

# 如何使用健保資料推估社經變數\*

連賢明\*\*

國立政治大學財政學系教授

近年來使用健保資料庫相關研究日益增加；然而，由於健保資料缺乏病患社經變數，阻礙健保資料在社會學科應用。在主管機關協助下，本文以 2005 年家庭收支調查家戶為樣本，串聯健保承保資料，利用家庭收支調查所記錄豐富社經變數，驗證使用承保資料推估病患戶籍縣市、全職就業、薪資所得以及兒童家庭結構等變數準確性。比較結果發現：(1)使用投保類別為地區人口，農民，與地方公職人員，且投保人和被保險人相同者，其投保單位縣市和戶籍所在縣市相同者達九成五；(2)使用投保人和被保險人相同樣本，在公、勞保投保人的就業準確度相當高，平均達九成以上；(3)使用就業準確樣本推估薪資所得，公保投保者其投保薪資為實際薪資 9 成，勞保民營受僱者投保薪資最低 1/3 投保人，其投保薪資較調查薪資低估 3 成，最高 1/3 投保人投保薪資高約 6%；(4)十五歲以下兒童投保人為父或母時，與投保人同居比率高達 93%；(5)使用承保資料推估十五歲以下手足數目，準確度達九成以上。最後，我們提供一個研究案例，利用承保資料來分析乳癌對就業婦女存活、就業和薪資的衝擊。

關鍵字：健保資料、就業、薪資、家庭結構

# How to Construct Social-Economic Variables from National Health Insurance Data

Hsien-ming Lien

Professor

Department of Public Finance, National ChengChi University

## ABSTRACT

Recently more and more researchers have conducted studies based on National Health Insurance Data (NHID). However, few studies are related to the social sciences due to the lack of social-economic variables in NHID. With the help from the Department of Health, which merged the sample of the 2005 Family Income and Expenditure Survey (FIES) with eligibility files of NHID, this study examines the validity of inferring four variables from eligibility files: residence county, full-time employment, working salary, and children's family structure. Our results show that: (1) The accuracy rate of inferring residence county based on one's county of enrolled units is above 95% for those obtaining coverage from a farmer's association, registration offices, or as public electorates; (2) The accuracy rate of inferring full-time employment status using self-insured status is above 90% for the sample of public employee insurance (PEI) or labor insurance (LI); (3) For PEI sample, the working salary inferred from one's insurance premium is approximately 90% of their reported working salary in FIES. For LI sample, the inferred salary is roughly 30% below the FIES working salary for low-paid workers (bottom 1/3), and 6% higher for high-paid ones (Top 1/3); (4) For children under 15 obtaining coverage through parents, more than 90% of them co-habit with at least one of their parents; (5) The accuracy rate of inferring sibling number and birth order based on eligibility files is at least 90% for children aged below 15. Finally, I use a research example to analyze how the incidence of breast cancer affects the survival, employment, and working salary of employed women.

Key Words: Taiwan National Health Insurance data, employment, working salary, family structure