화하였다. 이 경우 2013년 영업이익 외의 다른 변수들의 값은 일정하기 때문에 처치효과의 변동은 오로지 영업이익의 변화에서만 발생하게 되고 해당 변수의 변화가 처치효과의 크기에 어떤 영향을 끼칠 수 있는지 가늠할 수 있다.

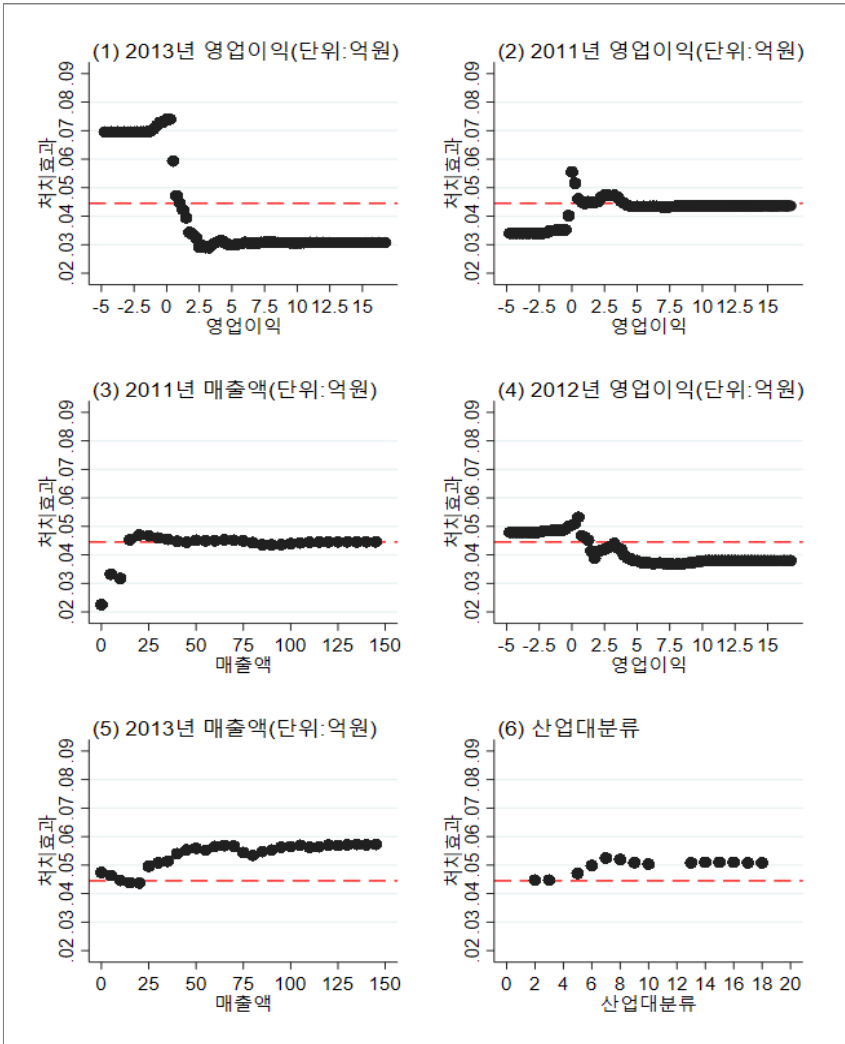
[그림 V-9]는 상자그림 분석에서 처치효과의 이질성에 중요한 역할을 할 것으로 보이는 6개의 설명변수(2013년 영업이익, 2011년 영업이익, 2011년 매출액, 2012년 영업이익, 2013년 매출액, 산업대분류)에 대한 비교정태분석 결과를 보여주고 있다. 여기서 X축은 각각의 설명변수이며 Y축은 추정된

처치효과이다. X축에 사용된 특정 설명변수를 제외한 나머지 변수들은 중위값으로 고정되어있기 때문에 특정 변수 값에 따라 하나의 처치효과를 가지게 된다.

먼저 2013년 영업이익을 살펴보면, 다른 설명변수들의 값이 해당 변수의 중위값으로 고정되었을 때 영업이익이 -5억원에서 0억원으로 증가할 때 처치효과는 약 6.95%에서 7.5%로 꾸준히 증가하였다. 하지만 영업이익이 5천만원 이상인 구간에서 처치효과는 급속히 감소하며 2.25억원 이상에서의 처치효과는 대략 3% 내외에서 안정적인 값을 유지하는 것으로 나타났다. 영업이익이 처치효과에 미치는 영향은 특정 구간에서 처치효과가 비선형적으로 변화하기 때문에 일반적인 선형회귀모형을 사용해서 처치효과를 분석하였다면 발견하기 어려운 패턴을 지닌 것으로 판단된다. 다음으로 2011년 영업이익과 2012년 영업이익의 변화에 따른 처치효과의 변화 양태를 살펴보면 2013년 영업이익에서 관측되는 패턴과는 다름을 알 수 있다. 다만 영업이익이 0억원 근방에서 처치효과가 큰 것으로 보이며, 이 부분에서는 2013년 영업이익이 처치효과에 미친 효과와 동일하였다.

2013년 매출액 변화에 따른 처치효과의 변화를 살펴보면 매출액 25억원을 기준으로 25억 이후 처치효과가 크게 추정되었다. 하지만 전 구간에서의 처치효과의 추정치가 평균처치효과보다 큰 것으로 나타나서 2013년 매출액의 변화가 처치효과의 이질성을 설명하지는 못하는 것으로 보인다. 상자그림 분석에서는 2013년 매출액이 10억원 이하인 사업체에서 처치효과들이 크게 관측되었는데, 비교정태분석에 의한 결과는 상자그림에서 살펴본 결과와 상반되었다. 이러한 차이는 처치효과의 이질성을 분석할 때 변수들 간의 상관관계를 통제하는 것이 중요함을 보여준다고 할 수 있다. 마지막으로 산업대분류에 따른 처치효과의 이질성을 살펴보면, 상자그림에서 확인하였듯이 도매 및 소매업과 운수 및 창고업에서 처치효과가 큰 것으로 확인되었다. 하지만 다른 산업대분류 등과 비교해서 눈에 띄는 차이는 관측되지 않아서 산업대분류 역시 처치효과의 이질성을 설명할 때 중요한 변수는 아닌 것으로 판단된다.

[그림 V-9] 예측변수별 처치효과도의 이질성: 비교정태접근법



자료: 「기업정보데이터」; 한국기업데이터 자료를 활용하여 저자 계산

한편 제Ⅲ장에서 신용보증기금 수혜효과의 이질성을 성향점수매칭을 활용하여 분석할 시 2013년 영업이익 중위값을 기준으로 이상과 미만으로 부 집단(subgroup)을 구성하고 이후 각각의 그룹에서 처치효과를 추정하였다.

추정 결과 영업이익 중위값 미만인 집단의 처치효과는 11.5%, 영업이익 중위값 이상의 집단의 처치효과는 4.0%로 추정되었다. 제Ⅲ장의 이질성 분석에서 중위값을 기준으로 집단 간의 처치효과 차이가 가장 큰 변수는 2013년 영업이익인 것으로 나타났으며 커절 포레스트를 활용한 이질성 분석에서도 2013년 영업이익 0억원을 기준으로 처치효과의 차이가 크게 나타났다. 즉 두 방법론에 의한 이질성 분석 결과는 일맥상통한다고 할 수 있다. 이러한 결과는 언뜻 생각하면 전통적인 선형모형을 활용한 이질성 분석 역시 처치효과 이질성을 설명할 수 있는 변수들을 잘 포착하는 것처럼 보일 수 있다. 하지만 이는 우연의 결과로 처치효과가 큰 것으로 나타나는 영업이익 1억원 이하를 기준으로 50% 사업체가 포함되었기 때문이다. 만약 영업이익 1억원 이하에 포함되는 사업체가 10%에 불과하였다면 중위값을 기준으로 부집단을 생성하여 처치효과의 이질성을 가늠하는 전통적인 방법은 영업이익에 따른 처치효과의 이질성을 포착하기 어려웠을 것으로 보인다.

물론 본 연구의 경우 기타 정책금융 지원 여부를 고려하지 못하고 신용보증기금 지원 이력만을 활용하여 정책을 평가하였기 때문에 해당 연구에서 파악한 처치효과의 이질성을 기반으로 강한 정책제언을 하는 것은 무리가 있다. 하지만 본 연구의 목적이 정책금융 정책 설계 측면에서 머신러닝 알고리즘의 활용 가능성의 탐색에 있기 때문에 이러한 한계점이 큰 문제는 아니라고 판단된다. 만약 정책연구자들이 사업체들의 정책금융 지원 이력을 정확히 파악하는 것이 가능하다면, 본 연구에서 소개한 머신러닝 알고리즘을 활용해서 연구자들이 생각하는 정책목표를 달성하기 위한 효과적인 정책금융 지원 정책을 설계할 수 있을 것으로 보인다. 예를 들어 정책금융의 지원 목표가 정책금융을 통한 사업체의 성장(매출액)에 있다면, 매출액을 종속 변수로 하는 커절 포레스트를 활용하여 처치효과의 이질적인 패턴을 파악하고 파악된 처치효과의 이질성에 따라 효과가 클 것으로 기대되는 대상을 정책 수혜자로 선정하게끔 정책을 재단한다면 더욱 효과적인 정책 집행을 실행할 수 있을 것이다.

마지막으로 이질성 분석에서 커절 포레스트의 또 다른 장점을 논의하면 연

구윤리 측면에서도 머신러닝 알고리즘을 활용해 처치효과의 이질성을 추정하는 것이 전통적인 이질성 분석 방법과 비교해 장점을 지닌다. 만약 처치효과의 이질적인 패턴에 대해 연구자가 경제이론 등에 기반하여 미리 가정하지 않고 블랙박스로 생각하고 있다면, 연구자들은 전통적인 선형회귀모형을 활용하여 처치효과의 이질성을 분석할 때 처치효과가 큰 집단을 포착하기 위해서 수많은 기준에서 모형탐색(model specification search)을 시도하게 된다. 이 경우 처치효과가 클 것으로 여겨지는 집단은 해당 표본에서의 과적합 문제로 인하여 처치효과가 크게 추정되는데, 이로 인해 다른 표본에서는 해당 처치효과를 찾을 수 없는 경우가 발생하게 된다(정재현 · 이환용, 2020).<sup>44)</sup> 물론 커질 포레스트 등의 알고리즘도 처치효과가 큰 변수를 찾기 위해 일종의 모형탐색을 하는 것으로 생각할 수 있다. 하지만 본 분석에서 사용한 Wager and Athey(2018)의 어니스티 접근법은 알고리즘의 훈련과 처치효과의 추정을 분리해서 모형의 과대적합을 방지해 주기 때문에 모형탐색에 따른 과적합 문제를 완화해 준다. 따라서 연구자들이 처치효과의 이질성에 대한 사전지식이나 명확한 이론을 가지지 못하였다면 전통적인 방법론을 활용하는 것보다 커질 포레스트를 활용하는 것이 연구윤리 측면에서도 보다 나은 대안이 될 것으로 판단된다.

#### 4. 머신러닝을 활용한 기업성과 예측

본 절에서는 정책금융을 지원받은 기업 및 지원받지 못한 기업을 대상으로 사전적으로 해당 기업들의 성과를 예측할 때 머신러닝 알고리즘을 통해 예측성고를 향상시킬 수 있는지를 탐색하였다. 구체적으로 앞서 선행연구에서 소개한 McKenzie and Sansone(2019)과 유사하게 신용보증기금을 수혜 받은 기업과 수혜받지 못한 기업으로 구분하여 각각의 집단에서 기업성고를 다양한 기업 관련 변수 등을 통해서 예측하였다.

44) 이러한 과적합 문제를 최소화하기 위해 전미경제학회와 많은 의학학회에서는 연구자들에게 연구 사전에 계획서를 제출하게 하고 계획서상에서 탐색하고자 한 이질성 분석에 한해서만 이질성 분석을 허용하고 있다.

McKenzie and Sansone(2019)의 분석은 커절 포레스트를 사용하여 정책 금융 수혜효과의 이질성을 살펴본 것과 비교하여 정책금융의 효과성을 향상시키는 데 제한적인 정보를 제공해 준다. 구체적인 설명을 위해 가상의 예시를 들면 정책금융의 수혜기업들의 결과값을 예측하는 모형의 성과가 머신러닝을 통해 향상되고 머신러닝에 기반한 예측모형에서 직전 연도 매출액이 해당 기업 성과 예측에 중요성을 갖는 경우를 생각할 수 있다. 이 경우 성과가 좋을 것으로 예측된 수혜기업은 정책금융의 수혜 여부와 관련 없이 직전 연도 매출액이 기업성과와 관련이 있어서 성과가 좋을 것으로 예측되었거나 혹은 직전 연도 매출액이 처치효과의 이질성과 깊은 연관이 있어서 성과가 좋을 것으로 예측되었을 것이다. 즉 기업의 성과가 좋을 것으로 예측되었다 하더라도 해당 기업의 성과가 정책금융에 의한 것인지 알 수 없게 된다.

이러한 한계점에도 불구하고 본 절에서는 머신러닝 알고리즘을 통해 기업 성과의 예측을 향상시킬 수 있는지 검토하였는데, 이는 선형회귀분석과 머신러닝의 성과를 정량화하여 비교하기 위함이다. 반면에 정책효과의 이질성을 머신러닝을 통해 예측한 경우 실제 처치효과를 관측할 수 없기 때문에 전통적인 방식과 비교하여 머신러닝의 성과를 정량화하여 파악하기가 어렵다. 따라서 본 절에서는 머신러닝을 활용하여 기업의 성과를 예측하고 전통적인 선형회귀모형과 예측성과를 비교하였다. 구체적으로 매출액, 영업이익, 고용량 등을 종속변수로 사용하고 NICE평가정보(주)에서 제공하는 「기업정보 데이터」를 활용하여 예측모형을 구축하고, 선형회귀모형과 머신러닝 알고리즘 간의 예측성과를 비교하였다.

단 상술했듯이 이와 같은 분석은 정책금융의 효과성 개선에 제한적인 정보를 제공하기 때문에 본 절에서는 머신러닝 알고리즘의 예측성과가 선형회귀분석과 비교하여 향상될 수 있는지 측면에서 검토하였다. 반면에 성과 예측에 어떤 변수가 중요하였는지 그리고 예측변수들의 값이 변화함에 따라 예측된 성과는 어떻게 변화하는지 살펴보는 것은 본 절의 분석 취지를 고려했을 때 중요성이 떨어지기 때문에 해당 결과는 부록에 간략하게 제시하였다.

## 가. 데이터

본 절에서는 머신러닝 알고리즘을 사용하며 미래의 실적(매출액/법인세차감전계속이익) 및 고용량을 예측하였는데, NICE평가정보(주)에서 제공하는 기업정보자료를 예측변수(설명변수)로 활용하였다. 구체적으로 신용보증기금 수혜 직전 3개년도의 기업 실적데이터 및 업체 현황 자료를 활용하되, 2014년도 신용보증을 지원받은 기업 집단과 2014년도 및 2015년도에 신용보증을 지원받지 못한 기업 집단으로 나누어 예측을 실시하였다.

모형구축에 사용된 NICE평가정보(주)에서 제공하는 기업정보데이터는 15개 파일을 포함하고 있는데, 활용한 데이터 리스트는 <부표 3>에서 제시하였다. 해당 자료에서 사용하지 못한 자료들을 언급하면 먼저 aa04(주식소유형태별 분포) 및 aa05(주식소유규모별 분포)는 구분별 주주수, 지분율 등을 포함하고 있는데, 해당 정보를 활용할 수 있는 기업의 수(총 업체 수 99개)가 적기 때문에 해당 자료는 사용하지 못하였다. 마찬가지로 aa18(주요제품가격동향) 및 aa19(원재료가격동향) 역시 각각 주요 제품의 가격 및 원재료 가격을 포함하고 있으나, 포함하고 있는 업체 수가 적어서(총 업체 수 81개) 예측변수로 활용하지 않았다. 마지막으로 aa20(영업소 및 현황), aa99(폐업일정보), ab00(재무KEY정보), az07(산업코드), az14(계열별 업체코드)는 기업체가 아닌 영업소들의 현황을 포함하는 자료이기 때문에 분석에 활용하지 않았다.

예측변수로 활용한 자료들을 살펴보면 먼저 aa17(생산 및 판매현황)은 생산실적금액, 생산실적수량, 수출금액, 수출수량, 내수금액, 내수수량, 총판매금액, 총판매수량 등의 정보를 포함하고 있는데 업체마다 품목이 다르기 때문에 수량 단위에 통일성이 없어 수량 데이터는 사용하지 않았으며, 생산실적금액, 수출금액, 내수금액, 총판매금액 등은 예측변수로 활용하였다(총 업체 수 1만 4,932개). aa21(장단기차입금 현황)은 구분별(당좌/단기/장기/외화 등) 차입금 등을 포함하고 있으며, 해당 정보를 포함하고 있는 기업체가 1만 4,932개이기 때문에 예측변수로 활용하였다.

ab08(약식재무제표)은 매출액, 순이익, 총자산, 부채총계, 자본금, 법인세비용차감전계속이익, 감사의견 등을 포함하고 있으며, 해당 정보를 포함하고

있는 총 업체 수는 3만 4,186개였다. 해당 데이터는 시점별 데이터로 연도 및 기업 단위에서 24만 7,504개의 데이터가 존재하기 때문에 예측모형에서 핵심적인 자료로 활용하였다. 마지막으로 az12(업체별 산업코드)는 기업의 산업코드 정보를 포함하고 있으며 총 3만 7,589개 업체의 정보를 제공하기 때문에 예측변수로 분석에 사용하였다.

예측변수로 활용할 변수들의 구축이 마무리되면 분석에 앞서 데이터의 전처리 과정(Preprocessing)이 필요한데 다음 세 가지 처리를 실시하였다. 먼저 커질 포레스트의 분석과 유사하게 직전 3개년 실적자료를 활용하여 성과를 예측하기 때문에 3개년 실적자료가 존재하지 않는 기업들은 데이터 분석 대상에서 제외하였다. 그다음 이상값이 모형 예측에 미치는 영향을 최소화하기 위해서 예측 모형마다 예측치 상위 1% 및 하위 1%를 포함하는 기업들의 데이터를 제거하였다. 예를 들면 매출액을 예측 대상으로 하는 경우 매출액 상위 1% 이상인 기업들과 매출액 하위 99% 이하인 기업들을 훈련 및 실험데이터에서 제거하였다. 마지막으로 모든 예측변수들의 크기를 일정하게 해서 변수의 크기가 미치는 영향을 최소화하기 위해 정규화(Normalization) 과정이 필요한데 본 분석에서는 최댓값을 1 최솟값을 0으로 하는 Min-Max Normalization을 수행하였다.

## 나. 머신러닝 알고리즘 및 실험방법

머신러닝 알고리즘과 선형회귀모형의 예측성과를 비교하기 위해 OLS, Boosting, 랜덤포레스트 알고리즘을 활용하였다. 먼저 벤치마크 케이스로 활용되는 선형회귀모형(OLS)은 모든 예측변수들을 설명변수로 활용하였다. 그다음 부스팅(Boosting)은 앞서 설명한 랜덤포레스트와 유사한 앙상블 기법 중 하나인데, 랜덤포레스트와 다르게 약한 학습기(weak learner)를 순차적으로 적용해 다음 단계의 학습기가 이전단계의 오류를 교정해 나가는 기법이다.<sup>45)</sup> 부스팅 알고리즘에는 Adaboost, Gradient Boosting, XGBoost 등이 존

---

45) Boosting에 대한 자세한 내용은 Friedman et al.(2001)의 pp. 299~302를 참조하라.

재하다.<sup>46)</sup> 본 분석에서는 분석에 따른 수행시간 등을 고려하여 Logit을 기본 추정치로 하는 Adaboost 알고리즘을 사용하였다.

사용할 알고리즘이 정해지면 훈련데이터에서 알고리즘을 적합하며 실험 데이터에서 성과를 평가하는데, 본 분석에서는 머신러닝에서 일반적으로 사용하는 5-fold cross validation 방법론을 사용하였다. 구체적으로 각 알고리즘의 성과를 살펴보기 위해, 전체 데이터를 임의적으로 5개의 폴드(fold)로 쪼갠 후에 그중 1개의 폴드는 실험데이터로 사용하고 그 외 나머지 4개는 훈련데이터로 사용하는 실험을 다섯번 수행하였다. 즉 모든 폴드는 총 4회에 걸쳐 훈련데이터로 사용되며, 단 1회만 실험데이터로 사용된다. 최종적으로 산출되는 모형의 성과는 RMSE(Root Mean Squared Error)이며, 5회의 실험 데이터에서 계산된 RMSE의 평균값을 통해서 구해진다.

한편 모형 구축 시 2014년 신용보증을 지원받은 기업과 2014~2015년 모두 신용보증을 지원받지 않은 기업으로 나누어서 구축하였으며 매출, 영업이익(법인세 차감 전 계속이익), 연간고용 예측을 위해 각각을 종속변수로 하는 모형을 별도로 구축하였다. 마지막으로 기업의 과거 실적데이터를 풍부하게 활용하는 것이 모형 성과에 영향을 미치는지 살펴보기 위해 ① 직전 3개년도의 실적데이터를 독립변수로 사용했을 경우 ② 직전 2개년도의 실적데이터를 독립변수로 사용했을 경우 ③ 직전 연도 데이터만을 독립변수로 사용했을 경우로 나누어 모형의 성과를 살펴보았다.

#### 다. 알고리즘별 성과 비교

직전 3개년도 실적데이터를 사용하였을 때 알고리즘별 퍼포먼스는 <표 V-6>에서 제시하였다. 참고로 모형 성과는 RMSE(Root Mean Squared Error)로 측정되었기 때문에 해당 값이 낮을수록 모형의 성과는 높아진다. 먼저 매출액을 예측하는 모형을 살펴보면, 신용보증을 지원받지 않은 기업들의

---

46) Adaboost와 Gradient Boosting의 가장 큰 차이는 예측오류를 개선하는 과정에서 어떤 관측치에 비중을 많이 부여하는지 여부이다. Adaboost의 경우 예측을 적절하게 수행하지 못한 관측치에 가중치를 보다 많이 부여하는 방법이다.

성과를 OLS 알고리즘을 활용하여 추정하면 RMSE는 0.1421인 것으로 나타났다. 반면에 AdaBoost와 랜덤포레스트 알고리즘을 활용하면 RMSE는 각각 0.0962, 0.0702를 기록하여 선형회귀모형을 활용한 예측과 비교하여 예측오차는 각각 32.30%, 50.60% 감소하는 것으로 나타났다. 반면에 신용보증 지원을 받은 기업을 대상으로 차년도 매출액을 예측한 경우 OLS와 비교하여 Adaboost는 예측오차가 증가하였으며 랜덤포레스트는 예측오차가 감소하였으나 감소폭은 적은 것으로 나타났다.

한편 영업이익과 고용을 예측대상으로 구축된 모형의 경우 OLS와 비교하여 머신러닝 알고리즘의 성과가 우수하다는 증거는 찾을 수 없었다. 즉 머신러닝의 예측성과는 무엇을 예측대상으로 하는지 그리고 어떤 집단을 대상으로 예측하는지에 따라 다를 수 있음을 확인하였다.

〈표 V-6〉 알고리즘별 예측 성과 비교: 직전 3개년도 실적데이터 사용

종속변수(금액기준)		매출	영업이익 <sup>1)</sup>	고용
신용보증 지원받지 않은 기업	OLS	0.1421	0.0230	0.2367
	AdaBoost	0.0962	0.0228	0.2324
	랜덤포레스트	0.0702	0.0204	0.2148
신용보증 지원받은 기업	OLS	0.0913	0.0243	0.2446
	AdaBoost	0.1137	0.0284	0.2501
	랜덤포레스트	0.0822	0.0225	0.2370

주: 1) 영업이익은 법인세 차감 전 계속이익을 뜻함  
 자료: NICE평가정보(주), 「기업정보데이터」를 활용해 알고리즘별 예측 성과를 추정함

한편 〈부표 5〉와 〈부표 6〉은 직전 2개년도 실적데이터 및 전년도 실적데이터만을 사용하였을 때의 알고리즘의 성과를 제시하였는데, 흥미롭게도 모형의 성과는 기업의 과거 실적데이터를 풍부하게 활용하는 것과 큰 영향이 없는 것으로 나타났다. 이에 대한 원인은 〈부표 4〉에서 확인할 수 있다. 〈부표 4〉는 매출액을 예측할 때 설명변수들의 중요도<sup>47)</sup>를 보여주는데, AdaBoost와

47) 본 분석에서는 Schonlau(2005)가 제시한 방법 중 특정변수가 예측에 끼치는 영향을 계산하는 방식을 활용해 중요변수를 검토하였다. Boosting을 최적화할 때 직관적으로 모든

랜덤포레스트 모두 직전 연도 매출액이 예측에 가장 중요한 것으로 나타났으며 해당 변수의 중요도는 상위 10개 변수 중에서 차지하는 비중이 88.85% 및 83.11%를 기록하였다. 반면 과거 2년 전과 3년 전 매출액의 중요도는 랜덤포레스트를 기준으로 각각 4.33%, 4.82%를 기록하였다. 즉 전년도 매출액이 차년도 매출액을 예측하는 데 다른 변수들과 비교하여 매우 중요하기 때문에 과거의 자료를 풍부하게 활용하는 것이 모형의 성과로 이어지지 않게 되는 것이다.

이상의 분석 결과를 종합하자면, 기업의 매출액을 예측할 때 머신러닝 알고리즘의 사용은 기존의 선형회귀모형을 사용하는 경우와 비교하여 예측오차를 최대 50% 이상 줄일 수 있는 것으로 나타나 기업성과의 예측에 머신러닝의 활용 가능성을 확인하였다. 이는 머신러닝 알고리즘을 사용하여 기업성과를 예측한 선행연구의 결과와 유사한 것으로 판단된다(Coad and Srhoj, 2019; Song et al., 2018; Miyakawa et al., 2017). 특히 Miyakawa et al.(2017)에서 분석한 것처럼 신용평가기관에서 사용하는 기업평가등급과 머신러닝 알고리즘의 성과를 비교 분석하는 것도 흥미로운 것으로 보이나 본 연구의 분석 범위를 벗어나기 때문에 해당 연구는 후속연구로 남겨 두고자 한다.

---

나무에서 특정 변수가 알고리즘에 의해 몇 번 선택되는지, 그리고 선택된 변수가 log-likelihood에 끼치는 영향이 어느 정도인지에 따라 중요도가 결정된다고 할 수 있다. 본 분석에서는 중요도 계산 시 특정변수의 포함 여부가 log-likelihood에 미치는 영향을 통해서 추정하였다. 구체적으로 각각의 설명변수들을 모형에 포함시킨 경우와 불포함시킨 경우의 RMSE의 차이를 통해서 추정하였으며, 동 테스트를 반복적으로 수행해 성과 차이의 평균값이 가장 큰 상위 10개 예측변수를 제시하였다.

---

## VI. 요약 및 정책적 시사점

---

### 1. 연구 결과 요약 및 한계

본 연구에서는 머신러닝 방법론이 정책금융의 효과를 추정하고 정책수혜자를 선정하는 데 있어 정책효과가 큰 집단을 식별하는 데 효과적인지 여부를 선행연구에서 많이 사용한 성향점수매칭 방법과 비교를 통해서 검토하였다. 본 연구의 기여는 정책금융 등 다양한 재정사업의 처치효과 추정 및 처치효과와 이질성 추정에 머신러닝 알고리즘의 사용 가능성을 검토하고 유용성을 확인하여 향후 재정사업의 성과평가 및 설계를 위한 새로운 기반을 마련했다는 것이다.

본 연구의 머신러닝 알고리즘을 활용한 처치효과 추정은 효과성 평가에 있어 머신러닝 알고리즘이 기존 방법론의 한계점을 보완할 수 있음을 보여주고 있다. 성향점수매칭과 커질 포레스트를 활용한 수혜효과 추정 결과는 질적으로 유사한 것으로 나타나 상호 보완적으로 사용할 수 있음을 확인하였다. 다만 성향점수매칭은 성향점수 추정을 위한 로짓모형의 변수 선택에 따라 정책효과와 추정 결과가 크게 의존하는 현상이 관측되었다. 즉 성향점수에 의한 추정치는 변수 선택에 취약한 특징이 있음을 의미하며, 커질 포레스트가 이러한 기존 통계모형의 취약성을 극복할 수 있는 좋은 대안이 될 수 있음을 보여주는 결과라 판단된다.

또한 본 연구에서는 성향점수매칭 방법론과 커질 포레스트를 통해 정책효과와 이질성을 분석하고 두 방법론을 비교하였다. 성향점수매칭 방법론을 활용한 정책효과와 이질성은 독립변수들의 중위값을 기준으로 부집단(subgroup)으로 구성한 후 각 집단의 정책의 효과를 분석하였으며, 분석 결과 사업체의 업력이 낮을수록, 매출액이 낮을수록, 영업이익이 낮을수록 매출액에 미치는

정책금융의 효과가 긍정적인 방향으로 더 큰 것으로 나타났다. 한편 커절 포레스트를 사용하는 경우에는 성향점수매칭 방법론보다 더욱 세밀한 이질성 분석이 가능함을 확인하였다. 매출액에 대한 이질적인 처치효과를 크기 순으로 나열했을 때 처치효과 크기 기준 상위 10%에 속하는 사업체들의 평균적인 처치효과는 8.82%, 하위 10%에 속하는 사업체들은 0.97%로 나타나 신용보증기금의 정책효과가 매우 이질적인 것으로 나타났다. 이러한 결과는 만약 본 연구의 분석 결과에 기반해 가장 낮은 처치효과를 가진 그룹들에서 정책금융 지원을 중단하고 가장 높은 처치효과를 가진 그룹에 정책금융을 제공하는 방법으로 정책금융을 타게팅하였다면 정책금융을 통한 사업체의 기대되는 매출액 증대효과를 향상시킬 수 있음을 의미한다. 그리고 커절 포레스트에 의한 이질성 분석은 매출액 증대효과가 사업체의 과거 영업이익의 크기의 변화에 따라 매우 크게 비선형적으로 변화함을 확인하였다. 전통적인 처치효과의 이질성 추정에서는 발견하기 어려운 이러한 비선형적 패턴의 발견은 해당 정보를 통해 수혜자를 선정하는 경우 정책금융의 효율성을 향상시킬 수 있을 뿐만 아니라 처치효과가 나타나는 과정에 대한 메커니즘을 이해하는 데 도움이 될 것으로 판단된다.

물론 본 연구의 분석에 사용한 자료는 신용보증기금 지원 이력 외에 기타 정책금융 지원 여부를 고려하지 못하기 때문에 본 연구에서 파악한 처치효과와의 이질성을 정책금융 지원대상자 선정에 즉각적으로 활용하는 데 한계가 존재한다. 하지만 본 연구의 목적이 정책금융의 수혜자 선정 측면에서 머신러닝의 활용 가능성을 탐색하는 데 있기 때문에 이러한 한계점은 구체적인 자료가 보강된다면 개선 가능한 한계점으로, 본 연구의 본질적인 시사점에는 영향을 미치지 않는다고 판단된다. 예를 들어 각 보증기금공단의 정책금융 지원이력 및 '중소기업 지원사업 통합관리시스템' 등을 활용해 중소기업 지원 금융·재정사업의 중복수혜 여부 등을 정확히 파악하는 것이 가능한데, 이 경우 본 연구에서 소개한 머신러닝 방법론을 사용한다면 분석 결과를 기반으로 구체화된 정책설계 방안을 도출할 수 있을 것으로 기대된다.

## 2. 정책적 시사점

본 연구는 머신러닝 방법론을 활용하여 정책금융의 효과를 추정하고 정책 수혜자를 선정하는 데 있어 정책효과가 큰 집단을 식별하는 데 머신러닝이 효과적인지 여부를 검토하였고, 이 점에서 머신러닝의 유용성을 확인하였다. 이러한 연구 결과의 정책적 시사점은 크게 두 가지로 제시할 수 있다. 첫째, 신용보증기금, 기술보증기금 등 중소기업을 대상으로 하는 정책금융 선정 과정에 머신러닝 방법론을 활용하는 것이며, 둘째, 비단 정책금융 분야에 국한되지 않고 향후 정부의 재정사업 효과성 평가에 머신러닝 방법론을 보완적으로 사용하는 것이다.

### 가. 정책금융 대상자 선정

신용보증기금의 『보증심사 운용요령』(2021)에 따르면 일반심사는 공통적으로 ① 연체대출금 보유 현황 ② 당좌부도, 신용관리등급 보유 및 신용보증부실 발생 현황 ③ 국세 및 지방세 체납 현황 ④ CB 연체 및 보험료 체납정보 보유 현황 ⑤ 매출액 대비 총차입금 및 운전자금차입금 규모 ⑥ 매출액 변동추이 및 자기자본 규모 ⑦ 사업장 및 거주주택에 대한 권리침해 현황 ⑧ 대표자 및 실제경영자 변동 내용 항목 등을 심사하며 전문심사는 전문심사위원회에서 경쟁력·수익성·성장성 등을 종합적으로 검토하여 보증취급 여부를 판단한다.<sup>48)</sup> 즉 보증심사 단계에서 지원업체의 재무적 정보와 비재무적 정보를 종합적으로 활용하여 심사하고 있음을 의미한다. 비재무적 정보 및 재무적 정보를 종합적으로 고려하는 것은 외부감사를 받지 않는 사업체가 중소기업의 대다수를 이루고 있는 점을 고려한다면 중소기업을 대상으로 하는 정책금융의 특수성을 반영하고 있는 것으로 판단된다.

다만 보증심사의 기준이 기금 건전성 측면에서 수혜기업들이 지원 이후에

---

48) 신기철(2013)은 중소벤처기업진흥공단의 정책자금 선정을 위한 의사결정을 연구하였는데, 중소기업의 비재무적 요소가 신용등급, 기술등급 등과 더불어 선정에 중요한 요소인 것으로 나타났다.

정책자금을 얼마나 애초의 목적대로 잘 활용하는지, 그리고 수혜기업들을 대상으로 자금회수가 얼마나 잘 이루어질지에 집중되어 있어 심사기준의 다양성 측면에서 미흡한 것으로 판단된다. 장우현 외(2013)가 중소기업 지원 정책을 생산적 중소기업 지원정책과 공공부양적 중소기업 정책으로 명시적으로 분류한 것과 유사하게 정책금융 정책도 생산성 중소기업을 위한 정책과 공공부양적 중소기업 정책으로 분류하여 보증심사 기준을 이에 걸맞게 보완하는 것이 바람직해 보인다.

정책금융 지원사업의 일부분을, 강소기업으로 성장 가능성이 있으나 시장 실패로 성장에 애로를 겪고 있는 사업체에 지원하는 것으로 운영한다면 본 연구에서 살펴본 머신러닝 알고리즘은 해당 지원사업의 효과성을 향상시킬 수 있을 것으로 기대된다. 예를 들어 강소기업의 기준이 매출액 혹은 영업이익의 성장률이라면 머신러닝 기법을 활용하여 사전적으로 처치효과가 클 것으로 기대되는 사업체를 식별하고 정책효과가 클 것으로 기대되는 기업군에 지원한다면 생산적 중소기업 지원정책의 관점에서 정책의 효과성은 향상될 것으로 보인다. 물론 정책효과가 뚜렷할 것으로 보이는 기업군에만 지원이 집중될 경우 정책금융의 본연의 목적성 달성이 저해될 가능성이 있다는 것도 염두에 두어야 하며, 그렇기에 더욱이 정책금융 지원사업을 생산적 기업을 위한 사업과 중소기업의 복지를 위한 사업으로 이원화하여 운용하는 것이 필요해 보인다.

머신러닝을 정책금융 대상자 선정에 활용할 때 발생할 수 있는 또 다른 문제점은 머신러닝 방법론에 의한 예측결과를 명확하게 설명하기 어렵다는 점이다. 예를 들어 [그림 V-7]의 결과에 기반해 처치효과가 클 것으로 기대되는 상위 10% 사업체들에 신용보증을 확대하는 정책을 생각할 수 있는데, 해당 사업체들의 처치효과가 어떤 요인에 의해서 클 것으로 기대되는지 파악할 수 없다면 지원 대상에서 탈락한 사업체들의 반발이 거세질 수 있고 사회적 갈등을 유발할 수 있음도 명심해야 한다. 그렇기에 머신러닝에 의해 추정된 처치효과와 발생 원인이 설명변수 중 어떤 변수에 의한 것인지 면밀히 살펴본 본 연구의 분석은 머신러닝 알고리즘을 대상자 선정에 활용함에

따라 유발될 수 있는 위험을 최소화하는 데 기여할 수 있을 것으로 판단된다.

#### 나. 재정사업 효과성 평가

「국가재정법」 제8조 제6항은 정부의 주요 재정사업에 대한 평가를 실시하고 그 결과를 재정운용에 반영할 수 있도록 하고 있다. 이러한 성과중심의 재정운용이 취지에 맞게 운용되기 위해서는 무엇보다도 효과성 평가가 신뢰성 있게 수행되어야 한다. 본 연구에서 검토한 것처럼 머신러닝 알고리즘은 기존의 효과성 평가에 사용되는 방법론의 한계점을 보완할 수 있기 때문에 재정사업 효과성 평가에 머신러닝 알고리즘을 적극적으로 활용하는 것도 필요할 것으로 보인다.

현재 재정사업의 효과성 평가는 재정사업 심층평가를 통해 일부 사업의 평가가 진행되고 있는데 「재정사업 심층평가 운용지침」 제13조에 따르면 심층평가 수행기관은 심층평가 수행지침에 따라 평가를 수행해야 한다. 따라서 재정사업 평가 시 머신러닝 알고리즘을 보다 적극적으로 활용할 공간을 넓히는 차원에서 효과성 평가 관련 수행지침에 머신러닝 알고리즘을 활용한 정책 평가를 준실험적 방법론과 함께 권고하는 방안을 고려할 수 있다. 이를 통해 참여 연구자들이 재정사업 효과성 평가에 머신러닝 알고리즘을 보다 적극적으로 도입하도록 장려할 수 있을 것으로 기대된다. 물론 머신러닝 방법론을 의미 있게 활용할 수 있는 자료 제공이 동반되어야 머신러닝 방법론의 활용에 따른 실익이 수반될 것이다. 재정사업 효과성 평가에 머신러닝 방법론의 사용은 재정사업 효과성 평가의 신뢰성을 향상시킬 수 있을 것으로 보이며, 이에 따라 재정사업 심층평가의 환류기능도 강화될 수 있을 것으로 기대된다.

---

## 참고문헌

---

- 권세훈, 「기술신용보증 기업의 특성과 성과에 관한 실증분석」, 『산업경제연구』, 제25권 제3호, 2012, pp. 2069~2087.
- 금융감독원, 『중소기업 금융지원제도 안내』, 2021.
- 기술보증기금, 『2019 기술보증기금 연차보고서』, 2020.
- 김상봉·김정렬, 「중소기업 신용보증지원 성과분석에 관한 연구」, 『산업경제연구』, 제26권 제3호, 2013, pp. 1381~1399.
- 노용환, 「중소기업 정책자금의 미시적 성과 분석과 역할에 관한 연구」, 『중소기업연구』, 제32권 제1호, 2010, pp. 153~175.
- 박노옥·강희우·이상현·표한형, 『중소기업 지원 재정사업 성과평가』, 한국조세재정연구원, 2016.
- 손상호, 「한국 정책금융의 평가와 분석 및 미래비전」, 『KIF 정책보고서(중간)』, 2013년 6월호, 2013, pp. 1~144.
- 신기철, 「중소기업 정책자금 대출을 위한 의사결정의 정확률 추정에 대한 연구」, 『행정논총(Korean Journal of Public Administration)』, 제51권 제2호, 2013, pp. 209~227.
- 신상훈·박정희, 「신용보증지원이 중소기업의 수익성과 성장성에 미치는 효과에 대한 패널분석」, 『중소기업연구』, 제32권 제1호, 2010, pp. 43~64.
- 신용보증기금, 『2015년 신용보증기금 연차보고서』, 2015.
- \_\_\_\_\_, 『2016년 신용보증기금 연차보고서』, 2016.
- \_\_\_\_\_, 『2017년 신용보증기금 연차보고서』, 2017.
- \_\_\_\_\_, 『2018년 신용보증기금 연차보고서』, 2018.
- \_\_\_\_\_, 『2019년 신용보증기금 연차보고서』, 2019.
- \_\_\_\_\_, 『보증심사 운용요령』, 2021.

- 우석진·빈기범, 「불완전 금융시장하에서 기술보증을 통한 정책금융이 중소기업 성과에 미치는 효과-분위회귀분석에 의한 처리효과 추정방법론의 응용」, *Journal of The Korean Data Analysis Society*, 제16권 제1호, 2014, pp. 279~295.
- 이기영·우석진, 「공적 신용보증이 한계기업의 생존에 미치는 효과」, 『재정학 연구』, 제8권 제4호, 2015, pp. 71~90.
- 이상엽·강경훈·남종오·박창균·송헌재·이경원·이기영·이명현, 『중소기업 금융지원 사업군 심층평가』, 한국조세재정연구원, 2014.
- 이준원, 「기술금융 중소기업과 일반 중소기업의 경영성과 비교분석-기술신용 대출을 받은 기술금융 중소기업을 중심으로」, 『한국혁신학회지』, 제14권 제1호, 2019, pp. 279~299.
- 이환웅, 『재정사업의 효과성 평가에 관한 소고: 스마트공장 구축 지원 사업을 중심으로』, 한국조세재정연구원, 2021.
- 장우현, 『중소기업지원정책의 개선방안에 관한 연구(Ⅱ)』, 한국개발연구원, 2014.
- 장우현·강희우·김빛마로, 『산업 및 기업재정정책 효율화를 위한 동태적 산업 재정정책 평가분석체계 구축에 관한 연구』, 한국조세재정연구원, 2020.
- 장우현·양용현·우석진, 『중소기업지원정책의 개선방안에 관한 연구(Ⅰ)』, 한국개발연구원, 2013.
- 장우현·우석진, 『중소기업지원정책의 개선방안에 관한 연구(Ⅲ)』, 한국개발연구원, 2015.
- 정재현·이환웅, 『머신러닝을 활용한 조세·재정 정책의 평가와 설계』, 한국조세재정연구원, 2020.
- 최현정·하봉찬, 「중소기업 정책금융지원 결정요인 분석」, 『산업경제연구』, 제28권 제3호, 2015, pp. 1195~1216.
- Abadie, A. and Kasy, M., “Choosing among regularized estimators in empirical economics: The risk of machine learning,” *Review of Economics and Statistics*, 101(5), 2019, pp. 743~762.

- Angrist, J. D. and Pischke, J. S., "Mostly harmless econometrics," Princeton university press, 2008.
- Angrist, J. and Hahn, J., "When to control for covariates? Panel asymptotics for estimates of treatment effects," *Review of Economics and statistics*, 86(1), 2004, pp. 58~72.
- Athey, S. and Imbens, G., "Recursive partitioning for heterogeneous causal effects," *Proceedings of the National Academy of Sciences*, 113(27), 2016, pp. 7353~7360.
- Athey, S. and Wager, S., "Estimating treatment effects with causal forests: An application," *Observational Studies*, 5(2), 2019, pp. 37~51.
- Athey, S., "21. The Impact of Machine Learning on Economics," *In The economics of artificial intelligence*, University of Chicago Press, 2019, pp. 507~552.
- Athey, S., Tibshirani, J., and Wager, S., "Generalized random forests," *The Annals of Statistics*, 47(2), 2019, pp. 1148~1178.
- Beattie, G., Laliberté, J. W. P., and Oreopoulos, P., "Thrivers and divers: Using non-academic measures to predict college success and failure," *Economics of Education Review*, 62, 2018, pp. 170~182.
- Bjerre-Nielsen, A., Kassamig, V., Lassen, D. D., and Lehmann, S., "Task-specific information outperforms surveillance-style big data in predictive analytics," *Proceedings of the National Academy of Sciences*, 118(14), 2021.
- Brieman, L., Friedman, J. H., Olshen, R. A., and Stone, C. J., "Classification and regression trees," Wadsworth Inc, 67, 1984.
- Caselli, S., Corbetta, G., Rossolini, M., and Vecchi, V., "Public credit guarantee schemes and SMEs' profitability: Evidence from Italy," *Journal of Small Business Management*, 57, 2019, pp. 555~578.
- Chalfin, A., Danieli, O., Hillis, A., Jelveh, Z., Luca, M., Ludwig, J., and

- Mullainathan, S., "Productivity and selection of human capital with machine learning," *American Economic Review*, 106(5), 2016, pp. 124~127.
- Chernozhukov, V., Demirer, M., Duflo, E., and Fernandez-Val, I., Generic Machine Learning Inference on Heterogeneous Treatment Effects in Randomized Experiments, with an Application to Immunization in India(No. w24678), National Bureau of Economic Research, 2018.
- Coad, A. and Srhoj, S., "Catching Gazelles with a Lasso: Big data techniques for the prediction of high-growth firms," *Small Business Economics*, 2019, pp. 1~25.
- Crump, Richard K., V. Joseph Hotz, Guido W. Imbens, and Oscar A. Mitnik, "Dealing with limited overlap in estimation of average treatment effects," *Biometrika* 96(1), 2009, pp. 187~199.
- Davis, J., and Heller, S. B., "Using causal forests to predict treatment heterogeneity: An application to summer jobs," *American Economic Review*, 107(5), 2017, pp. 546~550.
- Dehejia, R. H. and Wahba, S., "Causal effects in nonexperimental studies: Reevaluating the evaluation of training programs," *Journal of the American statistical Association*, 94(448), 1999, pp. 1053~1062.
- Friedman, J., Hastie, T., and Tibshirani, R., "The elements of statistical learning(Vol. 1, No. 10)," New York: Springer series in statistics, 2001.
- Fuster, A., Goldsmith-Pinkham, P., Ramadorai, T., and Walther, A., "Predictably unequal? the effects of machine learning on credit markets," *Journal of Finance*, The Effects of Machine Learning on Credit Markets, October 1, 2020.
- Genzko, M., Shapiro, J. M., and Taddy, M., "Measuring group differences in high-dimensional choices: method and application to congressional

- speech,” *Econometrica*, 87(4), 2019, pp. 1307~1340.
- Giannone, D., Lenza, M., and Primiceri, G. E., “Economic predictions with big data: The illusion of sparsity,” 2021.
- Glynn, A. N. and Quinn, K. M., “An introduction to the augmented inverse propensity weighted estimator,” *Political analysis*, 18(1), 2010, pp. 36~56.
- Gulen, H., Jens, C., and Page, T. B., “An application of causal forest in corporate finance: How does financing affect investment?,” Available at SSRN 3583685, 2020.
- Hahn, J., “On the role of the propensity score in efficient semiparametric estimation of average treatment effects,” *Econometrica*, 1998, pp. 315~331.
- Heckman, J. J., Ichimura, H., and Todd, P., “Matching as an econometric evaluation estimator,” *The review of economic studies*, 65(2), 1998, pp. 261~294.
- Imbens, G. W., “Matching Methods in Practice: Three Examples,” NBER Working Paper, w19959, 2014.
- Kang, J. W. and Heshmati, A., “Effect of credit guarantee policy on survival and performance of SMEs in Republic of Korea,” *Small Business Economics*, 31(4), 2008, pp. 445~462.
- King, G. and Nielsen, R., “Why propensity scores should not be used for matching,” *Political Analysis*, 27(4), 2019, pp. 435~454.
- Kleinberg, J., Lakkaraju, H., Leskovec, J., Ludwig, J., and Mullainathan, S., “Human decisions and machine predictions,” *The quarterly journal of economics*, 133(1), 2018, pp. 237~293.
- Kleinberg, J., Ludwig, J., Mullainathan, S., and Obermeyer, Z., “Prediction policy problems,” *American Economic Review*, 105(5), 2015, pp. 491~495.

- Knaus, M. C., Lechner, M., and Strittmatter, A., “Heterogeneous employment effects of job search programmes: A machine learning approach,” *Journal of Human Resources*, 0718-9615R1, 2020.
- Knittel, C. R. and Stolper, S., “Machine Learning about Treatment Effect Heterogeneity: The Case of Household Energy Use,” In *AEA Papers and Proceedings*, Vol. 111, May 2021, pp. 440~444.
- McKenzie, D. and Sansone, D., “Predicting entrepreneurial success is hard: Evidence from a business plan competition in Nigeria,” *Journal of Development Economics*, 141, 102369, 2019, pp. 1~15.
- Miyakawa, D., Miyauchi, Y., and Perez, C., “Forecasting firm performance with machine learning: Evidence from Japanese firm-level data,” RIETI, 2017.
- Mullainathan, S. and Obermeyer, Z., Diagnosing Physician Error: A Machine Learning Approach to Low-Value Health Care(No. w26168), National Bureau of Economic Research, 2019.
- Mullainathan, S. and Spiess, J., “Machine learning: an applied econometric approach,” *Journal of Economic Perspectives*, 31(2), 2017, pp. 87~106.
- Ordober, J. and Weiss, A., “Information and the law: Evaluating legal restrictions on competitive contracts,” *The American Economic Review*, 71(2), 1981, pp. 399~404.
- Orlov, G., McKee, D., Foster, I. R., Botton, D., and Thomas, S. R., “Identifying Students at Risk Using a New Math Skills Assessment,” In *AEA Papers and Proceedings*, 111, 2021, pp. 97~101.
- Robins, J. M., Mark, S. D., and Newey, W. K., “Estimating exposure effects by modelling the expectation of exposure conditional on confounders,” *Biometrics*, 1992, pp. 479~495.
- Rubin, D. B., “Estimating causal effects of treatments in randomized and nonrandomized studies,” *Journal of educational Psychology*, 66(5),

1974, pp. 688~701.

Sansone, D., "Beyond early warning indicators: high school dropout and machine learning," *Oxford bulletin of economics and statistics*, 81(2), 2019, pp. 456~485.

Schonlau, Matthias, "Boosted regression (boosting): An introductory tutorial and a Stata plugin," *The Stata Journal*, 5(3), 2005, pp. 330~354.

Schonlau, M. and Zou, R. Y., "The random forest algorithm for statistical learning," *The Stata Journal*, 20(1), 2020, pp. 3~29.

Shipman, J. E., Swanquist, Q. T., and Whited, R. L., "Propensity score matching in accounting research," *The Accounting Review*, 92(1), 2017, pp. 213~244.

Song, Y. G., Cao, Q. L., and Zhang, C., "Towards a new approach to predict business performance using machine learning," *Cognitive Systems Research*, 52, 2018, pp. 1004~1012.

Stiglitz, J. E., and Weiss, A., "Credit rationing in markets with imperfect information," *The American economic review*, 71(3), 1981, pp. 393~410.

Wager, S. and Athey, S., "Estimation and Inference of Heterogeneous Treatment Effects using Random Forests," *Journal of the American Statistical Association*, 113(523), 2018, pp. 1228~1242.

## 〈웹사이트〉

국가통계포털, <https://kosis.kr>, 검색일자: 2021. 11. 11.

기술보증기금, 「기술보증기금>보증이용안내>보증이용절차」, <https://www.kibo.or.kr/main/work/work010101.do>, 검색일자: 2021. 10. 30.

신용보증기금, 「신용보증기금>주요업무>신용보증>이용안내」, [https://www.kodit.co.kr/work/crdt\\_guar/crdt\\_guar/use\\_guid/user\\_outline.jsp](https://www.kodit.co.kr/work/crdt_guar/crdt_guar/use_guid/user_outline.jsp), 검색일자: 2021. 10. 30.

(주) 중소기업경영혁신연구원, 「신용보증기금 설명자료」, <http://smbc.re.kr/>

bbs/board.php?bo\_table=policyfund02&wr\_id=39&listsca=true, 검색일자:  
2021. 7. 17.

중소벤처기업진흥공단, 「지원사업>정책자금융자」, <http://kosmes.or.kr/sbc/SH/SBI/SHSBI001M0.do>, 검색일자: 2021. 10. 30.

\_\_\_\_\_, 「정책자금융자>세부사업>혁신창업산업화자금」, <https://www.kosmes.or.kr/sbc/SH/SBI/SHSBI004M0.do>, 검색일자: 2021. 10. 30.

한국은행, 「금융중개지원대출제도」, <https://www.bok.or.kr/portal/main/contents.do?menuNo=200303>, 검색일자: 2021. 10. 20.

\_\_\_\_\_, 「중소기업 금융지원제도-기본체계」, <https://www.bok.or.kr/portal/main/contents.do?menuNo=200302>, 검색일자: 2021. 10. 21.

e-나라지표, 「중소기업 금융지원 현황」, [https://www.index.go.kr/potal/stts/idxMain/selectPoSttsIdxMainPrint.do?idx\\_cd=2780&board\\_cd=INDX\\_001](https://www.index.go.kr/potal/stts/idxMain/selectPoSttsIdxMainPrint.do?idx_cd=2780&board_cd=INDX_001),  
검색일자: 2021. 10. 20.

#### 〈국내 기타자료〉

중소벤처기업부 공고 제2021-648호, 「2022년도 중소기업 정책자금 융자계획  
공고」

## 부 록

〈부표 1〉 강건성 검정 결과(2011~2015년 표본)

표본 유형	(로그)매출액		영업이익		연구개발비	
	계수	P값	계수	P값	계수	P값
GDP디플레이터 조정	0.076*** (0.013)	0.000	-6.539 (6.004)	0.276	-2.143 (1.840)	0.244
연구개발비 로그	해당 없음				-0.116* (0.065)	0.072

주: 1. 국가통계포털(<https://kosis.kr>, 검색일자: 2021. 11. 11.) 제공 GDP디플레이터 이용하여 조정  
 2. 금액 단위는 백만원  
 3. 통계적 유의성은 \*\*\* 1%, \*\* 5%, \* 10%와 같음  
 자료: 신용보증기금 내부자료; 한국기업데이터 자료

〈부표 2〉 영업이익에 대한 추가 이질성 분석 결과

구분	표본 크기	(로그)매출액		영업이익		연구개발비	
		ATT	P	ATT	P	ATT	P

Panel A: 2012년 영업이익

정책효과 (중위값 이상)	41,729	0.074*** (0.016)	0.000	-31.607*** (10,010)	0.002	-1.803 (3,138)	0.566
정책효과 (중위값 미만)	41,179	0.097*** (0.017)	0.000	-1.462 (4,917)	0.766	-1.859 (1,941)	0.338

Panel B: 2011년 영업이익

정책효과 (중위값 이상)	41,747	0.057*** (0.015)	0.000	-12.076 (9,764)	0.216	-2.501 (3,367)	0.458
정책효과 (중위값 미만)	41,583	0.114*** (0.017)	0.000	0.810 (5,130)	0.875	-3.673** (1,863)	0.049

주: 1. 금액 단위는 백만원  
 2. 통계적 유의성은 \*\*\* 1%, \*\* 5%, \* 10%와 같음  
 자료: 신용보증기금 내부자료; 한국기업데이터 자료

〈부표 3〉 NICE평가정보(주) 기업정보데이터 15개 파일

대분류	파일명	파일내용
개요	em01	업체개요
	em02	업체사업장정보
현황	aa04	업체현황-주식소유형태별분포
	aa05	업체현황-주식소유규모별분포
	aa17	업체현황-생산및판매현황
	aa18	업체현황-주요제품가격동향
	aa19	업체현황-원재료가격동향
	aa20	업체현황-영업소및공장현황
	aa21	업체현황-장단기차입금
	aa99	폐업일정보
재무	ab00	재무KEY정보
	ab08	약식재무
	az07	산업코드
산업정보	az12	업체별산업코드
	az14	업체별계열코드
	신규입수	국민연금

자료: 저자 작성

〈부표 4〉 매출 예측 관련 설명변수들의 중요도

구분	AdaBoost		RandomForest	
	변수명	중요도	변수명	중요도
신용보증 지원받지 않은 기업 집단	1	매출액_1 (전년도 매출액)	매출액_1 (전년도 매출액)	83.11%
	2	순이익_1 (전년도 순이익)	매출액_3 (3년 전 매출액)	4.82%
	3	aa21_11 (장기차입금 합계)	매출액_2 (2년 전 매출액)	4.33%
	4	순이익_2 (2년 전 순이익)	부채총계_2 (2년 전 부채총계)	1.52%
	5	aa17_4 (생산및판매현황_총판매금액)	총자산_3 (3년 전 총자산)	1.34%

〈부표 4〉의 계속

구분	AdaBoost		RandomForest		
	변수명	중요도	변수명	중요도	
신용보증 지원받지 않은 기업 집단	6	매출액_2 (2년 전 매출액)	1.09%	부채총계_1 (전년도 부채총계)	1.06%
	7	부채총계_2 (2년 전 부채총계)	1.07%	aa17_3 (생산및판매현황_ 내수금액)	1.01%
	8	aa21_8 (원화 장기차입금)	0.88%	부채총계_3 (3년 전 부채총계)	0.97%
	9	총자산_2 (2년 전 총자산)	0.79%	영업이익_1 (전년도 영업이익)	0.93%
	10	자본금_1 (전년도 자본금)	0.72%	순이익_1 (전년도 순이익)	0.92%
신용보증 지원 받은 기업 집단	1	매출액_1 (전년도 매출액)	78.08%	매출액_1 (전년도 매출액)	70.80%
	2	총자산_3 (3년 전 총자산)	5.75%	매출액_2 (2년 전 매출액)	9.90%
	3	aa21_16 (차입금 총계)	3.87%	부채총계_2 (2년 전 부채총계)	6.37%
	4	부채총계_3 (3년 전 부채총계)	2.56%	총자산_1 (전년도 총자산)	4.27%
	5	aa17_4 (생산및판매현황_ 총판매금액)	2.50%	매출액_3 (3년 전 매출액)	2.56%
	6	부채총계_2 (2년 전 매출액)	1.94%	aa17_3 (생산및판매현황_ 내수금액)	1.46%
	7	순이익_3 (3년 전 순이익)	1.74%	총자산_3 (3년 전 총자산)	1.45%
	8	총자산_2 (2년 전 총자산)	1.58%	자본금_3 (3년 전 자본금)	1.13%
	9	자본금_2 (2년 전 자본금)	1.10%	aa17_4 (생산및판매현황_ 총판매금액)	1.03%
	10	aa21_11 (장기차입금 합계)	0.86%	부채총계_1 (전년도 부채총계)	1.03%

자료: 저자 작성

〈부표 5〉 알고리즘별 예측 성과 비교: 직전 2개년도 실적데이터 사용

종속변수(금액기준)		매출	영업이익 <sup>1)</sup>	고용
신용보증 지원받지 않은 기업	OLS	0.1414	0.0230	0.2248
	AdaBoost	0.0976	0.0238	0.2344
	RandomForest	0.0703	0.0202	0.2144
신용보증 지원받은 기업	OLS	0.0924	0.0239	0.2408
	AdaBoost	0.1135	0.0290	0.2501
	RandomForest	0.0816	0.0233	0.2359

주: 1) 영업이익은 법인세 차감 전 계속이익을 뜻함  
 자료: NICE평가정보(주), 「기업정보데이터」를 사용하여 저자 계산

〈부표 6〉 알고리즘별 예측 성과 비교: 전년도 실적데이터 사용

종속변수(금액기준)		매출	영업이익 <sup>1)</sup>	고용
신용보증 지원받지 않은 기업	OLS	0.1472	0.0222	0.2270
	AdaBoost	0.0952	0.0233	0.2525
	RandomForest	0.0722	0.0213	0.2163
신용보증 지원받은 기업	OLS	0.0940	0.0237	0.2386
	AdaBoost	0.1065	0.0260	0.2454
	RandomForest	0.0838	0.0228	0.2346

주: 1) 영업이익은 법인세 차감 전 계속이익을 뜻함  
 자료: NICE평가정보(주), 「기업정보데이터」를 사용하여 저자 계산

## 머신러닝을 활용한 재정사업평가: 정책금융 사례를 중심으로

---

이환웅 · 고창수 · 배진수

본 연구는 정책금융 수혜자를 선정하는 데 머신러닝 방법론이 정책 효과성 측면에서 개선에 도움이 될 수 있는지를 탐색하였다. 머신러닝 방법론의 유용성을 검토하기 위해 신용보증기금의 정책효과를 선행연구에서 많이 사용한 성향점수매칭 방법론과 머신러닝 알고리즘인 커절 포레스트를 이용하여 분석하고 추정 결과를 비교하였다. 두 가지 방법론에 의한 추정 결과는 질적으로 유사해 두 방법론이 상호 보완적으로 사용될 수 있음을 확인하였으나, 성향점수매칭의 경우 성향점수를 로짓모형을 통해 추정할 때 추정에 사용된 변수에 따라 처치효과 크기가 많이 달라지는 현상이 관측되었다. 이는 성향점수에 의한 정책효과 추정치가 변수선택에 따라 강건하지 않음을 의미하며 머신러닝이 효과성 추정에서도 기존 통계모형의 한계점을 극복할 수 있는 좋은 대안이 될 수 있음을 함의한다.

또한 본 연구에서는 성향점수매칭과 커절 포레스트를 활용해 정책금융 지원효과 이질성을 분석하고 두 방법론을 비교하였다. 머신러닝을 활용한 처치효과 이질성 분석은 사업체들의 주어진 독립변수의 값에 따라 조건부 평균처치효과를 계산할 수 있기 때문에 성향점수매칭과 비교하여 보다 세밀한 이질성 분석이 가능하였다. 구체적으로 매출액에 대한 이질적인 처치효과를 처치효과 크기 순으로 나열했을 때 평균처치효과는 4.45%였지만, 처치

효과 기준으로 상위 10%에 속하는 사업체들의 평균적인 처치효과는 8.82%, 하위 10%에 속하는 사업체들은 0.97%로 나타나 신용보증기금의 정책효과가 매우 이질적임을 알 수 있었다. 또한 설명변수 중 하나인 사업체의 과거 영업이익의 크기 변화에 따라 매출액이 증대되는 효과가 매우 크게 비선형적으로 변화하는 것으로 나타났다. 머신러닝에 기반한 처치효과의 이질성 분석은 기존의 선형모형에서 발견하기 어려운 비선형적 패턴 발견이나 처치효과와의 차이를 크게 발생하는 설명변수가 무엇인지를 식별하는 데 유용할 것으로 보인다. 본 연구 결과에 따른 정책적 시사점은 크게 두 가지로 제시할 수 있다. 첫째, 신용보증기금, 기술보증기금 등 중소기업을 대상으로 하는 정책금융 선정 과정에 머신러닝 방법론을 활용하는 것이며, 둘째, 비단 정책금융 분야에 국한되지 않고 향후 정부의 재정사업 효과성 평가에 머신러닝 방법론을 보완적으로 사용하는 것이다.

## Evaluating and Designing the Public Policies with Machine Learning Algorithm: The Case of Korea Credit Guarantee Fund

---

Hwanoong Lee, Changsu Ko and Jinsoo Bae

This study explored whether machine learning(ML) approaches may help improve policy effectiveness when it comes to targeting policy recipients. To explore this, we analyzed the impact of receiving the Korea Credit Guarantee Fund(KCGF) on SMEs' revenues and profits by using the propensity score matching(PSM) and Causal Forest(ML algorithm) and the estimation results were compared to see if the ML was useful. The estimation results by the two methods were qualitatively similar, indicating that the two methods can be used complementary to each other. However, in the case of PSM, we observed that the treatment effects varied greatly depending on the variables used for propensity score estimation. This shows that the estimation of treatment effects via PSM is not robust to the modeling assumptions, and it implies that ML could be a viable alternative to existing statistical methods for estimating the treatment effects.

Next, We analyzed the heterogeneous treatment effects of the receiving credit guarantee policy using PSM and Causal Forest(CF). Compared to using PSM, a more detailed heterogeneity analysis was achieved using CF because the conditional average treatment impact could be estimated

based on the values of the given independent variables of SMEs. Specifically, When the heterogeneous treatment effects on sales were listed in order of size of treatment effect, the average treatment effect was 4.45 percent, but the average treatment effect of businesses belonging to the top 10% based on the treatment effect was 8.82 percent, and those belonging to the bottom 10% were 0.97 percent, indicating that the effects of the receiving KCGF are very heterogeneous. Also, it was found that the treatment effects of increasing revenues change very significantly and non-linearly according to the SMEs' past operating profit. The findings suggest that ML-based heterogeneity analysis could be effective in discovering non-linear patterns that are difficult to find in existing linear models, as well as identifying explanatory variables that are substantially responsible for disparities in treatment effects.

## ■ 저자약력

### 이환웅

연세대학교 경영학과 졸업  
미국 Michigan State University 경제학 박사  
현, 한국조세재정연구원 부연구위원

### 고창수

연세대학교 경제학과 졸업  
미국 University of California, Los Angeles 경제학 박사  
현, 한국조세재정연구원 부연구위원

### 배진수

연세대학교 경제학과 졸업  
미국 The Ohio State University 경제학 박사  
현, 한국조세재정연구원 부연구위원

연구보고서 21-07

## 머신러닝을 활용한 재정사업평가: 정책금융 사례를 중심으로

---

발행	행	2021년 12월 31일
저자	자	이환웅 · 고창수 · 배진수
발행인	인	김재진
발행처	처	한국조세재정연구원
주소	소	30147 세종특별자치시 시청대로 336
전화	화	(044)414-2114(대)
홈페이지	지	www.kipf.re.kr
등록	록	1993. 7. 15. 제2014-24호
정가	가	14,000원
조판 및 인쇄	쇄	일지사
I S B N		979-11-6655-099-7

---

© 한국조세재정연구원 2021 \* 잘못 만들어진 책은 바꾸어 드립니다.