

Are You Ready for the 4th Industrial Revolution?

2023. 4. 5 (수)

meritz Strategy Daily

전략 공감 2.0

Strategy Idea

미국 증시가 마주한 지난 실적시즌과 차이점

칼럼의 재해석

GPT(Generative Pre-trained Transformer)는 GPT(범용목적기술)인가?

본 조사분석자료는 제3자에게 사전 제공된 사실이 없습니다. 당사는 자료작성일 현재 본 조사분석자료에 언급된 종목의 지분을 1% 이상 보유하고 있지 않습니다.

본 자료를 작성한 애널리스트는 자료작성일 현재 해당 종목과 재산적 이해관계가 없습니다.

본 자료에 게재된 내용은 본인의 의견을 정확하게 반영하고 있으며, 외부의 부당한 압력이나 간섭 없이 신의 성실하게 작성되었음을 확인합니다.

본 자료는 투자자들의 투자판단에 참고가 되는 정보제공을 목적으로 배포되는 자료입니다. 본 자료에 수록된 내용은 당사 리서치센터의 추정치로서 오차가 발생할 수 있으며 정확성이나 완벽성은 보장하지 않습니다. 본 자료를 이용하시는 분은 본 자료와 관련한 투자의 최종 결정은 자신의 판단으로 하시기 바랍니다. 따라서 어떠한 경우에도 본 자료는 투자 결과와 관련한 법적 책임소재의 증빙자료로 사용될 수 없습니다. 본 조사분석자료는 당사 고객에 한하여 배포되는 자료로 당사의 허락 없이 복사, 대여, 배포 될 수 없습니다.

Strategy Idea



▲ 글로벌 투자전략 - 선진국

Analyst 황수욱

02. 6454-4896

soowook.hwang@meritz.co.kr

미국 증시가 마주한 지난 실적시즌과 차이점

- ✓ 미국 1Q23 실적시즌 은행 대출태도 악화에 따른 기업들의 가이드언스 추가 하향 가능성 경계
- ✓ 추가적인 경기둔화 확인이 통화정책 완화기대로 이어져 증시 반등 요인이 되긴 어려워 보임
- ✓ 이미 연내 50bp 인하기대까지 반영 중, 실물지표 둔화를 정책기대보단 실적에 반영할 것

1Q23 실적시즌 기업들의 가이드언스 추가 하향 가능성 경계

시장을 크게 두가지 관점에서 바라본다면 실적과 밸류에이션일 것이다. 4월은 1Q23 실적시즌이 시작하는 달로 4월 중순 금융주를 필두로 본격적인 실적발표가 시작된다. 이전에 반영되지 않았던 새로운 정보가 실적시즌에 업데이트 될 거리가 있다면 실적 시즌은 시장 방향성을 좌우할 수 있는 변수가 될 수 있다. Global Catch Mind 4월호에서는 이와 관련된 은행 규제 강화에 따른 대출태도 악화가 추가적으로 기업들의 실적 가이드언스에 반영될 가능성을 고려해야한다고 언급했다.

미국의 경우 밸류에이션은 장기 추세적으로 실적 추정치 방향과 동행했다. 다만 이번 국면의 경우 예외적으로 실적 하향에도 불구하고 밸류에이션 중심으로 상향되며 주가가 지지되고 있고, 그 배경에는 통화정책 긴축/완화 기대감이 냉탕/온탕을 오가며 월별로 증시 등락을 좌우하고 있다.

표1 4월 중순 이후 본격적으로 시작되는 미국 주요 기업 1Q23 실적발표 일정

4월 13일	4월 14일	4월 17일	4월 18일	4월 19일	4월 20일	4월 21일	4월 24일	4월 25일	4월 26일	4월 27일	4월 28일
TSMC	유나이티드 헬스 그룹	찰스슈왑	존슨앤존슨	테슬라	필립모리스	P&G	코카콜라	마이크로소프트	메타	애플	엑손모빌
델타항공	블랙록	스테이트 스트리트 은행	BofA	ASML	넥스트라 에너지	SAP		알파벳	퀄컴	아마존	뉴몬트
	JP모건		넷플릭스	모건스탠리	AT&T	슬럼버거		비자	보잉	마스터카드	
	웰스파고		록히드마틴	USB	아메리칸 익스프레스			텍사스 인스트루먼트	페이팔	일라이릴리	
	Citi		골드만삭스	램리서치	프리포트 맥모란			펍시코		맥도날드	
			IBM	베이커휴즈	노키아			HSBC		Linde	
					블랙스톤			UPS		캐터필러	
								GE		인텔	

자료: Nasdaq.com, 메리츠증권 리서치센터

경기 하강과 연동된 실적 하향은 1분기 중 이미 이어져왔음

1Q23 실적 추정치 하향폭은 과거 평균대비 큰 수준

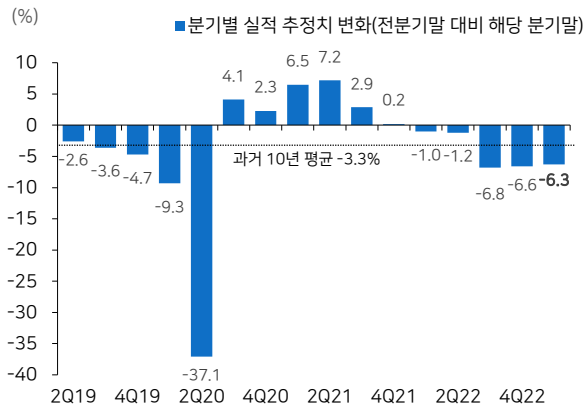
FactSet에 따르면, 분기 실적 추정치는 통상적으로 해당 분기를 통과하며 하향조정된다. 최근 10년간 S&P500 기업들의 평균적인 분기중 Bottom-up 실적 추정치 하락률은 -3.3%였다. 그런데 1Q23 실적 추정치는 3월 30일 기준 전분기말 대비 6.3% 하락하며 과거 평균 하락폭을 상회했다.

2023년 연간 실적 추정치도 전분기말 대비 3.8% 하락했다. FactSet은 통상적으로 매해 첫 분기에 당해년도 실적추정치의 하향조정이 나타나지만, 올해는 그 폭이 더 컸다고 분석했다. 최근 10년간 첫 분기의 당해년도 실적추정치 하락률은 평균 -1.4%였다.

3월 ISM 제조업지수는 COVID 제외 시 '09년 6월 이후 최저치, 경기하강 연동된 실적 하락 지속

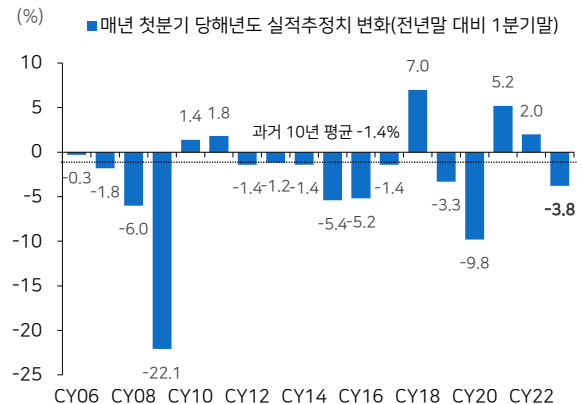
우리는 미국 경기의 수축국면이 지속되면서 실적 추정치 하락도 이어질 것으로 보았다. 경기국면의 주요 proxy인 3월 ISM 제조업지수는 지난달 발표된 지역 연은 제조업지수가 시사한대로 전월대비 1.4pt 하락한 46.3을 기록했다. COVID 시기('20.4월 41.8, '20.5월 43.5)를 제외하고 2009년 6월(46.3) 이후 최저치다.

그림1 S&P500 전분기말 대비 해당분기말 실적추정 변화



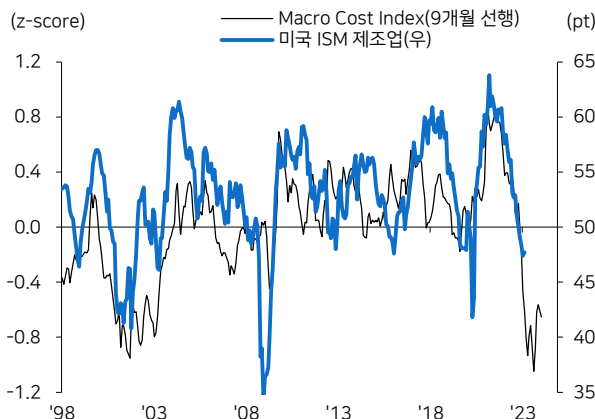
자료: FactSet, 메리츠증권 리서치센터

그림2 매년 첫분기 당해년도 실적 추정치 변화율



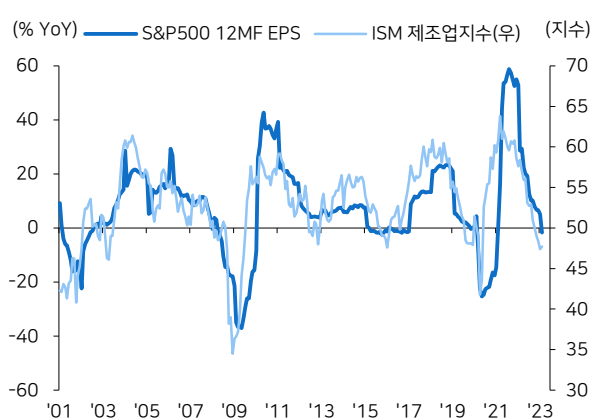
자료: FactSet, 메리츠증권 리서치센터

그림3 Macro cost에 후행하여 하방압력 이어지는 경기



자료: Bloomberg, US ISM, 메리츠증권 리서치센터

그림4 S&P500 12MF EPS 변화율과 ISM 제조업지수



자료: Bloomberg, US ISM, 메리츠증권 리서치센터

통화긴축의 후행적인 실물경제 파급효과 확인

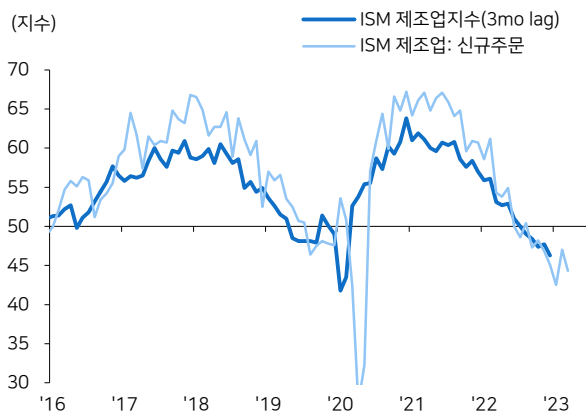
신규주문 하락, 헤드라인 하락 이어질 가능성 시사, 고용지수 기준선 하회, 3월 고용 발표에 긍정적

헤드라인 지수 하락을 견인한 세부지표는 신규주문과 고용이다. 통상적으로 헤드라인지수에 3개월 정도 선행한다고 알려져있는 ISM 제조업 신규주문지수는 1월 급락 이후 반등했다가 다시 하락하여 44.3(전월대비 -2.7pt)을 기록했다. 고용지수도 전월대비 2.2pt 하락한 46.9를 기록하며 기준선 하회폭을 키웠다. 서비스업 고용지수까지 확인해야겠으나, 최근 기업 감원 소식이 연달아 이어지는 가운데 이번주 확인될 3월 고용지표도 타이트하지 않을 가능성이 높아졌다.

기업들의 전형적인 경기 위축국면 대응, 물가압력 완화 징후도 확인

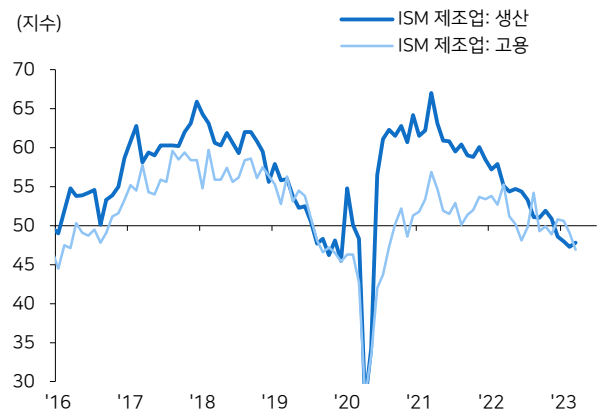
고용지표 외에서도 물가 압력이 완화되고 있는 징후가 확인되었다. 2월 잠시 반등하며 우려를 키웠던 ISM 제조업 물가지수는 다시 기준선을 하회(2월 51.3→49.2)했다. 공급자배송시간지수는 6개월 연속 기준선을 하회하며 공급측 물가압력은 계속 둔화되고 있음을 시사했다. 여기에 수주잔고, 재고/소비자 재고가 줄어 들고 있는 것에서 기업들의 전형적인 경기 위축국면 대응 모습을 보여주었다.

그림5 헤드라인에 선행하는 ISM 제조업 신규주문 하락 지속



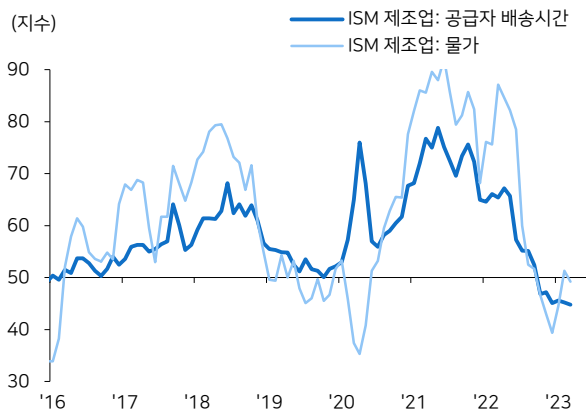
자료: US ISM, 메리츠증권 리서치센터

그림6 ISM 제조업 고용지수도 기준선 하회폭 심화



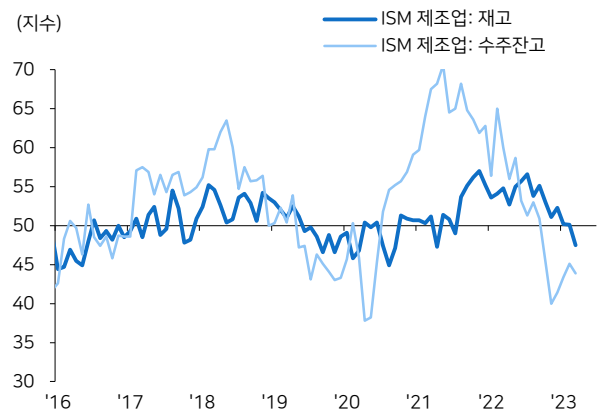
자료: US ISM, 메리츠증권 리서치센터

그림7 물가 압력 완화; 금리 인상의 실물 파급효과 확인



자료: US ISM, 메리츠증권 리서치센터

그림8 기업재고, 수주잔고 감소; 경기 위축국면 대응



자료: US ISM, 메리츠증권 리서치센터

은행 환경 영향은 아직 반영되지 않은 듯, 추가 반영 가능성 경계

세부 코멘트에서는 업종별 업황 차별화 확인

3월 ISM 제조업지수의 업종별 코멘트를 보면 업종별로 아직까지 호황인 산업과 그렇지 않은 산업이 구분된다. 컴퓨터/전자제품의 경우 그동안 리드타임 문제로 높았던 선주문에 대해 리드타임 안정으로 신규주문 축소 계약을 검토중이라는 코멘트가 있었다. 이외에 화학, 운송장비, 섬유/의류 산업은 업황 위축이 확인이 되었다. 반면 1차금속, 금속가공, 음식료 업황은 양호하다는 코멘트가 있었다.

아직 은행 규제 관련 추가 다운턴 가능성은 반영하지 않고 있는 듯. 향후 반영된다면 업황에 대한 기업들의 전망 추가 악화 가능성

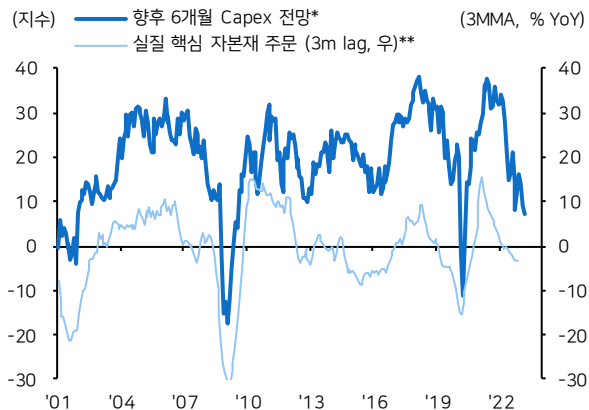
다만 지난달 글로벌 은행 리스크가 불거지면서 우리가 했던 우려는 감독 당국의 규제 강화에 따른 은행들의 대출 태도 타이트닝, 여기에 후행하는 기업들의 투자 활동 위축이다(3월 30일자 전략공감, 이승훈, '미국 금리인상 파급 효과: 기업투자 와 한국 수출 시사점', 3월 31일자 Global Catch Mind 4월호 참고). 3월 ISM 서베이에서 확인된 내용으로 미루어 볼 때, 기업들은 업황 전망에 이러한 영향력을 반영하지 않고 있는 듯하다. 우리의 우려대로라면 5월 이후 강화될 은행 규제 강화에 따라 기업들의 투자 심리가 추가 하향될 가능성이 남아있는 것으로 보인다.

표2 3월 ISM 제조업지수 서베이 업종별 코멘트

업종	내용
컴퓨터/전자제품	주문 및 생산량 전월대비 변화 없음. 리드타임 대부분 안정되어 일부 전략적인 구매를 제외하고 신규주문 계약 축소 검토 중
화학	매출 감소, 절약(savings)에 중점을 두고 예산 삭감되는 중
음식료/담배	투입 비용의 등락이 있는 가운데 전반적인 사업활동은 대체로 호조
운송장비	매출이 빠르게 감소 중이기 때문에 이월 주문을 빠르게 처리 가능
기계	리드 타임 개선 지속되나 인플레이션 압력 이어짐. 철강 제품 가격 상승 중. 우크라이나 전쟁과 중국발 불확실성에 재고수준을 높이고 있음
금속 가공	1분기 영업활동은 계획대로 진행 중, 매출은 예상 상회. 2분기는 어려울 것으로 예상되나 전망 고수 중
섬유/의류	업황은 전반적으로 부진, 아직 팬데믹 이전 수준의 주문을 받지 못함
기타 제조	1Q23은 전반적으로 2021-22년 대비 안정적. 고객 수요 예상대로 양호하게 증가, 전반적인 공급환경은 지난 2년대비 개선. 글로벌 은행 상황 주시 중이나 현시점에서는 어떠한 영향이 확인되지도, 예상되지도 않고 있음. 미-중 갈등도 관련 포인트
가전	신규주문 완화, 공급업체 배송시간 개선. 이 덕분에 일부 카테고리에서 수주잔량 줄이고 버퍼 구축 가능. 전자제품의 공급망 혼란은 아직 COVID 이전 대비 현저함(significant)
1차 금속	전반적인 비즈니스는 호조, 일부 간접 공급에서 공급망 문제 여전히 발생

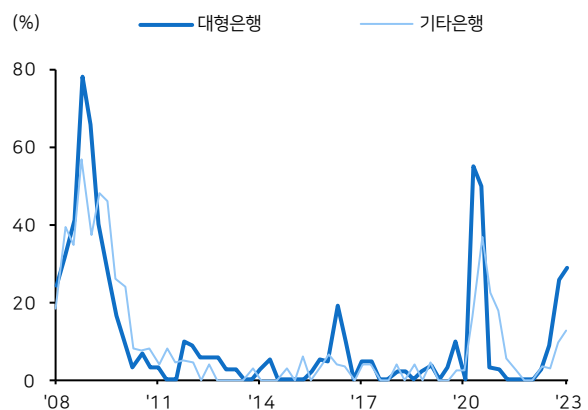
자료: US ISM, 메리츠증권 리서치센터

그림9 지역연은 서베이 capex 전망 vs 실질 자본재 주문



주: * 뉴욕, 필라델피아, 리치몬드, 캔자스, 달라스. ** 기계장비 PPI로 환인
자료: 각 지역연은, US Census Bureau, US BLS, 메리츠증권 리서치센터

그림10 은행의 대기업 대출태도(수치가 높을수록 타이트닝)



주: Tightened considerably + Tightened somewhat 응답비율의 합
자료: Federal Reserve Senior Loan Officer Survey, 메리츠증권 리서치센터

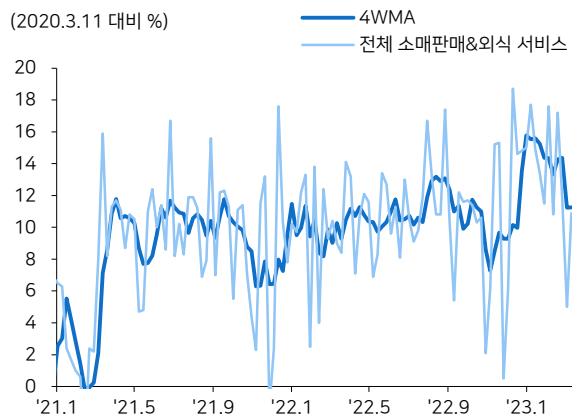
기업환경뿐만 아니라 미국 경기 근간인 소비도 둔화

소비 지표 중 High frequency data는 미국 3월 소비 추가 둔화 시사

4월 발표될 미국 3월 소비지표도 추가 둔화될 가능성이 높다. 미국 경제분석국(BEA)이 주간으로 발표하는 신용카드 매출 데이터는 COVID 이전('20년 3월) 대비 카드 매출이 얼마나 증감했는지 보여준다. 이를 기준으로 할 때, 1월 평균 신용카드 매출은 COVID 이전 대비 15.5% 많았다. 1월 소비 급등을 일시적 요인(3월 2일 전략공감 참고)이라고 분석한 바 있는데, 연초 높아졌던 소비 수준은 3월 들어 되돌리기 시작하여 1월 이전 수준으로 회귀하고 있다. 세부적으로는 리오프닝 관련 소비는 COVID 이전을 상회하나, 외식서비스나 재화소비는 감소 중이다.

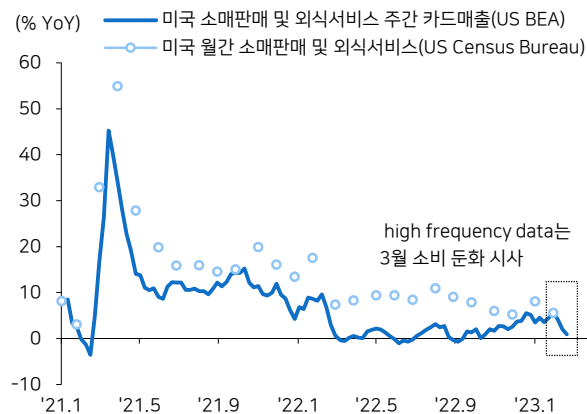
소비지표 중 시장이 가장 민감하게 반응하는 것은 미국 인구통계국(Census Bureau)이 월간으로 발표하는 소매판매금액이다. 그런데 BEA의 주간 카드매출액은 소매판매 지표를 잘 트래킹 하는 high frequency data로 활용할만하다. 주간 카드매출액을 기준점을 기준으로 지수화하여 YoY 증감율을 기준으로 월간 소매판매 데이터와 비교해 보면, 높은 상관성을 확인할 수 있다(그림 12).

그림11 미국 주간 신용카드 매출은 3월 소비 둔화 시사



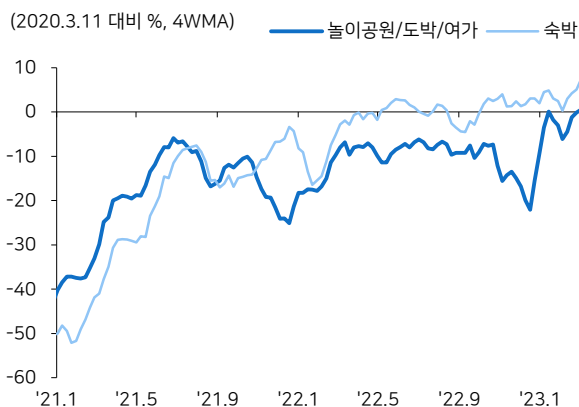
자료: US BEA, 메리츠증권 리서치센터

그림12 미국 월간 소매판매와 주간 신용카드 매출 상관성



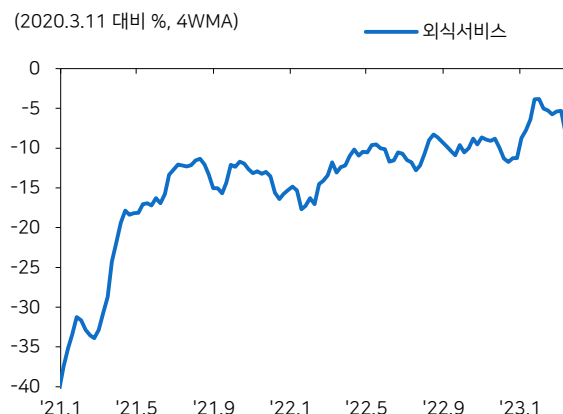
자료: US BEA, US Census Bureau, 메리츠증권 리서치센터

그림13 리오프닝 관련 신용카드 매출은 COVID 이전 상회



자료: US BEA, 메리츠증권 리서치센터

그림14 1월 반짝 올랐던 외식서비스 매출은 다시 둔화



자료: US BEA, 메리츠증권 리서치센터

이미 완화적으로 반영하고 있는 통화정책 기대

1월과 차이는 통화정책 기대가
쏟아지는 방향

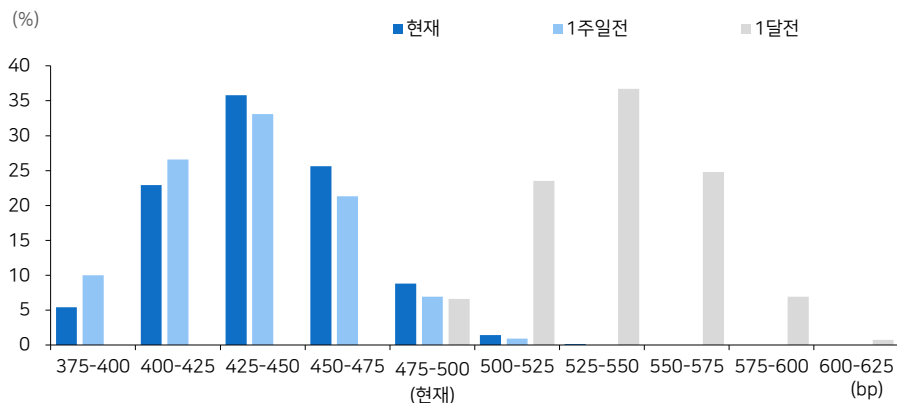
4월 증시가 마주하고 있는 상황을 1분기 초와 비교할 때 다른 점이 있다면 통화정책 기대가 쏟아지는 방향이다. 작년 12월 말 증시를 상기해보면 경기 하강 사이클 가운데 실적 시즌을 앞두고 있었고, 통화정책 기대는 긴축 우려를 더 반영하고 있었다. 1월 증시에서 12MF EPS가 지속적으로 하향했음에도 불구하고 증시가 반등한 배경에는 1) 통화정책의 완화기대 확산과 2) 일본발 유동성 효과가 있었다.

현재는 완화 기대로 쏟아져 있어 추
가인 경기 하방 우려가 완화 기대
강화로 이어지기 보다 실적 눈높이
하향으로 주가에 작용할 소지

1월과 지금 가장 큰 차이는 통화정책 기대가 쏟아지는 위치다. 현재는 이미 9월 FOMC에서 50bp 금리인하 가능성까지 기대를 녹이고 있다. 추가적인 경기 하방 압력과 인플레이션 완화가 확인되더라도 정책기대를 더 낮추면서 시장이 상승할 가능성은 높지 않아 보인다. 여기에 더해 이미 미국 경기가 위축 국면에 위치하고 있다는 사실을 다 알고 있지만, 은행 규제라는 새 변수가 추가적으로 기업들의 투자심리를 위축시키게 된다면 실적발표 기간 중 가이던스 하향으로 증시 하방압력으로 작용할 소지가 있다.

따라서 Global Catch Mind 4월호에서 제시하고 있는 바와 같이 4월 실적발표를 앞두고 유동성 우려가 상대적으로 적은 우량기업 중심의 안정적인 투자를 하는 것이 적절하다는 의견을 유지한다.

그림15 CME Fed Watch 12월 FOMC 정책금리 기대: 현재 대비 50bp 인하 수준까지 반영



주: 4월 4일 기준
자료: CME Group, 메리츠증권 리서치센터

칼럼의 재해석

장재혁 연구원

GPT(Generative Pre-trained Transformer)는 GPT(General Purpose Technology)인가? (OpenAI)

범용목적기술, GPT(General Purpose Technology)는 증기기관, 인터넷, 전기와 같이 여러 산업 분야에 걸쳐 혁신을 촉진하는 기술로, 경제 성장과 사회 변화의 핵심 동력이다. OpenAI는 LLM(Large Language Model)이 GPT로 분류되기 위한 세 가지 핵심조건(시간 경과에 따른 개선, 경제 전반에 걸친 보급, 보완적 혁신 유발)을 충족함을, LLM이 노동 시장에 미칠 수 있는 경제적 영향을 분석한 연구를 통해 증명한다.

연구에 따르면, 평균적으로 어떤 직업 내 작업의 15%가 LLM에 노출되어 있으며, 노출되어 있다는 것은 '해당 작업 수행능률의 100% 향상' 또는 '해당 작업 완료 소요시간의 50% 감소'를 의미한다. 직접적 LLM 접근에 더해 LLM 기반의 보완적 SW 접근을 허용한다면 노출도는 30~56%로 증가한다. LLM은 그 자체로도 노동을 일부 대체할 수 있으나, LLM으로부터 파생되는 SW의 혁신이 훨씬 더 광범위하고 지대한 영향을 미칠 수 있다는 의미이다. 또한, LLM 직업의 진입 장벽과 소득 수준이 높을수록 LLM 노출도가 높게 나타난다.

역사적으로 범용목적기술은 높은 수준의 생산성 향상을 이끌어내며, 경제 전반에 걸쳐 새로운 산업과 기회를 창출해 왔다. ChatGPT의 출시가 단 4개월 지난 이 시점, 장기적으로 일어날 변화들의 초기 신호에 촉각을 곤두세워야 한다.

GPT(General Purpose Technology)

범용목적기술의 조건

- 1) 시간 경과에 따른 개선
- 2) 경제 전반에 걸친 보급
- 3) 보완적 혁신 유발

범용목적기술(General Purpose Technology, 이하 GPT)은 여러 산업 분야에 걸쳐 혁신을 촉진하는 기술로, 경제 성장과 사회 변화의 핵심 동력이다. 대표적인 예로는 증기기관, 전기, 인터넷 등이 있다. 이러한 기술들은 높은 수준의 생산성 향상을 이끌어내며, 경제 전반에 걸쳐 새로운 산업과 기회를 창출한다. 새로운 GPT가 경제 변화와 장기적 경제 성장에 어떤 영향을 미치는지에 대한 이론과 연구를 다룬 저서 'Economic Transformations(Richard G. Lipsey 저)'에 따르면, 다음 세 가지 핵심조건을 충족시킨다면 GPT로 분류될 수 있다고 언급한다.

- 1) 시간 경과에 따른 개선 (Improvement over time)
- 2) 경제 전반에 걸친 보급 (Pervasiveness throughout economy)
- 3) 보완적 혁신 유발 (Ability to spawn complementary innovations)

OpenAI는 본문에서 소개할 논문 "GPTs are GPTs: An Early Look at the Labor Market Impact – Potential of Large Language Models"을 통해 거대 언어 모델(Large language models, 이하 LLM)이 노동 시장에 미칠 수 있는 경제적 영향력을 분석하고 GPT로 분류될 수 있음을 주장한다.

표1 과거의 GPT(범용목적기술), 경제 전반에 걸쳐 새로운 산업과 기회를 창출. LLM(거대언어모델)은 새로운 GPT가 될 것인가?

GPT (General-purpose technology)	Spillover Effects	Date	Classification
Electricity	centralized power generation, factory electrification, telegraphic communication	19 th century	product
Computer	Digital Revolution, Internet	20 th century	product
Internet	electronic business, crowdsourcing, social networking, information warfare	20 th century	product
Biotechnology	genetically modified food, bioengineering, gene therapy	20 th century	process
Nanotechnology	nanomaterials, nanomedicine, quantum dot solar cell, targeted cancer therapy	21 st century	product
GPT (Generative pre-trained transformer) & LLM (Large Language Model) ?		21 st century	

자료: Lipsey and Carlaw(2005), 메리츠증권 리서치센터

OpenAi, LLM이 노동 시장에 미칠 영향 연구

OpenAi의 연구,
직업별 LLM 노출도를
인간 응답자와 GPT4를 통해 측정

OpenAi는 LLM (해당 논문에서는 LLM과 GPT-generative pretrained transformer를 어느 정도 교환 가능한 의미로 사용, 이들을 ChatGPT 또는 OpenAI Playground를 통해 사용할 수 있는 GPT 계열 모델과 유사하게 간주) 이 노동 시장에 미칠 영향에 대해 연구하기 위해 Bottom-up 접근을 사용한다. 1,016개의 직업군과 각각의 직업에 대응하는 2,087종류의 DWA(Detailed Work Activities), 19,265종류의 task로 이루어진 데이터를 이용, 인간 응답자와 GPT-4에게 모든 task · DWA · 직업의 LLM 노출도(E)를 측정하게 한다.

표2 OpenAi의 연구에 사용된 직업 데이터셋 예시

Task ID	Occupation Title	DWAs	Task Description
14675	Computer Systems Engineers/Architects	Monitor computer system performance to ensure proper operation.	Monitor system operation to detect potential problems.
18310	Acute Care Nurses	Operate diagnostic or therapeutic medical instruments or equipment. Prepare medical supplies or equipment for use.	Set up, operate, or monitor invasive equipment and devices, such as colostomy or tracheotomy equipment, mechanical ventilators, catheters, gastrointestinal tubes, and central lines.
4668.0	Gambling Cage Workers	Execute sales or other financial transactions.	Cash checks and process credit card advances for patrons.
15709	Online Merchants	Execute sales or other financial transactions.	Deliver e-mail confirmation of completed transactions and shipment.
6529	Kindergarten Teachers, Except Special Education	-	Involve parent volunteers and older students in children's activities to facilitate involvement in focused, complex play.
6568	Elementary School Teachers, Except Special Education	-	Involve parent volunteers and older students in children's activities to facilitate involvement in focused, complex play.

자료: O*Net, OpenAi, 메리츠증권 리서치센터

노출의 종류는 아래와 같이 E0 · E1 · E2 세 가지로 구분되며, 응답자 또는 GPT4는 각각의 task · DWA · 직업에 대해 셋 중 하나로 라벨링한다.

E0: LLM 또는 LLM 기반 SW 사용으로 작업 효율을 전혀 높일 수 없는 경우

E1: LLM기반의 추가 SW(ex. MS Edge Chat 기능)가 아닌, LLM(ex. ChatGPT)에 직접 액세스하는 것만으로도, 그렇지 않았을 때와 비교해서 동등한 품질의 작업 완성도 달성에 소요 시간이 50% 감소하는 경우

E2: LLM에 직접 액세스하는 것만으로는 작업 완료에 걸리는 시간을 50%까지 줄일 수 없으나, LLM 기반의 추가 SW를 사용했을 때 작업 완료 소요 시간을 50% 이상 줄일 수 있는 경우

노출도 측정값, LLM에 대한 직접적 접근과 LLM 기반 SW에 대한 추가적 접근을 구분, 가중치를 적용

노출 측정값(α, β, ζ)은 E1 · E2의 가중치에 따라 달라진다. E1, LLM에 대한 직접적인 접근만을 고려한 α 는 한 직업 내에서 LLM에 노출된 업무 비율의 하한으로 해석할 수 있다. 반대로 ζ 는 E1과 E2의 합으로, LLM과 LLM 기반 SW에 대한 최대 노출을 고려한 LLM 노출도의 상한 수준으로 해석할 수 있다. OpenAi는 E2에 0.5의 가중치를 둔 β 를 기본값으로 사용한다. 즉, LLM에 직접 액세스하는 것과 비교했을 때, LLM 기반의 SW가 추가로 필요한 작업의 노출도는 절반 수준으로 간주하는 것이다.

인간 응답자와 GPT4의 직업 및 작업의 LLM 노출도에 대한 평가, 상당 부분 일치

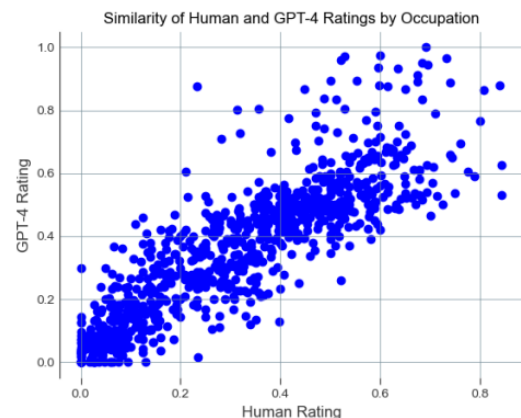
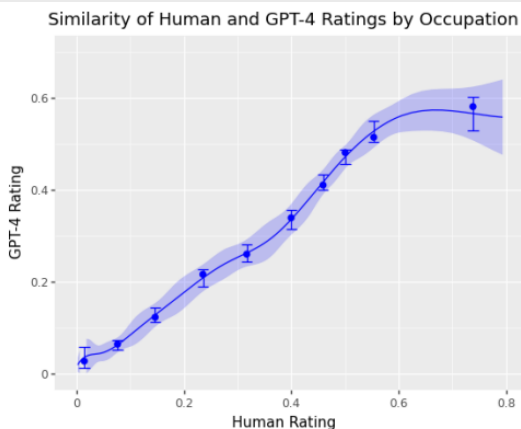
한 가지 흥미로운 사실은, 인간 응답자들과 GPT4의 설문 결과를 비교했을 때 특정 직업의 LLM 노출도에 대해 인간 응답자와 GPT4의 의견이 상당 부분 일치한다는 것이다. 또한, 노출도가 최상단에 가까울수록 평균적으로 인간 응답자가 노출도를 과대평가할 가능성이 높은 것으로 나타났다.

표3 인간 응답자와 GPT4의 task · DWA · 직업에 대한 LLM 노출도 평가는 대체로 일치

Comparison	γ	Weighting	Agreement	Pearson's
GPT-4, Rubric 1; Human	α	E1	80.8%	0.223
	β	E1 + .5*E2	65.6%	0.591
	ζ	E1 + E2	82.1%	0.654
GPT-4, Rubric 2; Human	α	E1	81.8%	0.221
	β	E1 + .5*E2	65.6%	0.538
	ζ	E1 + E2	79.5%	0.589
GPT-4, Rubric 1; GPT-4, Rubric 2	α	E1	91.1%	0.611
	β	E1 + .5*E2	76.0%	0.705
	ζ	E1 + E2	82.4%	0.680

자료: OpenAi, 메리츠증권 리서치센터

그림1 인간 응답자와 GPT4의 task · DWA · 직업에 대한 LLM 노출도 평가는 대체로 일치



자료: OpenAi, 메리츠증권 리서치센터

어떤 직업 내 작업의 최소 15%가 LLM에 노출

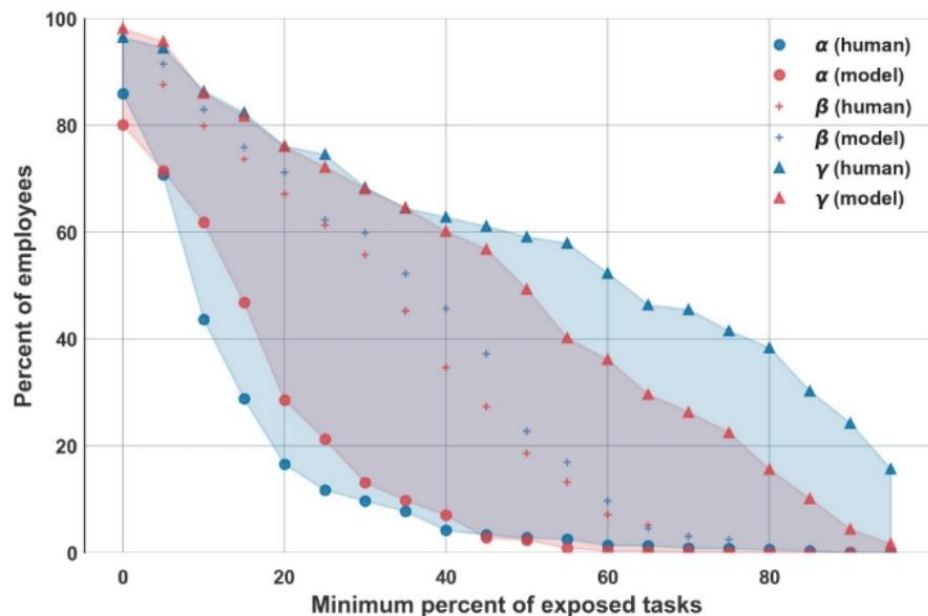
위의 방법으로 측정한 결과는 아래 그림4와 같다. 인간 응답자와 GPT4의 응답 모두 평균적인 직업의 LLM 노출 측정값 ' α '가 0.14~0.15임을 나타내며, 이는 평균적으로 어떤 직업 내 작업의 15%가 LLM에 노출되어 있음을 의미한다. 노출도는 ' β '의 경우 30% 이상으로 증가하고, ' ζ '의 경우 50%를 초과한다. ' β '를 기준으로 할 경우, 작업의 10% 이상이 LLM에 노출된 직종에 속하는 근로자는 전체의 80%, 작업의 절반 이상이 노출된 직종에 속하는 근로자는 전체의 19% 수준인 것으로 예상된다.

표4 0.15 ~ 0.56에 해당하는 노출 측정값(α , β , ζ), 평균적으로 각 직업의 작업들 중 15~56%가 LLM에 노출되었다는 의미

Occupation Level Exposure					Task Level Exposure				
	Human		GPT-4			Human		GPT-4	
	mean	std	mean	std		mean	std	mean	std
α	0.14	0.14	0.14	0.16	α	0.15	0.36	0.14	0.35
β	0.30	0.21	0.34	0.22	β	0.31	0.37	0.35	0.35
ζ	0.46	0.30	0.55	0.34	ζ	0.47	0.50	0.56	0.50

자료: OpenAi, 메리츠증권 리서치센터

그림2 β 와 α 의 차, 또는 ζ 와 α 의 차는 LLM 기반 SW의 접근성 획득으로 인해, 추가로 침투할 수 있는 노동 시장의 영역을 의미

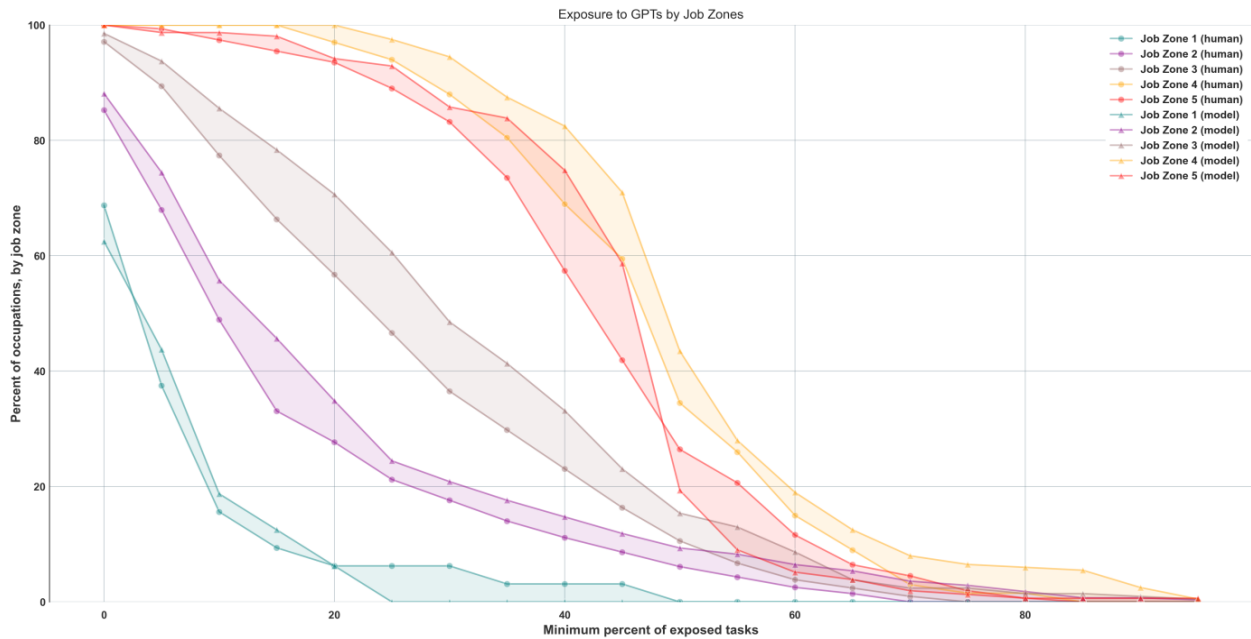


자료: OpenAi, 메리츠증권 리서치센터

직업의 진입 장벽이 높을수록,
소득 수준이 높을수록
LLM 노출도가 높게 나타남

추가적으로 수행된 직종 그룹별 노출도 분석은 특히 인상적이다. OpenAi는 직업 유형에 따른 노출도 차이 분석을 위해 직업별 진입 장벽(해당 직업을 얻기 위해 필요한 교육 수준, 경험의 양, 실무경험 기간)에 따라 [표5]와 같이 직종을 5개로 구분했다. 연구 결과에 따르면, [그림3]에서 확인할 수 있듯 직업의 진입 장벽이 높을수록 LLM 노출도가 높게 나타나며, 이 같은 결과는 소득 수준에도 같이 적용할 수 있다. 높은 수준의 학위와 오랜 기간의 준비가 필요하면서 고소득인 직업일 수록 LLM에 대한 노출도가 높은 것이다.

그림3 진입 장벽이 높은 직종 그룹(Zone1 < Zone5)에서 LLM 노출도가 더 높은 것으로 나타남



자료: OpenAi, 메리츠증권 리서치센터

표5 진입 장벽에 따라 분류한 직종 그룹, 진입 장벽과 소득 수준이 높을수록 LLM에 대한 노출도는 증가

Job Zone	Preparation Required	Education Required	Example Occupations	Median Income	Tot Emp (000s)	H α	M α	H β	M β	H ζ	M ζ
1	None or little (0-3 months)	High school diploma or GED (optional)	Food preparation workers, dishwashers, floor sanders	\$30,230	13,100	0.03	0.04	0.06	0.06	0.09	0.08
2	Some (3-12 months)	High school diploma	Orderlies, customer service representatives, tellers	\$38,215	73,962	0.07	0.12	0.16	0.20	0.24	0.27
3	Medium (1-2 years)	Vocational school, on-the-job training, or associate's degree	Electricians, barbers, medical assistants	\$54,815	37,881	0.11	0.14	0.26	0.32	0.41	0.51
4	Considerable (2-4 years)	Bachelor's degree	Database administrators, graphic designers, cost estimators	\$77,345	56,833	0.23	0.18	0.47	0.51	0.71	0.85
5	Extensive (4+ years)	Master's degree or higher	Pharmacists, lawyers, astronomers	\$81,980	21,221	0.23	0.13	0.43	0.45	0.63	0.76

자료: OpenAi, 메리츠증권 리서치센터

LLM, 대체한 만큼의 경제적 가치 창출

위 연구 결과에서 LLM에 대한 노출도는 ‘LLM 또는 LLM 기반 SW의 활용으로 능률을 100% 향상(또는 작업 완료 소요 시간을 50% 단축)시킬 수 있는지’에 대해 초점이 맞추어져 있다. 이를 ‘AI에 의한 대체 가능성’으로 해석할지, ‘AI 활용으로 인한 생산성 향상’으로 해석할지 선택하는 것은 각자의 몫이나, ‘LLM 또는 LLM 기반 SW’가 그만큼의 경제적 가치를 창출할 것이라는 사실은 변함없다.

GPT(Generative Pre-trained Transformer)는 GPT(General Purpose Technology)이다

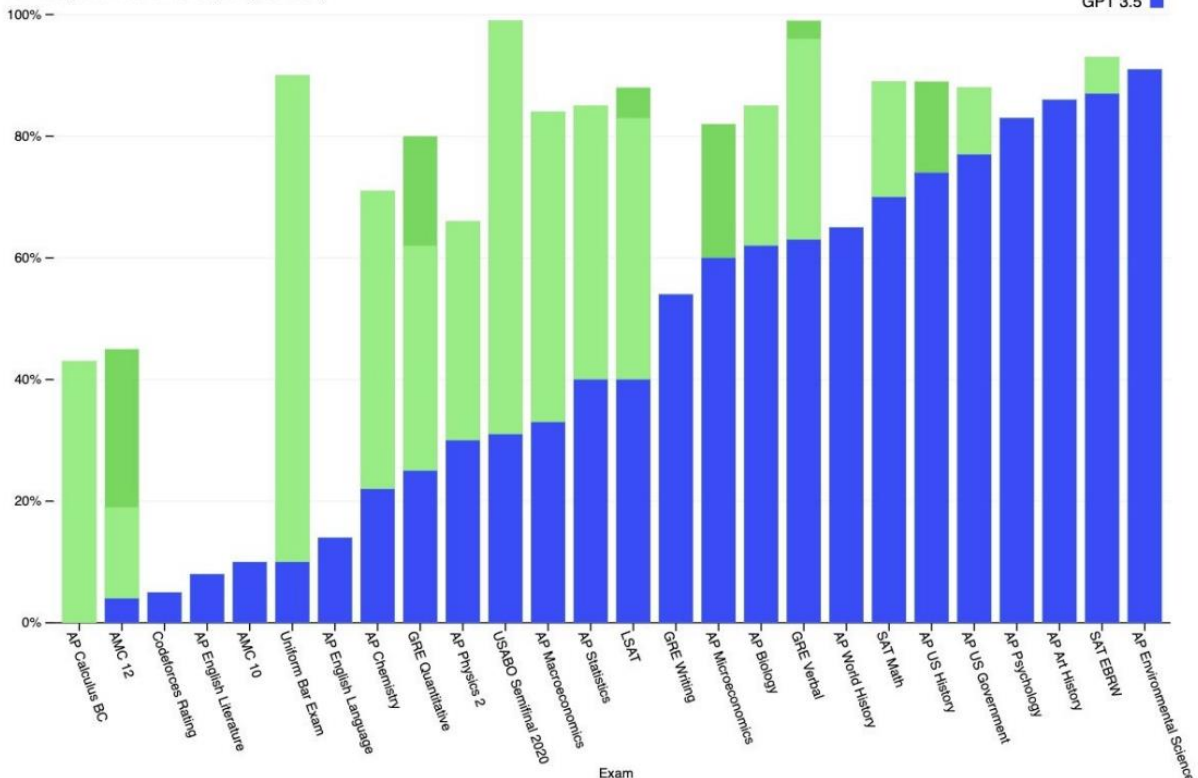
LLM, 시간 경과에 따른 개선

LLM은 서두에서 제시했던 범용목적기술(General Purpose Technology)이 되기 위한 세 가지 조건 중 첫 번째의 조건(시간 경과에 따른 개선)을 이미 충족한다. GPT(Generative pre-trained transformer)는 2018년 6월 11일 최초로 배포된 이래로(GPT-1), 2019년 11월 GPT-2, 2020년 5월 GPT-3로 업그레이드되며 성능 개선을 거듭했고, 2022년 11월 GPT-3.5는 ChatGPT의 형태로 대중에게 공개된 이후 2023년 3월 GPT-4를 공개하며 말 그대로 ‘General Purpose’를 위한 LLM임을 입증했다. [그림4]에서 확인할 수 있듯이, GPT-3.5보다 4개월 뒤에 공개된 GPT-4는 더 많은 영역에서 사람과 비교했을 때 상위 20% 이내의 성취도를 기록할 수 있었고, Vision 기능이 추가되었을 때는 더욱 우수한 성취도를 보여준다. 시간 경과에 따른 LLM의 개선 속도는 점점 가팔라지고 있다.

그림4 GPT-3.5가 출시(2022.11)된 이후 4개월만에 출시된 GPT-4, 다방면 영역에서 급격한 성능 개선

Exam results (ordered by GPT 3.5 performance)

Estimated percentile lower bound (among test takers)



자료: OpenAi, 메리츠증권 리서치센터

OpenAi, LLM 또는 GPT가 범용목적기술임을 주장

OpenAi는 이번 연구를 통해 LLM이 범용목적기술이 되기 위한 두 번째(경제 전반에 걸친 보급)와 세 번째(보완적 혁신 유발) 조건을 충족한다는 증거를 제시한다. 특히 LLM은 그 자체로도 경제 전반에 걸친 광범위한 영향을 미칠 수 있고, LLM 기반의 SW나 디지털 도구를 통한 보완적 혁신 또한 경제 활동에 광범위한 영향을 줄 수 있다.

LLM, 보완적 혁신 유발

그림 4와 5는 LLM 기반의 SW나 디지털 도구를 통한 보완적 혁신의 잠재적인 경제적 파급력을 보여주는 예시가 될 수 있다. 연구 결과에 따르면 직접적인 LLM 노출도($\alpha=0.14\sim0.15$)보다 LLM 기반의 SW를 활용했을 때의 노출도($\beta=0.46\sim0.56$)가 큰 폭으로 우세하며, 그 차이($\beta-\alpha=0.31\sim0.42$)는 LLM 노출도 그 자체보다도 두 배 이상 크다. 다시 말해서, LLM은 그 자체로도 노동을 일부 대체할 수 있으나, LLM으로부터 파생되는 SW의 혁신이 훨씬 더 광범위하고 지대한 영향을 미칠 수 있다.

LLM, 경제 전반에 보급

경제 전반에 걸친 보급은 기업 또는 사용자들의 채택 수준을 따른다. OpenAi GPT-3.5 기반의 ChatGPT는 출시 2개월만인 1월에 월간 MAU(monthly active users) 1억명을 돌파했으며, 역사상 가장 빠르게 MAU가 성장한 어플리케이션이 되었다. 앞선 연구에서 확인한 노동력 대체 가능성·생산성 향상을 고려한다면 향후 경제 전반에 걸친 보급이 이루어질 개연성은 충분하다.

GPT=GPT

따라서, OpenAi의 연구 결과는 ChatGPT가 범용목적기술로서 경제 변화에 영향을 줄 가능성이 높다는 것을 시사한다. 이러한 기술적 발전이 노동 생산성의 증가와 경제 구조의 변화를 초래할 수 있으며, 새로운 비즈니스 모델과 시장 창출, 사회 경제학적 변화 등의 장기적인 영향을 미칠 수 있다. ChatGPT의 출시가 단 4개월 지난 지금, 장기적으로 일어날 변화들의 초기 신호에 촉각을 곤두세워야 할 이유이다.

원문: GPTs are GPTs: An Early Look at the Labor Market Impact - Potential of Large Language Models (OpenAI)