

구인구직 플랫폼이 노동시장에 미친 영향

이름 : 김겨레

학번 : 2019314908

계량경제학 데이터 분석 보고서

요약

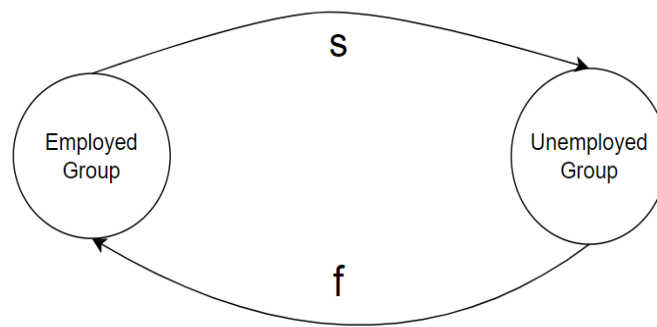
완전경쟁 노동시장 이론에서는 노동시장의 불균형이 존재하지 않는다. 하지만 실제 현실에서는 정보 불완전성 문제로 인해 노동시장의 비효율성이 나타난다. 21세기 들어 급격하게 확산된 정보화 흐름 속에서 구인구직 플랫폼이 만들어지고 확산되었다. 구인구직 플랫폼은 노동시장의 양면을 연결하여 노동시장의 정보 불완전성을 개선하기 때문에, 플랫폼의 확산이 노동시장을 보다 효율적으로 만들 수 있다는 의견이 지배적이다. 본 보고서는 실제로 구인구직 플랫폼의 확산이 노동시장의 효율성에 미치는 영향을 분석하였다. 구직기간이 1년 미만인 구직자 수를 1년 이상인 구직자 수로 나누어 노동시장 효율성 지표로 활용했고, 한국의 주요 구인구직 플랫폼의 트래픽 수를 구인구직 플랫폼의 규모 변수로 두었다. 또한, 구인구직 플랫폼의 확산 이외에 노동시장 효율성에 영향을 미칠 수 있는 요인인 경제 상황 관련 변수, 즉 고용률을 통제변수로 추가하여 분석을 진행했다. 분석결과, 구인구직 플랫폼의 확산이 모형 내에서 노동시장의 효율성에 양의 방향으로 영향을 준다는 것을 확인했지만, 통계적 유의성이 낮았다. 따라서 플랫폼 확산이 노동시장 효율성에 유의미한 영향을 미치지 않는다는 결과를 얻어냈다. 하지만 ‘노동시장 효율성 지표’가 본 보고서의 임의적 설정이라는 점과 구인구직에 들어가는 비용의 절감은 제대로 고려하지 못했다는 점에서 본 연구는 한계가 있다. 따라서 결과를 제한적으로 받아들이고 추가적인 연구를 수행할 필요가 있다.

1. 경제 모형 설정

완전경쟁적 노동시장이론은 노동력의 질이 균등하고 노동시장의 정보가 완벽하다는 완전경쟁을 가정한다. 이러한 기본적인 노동시장 모형에서는 실업과 빈 일자리의 공존 같은 노동시장의 불균형이 존재하지 않는다. 하지만 실제 현실에서 노동시장의 불균형은 빈번하게 나타난다. 이에 따라 현재 경제학에서는 노동시장에서 나타나는 실업의 종류를 크게 세 가지로 나눈다.

첫째, 경기적 실업률이다. 경기적 실업률은 ‘임금의 경직성’에 의해 나타난다. 경기가 악화되어 임금의 하방 압력이 작용할 때 실제로 임금이 그만큼 내려가지 못해 발생하는 실업을 의미한다. 둘째, 마찰적 실업이다. 마찰적 실업은 노동시장 참여자들의 특성과 선호 능력의 차이, 그리고 이런 정보에 대한 불완전성으로 인해 발생한다. 셋째, 구조적 실업이다. 구조적 실업은 산업의 구조가 변하면서 나타나거나, 지역 간의 불균형에 의해 발생하고, 넓은 의미에서 마찰적 실업으로 분류한다.

세 가지 종류의 실업을 중 경기적 실업은 임금 등의 경직성에 의해 자원이 비효율적으로 배분되어 생기는 반면, 나머지 두 가지는 자원이 효율적으로 배분되는 과정에서 노동시장의 일시적 수급 미스매칭에 의해 발생하는 실업이다. 이런 종류의 실업은 노동시장에서 나타나는 정보의 불완전성에 의해 각 참여자가 탐색과정을 거쳐야 하기 때문에 불가피하게 발생하는 실업이다. 노동시장 참여자가 취업자 그룹과 실업자 그룹을 이동하며 균형을 찾아가는 과정이다.



<그림 1> 마찰적 실업의 발생 (s는 실직률, f는 취업률)

하지만 정보통신기술이 발전하면서 일자리를 구하려는 사람과 노동자를 구하고자 하는 기업이 활용할 수 있는 정보가 확대되었다. 특히 구인구직플랫폼의 등장으로, 노동시장 참여자들은 적은 거래비용으로 계약을 성사시킬 수 있게 되었다. 구인구직플랫폼이 생겨나면서 노동시장에서 정보의 불완전성이 줄어든 것이다. 실업에 대한 경제학 이론에 따르면, 노동시장에서 발생하는 정보의 불완전성이 실업 혹은 구인구직 기간의 증가로 이어진다. 따라서 정보 불완전성의 감소는 노동시장의 효율성을 증가시킨다.

본 보고서에서는 정보통신기술의 발달로 인해 형성된 구인구직 플랫폼이 정보 불완전성을 개선하여 노동시장의 효율성 개선에 긍정적인 영향을 주었는지 분석할 것이다.

II. 계량경제학 모형

1. 주요 변수 설정과 데이터

본 연구는 구인구직 플랫폼의 확산이 노동시장의 효율성에 미친 영향을 탐구하고자 한다. ‘확산’을 확인하기 위해서는 횡단면 자료가 아닌 시계열 자료를 활용하는 것이 적절하다. 구인구직 플랫폼의 확산은 구인구직 플랫폼을 활용하는 사업체와 구직자의 규모를 통해 확인할 수 있을 것이다. 하지만 구인구직 플랫폼에서는 회원가입한 사업체 및 개인의 수를 연도별로 제공하지 않는다. 따라서 “SAMrush”에서 제공하는 트래픽 데이터를 사용하였다. 국내 구인구직 플랫폼 중 가장 높은 점유율을 차지하고 있는 업체는 ‘잡코리아’로 2021년 기준 42.97%의 점유율을 가지고 있다. 두 번째로 높은 ‘사람인’은 그보다 약 18% 낮은 점유율을 기록하고 있다. 두 업체의 점유율이 시장의 60%를 넘어가고 있는 것이다. 따라서 ‘잡코리아’와 ‘사람인’ 트래픽 수의 합을 구인구직 플랫폼의

규모 변수로 사용했다. 또한, 본 연구는 플랫폼의 확산을 모델에 반영하고자 하므로 트래픽 수의 변화율, 즉 로그스케일의 데이터를 독립변수로 활용했다.

노동시장의 효율성으로 활용할 수 있는 지표는 실업률이 있다. 하지만 실업률은 노동시장의 효율성보다 경제상황을 보여주는 지표 성격이 강하다. 따라서 실업률을 노동시장 효율성 지표로 사용하는 것은 적절하지 않다. 하지만 경제상황이 좋지 않아도 구인구직 플랫폼은 노동시장의 ‘수요 공급 매칭’을 개선해줄 수 있다. 이에 따라 구인구직의 기간에 영향을 줄 수 있다. 따라서 본 연구는 구인구직 기간 관련 변수를 종속변수로 활용했다. 통계청은 월별 구직 기간별 실업자 수 데이터를 제공하고 있다. 구직 기간은 3개월 미만, 3개월 이상 6개월 미만, 6개월 이상 12개월 미만, 12개월 이상 18개월 미만, 18개월 이상으로 나누어져 있는데, 전체 구직자 대비 12개월 미만의 비율을 데이터로 활용했다.

트래픽 데이터가 2017년 이후만 제공됨에 따라 본 연구에서 사용하는 모든 시계열 자료는 모두 2017년 1월 이후의 월별 데이터이다.

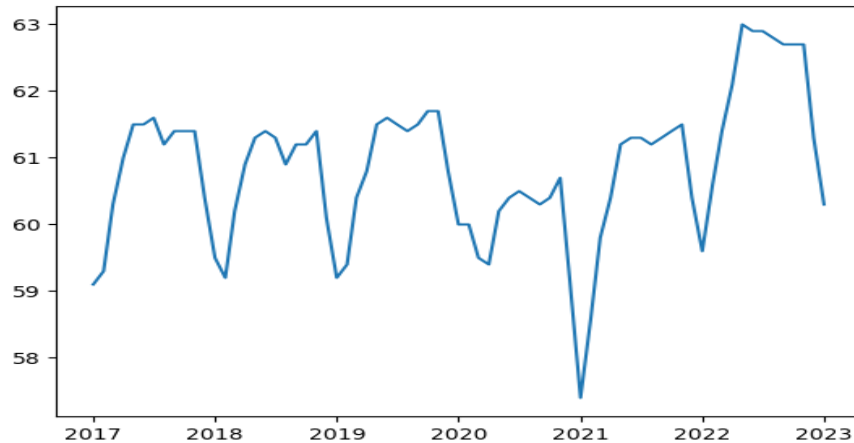
< 주요 변수와 활용 데이터 >

변수	데이터	출처
구인구직 플랫폼 규모 및 확산	잡코리아 및 사람인 트래픽	SAMrush
노동시장 효율성	전체 구직자 대비 구직기간 12개월 미만	통계청

2. 계량경제 모형

독립변수와 종속변수를 설정했지만, 경제 상황이 플랫폼 규모와 관계 없이 노동시장의 효율성 지표, 즉 구직기간 지표에 영향을 줄 수 있다. 경제 상황의 변화가 단기 구직자 수 및 단기 고용 경향에 영향을 줄 수 있기 때문이다. 따라서 모델에 경제상황에 대한 지표를 통제변수로 추가해야 한다.

앞서 밝혔다시피, 실업률은 경제상황을 나타내는 성격을 가진다. 따라서 경제상황을 보여주는 지표로 유용하다. 하지만 실업률은 경제활동인구만을 대상으로 한다. 따라서 전반적인 경제 여건을 보기 위해 본 연구는 경제활동인구뿐 아니라 생산가능인구 전체를 분모로 두는 고용률을 경제상황을 보여주는 지표로 사용했다. 고용률 지표는 경제상황을 보여주는 지표로서, 구인구직 플랫폼이 아닌 경제상황에 의해 단기 구직자 규모가 변하는 부분을 통제할 수 있다.

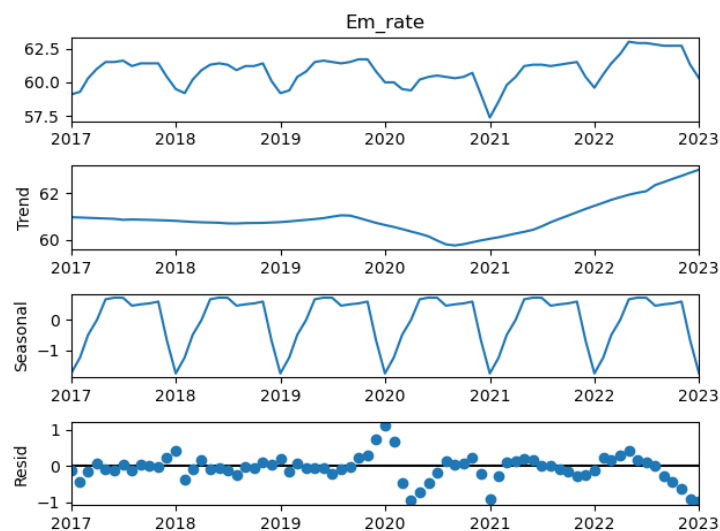


<월별 고용률 (2017년 1월 ~ 2023년 2월) >

고용률 데이터를 월별로 시각화하여 보면, 고용률은 ‘계절성’을 띠는 지표이다. 그렇기 때문에 고용률을 모델에 반영하기 위해서는 계절성을 조정해야 한다. 본 보고서에서는 계절성을 조정하기 위해 주 추세와 계절성을 분해하는 과정을 거쳤다. 변수 분해 모형과 분해 결과는 다음과 같다.

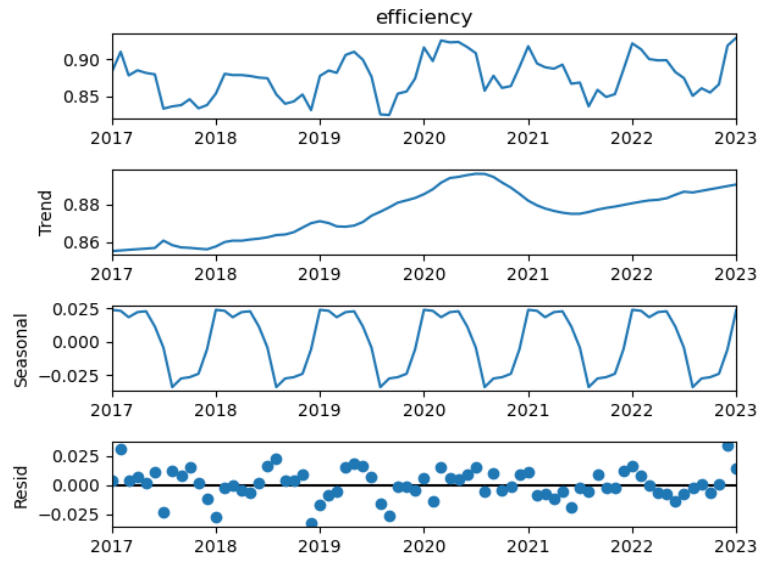
$$Emplment_t = intercept + trend_t + Season_t + v_t$$

$Emplment_t$ 고용률, $intercept$ 기초값, $trend_t$ 주 추세, $Season_t$ 계절성

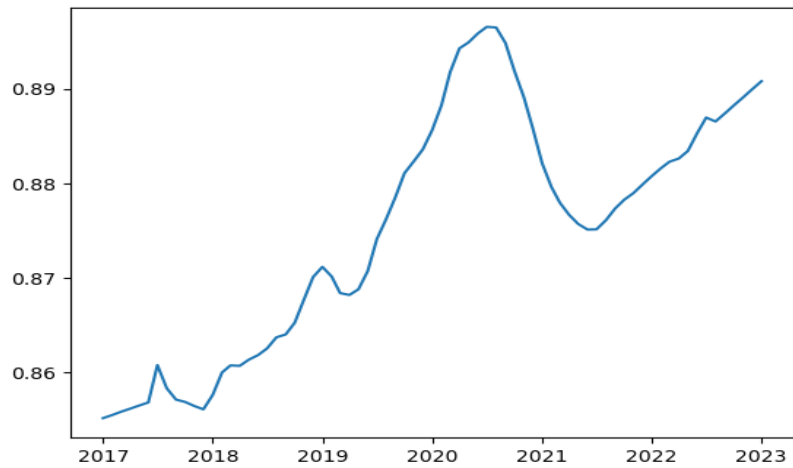


< 고용률 변수의 분해 결과 >

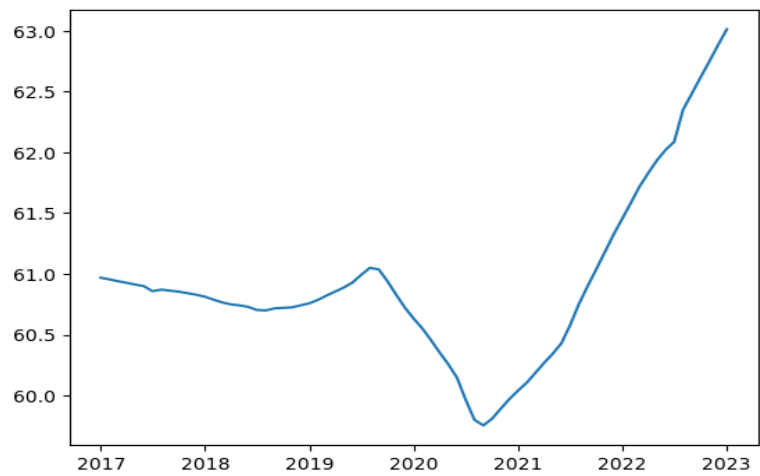
분해한 요소 중 ‘trend’, 즉 주 추세 변수는 계절조정 고용률이다. 본 보고서는 해당 변수를 경제 상황을 나타내는 지표로 활용하였다. 아래 그래프를 통해 볼 수 있듯이, 전체 대비 구직기간 12개월 미만 구직자 지표도 계절성을 보인다. 해당 데이터도 실업률과 마찬가지로의 방식으로 계절조정을 진행했다.



< 노동시장 효율성 변수 분해 결과 >



< 분해 후 노동시장 효율성 지표 >



< 분해 후 고용률 >

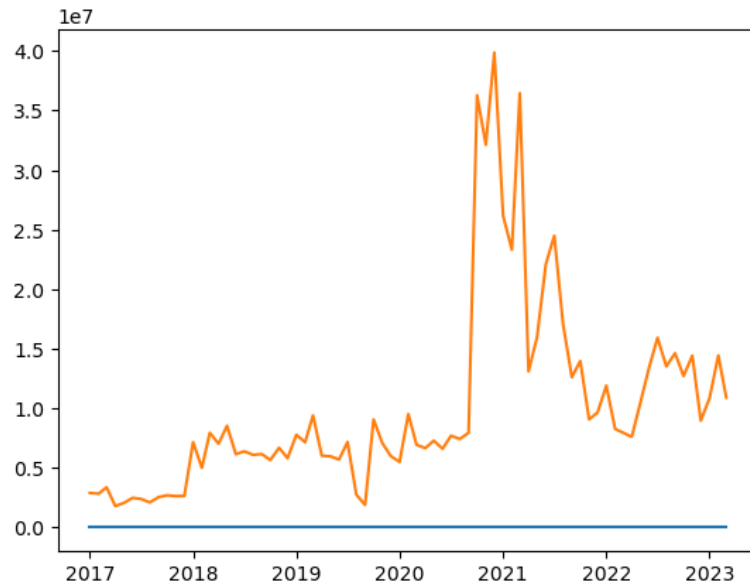
설정 한 독립변수, 종속변수, 통제변수로 구성된 계량경제학 모형은 다음과 같다.

$$\ln(\text{Volume}_t) = \alpha + \beta * \text{Efficiency}_t + \gamma * \text{Econ}_t + v_t$$

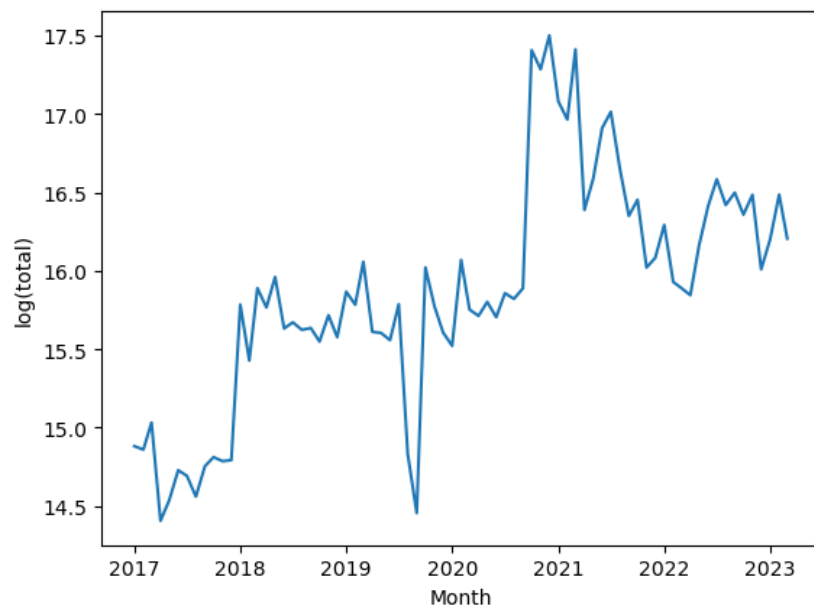
Volume_t t월의 구인구직 플랫폼 규모, Efficiency_t t월의 노동시장 효율성, Econ_t t월의 경제상황

III. 데이터 분석

1. 기초 데이터 분석



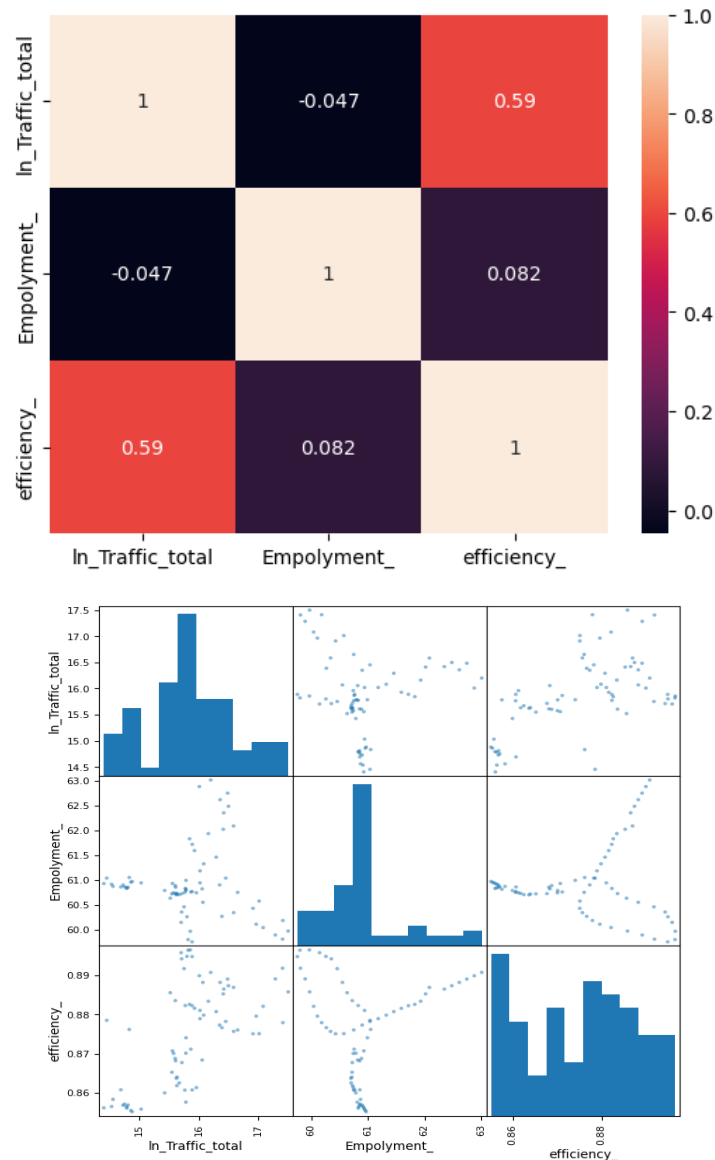
< 월별 트래픽 규모 (2017년 1월 ~ 2023년 2월) >



< 로그스케일의 월별 트래픽 규모 (2017년 1월 ~ 2023년 2월) >

위의 월별 트래픽 규모를 보면, 코로나 바이러스 발병 당시인 2020년에 트래픽 규모가 급격하게 증가했다. 경제상황에 따라 지표가 반응하는 모습을 보여준다. 해당 기간을 제외하면, 트래픽 수는 시간이 지남에 따라 증가하고 있다.

< 각 변수 간 상관관계 및 그래프 >



각 변수 간 상관계수를 보면, 고용률은 다른 변수와 약한 상관관계를 보인다. 경제 상황이 구인구직 플랫폼의 규모 및 노동시장 효율성 지표와 연동되어 움직이지 않는다고 추측해볼 수 있다. 반면, 플랫폼의 규모와 노동시장 효율성 지표 간에는 0.59의 높은 양의 상관관계를 보인다. 두 변수 간의 유의미한 관계가 있음을 추측할 수 있다

두 번째 그림을 통해, 세 변수의 분포와 각 변수 간 관계를 더 정확하게 확인할 수 있다. 주요 종속변수와 독립변수 간의 양의 상관성을 볼 수 있다. 또한 노동시장의 효율성과 경제 상황 지표 간의 관계가 시간에 따라 변했다는 것을 확인할 수 있다. 위의 분해 후 두 변수의 변화 양상을 보면 2021년 이전과 이후 해당 변수들 간의 관계가 바뀌었다. 2021년

3월을 기준으로 데이터를 나누어 상관계수를 구했을 때, 2021년 3월 이전에는 변수 간의 뚜렷한 음의 상관관계가 나타났다면, 이후에는 뚜렷한 양의 상관관계를 드러낸다.

구분	상관계수
2021년 3월 이전	-0.754
2021년 3월 이후	0.972

따라서 고용변수의 영향을 더 효과적으로 통제하기 위해 위의 계량경제 모형에 2021년 3월 이전 및 이후를 구분하는 더미변수를 추가하였다. 이에 따라 만들어진 최종적인 계량경제 모형은 아래와 같다.

$$Efficiency_t = \alpha + \beta * \ln(Volume_t) + \gamma * Econ_t * after202103_t + v_t$$

$after202103_t$ 는 2021년 3월 이전일 경우 0, 이후일 경우 1의 값을 갖는다.

2. 모형 분석

기초 통계, 즉 상관계수를 통해 모형 내 주요 변수 간 관계를 추측할 수 있었다. 이제 계량경제학 모형으로 회귀분석을 진행하고 구인구직 플랫폼이 노동시장 효율성이 미친 영향을 검증할 것이다.

< 계량경제 모형 분석 결과 >

	Coefficient	Standard Error	t-value	p-value
α	2.4341	0.274	8.478	0.000***
β	0.0018	0.002	0.822	0.411
γ_1 ($after202103_t$)	-1.9493	0.261	-7.862	0.000***
γ_2 ($Econ_t * after202103_t$)	0.0322	0.004	7.833	0.000***
γ_3 ($Econ_t$)	-0.0262	0.004	-6.012	0.000***
$R^2 = 0.650$ $Adj R^2 = 0.629$ $F - value = 31.56$ $Durbin - Watson = 0.083$				

계량경제 모형의 적합도(R^2)은 0.65로 높은 편이다. 하지만 시계열 자료라는 특성 때문에 *Durbin – Watson*는 0.083으로 양의 방향 시계열 자기상관을 보여준다. 따라서 Heteroskedasticity and Autocorrelation Consistent (HAC) standard error (Newey–West Covariance estimator)를 활용하여 모형을 다시 추정하였다.

< HAC 활용 분석 결과 >

	Standard Error	t-value	p-value
α	0.403	6.047	0.000***
β	0.003	0.702	0.483
γ_1 (<i>after202103_t</i>)	0.363	-5.373	0.000***
γ_2 (<i>Econ_t * after202103_t</i>)	0.006	5.359	0.000***
γ_3 (<i>Econ_t</i>)	0.006	-4.191	0.000***

0.65의 높은 적합도를 보면, 계량경제 모형은 잘 구성되어 있음을 알 수 있다. 또한, 회귀분석 결과 구인구직 플랫폼은 노동시장 효율성에 긍정적인 영향을 주는 것으로 나타났다. 하지만 구인구직 플랫폼의 규모의 회귀계수는 다른 통제변수에 따른 회귀변수와 다르게 높은 p-value 값을 보인다. HAC를 적용한 결과, p-value는 더 커졌다. 구인구직 플랫폼 규모 계수의 통계적 유의성이 낮은 것이다.

전반적인 모형의 분석 결과, 구인구직 플랫폼이 2017년부터 2023년 2월까지 기간 동안 노동시장의 효율성에 양의 영향을 준 것으로 보이지만, 통계적 유의성이 낮게 나타나는 것으로 보아 효율성에 미치는 영향의 강도가 강하지 않다. 노동시장의 효율성은 구인구직 플랫폼의 확산보다 경제 상황에 더 영향을 많이 받는 것으로 해석할 수 있다. 더 정확한 분석을 위해 가설검정을 진행하였다.

$$H_0 : \beta = 0$$

$$H_1 : \beta > 0$$

β 의 t-value는 0.702이고, p-value는 0.483이다. 5% 유의 수준에서 단측 검정의 귀무가설을 기각하기 위해서는 p-value가 0.05보다 작아야 하므로 귀무가설을 기각할 수 없다. 즉, 회귀계수가 유의하지 않다. 또한, β 를 0으로 두고 회귀분석을 진행하면 R^2 은

0.646인데, 이것을 통해 위의 모형과 F-test를 진행하면, F-value는 0.00009106으로, 통계적 유의성이 낮다.

계량경제 모형을 분석한 결과, 구인구직 플랫폼이 노동시장 정보불완전성을 개선해 노동시장 효율성을 효과적으로 개선했다고 보기 어렵다.

IV. 결론 및 제한점

본 보고서의 결과는 일반적인 통념과 다른 결과를 도출하였다. 정보통신 기술의 발전이 노동시장의 정보불완전성을 개선해 노동시장 효율성을 의미있게 개선할 것으로 기대했으나, 실제로 그렇지 못한 결과가 나타났다. 상관관계 분석에서 노동시장의 효율성을 나타내는 지표와 구인구직 플랫폼 간의 상관관계가 높은 수준으로 나타나는 것으로 보아 영향이 없었다고 단정할 수 없지만, 노동시장의 효율성은 정보 불완전성보다 전반적인 경제 상황에 의해 더 크게 좌우되는 경향이 있다고는 결론 내릴 수 있다.

하지만 본 보고서의 분석만으로 노동시장의 효율성과 구인구직 플랫폼의 확산 간의 관계를 확정지을 수 없다. 우선, 데이터 분석 과정에서 살펴본 바와 같이 고용률과 노동시장 효율성 간의 관계가 2021년을 기점으로 변한 것을 알 수 있는데, 이런 변화를 만든 요인이 무엇인지 제대로 파악하는 추가 연구가 있어야 본 보고서의 결과를 제대로 해석할 수 있다.

다음으로, 본 보고서의 분석에서 사용한 변수는 보고서 작성자의 주관이 크게 개입되었다. 만약 노동시장의 효율성을 더욱 잘 표현할 수 있는 다른 지표가 있고, 그것으로 데이터 분석을 한다면, 통계적으로 유의한 결과가 나올 가능성을 배제할 수 없다. 또한, 본 보고서는 구인구직 플랫폼의 규모 변수를 트래픽으로 설정했다. 하지만 실제로 구인구직 플랫폼을 이용하는 사람은 해당 플랫폼에 가입한 회원 및 기업이다. 따라서 회원수를 지표로 사용한다면, 얼마든지 다른 결과가 나올 수 있다.

마지막으로, 본 보고서는 트래픽, 구직기간 별 구직자 수, 고용률 만을 변수로 활용하였기 때문에 구인구직 플랫폼이 줄일 수 있는 거래비용, 즉 플랫폼이 존재하지 않았을 때 구인구직에 들어가는 시간 및 금전 비용을 반영하지 못했다. 실제로 이런 비용을 산출해 노동시장 효율성 지표로 활용한다면, 구인구직 플랫폼이 노동시장에 미친 영향은 본 보고서의 추정보다 더 커질 수 있다.

따라서 정보통신기술의 발전이 노동시장에 미치는 영향을 더욱 정확하게 분석하기 위해서는 포괄적인 자료를 활용한 추가적인 연구가 필요하다. 본 보고서에서 더욱 발전된 연구를 활용한다면, 다양한 정책과 비즈니스를 통해 노동시장을 더욱 효율적으로 만들 수 있는 방안도 고안할 수 있을 것이다.

데이터 분석에 활용한 코드

[202301_econometrics_lecture/data_analysis.ipynb](#) at [master](#) · [popper6508/202301_econometrics_lecture](#) · GitHub