

# 노인 빈곤에 관한 연구: 소득과 소비를 중심으로 출장보고서

2023. 8. 30

KDI 재정·사회정책연구부

## 1 출장 배경 및 목적

### □ (출장 배경)

- “노인 빈곤에 관한 연구 : 소득과 소비를 중심으로” 과제는 고령층의 물질적인 삶을 측정할 수 있는 소득과 소비를 이용하여 고령층의 빈곤 상태를 분석하여 정책적 함의를 도출하는 것을 목표로 하고 있음.
- 소득과 소비를 모두 이용하여 빈곤율을 산정하는 것에는 다양한 방법들이 존재할 수 있고 이 중 소득과 소비의 결합 분포로부터 빈곤율을 산정하는 방법을 과제에 적용하여 완성도를 제고하고자 함.

### □ (출장 목적)

- 일본 도쿄에서 개최되는 6th International Conference on Econometrics and Statistics (EcoSta2023) 국제학술대회에 참석하여 최신 통계학/계량경제학 분석 방법을 학습하고자 함.
- 학습하고자 하는 함수적 데이터 분석(functional data analysis)은 횡단면 데이터 분석과 시계열 분석에 모두 활용이 가능해 향후 연구 과제에 활용도가 높을 것으로 판단됨.

## 2 출장 개요

### □ 출장지 및 기간

- 출장지: 와세다대학교 (일본, 도쿄)
- 출장기간: 2023년 8월 1일(화) ~ 3일(목)

### □ 출장자 (총 1명)

| 소속             | 성명  | 직위    |
|----------------|-----|-------|
| KDI 재정·사회정책연구부 | 이승희 | 부연구위원 |

□ 주요활동

- 6th International Conference on Econometrics and Statistics (EcoSta2023) 학술대회에 참석하여 함수적 데이터 분석 방법과 Nowcasting을 포함한 최신 시계열 분석 방법 학습

**3 출장 일정**

| 일자     | 주요일정(안)   |
|--------|---|
| 8.1(화) | [09:00-11:10] (출국) 김포 → 일본 하네다(KE2101)  |
|        | [13:20-15:00] (세션 EO240) Advances in functional data analysis and their applications                                  |
|        | [15:10-16:00] (Keynote Talk1) Fusion learning: Combining inferences from diverse data sources with heterogeneous data |
| 8.2(수) | [16:30-18:35] (세션 EC317) Functional data analysis   |
|        | [08:15-10:20] (세션 EO169) Causal machine learning with high dimensional modelling                                      |
|        | [10:50-12:30] (세션 EO098) Dynamic topological data analysis on time series data  |
|        | [14:00-15:40] (세션 EO179) Snapshot on current functional data methodologies  |
|        | [16:10-17:50] (세션 EO200) Advances in time series econometrics   |
| 8.3(목) | [18:00-19:15] (세션 EO066) Recent advances in econometrics  |
|        | [07:30-09:10] (세션 EO182) Recent advances in nowcasting  |
|        | [09:20-11:00] (세션 EO229) New advances in functional data analysis   |
|        | [11:30-12:20] (Keynote Talk2) Network approaches in healthcare, business and social sciences                          |
|        | [13:40-14:55] (세션 EO058) Causal inference   |
|        | [15:25-17:05] (세션 EV274) Modelling and forecasting  |
|        | [19:45-22:05] (귀국) 일본 하네다 →김포(KE2104)   |

**4 활동 내용**

□ 과제 관련 주요 내용

- 소득과 소비의 결합 분포를 이용하여 함수적 데이터 분석의 방법들을 응용해서 노인빈곤율을 구하는 것은 유의미하지 않을 수 있음.
  - 우리나라의 노인 빈곤의 경우 소득 빈곤과 소비 빈곤간의 상관관계가 굉장히 높을 수 있는데 그렇다면 두 빈곤 집합 간의 교집합이 더 빈곤을 잘 나타낼 수 있음.
- 오히려 소득과 소비의 분포가 시간에 따라 어떻게 변하는지를 함수적 데이터 분석 기법과 시계열 분석 기법을 결합해서 분석해볼 수 있음.
  - 이와 같은 분석을 통해 우리나라의 과거 급격한 경제 성장으로 인한 양극화와 노인빈곤 간의 관계를 추론해 볼 수 있음.

□ 학술회의 참석 세션 내용

- 흥미롭고 앞으로 연구 과제에 활용도가 높을 것으로 생각되는 논문들 위주로 아래 정리함.

**1. 6th International Conference on Econometrics and Statistics (8월 1일)**

- [13:20-15:00] (세션 EO240) Advances in functional data analysis and their applications
  - Spatiotemporal factor models for functional data with application to population map forecast
    - 베이저안 요인 모형을 이용하여 시계열적 분석은 공통 요인들을 중심으로, 공간적 구조를 요인 적재값(loading)으로 모형화한 함수적 데이터 분석 모형을 활용하여 도쿄의 각 구역 인구가 시간마다 어떻게 변

화하는지를 예측

- A functional generalized additive model-based scan statistic for disease cluster detection

- 과거의 정보들을 여러 개의 함수적 데이터로 보고 이들을 활용하여 공간적 스캔 통계량을 구축하여 집단감염을 예측하고 이를 활용하여 집단감염 발생에 관한 변수와 다른 함수적 데이터들의 공간적 상관관계를 기반으로 통계량을 구축하고 예측에 활용

- [15:10-16:00] (Keynote Talk1) Fusion learning: Combining inferences from diverse data sources with heterogeneous data

- Liu (1990)의 data depth의 개념을 이용하여 data depth, bootstrap, confidence distribution을 결합한 depth confidence distribution을 제안하고 이를 바탕으로 하나의 데이터소스가 아닌 여러 개의 데이터소스를 결합하여 연구를 진행할 때 어떻게 추론을 할 수 있는지를 연구

- [16:30-18:30] (세션 EC317) Functional data analysis

- Additive regression with general imperfect variables

- 종속변수가 일반적인 힐버트 공간 상의 변수이고 설명변수가 일반적인 다변량 변수이지만 두 변수 모두 불완전하게 관측되는 경우 이들을 모형화하는 방법에 대해 연구

- Functional principal component analysis for partially observed elliptical process

- 함수적 데이터가 일부분만 관측가능할 때 함수적 주성분 분석을 관측이 되지 않은 부분의 데이터 관측치를 예측하는 방법을 제시하고 이를 공기 오염 데이터에 적용하여 분석함.

## 2. 6th International Conference on Econometrics and Statistics

(8월 2일)

- [08:15-10:20] (세션 EO169) Causal machine learning with high dimensional modelling

- Limit theorems for semi-discrete optimal transport maps

- 절대 연속형 분포로부터 타겟인 유한 이산형 분포로의 최적 운송(OT) 맵에 대한 통계적 추론을 연구함. 이 경우의 한계분포(limiting distribution)를 구하고 비모수 부트스트랩의 일관성을 밝힘.

- Optimal nonparametric inference with two-scale distributional nearest neighbors

- 배깅(bagging)은 자동으로 분포적 nearest neighbors(DNN) 추정을 하는 일반적인 방법이나 통계적 추론에 적용이 제한됨. 이를 극복하기 위해 서로 다른 샘플링 스케일을 가진 두 DNN을 선형 조합하는 방법을 제안하고 해당 방법에 대한 통계적 추론을 연구

- [10:50-12:30] (세션 EO098) Dynamic topological data analysis on time series data

- Dynamic topological data analysis of time-series data

- 토폴로지 데이터를 이용한 시계열 분석 방법을 연구함. 토폴로지 데이터는 주로 이미지와 같은 공간적 데이터에 많이 활용되어 왔는데, 접근 방법은 고차원 데이터 분석 방법과 유사한 점이 많은 것으로 보임. 연구에서 데이터를 푸리에 변환해서 적용하였는데 이는 계산적인 복잡성이 낮고 추론을 쉽게 할 수 있다는 장점이 있음.

- Topological state-space estimation of dynamically changing functional human brain networks

- 토폴로지 데이터를 시계열, 그 중에서도 상태 공간 모형에 적용하여 뇌의 자기공명영상(MRI)을 통해 뇌 네트워크의 동태를 분석함. 이 방법은 동적으로 변화하는 기능적 뇌 네트워크의 상태 공간을 정확하게

파악하여 기존 k-평균 군집 방법보다 우월함을 보임.

□ [14:00-15:40] (세션 EO179) Snapshot on current functional data methodologies

○ Functional predictor selection and its nonasymptotic behavior

- 종속변수가 스칼라이고 설명변수가 함수적 데이터(scalar-on-function)인 회귀 모형에서 group LASSO 타입의 별점을 이용한 추정 방법을 제시함. 특히 최근의 머신러닝 연구 분야에서는 한계 분포와 같은 점근적 성질보다 비점근적 특성을 많이 연구하는데 이와 유사하게 함수적 회귀 분석 추정치에 대한 비점근적 특성을 연구

○ Geometric and topological perspectives on unsupervised functional data analysis

- 군집 분석과 이상 탐지(anomaly detection)는 함수적 데이터 분석에서 많이 활용되고 있는 비지도 학습(unsupervised learning) 방법이고, 최신 군집 분석 알고리즘을 함수적 데이터에 적용하여 이상 탐지 수행력(performance)가 높음을 보임.

□ [16:10-17:50] (세션 EO200) Advances in time series econometrics

○ Instrumental factor models for high-dimensional functional data

- 도구적 요인 모형을 함수적 데이터에 적용하는 새로운 모형을 제시함. 기존의 함수적 주성분 분석의 경우 횡단과 시간 차원 모두에서 충분한 데이터가 필요하지만 도구적 요인 모형에서는 횡단 차원에서만 충분한 데이터만 있으면 추정치가 일치성을 보임. 이를 이용하여 지구온난화와 세계 GDP간의 관계를 분석

○ Inference on nonstationarity and common stochastic trends in high-dimensional or functional time series

- 단위근이나 공적분과 같이 비정상성 시계열 분석 방법을 함수적 데이터 분석에 적용하는 이론을 연구함. 함수적 데이터를 비정상성 시계열

에 적용하는 경우 확률적 추세의 개수가 몇 개인지를 파악하는 것이 중요한데, 점근적으로 확률적 추세 수에 대한 통계적 추론을 가능하게 하는 방법을 제시

□ [18:00-19:15] (세션 EO066) Recent advances in econometrics

○ A dynamic semiparametric characteristics-based model for optimal portfolio selection

- 최적 포트폴리오 선택 문제를 두 단계로 나누어서 분석함. 첫 번째 단계에서는 특성 기반 요인 적재값을 이용하여 포트폴리오를 구성하고, 두 번째 단계에서는 포트폴리오를 동적으로 결합하는 방법을 이용함. 이를 CRSP와 FRED 데이터에 적용하여 투자자의 위험 회피 수준과 일치하는지를 분석

○ Nonparametric range-based estimation of integrated variance with episodic extreme return persistence

- 고빈도 자료를 이용하여 단기간 내에서의 고, 저, 시가 및 종가 정보를 기반으로 하여 변동성을 추정하는 방법을 제시하고 이를 변동성을 실증적으로 분석하는 데 우수한 HAR(Heterogeneous AutoRegressive) 모형과 결합하여 변동성의 표본 외 예측에서 우수함을 보임.

### 3. 6th International Conference on Econometrics and Statistics (8월 3일)

□ [07:30-09:10] (세션 EO182) Recent advances in nowcasting

○ Panel data nowcasting: The Case of price-earnings ratios

- 혼합 주기로 구성된 패널 데이터의 nowcasting을 위해 기계 학습 회귀 분석 방법을 적용하여 기업의 실적을 예측하는 방법을 제시함. 구체적으로 최소 group LASSO를 이용하여 패널 데이터와 혼합 주기가 혼재된 상태의 회귀분석을 수행하였고 실증적으로 기존의 일반적인 회귀분석 방법이나 머신러닝 방법보다 예측력이 높음을 보임.

- Back to the present: Learning about the Euro Area through a now-casting model

- New York Fed의 Nowcasting 모형을 개발한 Domenico Giannone의 최근 연구로 유로 지역을 독일, 프랑스, 이탈리아, 그 외 국가들로 나누어 경제 상황을 동시에 nowcasting하는 모형을 제시함. 분석 방법은 일반적인 nowcasting과 유사하지만 서로 밀접한 연관이 있는 국가들의 경제상황을 한번에 분석하였을 때 예측력이 더 높았고 기업들을 대상으로 한 설문조사와 같은 연성변수들이 중요함을 강조함.

□ [09:20-11:00] (세션 EO229) New advances in functional data analysis

- Approximation, estimation and inferential theory for locally stationary functional time series

- 함수적 데이터를 시계열 분석 방법을 통해 분석하는 새로운 방법을 제시함. 구체적으로 국소 함수 자기회귀 모형을 제시하였는데 추정치의 분포가 정규분포에 가까워지고 곱셈 형태의 부트스트랩을 적용할 수 있음을 보임.

- Dynamic modelling for multivariate functional and longitudinal data

- 여러 확률 과정 간의 동적 상호 작용을 모형화하여 다변량 확률 과정과 그 도함수의 동적 관계 분석, 예측함. 특히 다변량 데이터를 처리하기 위해 이러한 다변량 과정의 도함수를 이용하여 동적 모형화함.

□ [11:30-12:20] (Keynote Talk2) Network approaches in healthcare, business and social sciences

- 경제학에서도 네트워크 구조를 이용하여 현실 문제를 분석하는 것에 관한 관심이 증가하고 있는데 구체적으로 적용 가능한 분야는 일반적인 개인들의 관계 구조와 관련된 네트워크를 분석하거나 기업 간 관계를 통한 위험 추정, COVID-19 위험 등이 있음. 특히 동적 모형화를 위해서는 인간 행동, 금융 및 팬데믹 위험을 추적하여 모형화할 수 있고 이러한 분석 과정들에서 계량경제학과 통계학이 결합됨.

□ [13:40-14:55] (세션 EO058) Causal inference

- Sparse quantile regression

- 분위수 회귀(quantile regression)에 희소성을 도입하기 위해 best subset regression의 L0 벌점이나 LASSO의 L1 벌점을 적용하는 방법을 연구함. 실증 분석으로는 신생아 출생 체중의 예측과 관련하여 모형의 유용성을 입증함.

- Separable effects under semicompeting risks

- 최근 경쟁 위험 관련 연구들에서는 인과 효과를 보기 위해 효과가 분리가능하다고 가정하는데, 주로 첫 번째 부분은 주요 결과에 직접 영향을 미치는 것이며, 다른 부분은 중간 결과를 통해 주요 결과에 영향을 미치는 것을 가정함. 이러한 가정 하에서 원인별 비모수 및 준모수 추정 방법을 제안하여 인과 효과를 추정하고 그 점근적 특성을 연구함.

□ [15:25-17:05] (세션 EV274) Modelling and forecasting

- Forecasting economic activity with a neural network in uncertain times: Application to German GDP

- 일반화 회귀 신경망 (Generalized Regression Neural Network)을 이용하여 독일 GDP를 예측함. 특히 분기별 독일 GDP 성장을 예측할 때, 일반적인 "보통" 기간과 COVID-19 침체 및 회복기와 같이 시계열적인 동태가 크게 다른 상황을 구별하여 예측하여 COVID-19 위기 중에 기존의 nowcasting 모형 (AR(1), DFM 등)을 능가하는 높은 예측 능력을 보임.

- Step by step: A quarterly evaluation of EU commissions' GDP forecasts

- 유럽 위원회(EU commission)의 연 성장 예측은 정책 결정에 중요한 수치로 분기 예측을 기반으로 하며 EU, 유로 지역 및 여러 EU 회원국에 대한 다가분기별 GDP 성장 예측에 대한 자세한 분석이 제공됨.

이와 같은 예측에 구조적인 결함이 있음을 주장하고 개선의 여지가 있음을 강조함.