

모형명세화 오류와 소표본에서 구조방정식모형 모수추정 방법들 비교: 모수추정 정확도와 이론모형 검정력을 중심으로[†]

서동기¹ · 정선호²

¹한림대학교 심리학과 · ²경희대학교 경영학과

접수 2017년 6월 9일, 수정 2017년 9월 20일, 게재확정 2017년 9월 21일

요약

구조방정식모형은 사회과학 및 행동과학 연구 분야에서 이론검정을 위해 주로 사용되는 통계방법이다. 최근 이 통계기법에 대한 방법론적 이슈로서 모형명세화 오류와 소표본 문제가 부각되고 있다. 그런데 이 문제들이 구조방정식모형의 대표 추정 방법인 최대우도법에 의한 이론검정에 어떤 영향을 주는지에 대해 여전히 명확하지 않다. 따라서 본 연구에서 최대우도법 그리고 이에 대한 대안으로 개발된 2단계최소자승법과 2단계능형최소자승법을 정확도와 검정력 관점에서 시뮬레이션을 통해 체계적으로 비교해 본다. 이 실험 결과에 따르면, 모형이 정확하게 설정된 경우, 정확도 기준에서 추정방법들 간의 차이는 미미했다. 하지만 모형오류가 발생한 경우, 2단계능형최소자승법은 다른 방법들보다 표본 크기가 작을 때 훨씬 더 정확한 모수추정치를 산출해 내었다. 그리고 이 방법은 명세화 오류에 관계없이 표본 크기가 작을 때에도 제 2종 오류 (Type II error) 수준이 상대적으로 작거나 만족할 만한 수준의 검정력을 보여주었다. 이에 반해 다른 두 방법들은 표본이 작은 경우 또는 명세화 오류가 있는 경우 상당히 높은 수준의 제 2종 오류를 나타내었다.

주요용어: 구조방정식 모형, 모형명세화 오류, 소표본, 2단계최소자승법, 2단계능형최소자승법, 최대우도법.

1. 서론

구조방정식모형 (structural equation modeling; SEM)은 다양한 연구 분야에서 이론검정 (theory testing)을 위한 분석도구로서 가장 널리 사용되고 있다 (Choi 등, 2013; Hong 등, 2016; Min과 Choi, 2016; Steenkamp와 Baumgartner, 2000). 이와 같은 구조방정식의 대중적 인기의 이유는 응용연구자들이 이론기반의 연구모형 (conceptual model)에 따라 잠재변수 간의 인과관계를 나타내는 잠재변수모형을 유연하고 다양하게 설정할 수 있을 뿐만 아니라 잠재변수와 관찰변수와의 구조적 관계를 나타내는 측정모형을 통해 측정오차 (measurement error)를 효과적으로 통제할 수 있기 때문이다 (Kaplan, 2009).

구조방정식에 대한 응용연구자들의 지속적인 관심과 높은 활용도는 새로운 통계모델의 개발을 촉진해 왔는데, 대표적으로 마케팅 분야에서 소비자 특성의 이질성을 모형에 반영하는 기법 (Ansari 등,

[†] 이 논문은 2014년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(NRF-2014S1A5A2A03064999).

¹ (24252) 강원도 춘천시 한림대학길 1 한림대학교 심리학과, 조교수.

² 교신저자: (02447) 서울특별시 동대문구 경희대로 26 경희대학교 경영학과, 부교수.

E-mail: sunho.jung@khu.ac.kr

2000)이 높은 관심을 받고 있다. 이런 연구뿐만 아니라 부적해 (improper solution)와 비수렴 문제 (Chen 등, 2001)처럼 구조방정식모형의 활용에서 발생하는 기술적 문제를 다루는 연구도 활발하게 진행 중이다.

최근에 가장 활발하게 다루어진 방법론적 이슈들 중 첫 번째는 구조방정식모형에서의 모형명세화 오류 (model misspecification)에 관한 것이다 (Bagozzi와 Yi, 2012; Bollen 등, 2007; Fan과 Sivo, 2005; Jarvis 등, 2003; Olsson 등, 2000; MacKenzie 등, 2005). 이 문제가 중요하게 다루어진 이유는 구조방정식모형이 연구가설 검정을 위한 확인적 분석기법인데 (James 등, 2006) 일반적으로 연구자들이 이와 같은 확인적 기법을 사용할 때 모형명세화의 오류를 사전에 쉽게 파악하기 어렵고 (Jarvis 등, 2003), 사후적으로 구조방정식모형 통계프로그램에서 산출하는 모형 합치도 지수 (goodness-of-fit index) 또한 모형이 자료에 잘 적합한지를 평가할 뿐 모형의 정확성에 대한 정보를 제공하지 못하기 때문이다 (Hayduk 등, 2007). 실제로 모형이 잘못 설정되더라도 수용수준 이상의 높은 합치도 지수를 얻을 수 있다 (Bollen 등, 2007). 그리고 동일한 합치도를 보이는 모형이 존재하는 경우도 있다 (Lee, 2007).

두 번째로, 구조방정식모형을 이용한 실증연구들이 새로운 응용분야로 확장되어 나가면서 기존의 연구에서 심각하게 부각되지 않던 문제들이 생겨나게 되었다. 대표적인 새로운 문제 중의 하나가 소표본에 관한 문제이다 (Tenenhaus, 2008). 구조방정식모형 활용에서 이 문제에 직접적으로 영향을 받는 분야로서 뇌영상 연구 (Sideridis 등, 2014), 신제품개발 연구 (Tenenhaus 등, 2005), 동물행동학 (Budaev, 2010), 생태학 (Grace 등, 2010) 등을 예로 들 수 있다.

구조방정식모형을 활용한 실증연구에서 가장 대표적인 모수 추정방법은 최대우도법 (maximum likelihood; ML)이다. 하지만 ML은 완전정보 (full information) 접근에 따라 모수를 추정하기 때문에 명세화 오류에 취약하다 (Olsson 등, 2000). 따라서 Bollen (1996)은 그 대안으로서 2단계최소자승법 (two stage least squares; 2SLS)을 제안하였다. Jung (2013)은 표본이 작은 경우 구조방정식모형의 모수들을 정확하게 추정할 수 있는 방법으로서 2단계능형최소자승법 (two stage ridge least squares; 2SRLS)을 최근에 개발하였다.

하지만 현재까지 명세화 오류와 소표본 문제가 결합되어 이 세 가지 모수추정 방법에 의한 이론검정에 어떤 영향을 주는지, 언제 심각한 검정 오류가 발생하는지를 평가한 연구를 찾아볼 수가 없다. 따라서 본 연구에서는 각 추정방법을 통해 도출되는 경로계수 (path coefficient) 추정치를 정확도 관점 (Olsson 등, 2000)과 검정력 관점 (Grewal 등, 2004)에 따라 시뮬레이션 기법을 통해 체계적으로 비교해 보고자 한다. 이를 통해 연구자들이 명세화 오류와 소표본 문제 상황 하에서 각각의 추정방법을 정확도와 검정력 관점에 따라 선택할 수 있는 방법론적 가이드라인을 마련하고자 한다.

2. 구조방정식모형에서 2단계최소자승법

현재까지 상당히 많은 구조방정식모형 관련 논문과 서적에서 모수추정 방법인 ML에 대한 이론적 배경과 원리를 자세하게 설명해 왔다 (Bollen, 1989). 그런데 2SLS 추정방법은 현재 구조방정식 프로그램 중 하나인 LISREL에서 사용할 수 있음에도 불구하고 여전히 응용연구자들에게 낯선 방법이다. 따라서 여기서는 2SLS와 2SRLS의 공통된 이론적 배경에 대해서 먼저 중점적으로 논의하고자 한다. 그리고 나서 소표본 문제를 해결하기 위해 개발된 2SRLS의 방법론적 특성에 대해서 다루고자 한다.

2SLS는 ML과 달리 부분정보 (limited information) 추정방법이다. 구조방정식모형은 잠재변수모형과 측정모형으로 구성되어 있는데 2SLS는 이 추정 특성에 따라 두 개의 모형을 각각 따라 추정하게 된다. 응용연구에서 주로 잠재변수모형을 기반으로 가설검정을 다루므로 여기서는 2SLS이 잠재변수모형에서 모수를 추정하는 방법에 대해 논의한다. 하지만 2SLS의 분석방법은 잠재변수모형과 측정모형 모두 동일하다. 먼저 잠재변수모형은 다음과 같이 나타낼 수 있다.

$$\eta = B\eta + \Gamma\xi + \zeta, \quad (2.1)$$

여기서 η 는 내생잠재변수, ξ 는 외생잠재변수, ζ 는 잔차항을 나타낸다. 그리고 B 는 내생잠재변수 간의 구조적인 관계, Γ 는 외생잠재변수가 내생잠재변수에 미치는 영향의 크기와 방향을 뜻하는 경로계수 모수이다. (2.1)에서 모형식별을 위해 일반 구조방정식 프로그램에서 하듯이 먼저 잠재요인을 척도화 하는 것이 필요하다. 이를 위해 각 요인에 해당되는 측정변수 하나를 골라 이에 대한 요인계수를 1로 고정한다. 이 측정변수 (marker variable)의 벡터를 외생변수인 경우 x_1 , 내생변수인 경우 y_1 로 나타낸다. 이는 다음과 같은 관계식으로 정리할 수 있다.

$$\begin{aligned} y_1 &= \eta + \epsilon_1, \\ x_1 &= \xi + \delta_1. \end{aligned} \quad (2.2)$$

그리고 식 (2.2)를 식 (2.1)에 대입하면 이제 잠재변수가 아닌 측정변수 수준에서 잠재변수모형의 모수를 추정해 볼 수 있는 모형으로 변환된다. 이에 따라 식 (2.1)에 해당하는 새로운 연립방정식을 도출할 수 있는데 여기에는 내생잠재변수의 개수만큼의 방정식이 존재한다. 예시로서 내생변수 i 에 해당하는 방정식을 표현하면 다음과 같다.

$$y_i = B_i y_1 + \Gamma_i x_1 + u_i. \quad (2.3)$$

이 (2.3)은 다중회귀모형과 유사한 형태를 가지고 있는데 관찰변수를 통해 잠재변수모형의 모수를 추정해 보는 간단한 접근법에 대해 생각해볼 수 있다. 하지만 문제는 오차항 u_i 에는 x_1 과 y_1 의 측정오차가 포함되어 있고 이 내생성 문제로 인해 최소자승법을 직접 적용하지 못한다. 이 문제를 극복할 수 있는 대안으로서 Bollen (1996)은 구조방정식모형을 위한 2단계최소자승법을 제안하였다. 모수추정을 위한 표본 단위의 형식으로 (2.3)을 다시 표현하면 다음과 같다.

$$w_1 = Z_i a_i + u_i, \quad (2.4)$$

여기서 w_1 는 종속변수로서 y_i 에 해당하는 표본으로 이루어진 $n \times 1$ 벡터이다. B_i 과 Γ_i 에는 r 개의 자유 모수 (free parameter)가 있다. Z_i 는 이 모수에 대응되는 y_1 과 x_1 의 요소에 해당하는 표본으로 구성된 $n \times r$ 설명변수 행렬이다. 그리고 a_i 는 추정해야 할 경로계수로 이루어진 벡터를 나타낸다. 2SLS 방법 특성상 모수추정을 위해 도구변수 (instrumental variable)가 필요한데 식 (2.6)에서 V_i 로 표현된다. 이 변수는 Bollen과 Bauer (2004) 기법을 이용해서 관찰변수 중에서 잠재변수모형에 따라 자동적으로 결정된다. 이제 a_i 에 대한 Bollen (1996)의 2SLS 추정량 (estimator)은 다음과 같다.

$$\hat{a}_i = (\hat{Z}_i' \hat{Z}_i)^{-1} \hat{Z}_i' w_i, \quad (2.5)$$

$$\hat{Z}_i = V_i (V_i' V_i)^{-1} V_i' Z_i. \quad (2.6)$$

통계학 분야에서 소표본 문제를 다루기 위한 효과적인 방법으로서 정형화 (regularization) 기법에 대해 논의해 왔다 (Hastie 등, 2001). 가장 대표적인 예로서 Hoerl과 Kennard (1970)의 능형회귀 (ridge

regression) 분석법을 들 수 있다. 이는 또한 다중공선성 문제를 다루는데도 효과적인 방법으로 알려져 있다. 그동안 다중회귀모형을 넘어서 다양한 다변량 통계모형에 정형화 원리를 적용하는 연구가 활발하게 진행되어 왔다. 그 중 Jung (2013)은 구조방정식모형에 처음으로 정형화 기법을 적용하여 능형회귀와 유사한 방식의 2단계능형최소자승법 (2SRLS)을 제안하였다. 이에 따라 에 대한 2SRLS 추정량 (estimator)은 다음과 같다.

$$\hat{a}_i(\lambda) = (\hat{Z}'_i \hat{Z}_i + \lambda I)^{-1} \hat{Z}'_i w_i, \quad (2.7)$$

여기서 λ 는 능형모수 (ridge parameter)이다. λ 는 보통 작은 크기의 양의 값으로 선택하는데 이에 따른 편익은 모수추정치에 효율성이 크게 증가한다는 점이고 비용은 작은 정도의 편향이 발생한다는 점이다 (Hastie 등, 2001). 그래서 편익은 최대한 늘리고 비용은 최소한 줄이기 위한 최적의 λ 값을 선택해야 하는데 이는 교차검증법 (cross validation) 통해 결정된다 (Jung, 2013). 이 기법에서 표본을 무작위로 k 개의 하위집단으로 나눈 뒤, $k - 1$ 개의 집단은 모형의 모수를 추정하는 표본으로 사용하고 남은 한 집단은 여기서 얻은 모수추정치의 예측성능을 검증하는 목적으로 사용된다. 사전에 설정된 능형모수 값의 범위 ($\lambda = 0, 0.2, 0.5, 1, \dots, 5, 10$)를 대상으로 이 과정을 체계적으로 k 번 반복한 후 얻은 누적된 예측오차 (prediction error)를 비교해서 가장 낮은 값을 산출한 능형모수를 최종 선택한다. 또한 각 모수에 대한 가설검정을 위해 필요한 모수추정치의 표준오차는 부트스트래핑 기법을 통해 추정된다 (Efron, 1982). 본 연구에서는 500개의 부트스트랩 표본을 추출한 뒤 각 모수추정치의 표준오차를 비모수적인 방식으로 추정한다.

3. Monte Carlo study

앞에서 2SLS와 2SRLS의 방법론적 배경과 특성을 살펴보았다. 이제 이 방법들을 몬테카를로 시뮬레이션을 통해 구조방정식모형의 대표적인 모수추정 방법인 ML과 체계적으로 비교해보고자 한다. 특히, 모형명세화 오류와 소표본 상황에서 경로계수 추정의 정확도와 가설검정의 오류를 중심으로 실험결과에 대해서 논의하고자 한다. 마지막으로 이 결과를 바탕으로 이론검정을 위해 구조방정식모형을 사용하는 응용연구자들에게 주는 시사점을 도출하고자 한다.

3.1. Simulation design

본 연구에서 시뮬레이션 스타디는 총 3개의 실험조건을 설정하여 실시하였다. 첫 번째 요인은 모수 추정방법인데 3개의 수준 (2SLS, 2SRLS, ML)을 가지고 있다. 두 번째 요인은 표본크기이다. 표본수 (n)는 50, 70, 100, 150, 200, 300로 여섯 개의 수준으로 나누었다. 특히, 표본수 50과 70은 소표본을 나타내는 실험 수준이다 (Boomsma와 Hoogland, 2001). 마지막 요인은 모형명세화이다. 이는 모형명세화 오류 유무를 나타내는 두 개의 수준으로 구성되었다.

본 시뮬레이션 연구에서 설정된 구조방정식모형은 Figure 3.1에 제시되어 있다. 이는 그동안 다양한 구조방정식 시뮬레이션 연구에서 사용된 가장 대표적인 모형이다 (Henseler, 2012). Bollen 등 (2007)은 처음으로 이 모형을 명세화 오류에 따른 2SLS와 ML의 성능 차이를 비교하기 위한 연구에서 사용하였다. 하지만 이 연구의 한계점은 측정모형에서의 명세화 오류만을 설정하였다는 것이다. Figure 3.1은 잠재변수모형의 명세화 오류를 포함하고 있다. 실제 응용연구에서 주로 다루는 연구가설은 잠재변수들 간의 구조적 관계로 표현되므로 잠재변수모형에서의 명세화 오류를 포함하는 것은 이론검정 관점에서 훨씬 더 중요하다.

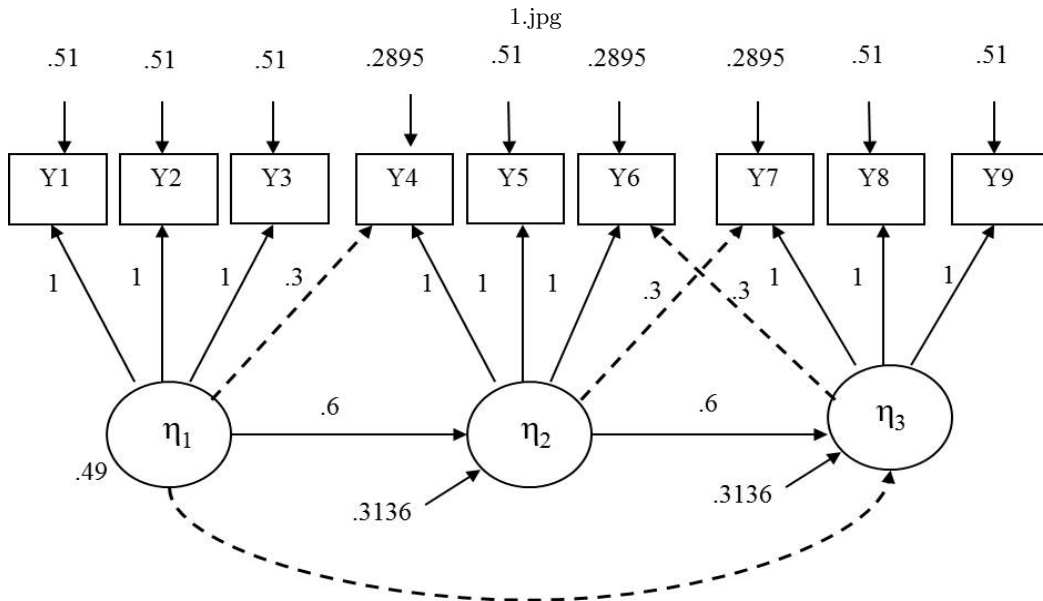


Figure 3.1 The specified model for the simulation study. The model misspecification indicates omission of dashed loadings and inclusion of the dashed path coefficient.

또한 Figure 3.1은 사회과학에서 가장 대표적인 연구모형 중 하나인 Baron과 Kenny (1986)의 매개모형을 나타낸다. 명세화 오류가 없는 모형은 완전매개모형을, 명세화 오류가 있는 모형은 부분매개모형을 나타낸다. 구조방정식 접근법에 따라 매개모형을 검정하는 경우 사전에 완전매개모형 또는 부분매개모형 중 하나를 설정해야 한다 (James 등, 2006). 충분한 선행연구와 이론적 배경이 뒷받침되는 경우 어떤 모형이 정확한 모형인지 사전에 결정할 수 있겠지만 그렇지 못한 경우 연구자들은 현실적으로 직접 경로와 간접경로 모두 존재하는 부분매개모형을 선택할 수밖에 없다 (Baron과 Kenny, 1986). Figure 3.1은 이처럼 구조방정식모형을 사용하는데 있어서 실제 응용연구자들이 접하는 연구 상황을 잘 반영하고 있는 모형이다.

Figure 3.1에 제시된 구조방정식모형을 McArdle와 McDonald (1984)의 RAM (reticular action model) 모형을 바탕으로 모집단 공분산행렬을 도출하고 이로부터 각 실험조건에 수준에 500개의 표본을 생성하였다 (Fox, 2006). 결과적으로 각 추정방법이 분석한 총 표본 수는 6000개이다 (6 sample sizes × 2 model specifications × 500 replications). 모수추정방법은 각각의 표본을 사용해서 동일한 경로계수를 추정하였다. 이 추정치를 정확도 관점과 검정력 관점에서 평가한 결과를 바탕으로 추정방법들을 비교하였다.

3.2. Simulation results

응용연구자들이 이론검정을 위해 구조방정식모형 분석결과를 확인할 때 일반적으로 두 가지를 살펴본다. 먼저 경로계수 추정치이다. 이를 통해 잠재변수 간 관계의 방향과 강도가 이론적으로 잘 부합하는지 파악한다. 따라서 모수추정 방법이 경로계수 모수를 얼마나 정확하게 복원하는지 평가해 보고자 한다. 다음으로 경로계수에 대한 가설검정 결과이다. 응용연구자들은 이론검정을 위해 사전에 설정된 연구모형을 뒷받침하는 경로계수의 통계적 유의성에 높은 관심을 가진다. 이에 따라 연구가설 검정의 오

류인 제 2종 오류를 평가해 보고자 한다.

모수추정방법이 정확하게 모수를 복원했다고 하더라도 표준오차를 정확하게 계산하지 못한다면 제 2종 오류는 높아질 수밖에 없다. 즉, 이는 각 추정방법이 얼마나 안정적으로 표준오차를 제공할 수 있는가의 문제이기도 하다. 본 연구에서 2SLS와 2SRLS는 부트스트래핑 방식에 의해서 표준오차를 추정하고, ML은 모수추정치와 점근공분산행렬 (asymptotic covariance matrix)로부터 계산한다 (Bollen, 1996). 마지막으로, 아래 시뮬레이션 결과를 보고할 때 부적절해 또는 비수렴이 발생한 ML 모수추정치는 모두 포함하지 않았다. 이는 실증연구에서 이런 문제가 발생할 경우 연구결과를 보고하거나 논의하지 않기 때문이다.

본 시뮬레이션 연구에서 모수추정의 정확도는 평균제곱오차 (mean square error; MSE)를 사용해서 평가하였다. 모수추정치가 모수로부터 얼마나 가까운 값인지를 측정하는데 모수에 근접할수록 MSE의 값은 작아진다. 이는 아래 식으로 표현되는데 이 식 맨 오른쪽에 있는 두 개의 요소는 각각 모수추정치의 분산과 편향을 나타낸다.

$$\text{MSE}(\hat{\theta}_j) = E \left[(\hat{\theta}_j - \theta_j)^2 \right] = E \left[(\hat{\theta}_j - E(\hat{\theta}_j))^2 \right] + (E(\hat{\theta}_j) - \theta_j)^2. \quad (3.1)$$

Figure 3.2는 구조방정식모형이 정확하게 설정된 경우, 모수추정방법이 산출한 경로계수 추정치의 정확도를 다양한 표본크기에 따라 비교한 평가결과를 보여준다. 표본크기가 작은 경우 ($n < 100$) 상대적으로 2SRLS이 다른 두 방법보다 MSE가 가장 작고 반대로 표본크기가 커지면서 ML이 가장 낮은 MSE를 산출하는 것으로 해석할 수 있다. 하지만 추정방법 간 실질적인 차이는 미미하다고 볼 수 있다.

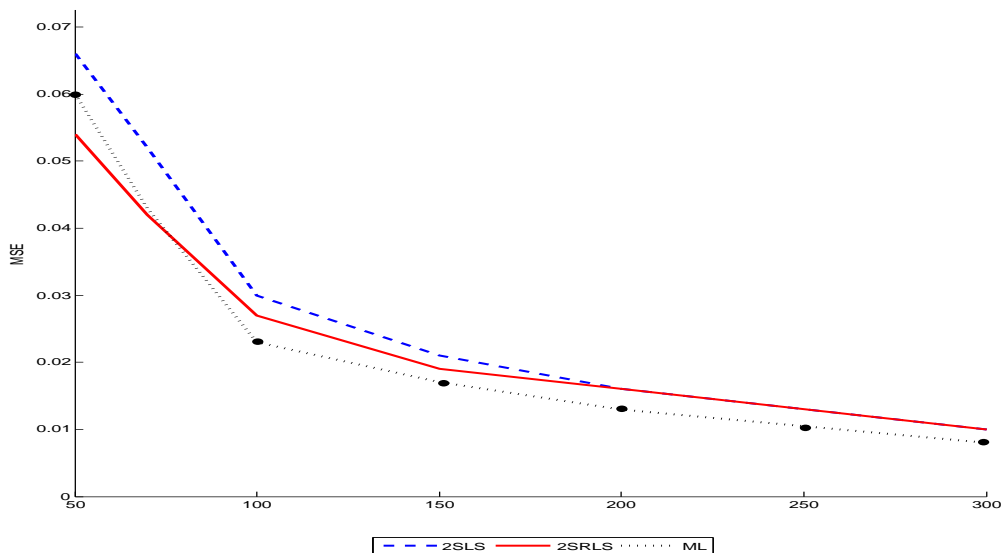


Figure 3.2 Mean square error of the parameter estimates of the three estimation methods under correct specification across different sample sizes (2SLS = dashed line, 2SRLS = solid line, ML= dotted line)

Figure 3.3은 모형명세화 오류가 발생한 경우, 모수 추정치의 정확도 패턴 차이를 보여주고 있다. 앞서 논의한 이론적 배경에 상응하는 결과로서, ML 방식이 전체 표본 크기에 걸쳐서 가장 높은 MSE 수치를 보여주고 있다. 표본크기가 100 이상인 경우 ML은 다른 방법들과 상대적으로 유사한 정확도를 보이지만 그 보다 작은 경우 매우 불안정한 패턴을 보이고 있다. Bollen 등 (2007)의 연구에서 검증되었듯이, 2SLS는 명세화 오류의 영향을 상대적으로 적게 받는 모습을 나타내고 있다. 하지만 표본크기가 작은 경우 2SRLS에 비해 더 큰 MSE 수치를 보이고 있다. 결론적으로, 2SRLS가 모형 오류가 있고 표본 크기가 작은 경우 가장 정확한 모수추정치들을 산출해 내는 방법이라고 말할 수 있다.

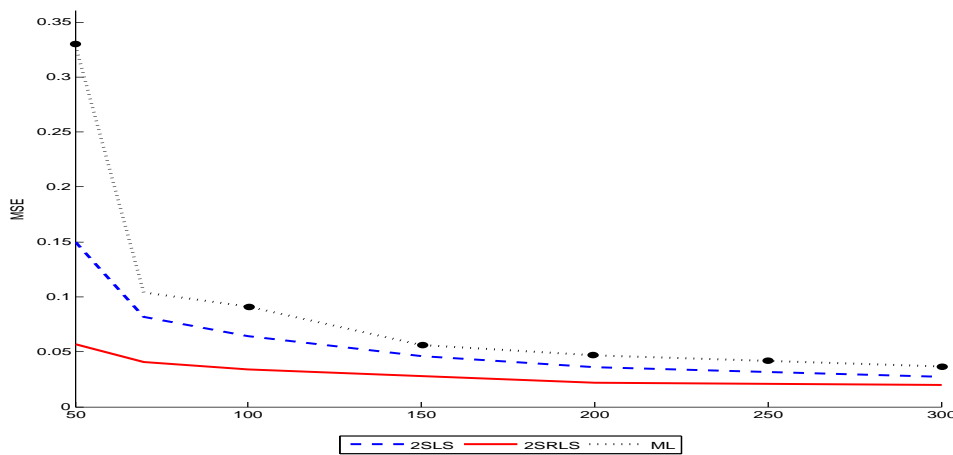


Figure 3.3 Mean square error of the parameter estimates of the three estimation methods under misspecification across different sample sizes (2SLS = dashed line, 2SRLS = solid line, ML= dotted line)

Figure 3.2와 Figure 3.3을 종합해 보면, 다른 방법들과 달리 2SRLS는 명세화 오류에 가장 적게 영향을 받는다고 볼 수 있다. 그리고 이는 표본 크기가 작은 경우에도 안정된 MSE 패턴을 보여주고 있다. 즉, 소표본 상황에서 가장 적합한 추정방법으로서 2SRLS를 선택해야 함을 알 수 있다. 그런데 여기서 한 가지 짚고 넘어갈 점은 능형회귀 기법 특성상 모수추정 편향이 발생할 수밖에 없다는 사실이다. 비록 편향이 발생하더라도 이보다 훨씬 더 큰 크기로 분산이 줄어들기 때문에 식 (3.1)에서 보듯이 전체적으로 MSE는 줄어들게 되는 것이다. 하지만 일반적으로 통계 연구에서 모수추정 방법을 평가할 때 불편향성을 고려하므로 여기서도 이를 별도로 분석해 보고자 한다. Bollen 등 (2007)은 구조방정식모형에서 상대편향 (relative bias) 비율이 10%를 넘지 않으면 불편향성을 가진다고 해석기준을 제시하였다. 이에 따라 세 가지 모수추정 방법을 평가하였는데 이에 대한 결과는 Figure 3.4와 Figure 3.5에 제시되어 있다.

Figure 3.4는 정확한 모형명세화 조건에서 모든 추정방법이 상대편향 임계치 내에 속한다고 볼 수 있다. 그런데 2SRLS는 표본 크기가 50일 때 대략 13% 정도의 부적 편향을 보이므로 이 경우 제한적으로 (marginally) 불편향성을 나타낸다고 해석할 수 있다. 하지만 2SRLS는 표본 크기가 증가함에 따라 편향성이 큰 폭으로 감소하는 패턴을 보이고 있다. 이와 같은 2SRLS의 특성은 모형명세화 오류가 있는 조건에서도 마찬가지로 관찰된다. Figure 3.5에서 표본이 50인 경우 상대편향이 대략 14%로 제한적

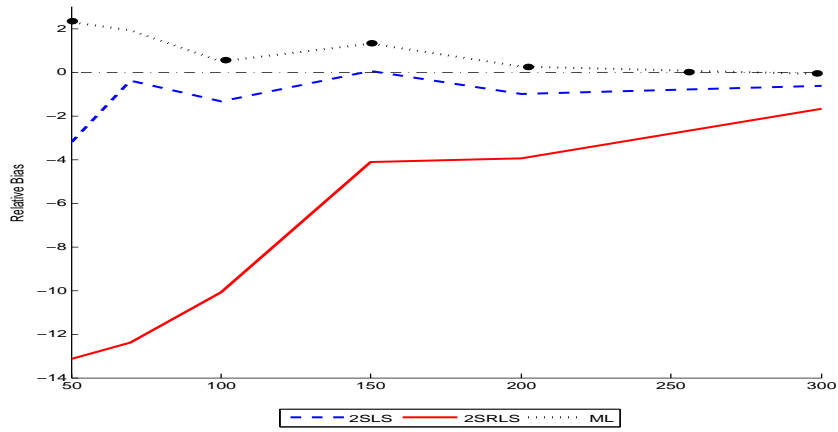


Figure 3.4 Relative bias of the parameter estimates of the three estimation

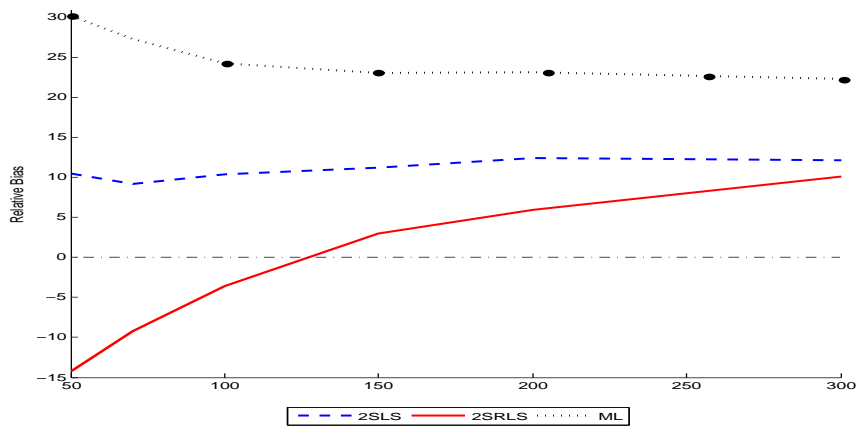


Figure 3.5 Relative bias of the parameter estimates of the three estimation methods under misspecification across different sample sizes. A dash-dot line indicates no relative bias (2SLS = dashed line, 2SRLS = solid line, ML= dotted line).

불편향성을 보이는 경우를 제외하고는 전체 표본 범위에 걸쳐 불편향성을 보이고 있다. 그런데 문제는 ML이 모형명세화 오류의 영향으로 인해 전체 표본크기 범위에 걸쳐 상대편향 정도가 전반적으로 높은 경향을 보인다는 점이다. 이를 통해 명세화 오류가 ML을 통한 추정치의 편향성에 매우 부정적인 영향을 주고 있음을 확인할 수 있다.

Table 3.1 Percentage of type II error rates under correct model specification (β_1 : the path coefficient from η_2 to η_1 , β_2 : the path coefficient from η_3 to η_2)

	2SLS		2SRLS		ML	
	β_1	β_2	β_1	β_2	β_1	β_2
$n = 50$	34%	41%	31%	28%	51%	40%
$n = 70$	19%	23%	17%	12%	31%	22%
$n = 100$	6%	8%	6%	7%	7%	1%
$n = 150$	0%	2%	0%	1%	1%	0%
$n = 200$	0%	0%	0%	0%	0%	0%
$n = 300$	0%	0%	0%	0%	0%	0%

Table 3.2 Percentage of type II error rates under misspecification (β_1 : the path coefficient from η_2 to η_1 , β_2 : the path coefficient from η_3 to η_2)

	2SLS		2SRLS		ML	
	β_1	β_2	β_1	β_2	β_1	β_2
$n = 50$	60.4%	47.4%	31.8%	34.8%	58.0%	55.4%
$n = 70$	42.8%	24.8%	19.8%	17.0%	20.8%	28.8%
$n = 100$	22.4%	9.4%	10.4%	7.4%	3.9%	6.1%
$n = 150$	8.4%	2.0%	2.6%	2.0%	0.0%	0.5%
$n = 200$	2.2%	0.0%	0.4%	0.0%	0.0%	0.0%
$n = 300$	0.2%	0.0%	0.0%	0.0%	0.0%	0.0%

마지막으로 구조방정식모형을 사용하는 주된 목적인 이론검정을 평가기준에 놓고 추정방법들의 검정력을 평가해보고자 한다. 앞에서 설명하였듯이 응용연구자들은 잠재변수 간의 이론적 인과관계를 실증적으로 확인하기 위해 잠재변수모형에 이를 설정한다. 제 2종 오류는 이 관계에 대한 통계적 유의성을 확인하는데 실패할 확률을 뜻하므로 구조방정식모형에서 가장 중요하게 다루어져야 한다. 통계적 검정력 기준에서 허용될 수 있는 제 2종 오류의 임계치는 보통 20%로 정해진다 (Cohen, 1988). Table 3.1을 보면 모형명세화 오류가 없는 실험조건에서 표본 크기가 100 이상일 때 모든 추정방법들은 높은 검정력을 제공한다. 하지만 표본크기가 작을 때 2SRLS은 상대적으로 만족할만한 수준 또는 허용 범위 내 제 2종 오류 크기를 보이고 있다. Table 3.2에서 보이듯이 명세화 오류가 발생한 경우에도, 2SRLS는 표본 크기가 작을 때에도 여전히 제 2종 오류가 20% 보다 작거나 제한적으로 허용할 수 있는 범위의 검정력을 보이고 있다. 종합하자면, 2SRLS는 구조방정식모형에서 표본크기가 작은 경우에 상대적으로 더 뛰어난 검정력을 가진 추정방법으로 볼 수 있다. 이에 반해, ML과 2SLS는 표본크기가 작은 경우 상당히 높은 수준의 제 2종 오류 비율을 보이므로 이론검정을 목적으로 사용하는데 있어 상당한 주의를 기울일 필요가 있다.

4. 결론

구조방정식모형은 다양한 분야에서 연구논문 작성을 위해 가장 널리 사용되는 통계방법 중 하나이다. 이는 구조방정식모형이 연구모형에 따른 이론검정을 수행하기 위한 효과적인 분석방법이기 때문이다.

이 구조방정식모형의 추정방법으로서 ML은 연구모형에 따라 설정된 잠재변수모형과 측정모형을 관측 공분산행렬에 담긴 모든 정보를 이용하여 모수를 추정하고 연구모형의 전반적 적합도를 다양한 지수를 통해 평가할 수 있는 장점을 가지고 있다. 하지만 그동안 다양한 연구들에서 이 ML이 모형명세화 오류에 영향을 크게 받아 모수추정치에 편향성이 증대되거나 정확한 모수추정이 어려운 문제들이 보고되어 왔다. 이런 문제를 다루고자 Bollen (1996)은 2SLS 추정방법을 제안하였다. Jung (2013)은 최근 구조방정식 연구영역의 확장으로 인한 소표본 문제가 대두됨에 따라 구조방정식에 정형화 기법을 도입하여 능형회귀와 유사한 추정방법인 2SRLS를 제안하였다.

실제 응용연구자들은 구조방정식모형을 통한 이론검정을 수행할 때 경로계수 추정치와 이에 대한 통계적 검정결과를 보고하는데 현재까지 추정치의 정확성 정도와 연구가설 검정오류 관점에서 이 세 가지 추정방법을 체계적으로 비교 평가한 연구를 찾아볼 수 없었다. 특히 본 연구에서는 모형명세화 오류와 소표본 문제 상황에서 구조방정식모형 추정방법의 선택을 위한 가이드라인을 마련하기 위해 체계적으로 이들을 비교하였다.

연구결과에 따르면, 모형이 정확하게 설정된 경우, 모수추정의 정확성 기준에서는 세 가지 추정방법 간에 주목할만한 차이를 찾을 수 없었다. 하지만 명세화 오류가 있는 경우, 2SRLS는 다른 두 방법과 비교해서 가장 정확한 모수추정치를 산출해 내었고 특히 그 차이는 표본 크기가 작을 때 ($n < 100$) 가장 큰 폭으로 두드러졌다. 또한 검정력 관점에서 살펴보면, 세 추정방법 모두 적어도 중간 이상의 표본 크기 ($n > 100$)에서는 명세화오류와 관계없이 제 2종 오류가 임계치(20% 기준)보다 낮아 결과적으로 높은 검정력을 가지고 있음을 확인할 수 있었다. 하지만 표본크기가 작은 경우, 2SRLS가 유일하게 임계치 이하 또는 제한적으로 허용 가능한 수준의 제 2종 오류를 나타내었다. 결론적으로, 2SRLS는 모형명세화 오류에 관계없이 소표본 상황에서 가장 정확한 추정치를 산출하고 상대적으로 높은 검정력을 가진 추정방법이라고 말할 수 있다.

사회과학 영역에서 흔히 사용하는 구조방정식모형, 경로분석, 확인적 요인분석 등은 확인적 기법(confirmatory technique)에 해당한다. 이는 연구방법 특성상 연구자들이 사전에 모형을 정확하게 설정하는 것을 필요로 한다. 응용연구자들은 종종 사후적으로 모형의 합치도 지수를 토대로 본인들의 연구모형이 정확하다고 말하지만 이는 사실 타당한 접근법이 아니다. 동일한 합치도를 보이는 동치(equivalent)모형(Lee, 2007)이 항상 존재한다. 모형적합도 평가를 위한 카이제곱 검증은 표본 크기가 작은 경우 검정력이 상당히 떨어지는 문제를 안고 있다. 실증연구자들은 종종 유의미한 카이제곱 검증결과를 무시하고 수용수준을 넘긴 다른 합치도 지수를 연구논문에서 보고함으로써 본인들이 설정한 모형이 정확하다고 간주해 버리곤 한다. Bollen 등 (2007)의 실증연구에 따르면 모형명세화 오류가 있는 구조방정식모형이더라도 CFI(comparative fit index)와 같은 합치도지수가 수용수준 이상의 높은 값을 가질 수 있다고 보고하고 있다.

따라서 본 연구는 다음과 같은 중요한 시사점을 제공한다. ML은 표본 크기가 큰 경우 모형명세화 오류의 유무에 관계없이 높은 검정력을 가지고 있다. 하지만 모형명세화 오류가 발생한 경우, 모수추정의 정확성이 떨어지고 특히 표본크기가 충분히 크더라도($n > 200$) 여전히 높은 편향성을 보이는 문제점을 가지고 있다. 이와 달리, 2SRLS는 모형명세화 오류에 관계없이, 표본 크기가 작은 경우에도 가장 높은 이론검정력과 우수한 모수추정 정확도를 보여주었다. 따라서 2SRLS는 (1) 이론적 배경과 선행연구가 충분하지 못한 경우, (2) 모집단의 크기가 작은 경우, (3) 시간과 비용 등의 문제로 많은 수의 표본을 수집하기 어려운 경우, 구조방정식모형을 통한 이론검정을 위해 선택할 수 있는 효과적인 추정방법으로 볼 수 있다. 이런 문제를 접하는 경우, 실증연구에서는 보통 PLS(partial least squares) 구조방정식모형을 사용한다(Hair 등, 2011). 그런데 이는 성분형 구조방정식모형으로서 우리가 흔히 말하는 공분산구조분석(covariance structure analysis) 접근법과 다르다. 결론적으로 본 연구는 2SRLS 방법이 모형 명세화오류와 소표본 상황에서 공분산구조분석의 모수추정을 위해 선택할 수 있는 방법임을 보여

주고 있다. 그러나 ML은 다양한 모형합치도 지수와 통계적 검정방법을 제공하는 유용한 방법이다. 본 연구는 모수추정의 정확도와 검정력 관점에서 각 추정방법을 비교하였으므로 모형합치도에 대한 논의는 추후 연구에서 다루어져야 한다.

References

- Ansari, A., Kamel J. and Sharan J. (2000). A hierarchical Bayesian methodology for treating heterogeneity in structural equation models. *Marketing Science*, **19**, 328-347.
- Bagozzi, R. P. and Yi, Y. (2012). Specification, evaluation, and interpretation of structural equation models. *Journal of the Academy of Marketing Science*, **40**, 8-34.
- Baron, R. M. and Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, **51**, 1173-1182.
- Bollen, K. A. (1989). *Structural equations with latent variables*, John Wiley & Sons, New York.
- Bollen, K. A. (1996). An alternative two stage least squares (2SLS) estimator for latent variable equations. *Psychometrika*, **61**, 109-121.
- Bollen, K. A. and Bauer, D. J. (2004). Automating the selection of model-implied instrumental variables. *Sociological Methods and Research*, **32**, 425-452.
- Bollen, K. A., Kirby, J. B., Curran, P. J., Paxton, P. and Chen, F. (2007). Latent variable models under misspecification: Two-stage least squares (2SLS) and maximum likelihood (ML) estimators. *Sociological Methods and Research*, **36**, 48-86.
- Boomsma, A. and Hoogland, J. J. (2001). *The robustness of LISREL modeling revisited*. In *Structural Equation Modeling: Present and Future (Cudeck, R., Du Toit, S. & Sorbom, D., eds)*, SSI Scientific Software, Chicago, 139-168.
- Budaev, S. V. (2010). Using principal components and factor analysis in animal behaviour research: Caveats and guidelines. *Ethology*, **116**, 472-480.
- Chen, F., Bollen, K., Paxton, P., Curran, P. J. and Kirby, J. (2001). Improper solutions in structural equation models: Causes, consequences, and strategies. *Sociological Methods and Research*, **29**, 468-508.
- Choi, H. S., Kwon, Y. J. and Ha, J. C. (2013). Study of university students' perceptions on participation in elections via structural equation model-Focusing on K university students. *Journal of the Korean Data & Information Science Society*, **24**, 379-390.
- Cohen, J. (1988). *Statistical power for the behavioral sciences*, 2nd ed., Lawrence Erlbaum, New Jersey.
- Efron, B. (1982). *The jackknife, the bootstrap and other resampling Plans*, SIAM, Philadelphia.
- Fan, X. and Sivo, S. A. (2005). Sensitivity of fit indices to misspecified structural or measurement model components: Rationale of two-index strategy revisited. *Structural Equation Modeling*, **12**, 343-367.
- Fox, J. (2006). Structural equation modeling with the sem package in R. *Structural Equation Modeling*, **13**, 465-486.
- Grace, J. B., Anderson, T. M., Olf, H. and Scheiner, S. M. (2010). On the specification of structural equation models for ecological systems. *Ecological Monographs*, **80**, 67-87.
- Grewal, R., Cote, J. A. and Baumgartner, H. (2004). Multicollinearity and measurement error in structural equation models: Implications for theory testing. *Marketing Science*, **23**, 519-529.
- Hair, J. F., Ringle, C. M. and Sarstedt, M. (2011). PLS-SEM: Indeed a silver bullet. *Journal of Marketing Theory and Practice*, **19**, 139-151.
- Hastie, T., Tibshirani, R. and Friedman, J. (2001). *The elements of statistical learning; Data mining, inference, and prediction*, Springer-Verlag, New York.
- Hayduk, L., Cummings, G. G., Boadu, K., Pazderka-Robinson, H. and Boulianne, S. (2007). Testing! testing! one, two three-Testing the theory in structural equation models. *Personality and Individual Differences*, **42**, 841-850.
- Henseler, J. (2012). Why generalized structured component analysis is not universally preferable to structural equation modeling. *Journal of the Academy of Marketing Science*, **40**, 402-413.
- Hoerl, A. E. and Kennard, R. W. (1970). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, **12**, 55-67.
- Hong, Y., Jang, G. and Choi, C. (2016). Life satisfaction and self-esteem of children from low-income class: Testing mediation model of depression. *Journal of the Korean Data & Information Science Society*, **27**, 179-189.

- James, L. R., Mulaik, S. A. and Brett, J. M. (2006). A tale of two methods. *Organizational Research Methods*, **9**, 233-244.
- Jarvis, C. B., MacKenzie, S. B. and Podsakoff, P. M. (2003). A critical review of construct indicators and measurement model misspecification in marketing and consumer research. *Journal of Consumer Research*, **30**, 199-218.
- Jung, S. (2013). Structural equation modeling with small sample sizes using two-stage least-squares estimation. *Behavior Research Methods*, **45**, 75-81.
- Kaplan, D. (2009). *Structural equation modeling: Foundations and extensions, 2nd ed.*, Sage, California.
- Lee, S. (2007). Problems caused by model equivalence in developing and testing models. *Journal of Educational Evaluation*, **20**, 125-146.
- McArdle, J. J. and McDonald, R. P. (1984). Some algebraic properties of the reticular action model. *British Journal of Mathematical and Statistical Psychology*, **37**, 234-251.
- MacKenzie, S., Podsakoff, P. and Jarvis, C. (2005). The problem of measurement model misspecification in behavioural and organizational research and some recommended solutions. *Journal of Applied Psychology*, **90**, 710-730.
- Min, D. K. and Choi, M. K. (2016). How depression affects girls who experienced violence in home or at school: Using mixed model. *Journal of the Korean Data & Information Science Society*, **27**, 101-110.
- Olsson, U. H., Foss, T., Troye, S. and Howell, R. (2000). The Performance of ML, GLS, and WLS estimation in structural equation modeling under conditions of misspecification and non-normality. *Structural Equation Modeling*, **7**, 557-595.
- Sideridis, G., Simos, P., Papanicolaou, A. and Fletcher, J. (2014). Using structural equation modeling to assess functional connectivity in the brain: Power and sample size considerations. *Educational & Psychological Measurement*, **74**, 733-758.
- Steenkamp, J. B. E. M. and Baumgartner, H. (2000). On the use of structural equation models for marketing modeling. *International Journal of Research in Marketing*, **17**, 195-202.
- Tenenhaus, M., Pages, J., Ambroisine, L. and Guinot, C. (2005). PLS methodology to study relationships between hedonic judgments and product characteristics. *Food Quality and Preference*, **16**, 315-325.
- Tenenhaus, M. (2008). *Structural equation modeling for small samples*, HEC school of management (GRECHEC).

A study on the performance of three methods of estimation in SEM under conditions of misspecification and small sample sizes[†]

Dong Gi Seo¹ · Sunho Jung²

¹Department of Psychology, Hallym University

²School of Management, Kyung Hee University

Received 9 June 2017, revised 20 September 2017, accepted 21 September 2017

Abstract

Structural equation modeling (SEM) is a basic tool for testing theories in a variety of disciplines. A maximum likelihood (ML) method for parameter estimation is by far the most widely used in SEM. Alternatively, two-stage least squares (2SLS) estimator has been proposed as a more robust procedure to address model misspecification. A regularized extension of 2SLS, two-stage ridge least squares (2SRLS) has recently been introduced as an alternative to ML to effectively handle the small-sample-size issue. However, it is unclear whether and when misspecification and small sample sizes may pose problems in theory testing with 2SLS, 2SRLS, and ML. The purpose of this article is to evaluate the three estimation methods in terms of inferences errors as well as parameter recovery under two experimental conditions. We find that: 1) when the model is misspecified, 2SRLS tends to recover parameters better than the other two estimation methods; 2) Regardless of specification errors, 2SRLS produces small or relatively acceptable Type II error rates for the small sample sizes.

Keywords: Maximum likelihood estimation, misspecification, small sample sizes, structural equation models, two-stage least squares, two-stage ridge least squares.

[†] This work was supported by a National Research Foundation of Korea (NRF) grant funded by the Korean Government (NRF-2014S1A5A2A03064999).

¹ Assistant professor, Department of Psychology, Hallym University, Chuncheon, Gangwon-do, 24252, Korea.

² Corresponding author: Associate professor, School of Management, Kyung Hee University, Seoul 02447, Korea. E-mail: sunho.jung@khu.ac.kr