

## 연구의 인과성 제고: 통제변수의 의의, 활용 현상 분석 및 제언

박 원 우\*  
고 동 운\*\*  
윤 은 성\*\*\*

### 〈목 차〉

- |                        |                       |
|------------------------|-----------------------|
| I. 서론                  | IV. 구조 방정식 모형에서의 통제변수 |
| II. 통제 및 통제변수의 개념 및 의의 | 적용                    |
| III. 국내의 통제변수 활용 실태 분석 | V. 결 론                |

사회과학 연구는 대부분 인과관계를 확인하기 위한 목적으로 이루어지지만, 인과성을 입증하기 위한 노력은 부족한 편이다. 본 연구는 인과관계의 한 조건인 '외생변수의 통제'에 초점을 두고 '통제' 및 '통제변수'의 개념과 의의에 대해서 살펴보았다. 이를 통해, 저자들은 비실험연구에서 연구의 인과성을 높이기 위해서는 통제변수의 도입이 필요함을 역설하였다.

통제변수 활용실태를 파악하기 위해, 본 연구는 국내 주요 저널의 인사조직관련 연구(179편)를 대상으로 하여 질적 및 양적 분석을 실시하였다. 이를 통해, 통제변수 활용의 모범 사례와 문제점을 도출하여 바람직한 통제변수의 활용에 대한 구체적인 기준과 절차를 제안하였다. 특히, 통제변수를 전혀 사용하지 않거나 아무런 근거 없이 선정되는 현상은 시급히 개선되어야 할 문제점으로 확인되었다.

통제변수의 활용 현상에 있어 특히 주목할 점은 구조방정식 연구에서 통제변수의 사용이 극히 저조하다는 것이다. 본 연구에서 이러한 문제에 대해 지적하고 구조방정식모형에서도 통제변수가 활용되어야 함을 강조하였다.

결론적으로 본 연구에서 중점을 두어 강조한 것은 인과관계를 보다 확신하기 위해서는 적절한 통제변수를 포함하는 것이 중요하며 통제변수도 독립변수나 종속변수와 마찬가지로 동일하게 취급되어야 한다는 점이다. 통제변수의 적절한 사용은 연구자의 연구 결과를 보다 가치 있게 하는데 기여할 것이다.

\* 서울대학교 경영대학 교수

\*\* 서울대학교 대학원 경영학과 박사과정

\*\*\* 서울대학교 대학원 경영학과 석사과정

## I. 서론

대다수 사회과학 연구는 인과관계를 확인하기 위한 목적에서 비롯되며 이를 위해서 과학적 연구방법을 적용한다. 과학적 연구(scientific research)는 특정 현상을 설명하기 위한 체계적이고, 통제적이며, 경험적 및 비판적인 탐구활동이며, 여기서 '체계적이고 통제적이어야한다.'라는 말은 다른 대체적 설명들이 체계적으로 배제되고 통제되어야 한다는 의미로, 이러한 조건이 충족되어야 비로소 연구 결과에 대한 인과관계를 확신할 수 있다(Kerlinger & Lee, 2000). 사회과학의 특성이 결과변수에 대해 하나 이상의 다수의 원인변수가 존재하는 다변수적(multivariable) 특성임(Kerlinger & Lee, 2000; Marini & Singer, 1988; Schwab, 2005)을 고려한다면, 두 현상 간의 인과관계를 확신하기 위해서는 반드시 대체적인 설명이 통제되어야 한다. 따라서 사회과학 연구의 수행은 이러한 대체적인 설명을 배제하기 위한 통제의 과정이라고 할 수 있다. 본 연구는 타당성 있는 연구결과를 얻기 위한 이러한 '통제' 개념에 대한 관심에서 출발한다.

사회과학의 다변수적 특성은 일반적인 통계적 방법으로 인과적 추론(causal inferences)을 하는데 있어서 타당성의 문제를 낳고 있는데, 이러한 타당성 문제의 원인은 외생변수(extraneous variable)를 전적으로 통제하지 못한다는 한계에서 비롯된다(Frank, 2000; Pan & Frank, 2003). 인과관계를 정확하게 추론하기 위해서는 외생변수를 통제하여야 하는데 일반적인 방법은 실험에 의해서 통제하는 방법과 비실험적(nonexperimental) 연구에서의 통계적인 통제방법(statistical control)이 있다(Becker, 2005). 일반적으로 무작위할당 실험(random assignment experiment)이 인과관계를 추론하는데 이상적이라는 데 인식을 같이 하지만, 논리적인 문제나 윤리적 및 정치적인 문제로 인해 사회과학에서는 실용적이지 못한 문제점(Frank, 2000; Pan & Frank, 2003; Rubin, 1974; Winship & Morgan, 1999)과 자연스러운 현실 상황을 반영하지 못하고 일반화하기에는 제한되는 단점이 있다(Cook & Campbell, 1979; Kish, 1959; Martin & Sell, 1979; Singer & Marini, 1987; Smith, 1990). 이러한 이유로 현재의 사회과학 연구에서는 실험 방법보다는 비실험적 연구 방법이 많이 활용되고 있다.

설문 조사와 같은 비실험적 연구는 통제변수(control variable)의 도입과 같은 통계적인 통제방법을 통해서 외생변수를 통제하고 있다(Blalock, 1961a; Frank, 2000; Rubin, 1974; Singer & Marini, 1987; Smith, 1990). 이는 하나의 종속변수를 설명하기

위해 여러 독립변수를 고려해야 하는 다변수적 특성으로 인해 연구자가 규명하고자 하는 변수 간의 관계가 혼재되어 왜곡된 결과를 나타내기 때문에 필요한 조치이다.

예를 들어, 연구자가 직무만족의 영향요인으로 나이를 고려하고 두 변수 간의 관계를 회귀분석을 통해서 알고자 한다면, 직무만족에 영향을 줄 수 있는 다른 요인은 통제되어야 정확한 관계를 알 수 있을 것이다. 그야말로 숫자에 불과한 나이가 직무만족에 영향을 준다는 해석은 논리적으로 맞지 않다. 나이가 많아짐에 따라 긍정적인 직업관이나 가치관이 형성될 수 있고 지위, 보상수준 등의 변수들이 직무만족에 영향을 줄 수 있을 것이다. 즉, 나이와 공변하는 속성들이 통제변수로 분석에 포함되지 않으면 왜곡된 결과를 도출할 수 있다. 대체적으로 조직에서 나이가 많아질수록 지위가 높아지며 보상수준이 증가하기 때문에 지위나 보상수준의 증가로 인한 직무만족의 증가가 나타날 수 있다. 따라서 오히려 나이와 직무만족의 관계는 잘못 해석되었을 가능성이 높으며, 이러한 현상을 방지하기 위해서는 지위 및 보상수준 등의 변수를 통제변수로 도입하여야만 한다. 통제변수를 제대로 사용하지 않아 해석을 잘못하는 현상은 실제로 많은 연구에서 나타나고 있다.

따라서 타당성 있는 연구 결과를 위해서는, 이러한 통계적인 통제방법이 반드시 필요함에도 불구하고 많은 연구에서 간과되고 있고 통제변수를 포함한 연구에서도 필요한 절차나 과정이 무시되고 있는 실정이다(Becker, 2005). 이러한 현상은 국내의 연구에서 특히 두드러진데 이는 인과관계 추론의 엄밀성에 관한 인식 부족과 통계적인 통제방법에 대한 논의가 부족한 실정에서 비롯된다. 이러한 문제 인식하에서 본 연구는 이루어졌다.

따라서 본 연구는 먼저, 연구의 인과성(causality)을 제고하기 위한 통제 및 통제변수의 개념과 의의를 고찰하고, 통제변수 활용의 필요성을 강조하겠다. 이어 국내 인사/조직 관련 주요 조사 연구들을 대상으로 통제변수 활용 실태를 분석하고, 통제변수 활용에 있어 적절한 제언과 논의 필요성이 있는 주요 이슈들을 제시할 것이다. 마지막으로 연구의 시사점과 한계를 논의 하였다.

## II. 통제 및 통제변수의 개념 및 의의

### 1. 연구의 인과성 및 통제의 개념 및 의의

#### 1) 인과성의 조건

사회과학은 다변수적 특성을 가지고 있다(채서일, 2002; Kerlinger & Lee, 2000; Marini & Singer, 1988; Schwab, 2005). 자연과학의 폐쇄시스템(closed system)과는 달리 사회과학 연구는 개방시스템에서 이루어지기 때문에 인과관계에서 결과에 영향을 주는 원인이 여러 가지 있을 수 있다(채서일, 2002). Kerlinger & Lee(2000: 32)는 심리학, 사회학, 교육학 등 인간 행동에 관심을 둔 연구에 있어서, 'x이면 y이다' 식의 두 변수만을 고려한 가설은 정확한 현상을 이해하는데 한계가 있으며 기본적으로 행동 과학이 다변수적 특성임을 고려해야 함을 강조하고 따라서 "p이면 q이다. 이것 보다는 'p1, p2, ..., pk이면 q이다' 또는 'r, s, t 조건에서는 p이면 q이다'와 같은 가설이 보다 적절하다."라고 하였다. 이는 특정한 현상(결과)에 영향을 미치는 변수가 다수임을 뜻하고, 이로 인해 독립변수 이외의 기타 변수의 영향을 통제하는 절차가 필요함을 암시한다.

대부분의 연구는 인과관계를 검증할 목적으로 실시되어지고 있고, 실제로 많은 연구에서 검증 결과를 바탕으로 인과관계를 주장하고 있지만, 주지하다시피 인과관계를 증명할 수 있는 연구는 오로지 무작위할당 실험연구이다(Holland, 1986; Rubin, 1974). 인과성을 밝히려는 연구에 비해 인과성의 조건 및 이를 달성하기 위한 방법에 대한 관심은 부족하였다. 따라서 '인과성'에 관한 고찰을 통하여 무엇보다 인과관계가 무엇이며, 이를 달성하기 위한 인과조건과 통제의 개념에 대해서 살펴보고자 한다.

인과관계는 두 현상(X, Y)의 관계에서, X가 Y의 원인이 되는 것으로, X의 변화가 Y의 변화를 가져올 때의 관계를 의미하는 것이다. 이는 단순히 X가 변화할 때 Y의 변화가 나타나거나 관련되어 있음을 의미하는 것은 아니다(Blalock, 1961a). '닭이 울면 해가 뜬다.'라고 해서 닭의 울음이 원인이 될 수 없는 것과 마찬가지로 인과관계는 단순히 관련성을 의미하는 것이 아니다.

독립변수 X와 종속변수 Y의 인과관계가 성립하기 위해서는 세 가지 기준이 만족해야 한다. 첫째, X와 Y의 의미 있는 관련성이 실증되어야 하며, 둘째, 독립변수(X)의 변동(variation)은 종속변수(Y)의 변동보다 선행하거나 동시적이어야 한다. 셋째, 두 변수

의 관계에 대한 적절한 원인적 설명을 할 수 있고 이 외의 설명의 여지가 없어야 한다 (Cohen et al., 2003; Cook & Campbell, 1979; Schwab, 2005). 첫 번째 기준은 일반적인 통계적인 방법으로 검증이 가능한 것으로 두 변수 간에 의미 있는 상관관계가 있어야 한다는 것이다. 두 번째 기준은 인과관계의 방향을 확인하기 위한 기준이다. 본 논문에서 관심을 두고 있는 것은 세 번째 기준으로, 두 변수 간의 관계에서 다른 대체적 설명이 없어야 한다는 것은 다른 변수의 영향, 즉 외생변수의 영향이 통제되어야 한다는 의미이다(이근희, 2005; 채서일, 2002; Cook & Campbell, 1979; Schwab, 2005).

이와 같이, 가설로서 주장하는 인과관계 외의 다른 대체적 설명을 배제하기 위해서 통제의 개념이 도입된다. 과학적 연구에 있어서 통제란 무엇인지, 통제방법에는 어떤 것이 있는지 살펴봄으로써 본 연구의 초점인 통제변수에 대한 이해를 돕고자 한다.

## 2) 통제 개념 및 통제 방법

과학적 연구에서 통제(control)는 연구자의 가설적 관계 외의 다른 설명들을 배제하기 위한 노력, 즉 외생변수의 영향을 제거하기 위한 활동(Kerlinger & Lee, 2000)이라고 할 수 있다. Boring(1954)은 실험연구에서의 통제 개념을 몇 가지로 구분하여 제시하고 있다. 먼저, 통제는 억제(restraint)의 의미를 내포하고 있는데, 실험 조건을 실험 내내 일정한 상태로 유지하는 것을 뜻한다. 둘째, 유도(guidance)의 의미로서 독립변수를 의도대로 정확하게 변화시키거나 유지, 제어하는 것을 의미한다. 만약 독립변수를 실험자가 의도한 대로 조작(manipulation)할 수 없다면, 이미 실험 조건을 만족할 수 없게 되므로 유도(guidance)는 실험의 기본적인 요건임을 알 수 있다. 셋째, 점검(check) 혹은 대조(comparison)의 개념으로 통제의 가장 본질적인 의미라고 할 수 있다. 특정의 현상에 대해서 사실을 발견하기 위해서는 차이나 관련성을 기준과 비교함으로써 가능한데, 이를 위해서 실험에서 통제집단(control group)의 개념이 들어오게 된다. 통제의 개념은 오랜 역사를 가지고 있지만 실험연구에서 통제집단의 도입은 20세기에 이르러서야 비로소 이루어졌다(Boring, 1954; Solomon, 1949).

Campbell & Stanley(1966) 및 Cook & Campbell(1979)의 실험연구에서 내적 타당성을 저해하는 외생변수들<sup>1)</sup>이 망라되어 있는데 이러한 외생변수들은 적절한 통제

1) Campbell & Stanley(1966: 5): 우발적 사건(history), 성숙효과(maturation), 시험효과(testing), 측정도구의 변화(instrumentation), 통계적 회귀(statistical regression), 표본의 편중(selection),

조치로서 제거될 수 있다. Boring의 통제개념에서 실험조건을 일정하게 유지함, 즉 억제조치는 우발적 사건, 성숙효과, 시험효과, 측정도구의 변화, 그리고 실험대상의 소멸과 같은 외생변수를 통제하는 것을 의미한다. 유도는 실험변수를 연구자가 의도한 대로 정확하게 조작하는 것을 의미하므로 실험변수의 확산 또는 모방, 실험변수의 효과를 상쇄하는 보정, 열등실험변수를 받은 집단의 보정적 경쟁, 그리고 열등실험변수를 받은 경우의 자포자기 등의 외생변수가 발생하는 것을 억제하는 개념이다. 점검 혹은 대조의 개념은 통제집단과 실험집단을 대조함으로써 종속변수에 영향을 미치는 외생변수의 영향을 제거할 수 있다. 이러한 통제집단의 사용으로 대부분의 외생변수를 통제할 수 있으나, 실험변수의 확산 또는 모방, 실험변수의 효과를 상쇄하는 보정, 열등실험변수를 받은 집단의 보정적 경쟁, 열등실험변수를 받은 경우의 자포자기 등과 같이 실험집단 및 통제집단의 처치 실패에서 오는 효과는 제거할 수 없다. 따라서 실험 전 과정에 걸쳐서 위에서 제기한 세 차원의 통제 개념이 동시에 엄격하게 적용되어야 할 것이다.

연구설계(research design)는 외생변수의 통제가 주요한 목적이다(Kerlinger & Lee, 2000). 외생변수를 통제한다는 의미는 영향을 최소화하고, 무의미하게 하거나, 분리함을 뜻하는 것으로 다음과 같은 방법들이 있다. 채서일(2002), Kerlinger & Lee(2000), 그리고 Schwab(2005) 등의 연구에서 제시된 통제의 구체적인 수행방법을 살펴보았다.

먼저, 제거는 외생변수가 될 수 있는 변수의 속성을 동일(homogeneous)하게 하거나 범위를 제한(restriction of range)함으로써 영향을 제거하는 방법을 의미한다. 성별 변수가 외생변수로서 우려된다면, 남성 및 여성 중에서 하나의 성으로만 표본을 구성하여 성별의 효과를 제거할 수 있을 것이다. 이러한 제거 방법은 일반화(generalization)를 제한하는 단점이 있다(Kerlinger & Lee, 2000; Schwab, 2005).

---

실험대상의 소멸(experimental mortality), 표본편중과 다른 요인과의 상호작용(interactions with selection) 등의 8종류의 외생변수가 있다.

Cook & Campbell(1979: 51-55): 인과방향의 모호성(ambiguity about the direction of causal influence), 실험 변수의 확산 또는 모방(diffusion or imitation of treatments), 실험변수의 효과를 상쇄하는 보정(compensatory equalization of treatments), 열등실험변수를 받은 집단의 보정적 경쟁(compensatory rivalry by respondents receiving less desirable treatments), 열등실험변수를 받은 경우의 자포자기(resentful demoralization of respondents receiving less desirable treatments) 등의 5종류가 Campbell & Stanley(1966)의 8종류에 추가하여 설명되고 있다. 각각의 외생변수에 대한 구체적인 설명은 본 연구의 범위를 벗어나기 때문에 논외로 한다.

둘째, 무작위화(randomization)는 실험 및 통제집단에 표본을 무작위로 배분하고 무작위로 실험집단을 선정하는 방법이다. 이 방법은 모든 가능한 외생변수의 영향을 통제할 수 있는 가장 최선의 방법으로 실험집단 및 통제집단의 속성을 확률적으로 이상적인 균형(equivalence) 상태를 이루게 하는 것이다. 무작위화를 달성하면 연구자가 인지하고 있는 외생변수의 영향뿐만 아니라 연구자가 미처 생각하지 못한 외생변수의 효과를 제거할 수 있는 장점과 이로 인해 특정의 다른 외생변수를 측정하지 않아도 되는 이점이 있다(이근희, 2005; Blalock, 1961a; Kerlinger & Lee, 2000; Schwab, 2005). 반면에 소수의 표본일 경우, 집단 간 불균형이 발생할 수 있으며(Kerlinger & Lee, 2000) 외생변수의 효과를 실험효과와 분리시켜 오차에 들어가도록 함으로써 실험의 정확도가 떨어질 수 있다는 단점과 현실적으로 적용이 불가능한 한계가 있다(이근희, 2005; Bollen, 1989; Kish, 1959).

셋째, 균형화(matching)의 방법이 있다. 균형화는 종속변수에 영향을 주는 외생변수를 계층별로 대상자들을 분류하고 실험집단과 통제집단에 무작위로 배정하는 방법이다. 즉, 예상되는 외생변수를 실험집단과 통제집단에 계층별로 동일하게 배정함으로써 해당 외생변수의 영향을 제거할 수 있다. 균형화의 단점은 통제하여야 할 외생변수의 수가 증가하면 필요한 표본의 수가 급격하게 늘어난다는 점이다(Breaugh & Arnold, 2007; Kerlinger & Lee, 2000; Schwab, 2005). 성별을 통제하고자 한다면 남성 및 여성 각각의 실험집단과 통제집단 4개의 집단만 필요하지만, 여기에 상·중·하로 구분되는 사회적 계층을 또 다른 외생변수로 추가한다면 집단의 수는 12개로 대폭 증가하게 된다. 따라서 3, 4개 이상의 외생변수를 통제하는 것은 사실상 극히 제한적이다.

넷째, 외생변수를 독립변수로 설계에 포함하여 통제하는 방법, 즉 통제변수를 활용하는 방법이 있다. 이 방법은 종속변수에 대한 해당 외생변수의 영향 분산을 총분산에서 분리함으로써 가능한데, 이 방법은 통제 효과뿐만 아니라 해당 외생변수가 종속변수에 미치는 영향을 추가적으로 확인할 수 있고 다른 독립변수와의 관계도 확인가능하다는 장점이 있다(Kerlinger & Lee, 2000).

마지막으로, 통계적 방법은 외생변수의 효과를 통제하기 위하여 통계적인 방법을 사용하는 것으로 공분산분석(ANCOVA) 등의 방법이 있다.

이제까지 인과성을 검증하기 위해서는 외생변수의 통제가 중요하다는 점과 통제의 개념과 통제 방법에 대해서 고찰하였다. 이를 통해서 통제변수의 의의 및 필요성을 간접적

으로 확인하였는데, 다음 절에서 보다 구체적으로 통제변수의 개념 및 의의에 대해서 살펴해보도록 하겠다.

## 2. 통제변수의 개념 및 의의

### 1) 통제변수의 개념

전술한 바와 같이, 통제변수는 종속변수에 대한 영향을 분리하기 위해 분석모형에 포함시켜서 분석하는 외생변수를 의미한다. 통제변수에 대한 개념을 이해하기 위해서는 외생변수에 대한 이해가 선행되어야 할 것이다.

외생변수는 종속변수에 영향을 미치는 원인변수 중에서 독립변수를 제외한 모든 다른 변수를 의미하는데(이근희, 2005; 채서일, 2002; Kerlinger & Lee, 2000; Kish, 1959; Rosenberg, 1968), 혼란변수<sup>2)</sup>(confounding variable)로 사용되거나(이근희, 2005; Kerlinger & Lee, 2000), 외생독립변수(Kerlinger & Lee, 2000), 장애변수<sup>3)</sup>(nuisance variable) 등으로 혼용(Breaugh, 2006; Breaugh & Arnold, 2007; Kerlinger & Lee, 2000; Schwab, 2005)되기도 한다. 따라서 여러 가지로 혼용되어 쓰여 지는 용어들의 개념을 정리하고, 이를 통해 통제변수에 대한 개념을 정확히 하기 위해서 용어 구분에 관한 고찰이 필요하였다.

Kish(1959)는 종속변수에 영향을 주는 변동의 원인(sources of variation)으로 설명변수(explanatory variable)와 외생변수를 구분하여 설명하고 있다. 설명변수는 이론체계를 바탕으로 연구의 목적에 맞게 도입된 변수라고 할 수 있는데 '실험변수(experimental variable)'라고 하였다.<sup>4)</sup> 설명변수를 1부류의 변수로 하고 이를 제외한 나머지 변동의 원천인 외생변수를 세 가지 종류로 제시하였다. 먼저, 2부류의 변수는 표본의 선택이나 평가 과정을 거쳐서 통제가 이루어진 외생변수이다. 3부류의 변수는 설

2) 한국통계학회에서는 '중첩변수'로 번역되어 사용되고 있고, 이근희(2005: 215)에서는 '혼란변수'로 사용되었다. 저자들은 외생변수의 효과가 독립변수와 혼재되어 있다는 의미를 강조하여 혼란변수의 용어가 보다 적절하다고 생각하여 본 연구에서는 이를 사용하였다.

3) 현재까지 사회과학 연구에서 'nuisance variable'의 번역된 용어의 사용 사례가 보이지 않기 때문에 한국통계학회의 통계학용어집을 준용하였다.

4) Kish(1959)는 설명변수를 독립변수(independent variable)뿐만 아니라 종속변수(dependent variable)까지 모두 포함한다고 하였는데 이 주장은 현재의 개념 구분상 적절하지 않은 것 같다. 일반적으로 설명변수는 독립변수 또는 예측변수와 같은 의미로 사용되고 있다(이근희, 2005).

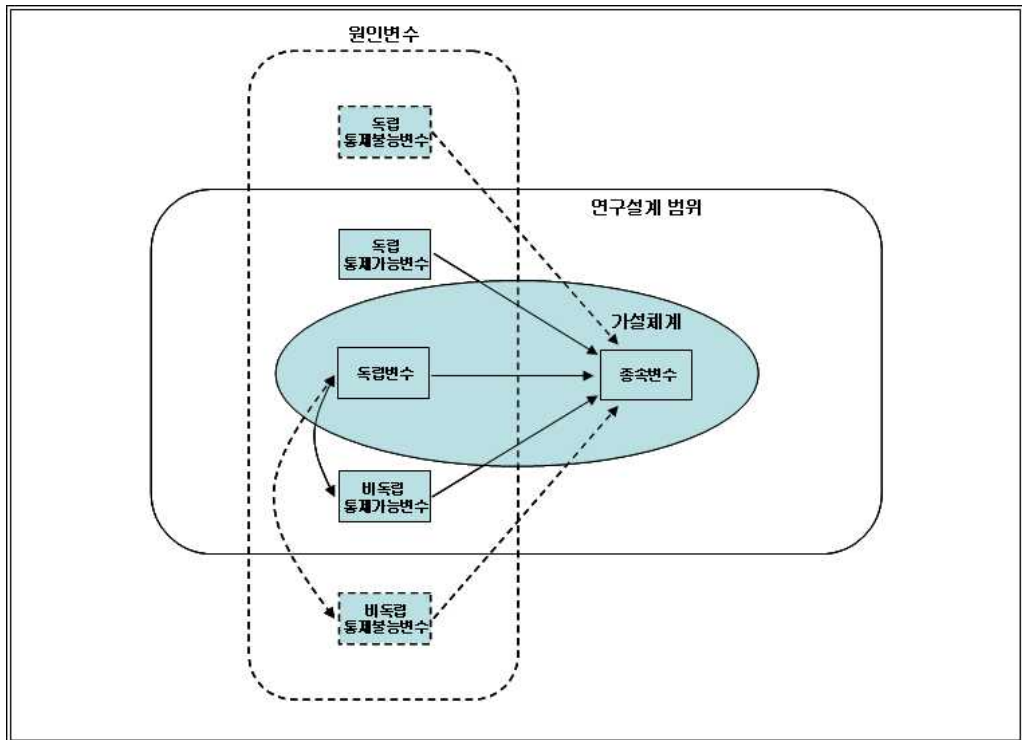


명변수와 혼재(confounded)되어 있어 통제가 되지 않은 외생변수, 4부류는 무작위오차(randomized errors)로 간주되는 통제되지 않는 외생변수이다. Kish(1959)는 효과적인 연구 설계를 위해서는 외생변수를 가능한 많이 2부류의 통제변수로 전환시키고, 3부류의 혼란변수를 설명변수와 분리해 내는 것이라고 하였다. 따라서 이상적인 실험연구는 이러한 3부류의 변수가 전혀 없는 것을 의미하며, 조사연구의 통제는 3부류의 변수를 1부류의 설명변수에서 분리하여 영향을 제거하는 것이다(Kish, 1959). 또한, 3부류의 혼란변수는 무작위화(randomization)를 통해서 4부류의 무작위오류로 전환하여 혼란변수의 영향을 제한할 수 있다고 하였다. Kish(1959)의 이러한 외생변수에 대한 구분의 특성은 표본의 선택이나 절차를 통해서 통제 가능한 변수, 설명변수와 혼재되어 통제변수로 도입하여 통계적인 분석을 통해서 분리할 수 있는 변수, 무작위오차로 구분하였다는 점이다.

본 연구에서는 통제변수에 대한 이해가 용이토록 외생변수를 두 차원으로 구분하였다. 먼저, 독립변수와의 관계에서 '독립성' 여부를 한 차원으로 하여 '독립적 외생변수' 및 '비독립적 외생변수'로 구분하였고, 분석 시에 통제변수로 포함하여 통계적인 방법으로 통제 가능한 변수를 '통제가능 외생변수'로 하고, 통계적인 방법으로 통제가 불가능한 외생변수를 '통제불능 외생변수'로 구분하였다.

인과관계에서 다른 변수에 영향을 미치는 변수를 원인변수라 하고 영향을 받는 변수를 결과변수라고 한다(이군희, 2005). 여기에서 원인변수는 독립변수 및 외생변수를 모두 포함하는 개념으로 이해 될 수 있다. 연구자의 가설체계 안에 포함되고 연구 목적에 부합되는 원인변수를 독립변수라 할 수 있고 이를 제외한 모든 원인변수는 외생변수가 되는 것이다. <그림 1>은 외생변수를 저자들의 분류대로 도식한 것이다.

이러한 외생변수는 독립변수와 독립적(independent) 관계에서 종속변수에 영향을 미칠 수 있고 비독립적(interdependent) 상태에서 영향을 미칠 수 있다. 이를테면 독립적 외생변수는 독립변수와는 상관관계가 없지만 종속변수에는 영향을 미치는 변수들을 의미한다. 예를 들어, '사용하는 인터넷의 속도가 빠를수록 직무만족이 증가 한다.'라는 가설을 세웠을 경우, 여기서 독립변수는 '인터넷 속도'이고 종속변수는 '직무만족'이다. 직무만족에 영향을 주는 여러 요인들 중에서 '긍정적 정서(positive affect)'는 독립변수인 '인터넷 속도'와는 전혀 관계가 없지만 '직무만족'에는 영향을 주고 있다. 따라서 이러한 '독립적 외생변수'들은 독립변수와 종속변수의 관계의 강도(strength)를 변화시키지는



〈그림 1〉 외생변수의 구분

않지만 통제변수로 포함하지 않았을 경우 독립변수가 설명하는 종속변수의 분산을 정확하게 평가할 수 없게 된다. '비독립적 외생변수'는 대부분의 사회과학 연구에서 다루어지는 외생변수로서 독립변수와 관련성이 있으면서 종속변수에 영향을 주는 것으로, 이러한 외생변수의 영향을 통제하지 않았을 경우 독립변수와 종속변수의 관련성을 잘못 해석 (misrepresentation)할 수 있는 문제점이 있다. 예를 들어, '아이스크림의 소비가 증가할수록 냉면의 소비가 증가한다.'라는 가설의 인과성을 검증하고자 한다면, '기온'을 통제하여야 할 것이다. 기온이 높을수록 아이스크림 및 냉면의 소비는 증가한다. 따라서 아이스크림과 냉면의 소비는 높은 관련성을 나타낼 수 있지만 인과관계로 볼 수는 없다. 아이스크림과 냉면 소비의 사실적 관계를 확인하고 싶다면 기온을 통제변수로 포함시켜서 분석하여야 한다.

외생변수를 구분하는 또 다른 하나의 차원은 통계적인 방법으로 통제 가능여부를 기준

으로 삼는 것으로 ‘통제가능변수’ 및 ‘통제불능변수’로 구분이 된다. 대부분의 외생변수는 인지할 수 없거나 측정이 제한되기 때문에 통제가 불가능한 경우가 많다. 위의 예와 같이 ‘긍정적 정서’ 및 ‘기온’ 등은 연구자가 외생변수로 인지할 수 있고 측정이 가능하기 때문에 통계적인 분석을 통해서 통제가 가능하다. 그러나 Campbell & Stanley(1966)에서 제시된 우발적 사건(history)과 같은 외생변수는 독립변수와 독립적이면서 통계적인 방법으로 통제가 불가능한 특성을 가지고 있다. 한편, 비독립적이면서 통계적 방법으로 통제가 불가능한 변수의 예를 생각해보면, 우유와 성장과의 관계에서 자연적인 성장효과(maturation)는 우유로 인한 성장효과와 관련성을 가지면서 전체적인 성장에도 영향을 미치기 된다. 따라서 자연적인 성장효과를 제거함으로써 성장에 대한 우유의 영향을 정확하게 확인할 수 있지만, 자연적인 성장효과를 분리할 수 없기 때문에 이러한 경우 자연적인 성장효과는 비독립 통제불능 외생변수가 된다. 이러한 통제불능 외생변수는 실험연구에서 통제집단과의 비교를 통해서 통제할 수 있거나 연구 진행 간에 정교한 통제 노력으로 그 영향을 제거할 수 있다. <표 1>은 본 연구에서 사용된 외생변수의 구분을 제시하고 있다.

<표 1> 외생변수의 구분 및 특성

구 분	독립성 여부	
	독립	비독립
통계적 통제 가능성	<ul style="list-style-type: none"> <li>· 독립변수와 상관관계가 없고 통제변수로 통계적 분석이 가능한 외생변수</li> <li>· 통제하지 않을 경우, 독립변수의 설명력을 정확하게 확인할 수 없음</li> <li>· 예) 잡음변수(noise)</li> </ul>	<ul style="list-style-type: none"> <li>· 독립변수와 상관관계가 있고 통제변수로 통계적 분석이 가능한 외생변수</li> <li>· 통제하지 않을 경우, 독립변수와 종속변수와의 관계를 왜곡할 수 있음</li> <li>· 예) 허위변수(spurious variable)</li> </ul>
	<ul style="list-style-type: none"> <li>· 독립변수와 관련성이 없으며 통제변수로 통계적 분석이 불가능한 외생변수</li> <li>· 인지 혹은 측정이 불가능하기 때문에 통제변수로 도입될 수 없음</li> <li>· 통계적 분석은 불가하나 통제집단과 대조하는 방법으로 통제할 수 있음</li> <li>· 통제하지 않을 경우, 독립변수의 설명력을 정확하게 평가할 수 없음</li> <li>· 예) 우발적 사건(history)</li> </ul>	<ul style="list-style-type: none"> <li>· 독립변수와 관련성이 있으며 통제변수로 통계적 분석이 불가능한 외생변수</li> <li>· 인지 혹은 측정이 불가능하기 때문에 통제변수로 도입될 수 없음</li> <li>· 통계적 분석은 불가하나 통제집단과 대조하는 방법으로 통제할 수 있음</li> <li>· 통제하지 않을 경우, 독립변수와 종속변수의 관계를 왜곡할 수 있음</li> <li>· 예) 호손효과(hawthorne effect)</li> </ul>

이러한 외생변수의 구분을 통해서 다음과 같은 사실을 이해할 수 있다. 첫째, 비록 인과관계를 검증하기 위해서 가설 관계 외의 설명될 수 있는 외생변수를 통제하는 것이 중요하지만, 통계적 분석방법을 통해서 통제 가능한 외생변수가 있는 반면에 통제 불가능한 외생변수도 있다는 것이다. 이러한 통제불능 외생변수는 본 연구의 초점인 통제변수와는 거리가 있다. 이러한 변수들은 실험연구에서 통제집단과 대조함으로써 제거하거나 무작위화 등 연구 설계를 정교하게 통제하여 달성하여야 할 것이다. 둘째, 외생변수는 독립변수와 독립적 및 비독립적인 관계로 존재하며, 독립적 외생변수는 통계적 분석에 포함되어 통제변수의 역할을 수행함으로써 독립변수가 설명하는 종속변수의 분산을 정확하게 측정하게 한다. 비독립적 외생변수는 독립변수와 종속변수와의 관계를 사실적으로 설명하는데 의의가 있다. 이러한 통제변수의 의의와 필요성을 비실험연구 상황에서 보다 구체적으로 살펴보자.

## 2) 비실험연구에서 통제변수의 의의 및 필요성

인과관계를 증명할 수 있는 연구는 무작위할당 실험연구에서만 가능하다<sup>5)</sup>(Holland, 1986; Rubin, 1974). 그렇지만 실험연구는 실험환경과 실제 환경 사이에는 괴리가 있어 현실에 적용하기 힘들고 일반화가 제한되는 단점이 있다(Kish, 1959; Martin & Sell, 1979). 또한 무작위화는 비용이 너무 많이 든다는 한계와 윤리적인 이유로 제한되며(예: 마약중독이 지능에 미치는 영향), 실험의 결과를 확인하기까지 수년이 소요될 수 있다(예: 콜레스테롤 섭취가 수명에 미치는 영향)는 점이 무작위화의 한계라고 하였다(Rubin, 1974). 따라서 무작위할당 실험 연구는 실제 사회과학 연구에 적용하기에는 제한되기 때문에 실험연구 보다는 비실험연구 및 관찰조사연구(observational study)를 하게 되고 통계적인 통제방법을 이용하여 외생변수를 통제한다(Blalock, 1961a; Frank, 2000; Rubin, 1974; Singer & Marini, 1987; Smith, 1990). 따라서 비실험연구에서의 통제변수의 사용은 독립변수와 종속변수의 관계를 사실적으로 확인하기 위해서 반드시 거쳐야 하는 과정이고, 나아가 인과관계를 추론하는데 반드시 필요한 조치이다.

통제변수 활용의 중요성과 필요성을 이해하기 위해서는 통제변수를 사용하지 않았을

5) Smith(1990)은 비록 무작위할당 실험연구를 하더라도 인과관계를 증명할 수 없는 예외적인 경우를 제시하고 있다. 실험처치(treatment)와 실험참가자의 특성변수(subject characteristic)가 상호작용 할 때 무작위화를 달성하더라도 인과관계를 검증할 수 없다고 하였다.

경우 어떻게 사실이 왜곡되는지 살펴보는 것이 중요하다. 연구에서 흔히 간과하기 쉬운 현상들을 살펴보았다.

인과모형(causal model)은 연구자가 실제 관심을 가지고 있는 독립 및 종속변수뿐만 아니라 영향을 미치는 관련 변수들을 포함하여 구성하여야 된다. 예를 들어, 보상수준이 직무만족에 미치는 영향을 검증하고 싶다면 직위, 직무특성, 리더십 등의 변수들이 고려되어야 한다. 만약 이러한 변수가 연구모형에 포함되지 않는다면 상술오류(misspecification)를 범하게 된다(Schwab, 2005). 장애변수는 모형에 포함하지 않으면 상술오류를 초래하는 변수를 의미하며 편의(bias), 허위변수(spurious variable), 억제변수(suppressor variable), 잡음(noise)으로 구분된다(Schwab, 2005).<sup>6)</sup>

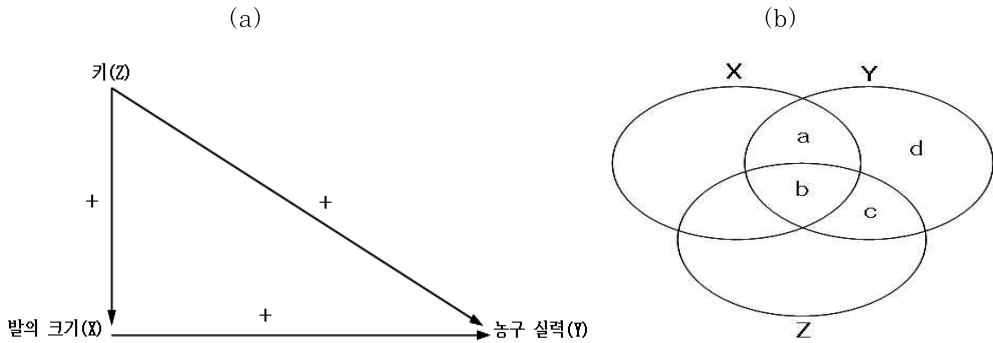
편의는 한 변수 Z가 종속변수 Y와 인과적인 관계를 가지고 있고 독립변수 X와도 관련성이 있을 경우, 만약 변수 Z를 통제하지 않으면 독립변수 X와 종속변수 Y와의 관계를 과소 및 과대평가하게 된다. 예를 들어 <그림 2>에서, 연구자가 '밭이 클수록 농구를 잘 할 것이다.'라는 가설을 세웠다면, 밭(X)이 크면 몸의 균형 유지가 용이하고 안정적이기 때문에 농구(Y)를 잘 할 수 있을 것이라 기대할 수 있다. 그렇지만 농구 실력에 중요한 키(Z)를 고려하지 않았다면 연구는 심각한 오류를 범할 수 있다. 일반적으로 키가 크면 밭도 크기 때문에 키와 밭의 크기는 관련성이 높을 것이다. 이러한 상황에서 만약 키를 통제하지 않는다면 밭의 크기와 농구실력의 관계를 과대평가하게 된다. 따라서 키는 밭의 크기와 농구 실력의 편의가 되는 것으로 반드시 통제변수로 포함하여 분석하여야 한다.

<그림 2>의 (b)는 편의관계를 일반적인 회귀분석 맥락에서 도식화하였다. 먼저, X와 Y만을 고려한 단순회귀분석에서 중첩 부분인 a, b는 X에 의해 설명되는 Y의 분산을 의미하며 따라서  $r^{27}$  값은  $[a+b/a+b+c+d]$ 이다. 편의 변수인 Z를 통제한 상태에서 독립변수 X에 대하여 종속변수 Y를 다중회귀분석하면 독립변수 X에 의해서 설명되는 Y의 고유분산(unique variance)을 도출할 수 있다. 고유분산은  $R^2$ 변화량(incremental  $R^2$ )을 의미하며  $R^2$ 변화량의 제공근은 부분상관계수(semipartial correlation coefficient)와 일치한다. 여기서  $R^2$ 변화량은  $[a/a+b+c+d]$ 로서 장애변수 Z와 중첩된 X의 부분

6) Schwab(2005)가 제시한 장애변수에는 매개변수(mediators)와 조절변수(moderators)도 포함하지만, 일반적으로 이 변수는 연구자의 가설에 포함되어 분석되기 때문에 통제된다고 볼 수 있다. 따라서 본 연구에서는 제외하였다.

7) 일반적으로 단순결정계수는 소문자로 표기함으로  $r^2$ 으로 하였고 다중결정계수( $R^2$ )는 대문자로 표기한다.

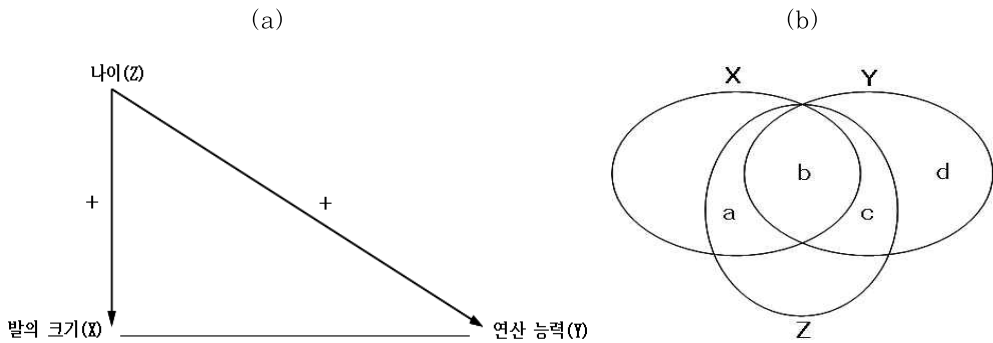
b는 제거되어진다. 따라서 편의변수인 키를 통제하지 않을 경우 발의 크기와 농구실력의 관계는 b부분만큼 과대평가되어 나타난다.



〈그림 2〉 편의 관계

통제되지 않은 변수 Z가 독립변수 X 및 종속변수 Y와 인과적 관계가 있고 독립변수 X와 종속변수 Y는 실제로 아무런 관련성이 없지만 상관관계가 나타날 때 X는 허위 변수일 가능성이 높으므로 Z를 모형에 포함하여 통제하여야 한다(Blalock, 1961b; Blalock, 1962; Schwab, 2005; Simon, 1954). 예를 들어, 학생들을 대상으로 한 연구에서 연구자가 가설로서 ‘발의 크기가 클수록 연산을 잘 할 것이다.’라고 설정하였을 경우 실제로는 독립변수와 종속변수는 아무런 관련성이 없는 데도 불구하고 관련성이 나타날 수 있다. 〈그림 3〉에서, 나이(Z)가 늘어나면 학년이 올라가고 연산을 배움에 따라 연산(Y)을 잘 할 수 있을 것이다. 또한 나이가 늘어감에 따라 체격이 커지면서 발(X)도 커질 것이다. 따라서 ‘나이’가 허위변수의 역할을 하여 두 변수 X, Y에 동일하게 영향을 줌으로써 X, Y의 관련성이 나타났다고 볼 수 있다. 그러므로 발(X)과 연산(Y)의 순수한 관계를 검증하기 위해서는 나이(Z)를 통제하여야 할 것이다. 이와 같이 허위변수를 통제함으로써 거짓인 가설을 채택하는 I형 오류를 예방할 수 있다(Rosenberg, 1968).

〈그림 3〉의 (b)는 허위관계를 회귀분석 맥락에서 도식화하였다. X와 Y의 단순회귀분석에서 중첩 부분은 b로서  $r^2$ 값은  $[b/b+c+d]$ 이다. 나이 Z를 통제한 상태에서 발의 크기 X에 대하여 연산능력 Y를 다중회귀분석하면  $R^2$ 변화량은  $[0/b+c+d]$ 로서 0이 된다. 그림에서 X에 의해 설명되는 Y의 부분 b는 Z와 X의 중첩부분 a+b에 포함되어 있기 때문에 이를 제거하면 X와 Y의 고유분산은 사라지게 된다. 따라서 발의 크기와 연산



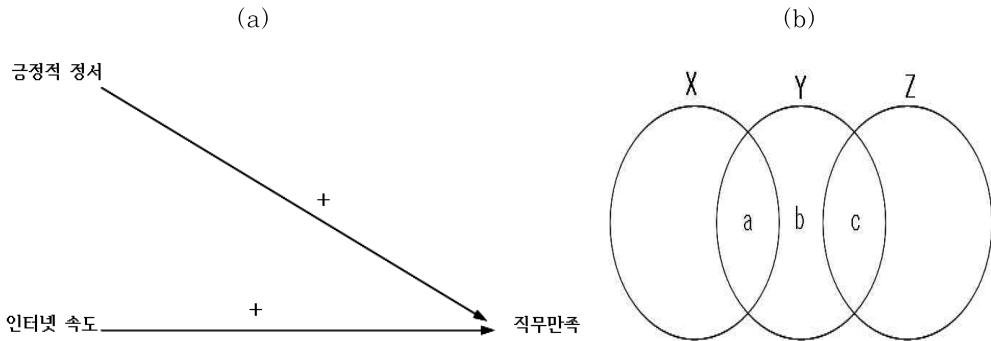
〈그림 3〉 허위 관계

능력은 비록 단순상관관계가 나타나더라도 허위관계임을 이해할 수 있다.

종속변수(Y)와는 아무런 관련성이 없거나 약한 부의 상관관계를 가지고 있고, 독립변수(X)와는 유의미한 정의 상관관계가 있을 경우 이러한 변수들은 억제변수의 역할을 하는데 통제하지 못한다면 독립변수(X)와 종속변수(Y)와의 실제 관계를 과소평가하게 된다. 그러므로 정확한 관련성을 확인하기 위해서는 모형에 포함하여 통제를 해야 한다(Cohen et al., 2003; Kerlinger & Lee, 2000, Maassen & Bakker, 2001; McNemar, 1945; Pedhazur, 1982; Rosenberg, 1968; Schwab, 2005). 억제변수는 독립변수에서 관련성이 없거나 부적절한 분산을 제거함으로써 참인 가설을 기각하는 사태(Ⅱ형 오류)를 예방할 수 있다(Rosenberg, 1968). 예를 들어, 연구자가 기계조작 이론 시험(X)과 실기능력(Y) 간의 관계를 연구하고자 한다면 억제변수로 '독해능력(Z)'을 통제변수로 포함하여 분석하여야 한다. 여기서 '독해능력'은 이론시험(X)과는 관련성이 높지만 종속변수인 실기능력(Y)과는 관련성이 없을 것으로 기대된다. 이론 시험은 독해능력(Z)의 영향을 많이 받기 때문에 '독해능력'을 포함하지 않으면 실질적인 이론능력을 측정한 것으로 볼 수가 없어 이론시험(X)과 실기능력(Y)의 관계는 아주 약하게 나타나게 된다. 따라서 '독해능력'은 억제변수의 기능을 하게 되는데, 이를 모형에 포함시켜 독해능력의 차이에서 생긴 이론시험의 분산을 제거함으로써 이론시험과 실기능력 간의 관계를 증가시킬 수 있다.

잡음은 종속변수와는 관련성이 있지만 독립변수와는 아무런 관련성이 없는 변수이다. 통제하지 않을 경우 독립변수와 종속변수의 관련성에는 영향을 미치지 않지만, 종속변수의 변동에 영향을 미치기 때문에 만약 통제하지 않게 되면 실제적인 종속변수의 변동을

측정할 수 없다는 문제점이 발생한다. 앞의 예에서 제시된 것처럼, 인터넷의 속도와 직무만족 간의 관계에서 긍정적 정서는 잡음로서 독립적 통제가능변수의 역할을 한다고 볼 수 있다.



〈그림 4〉 잡음 관계

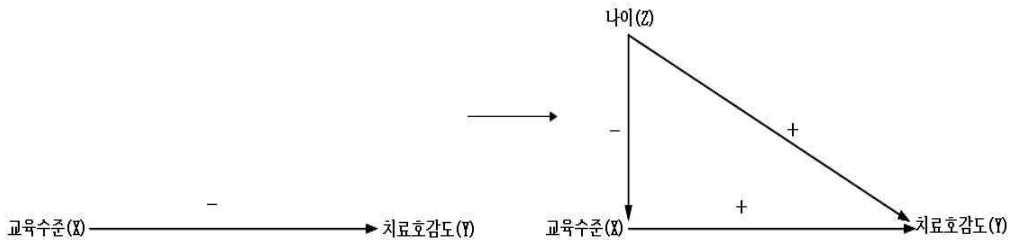
〈그림 4〉의 (b)에서 X와 Y의 단순회귀분석에서  $r^2$  값은  $[a/a+b+c]$ 이다. X에 독립적인 변수 Z를 포함하여 이 변수를 통제한 상태에서 다중회귀분석을 하더라도 고유분산  $R^2$  변화량은  $[a/a+b+c]$ 로서 단순결정계수와 동일하다.

잡음 변수와 같이 독립변수와 독립적인 변수일 경우 통제변수로서 포함하여야 할 필요성이 간과될 수 있지만, 다음 몇 가지 측면에서 의의를 찾을 수 있다. 첫째, 종속변수에 영향을 미치는 변수임에도 불구하고, 관심변수가 아닌 이유로 누락하면 상술오류를 범할 수 있다(Schwab, 2005). 인과관계를 규명하기 위해서는 종속변수에 영향을 미치는 다른 요인들을 독립변수의 영향과 분리하는 것이 필요하며 이를 위해서는 독립적인 외생변수의 효과 또한 통제되어야 한다(이군희, 2005; 채서일, 2002; Cook & Campbell, 1979; Schwab, 2005). 둘째, 독립외생변수를 모형에 포함함으로써 연구자가 관심을 가지고 있는 독립변수의 상대적인 효과를 확인할 수 있다. 다중회귀분석을 수행한다면, 표준화된 회귀계수( $\beta$  coefficient)를 비교함으로써 상대적인 영향력을 확인할 수 있다. 셋째, 종속변수 값을 예측하는데 필요하다. 다중회귀분석에서 회귀방정식을 통해 종속변수를 예측할 수 있는데 이를 위해서는 독립변수뿐만 아니라 외생변수의 설명력까지 포함하여 모형을 구체화하여야 한다. 독립변수와 관련성은 없지만 종속변수와는 상관이 있는 변수를 만약 생략한다면 그 모형에서 전반적인 추정은 편의가 발생하게 된다(Schwab,



2005).

Rosenberg(1968)는 독립변수(X)와 종속변수(Y)의 관계를 사실적으로 분석하기 위해서 제 3의 변수를 도입이 필요함을 설명하면서 이를 검증요인(test factor)<sup>8)</sup>이라고 하였다. Rosenberg의 검증요인 중에서 왜곡변수(distorter variable)는 독립변수(X)와 종속변수(Y)의 관계를 역으로 변화시키는 역할을 한다. 따라서 이러한 왜곡변수를 분석에 포함하지 않을 경우 결과는 정반대로 나타나게 된다. 예를 들면, 연구자가 ‘교육수준(X)이 높을수록 치료호감도(Y)가 높을 것이다’라는 가설을 세웠다고 하자. 자료를 분석하였을 때 연구자의 예측과 반대로 결과가 나왔다면 가설은 기각될 것이다. 그러나 이러한 결과는 ‘나이(Z)’라는 왜곡변수를 통제하지 않아 초래된 것이라 볼 수 있다. 일반적으로 나이(Z)가 많으면 교육수준(X)이 낮고 반면에 의료치료에 대한 호감도(Y)는 높을 것이다. 따라서 나이를 통제하면 순수하게 교육수준과 의료치료 호감도 간의 정(+)의 관계를 확인할 수 있다. <그림 5>에서 보는 바와 같이, 왜곡변수와 독립변수 및 왜곡변수와 종속변수 간 관계 방향이 서로 다를 경우 독립변수와 종속변수의 관계는 초기의 두 변수(X, Y) 간의 관계와 반대방향으로 나타난다. 따라서 통제를 하지 않을 경우 결과 자체가 반대로 해석된다는 점에서 다른 장애변수에 비해 중요한 변수이다.



<그림 5> 왜곡 관계

8) Rosenberg(1968)는 검증요인(test factor)으로 6가지 요인(extraneous variable, component variable, suppressor variable, distorter variable, intervening variable, antecedent variable)을 제시하고 있다. 이러한 제3의 요인들을 분석에 포함함으로써 독립변수와 종속변수의 관계를 정확하게 예측할 수 있다고 하였다. 여기서 component variable, intervening variable, 그리고 antecedent variable은 본 연구의 중점인 ‘통제변수’의 역할과는 차이가 있다. component variable은 개념의 복잡성으로 인해 관계를 정확히 해석하지 못할 상황에서 필요한 검증요인이다. intervening variable 및 antecedent variable은 각각 매개변수와 독립변수의 선행요인에 대한 설명으로 일반적으로 가설에 포함시켜 연구가 이루어지기 때문에 본 연구에서 제외되었다. extraneous variable은 Schwab(2005)의 허위(spurious)와 동일한 개념으로 설명되고 있다.

이제까지 살펴본 대로 통제변수의 사용은 독립변수와 종속변수 간의 인과관계에 대한 확신을 증가시키며, 관계를 잘못 해석하는 것을 방지할 수 있고 정확하고 구체적인 관계를 이해하는 데 도움이 된다(Rosenberg, 1968).

또한, 연구의 일반화에도 기여하게 된다. 사회과학 연구는 단 한 번의 연구결과만으로 연구문제에 대한 답을 제시할 수 없고, 여러 연구로부터 반복적인 동일한 연구결과를 축적함으로써 문제 해결이 가능하다(Schwab, 2005). 특정 현상을 설명하기 위한 연구에서, 한 연구에서의 독립변수가 다른 연구에서 통제변수로 사용될 수도 있다. 만약 이러한 연구들이 축적된다면 메타분석(meta analysis)을 통해서 양적인 측정이 가능하므로 연구의 일반화가 가능하게 된다(Schwab, 2005).

### 3) 통제변수 활용에서의 고려사항

Rosenberg(1968)는 통제변수를 선택하는 것은 다분히 논리적인 과정이라고 설명하고, 통제변수 포함여부를 결정하는 데 있어 몇 가지 고려할 점을 제시하였다. 첫째, 관련성에 대한 이론 및 실증적 근거가 있어야 한다. 둘째, 독립변수 및 종속변수와의 상관관계를 확인하여 각각 관련성 있는 변수를 포함하여야 하며 관련성이 없음을 확신할 수 없을 때 포함한다. 종속변수 및 독립변수와 아무런 관계도 없고 이론적인 근거도 부족한 통제변수의 사용은 오히려 독립변수의 설명력을 약화시키고 'II형 오류'를 낳게 되므로(Becker, 2005) 이런 변수의 포함은 부적절하다. 따라서 어떤 통제변수를 선택하여 모형에 포함하는지가 해당 연구 모형을 정확하게 해석하는데 중요한 문제이다.

통제변수의 활용은 연구의 일반화에 기여하고 리뷰 및 메타분석 시에 제 3의 연구자가 결과를 해석하는데 기여한다. 따라서 연구자는 구체적인 통제변수의 측정방법과 기술통계, 분석결과 효과크기 및 유의수준을 기술하여야 한다(Becker, 2005).

본 연구는 이러한 통제변수 활용의 주요 고려사항을 바탕으로 실태 분석의 초점을 '통제변수의 포함여부', '적절한 통제변수의 사용', 그리고 '통제변수의 측정 및 기술통계', '분석결과의 제시'에 두었다. 통제변수 활용현상에 대한 분석을 통해, 연구자의 통제변수에 대한 인식수준을 가늠할 수 있으며 문제점을 도출하고 이에 따른 적절한 제언을 제시할 수 있을 것이다.

### Ⅲ. 국내의 통제변수 활용 실태 분석

#### 1. 분석방법

##### 1) 분석대상 논문

분석 대상 논문은 경영학연구 및 인사·조직연구에 실린 인사조직 관련 논문이다. 이들 저널에서 지난 10년 간(1997~2006년) 출판된 논문으로서 경영학연구는 인사·조직관련 연구논문만을 한정하였다. 연구 기간을 지난 10년으로 한정 한 것은 비록 임의적인 측면이 있지만, 10년이란 기간은 연구에서 방법적인 측면의 변화를 충분히 설명할 수 있는 시간으로 판단했기 때문이다. 또한, 이 기간의 연구에 대한 전수조사는 모든 형태의 연구를 충분히 망라했다고 볼 수 있다. 통제변수의 활용 실태를 분석하는 연구의 목적에 적합하지 않은 논문들은 제외되었는데, 계량적 분석 방법(quantitative analysis)이 아닌 질적(qualitative)인 연구들과 리뷰 논문 및 실험설계를 통한 외생변수를 통제 한 연구는 분석대상에서 제외되었다. 전체 231편에서 52편의 논문이 이러한 이유로 제외되어 분석 대상 연구 사례는 179편이었다.

〈표 2〉 연도별 대상 논문

	계(%)	97년	98년	99년	00년	01년	02년	03년	04년	05년	06년
계	179	14	14	10	12	19	17	21	25	20	27
경영학연구	56(31.3)	4	4	3	4	3	6	9	5	6	12
인사·조직연구	123(68.7)	10	10	7	8	16	11	12	20	14	15

##### 2) 분석 방법 및 평가 기준

계량적 분석을 위한 데이터는 연구자들이 179편의 논문들을 리뷰하면서 각각의 분석 범주 및 세부 항목별로 통제변수 사용 실태를 질적 분석함으로써 도출하였다. 이 연구의 목적이 통제변수의 사용에 있어서 구체적인 문제점과 모범사례를 도출하는 것이므로, 이러한 연구 목적을 위해 통제변수를 두 종류로 구분하였다. 추상적 개념수준이나 조작적 측정수준이나에 따라 변수를 구분할 수 있는데 전자를 개념적 변수(conceptual variable), 후자를 조작적 변수(operational variable)라 한다(Schwab, 2005). 이를 바탕으로

Becker(2005)는 통제변수를 CCV(conceptual control variable) 및 MCV(measured control variable)로 구분하였는데 본 연구에서는 이 구분법을 준용하였다. CCV는 연구자가 통제하고자 하는 개념적 변수이며 MCV는 이러한 CCV를 조작적 정의를 통해서 측정 가능하도록 만든 변수를 말한다. 예를 들어 '조직의 규모' 변수를 통제하고자 한다면, 여기서 '조직의 규모' 변수 그 자체는 CCV가 되며 조직의 규모를 조직에 속한 '종업원의 수'로 조작적 정의를 하였다면 '종업원의 수'는 MCV가 된다.

CCV 및 MCV로 구분함으로써 통제변수의 사용 실태를 하위범주까지 구체적으로 분석할 수 있다. 분석의 기준은 통제변수도 독립변수나 종속변수처럼 분석에 있어서 동일하게 취급(Becker, 2005)되는지 여부로써 구체적인 분석기준은 아래와 같다.

1. CCV 포함여부
  - CCV를 포함하였는가?
  - CCV를 포함하게 된 이유를 설명하였는가?
  - CCV를 포함하게 된 이유를 인용이나 구체적인 근거를 제시하였는가?
2. CCV 측정방법의 구체화
  - CCV를 측정하기 위한 방법을 구체화 시켰는가?(MCV 제시 여부)
3. MCV 기술통계
  - MCV에 대한 평균과 표준편차는 제시되었는가?
  - MCV와 다른 관심변수들 간의 상관계수 크기와 유의수준이 제시되었는가?
4. 분석결과
  - MCV와 종속변수간의 분석결과, 효과크기 및 유의수준 등이 기술되었는가?
5. 포함된 통제변수의 적절성
  - 포함된 통제변수는 종속변수와의 관련성이 구체적인 근거로써 뒷받침되고 분석결과 관련성이 있었는가?
  - 이전의 연구 결과에서 종속변수와의 관련성도 없으며 분석결과에서도 아무런 관련성이 없는 변수가 통제변수로서 활용되지 않았는가?  
(불필요한 통제변수를 도입함으로써 연구결과를 과대/과소 해석)
6. 분석결과의 논의
  - 분석결과에 대한 논의에 통제변수도 포함하였는가?

본 연구는 분석의 구체적인 항목에 있어서 Becker(2005)의 연구를 참조하였지만 약간의 차이가 있다. Becker(2005)의 분석 항목에 추가해서 '포함된 통제변수의 적절성' 및 '분석결과의 논의'를 포함하였다.

통제변수를 포함하지 않는 것도 문제지만 아무런 관련 없는 통제변수의 활용은 과대통

제의 문제를 야기한다(Bandura, 1997; Breugh, 2006; Gordon, 1968; Pedhazur, 1982). 따라서 포함된 통제변수들이 실제로 적절한 관련성을 나타내고 있는지 상관분석 결과를 분석하여 그 적절성을 확인하였다.

분석 결과를 바탕으로 통제변수에 대해서 논의하는 것은 단순히 통제변수와 종속변수의 관계만을 이해하는 그 이상의 의의를 제공한다. 이러한 과정은 다수의 인과관계 증거들을 축적하게 함으로써 인과관계를 추론하는데 이점을 제공하며 또한 일반화의 이점을 제공한다(Marini & Singer, 1988). 어떤 특정 연구자 A의 통제변수가 B 연구자의 연구에서는 독립변수가 될 수 있고 B의 통제변수가 A의 독립변수가 될 수 있다. 연구의 맥락은 동일할 수 없으므로 A 및 B 연구자의 연구결과를 통해서 다양한 상황에 적용된 연구 결과를 얻을 수 있어 이론의 일반화에 기여할 수 있고 또한 향후 메타 분석과 같은 제 3자의 연구에 적절한 자료가 제공될 수 있다. 따라서 이러한 측면을 고려해서 '통제변수 분석 결과 논의 여부'를 통제변수 사용의 평가 기준에 포함시켰다.

### 3) 코딩 및 양적 분석

질적 분석 자료를 변환하여 계량적 분석에 활용하였다. 저널명, 출판연도, 페이지 수, 데이터 특성, 분석방법, 주요 변수의 수, CCV 및 MCV 개수 등과 통제변수 사용에 대한 분석 결과가 포함되었다.

저널명은 경영학연구 0, 인사·조직 연구 1로 구분하고 연도는 해당 논문이 게재된 저널의 발행 연도이다. 페이지 수는 대상 논문이 차지하고 있는 저널의 페이지 수를 의미하며, 데이터 특성의 구분은 현장조사연구 0, 문헌자료 사용 1로 코딩하였다. 분석방법은 회귀분석, 구조방정식(structural equation model), 분산분석 그리고 기타로 구분하였다. 주요 변수의 수는 연구에서 사용된 전체 변수의 수에서 통제변수를 제외한 변수들을 의미하며 하위개념별로 구분될 경우 개념별로 구분하여 수를 환산하였다. CCV의 개수는 연구자가 통제하고자 하는 개념적 통제변수의 수를 의미하고 MCV는 실제로 측정될 수 있는 CCV의 수를 의미한다. MCV는 개념적 CCV의 수와 대부분의 경우에 그 수가 일치하지만 CCV의 개념이 하위개념으로 구분될 때에는 하나의 CCV에 두 개 이상의 MCV가 도출될 수도 있다. 세부적인 항목별 코딩 방법은 부록에 제시되었다(부록 1 참조).

## 2. 분석 결과

### 1) 모범적인 사례

각각의 범주별로 통제변수의 모범적인 활용 사례를 도출하였고, 그 중 일부를 제시하겠다. 이러한 모범적인 사례는 향후 연구자에게 적절한 사용 방법을 제시할 수 있을 것이다.

#### (1) 통제변수 선택

통제변수를 선택하고 모형에 포함하게 된 이론적 근거를 제시하거나 실증 결과를 인용하였는가에 관한 모범 사례이다. 분석 결과, 각 연구에 포함된 모든 통제변수에 대해서 합리적인 설명이나 인용이 뒷받침 된 연구 사례는 179편 중 19편(10%)이었다. 대표적인 연구로서, 김해룡(2003)은 기업집단 본부의 통제가 계열사의 연구개발투자에 미치는 영향에 대한 실증연구를 하면서 바람직한 통제변수의 선택 과정을 보여 주고 있다.

본 연구의 결과변수인 연구개발투자는 기업집단 본부의 통제이외의 다른 변수들에 의해 영향을 받을 수 있으므로 그러한 변수들의 방해효과를 통제할 필요가 있다. 기존 연구들에서 사용된 기업규모, 유동비율, 그리고 기업성과 등 세 가지 변수들이 통제변수로 가설검정 모형에 포함되었다. 우선, 기업규모는 연구개발투자에 영향을 미치는 주요 요인들 가운데 하나로 인식되고 있다. 규모의 경제를 누릴 수 있는 규모의 기업은 연구 개발에 드는 비용이나 혜택을 빠르게 회수하고 실현할 수 있기 때문이다(Rothwell, 1984). ..... 기업성과 역시 연구개발투자에 영향을 주기 쉽다. 높은 수준의 기업성과는 연구개발투자와 신제품 도입에 영향을 주는 것으로 보고된다(Chaney & Devinney, 1992). ..... 마지막으로 유동비율이 포함되었는데, 기업의 유동성은 연구개발에 사용할 수 있는 가용자금에 영향을 준다고 보고된다(Baysinger & Hoskisson, 1989).(p. 55)

#### (2) 통제변수 측정의 방법

개념적 통제변수(CCV)를 측정하기 위해 MCV를 구체화했는지에 관한 분석으로, 43편(24%)의 연구가 적절한 설명을 제시한 것으로 나타났다. 김양민(2006)은 외국 기업 인수 시에 인수를 행하는 기업의 가치를 올리는 요인을 확인하기 위한 연구에서 통제변수로 사용된 '상대적인 인수의 규모' CCV를 '발표된 인수 가격과 인수하는 기업의 매출액의 비'로 정의하여 측정하고 있다. 그리고 인수규모 및 기업의 매출액에 대한 측정방법

을 구체적으로 기술하고 있다.

독립 변수 이외에 종속변수에 영향을 미칠 수 있는 변수들을 통제하기 위하여 상대적인 인수의 규모, 다른 인수희망자의 존재여부, 인수를 행하는 해당 기업의 사전 기업성과, 인수대상국가에 해당 기업이 이미 진출해 있는지의 여부, 해당 기업의 부채 비율 등을 측정하였다. ....인수를 행하는 기업이 비해서 인수 대상 기업이 크다면 주식 시장에서의 반응은 더욱 클 것이다. .... 상대적인 인수의 규모는 발표된 인수 가격을 인수를 행하는 기업의 매출액으로 나누어 측정 하였다. 인수 규모는 Mergers & Acquisitions에서 그리고 인수를 행하는 기업의 매출액 규모는 COMPUSTAT를 통해 측정하였다.(p. 129)

양동훈(2005)은 보상제도가 협동행동에 미치는 영향에 관한 연구에서 통제변수로서 '집단주의 성향'을 고려하고 있는데 사용한 측정도구를 기술하고 이러한 다항목 매트릭 척도에서 요구되는 신뢰도와 타당도의 정화과정을 구체적으로 제시하고 있다.

집단주의(collectivism)는 Traiandis(1990)의 설문 8개항을 활용하였으며 ..... 집단주의는 지각에 기초한 변수이므로 이들 변수에 대한 신뢰성과 타당성 그리고 각 변수들의 단일 차원성을 조사하였다. 신뢰성에 대해서는 Nunnally(1978)의 항목모집단 추출모형(Domain Sampling Model)에 근거한 정화절차(purification process)를 적용하여 일부 신뢰성을 낮추는 항목들을 제거하였다. 그 결과로서 신뢰도를 낮추는 항목으로 집단주의에서 2개 설문항목 .....제거되었다.(p. 20)

### (3) 기술통계량 및 분석결과의 보고

통제변수로서의 적절성을 판단하고 분석 결과에 대한 논의를 위해서는 포함된 통제변수의 평균 및 표준편차와 같은 분포에 관한 정보와 변수 간의 상관관계가 제시되어야 한다. 또한 가설 검증을 위한 분석에서 통제변수의 효과크기(effect size) 및 유의수준(significance level)이 보고되어야 한다. 분석 결과, 대부분의 연구에서 사용된 MCV의 평균 및 표준편차, 상관관계 계수 및 유의수준, 분석결과 효과크기 및 유의수준이 모두 잘 나타나고 있었다. 예를 들어, 임준철·윤정구(1998)의 논문에서 연구자들은 사용된 8개의 통제변수에 대해서 모든 변수의 평균 및 표준편차, 상관분석 결과를 제시하였고 회귀분석 결과에도 회귀계수 및 유의수준을 제시하고 있다. 또한 이들은 통제변수가 종속변수에 미치는 영향을 분석 결과 논의에 포함시키고 있어 바람직한 모습을 보여 주고 있다.

직무관련 통제변수 및 인구통계변수들 중에서는 근무기간과 나이만이 유의하게 직무만족에 영향을 주고 있는 것으로 나타났다. 즉 나이와 직위, 그 밖의 다른 변수들의 영향력을 통제하더라도 근무기간이 길수록, 그리고 나이가 많을수록 현 직무에 만족하는 정도가 높은 것으로 나타나고 있다. 이중에서 근무기간은 4.5년을 기준으로 직무만족과 정(+)의 비선형 관계가 있음이 판명되었다. 즉 4년 반이 될 때까지 종업원들의 직무에 대한 만족은 완만하게 감소하다가 이 시점을 기준으로 다시 증가하는 추세를 보이고 있다. 나이와 근무연한을 중심으로 하는 연공이 올라갈수록 종업원들의 직무만족도 증가한다는 것은 지금까지 한국기업이 연공중심의 경영체계를 유지해오고 있었음을 간접적으로 뒷받침한다고도 볼 수 있다.(p. 101)

## 2) 통제변수 활용의 문제점

국내 연구에 있어서 통제변수의 활용은 비록 앞에서 제시한 모범적인 사례에도 불구하고 많은 문제점을 내포하고 있었다. 여기에서는 개별적인 사례보다는 현상에 대한 계량적 분석을 통해 문제점을 도출하였다.

### (1) 통제변수 포함 여부

각 연구들에서 기본적으로 통제변수의 포함여부를 확인하였다. 전체 179편의 논문에서 83편(46.4%)이 통제변수를 사용하지 않고 연구가 이루어졌다. 이를 연도별 및 분석방법별로 구분하여 살펴보았다. 연도별 변화 추이는 설명할 수 있는 일정한 형태가 없었다. 연도별 CCV를 포함하지 않은 논문의 수는 아래 <표 3>에서 제시되고 있다.

<표 3> 연도별 CCV를 포함하지 않은 논문 수

계	97년	98년	99년	00년	01년	02년	03년	04년	05년	06년
83	10	8	6	10	10	4	8	10	5	12

<표 4>는 분석방법별 포함된 통제변수(CCV)의 수를 분산분석을 통해 확인한 결과로써 유의미한 차이가 있었다. 사후 검증 결과 회귀분석은 구조방정식 및 분산분석과 유의미한 차이가 있었으며 구조방정식과 분산분석은 차이가 없었다. 두 방법은 회귀분석에 비해 통제변수를 거의 사용하지 않는 것으로 나타났다. 연구에서 통제변수를 아예 사용하지 않은 것은 인과관계를 확신하는데 상당한 위험을 내포하고 있으므로 이러한 경향에 대한 연구자의 주의가 필요하다.



〈표 4〉 분석방법별 CCV 수 분산분석 결과

	사례수	평균	표준편차	분산분석 결과
회귀분석	112	3.321	2.9598	
구조방정식	36	.056	.03333	df: 3
분산분석	20	.200	.6156	F: 20.194
기 타	11	3.455	4.1800	p<.000
계	179	2.324	2.9597	

### (2) 통제변수 설정 근거

통제변수를 연구를 포함하는 데 있어서 주의가 요망되는데 이는 부적절한 통제변수의 포함은 II종 오류를 낳을 수 있고, 필요한 통제변수를 포함하지 않으면 결과를 왜곡 해석하는 I종 오류를 초래하기 때문이다(Becker, 2005). 따라서 연구자는 통제변수를 연구에 도입함에 있어서 선행연구나 구체적인 이론적 근거를 바탕으로 포함하여야 할 것이다.

모든 통제변수에 대해서 구체적인 근거를 들어 포함시킨 경우는 11.2%였으며 전혀 설명을 하지 않은 경우가 24%에 이르렀고, 1개 이상의 통제변수에 대해서 아무런 설명을 하지 않은 경우를 포함하면 77.1%로서 절반 이상의 연구에서 통제변수로 포함함에 있어 적절한 주의를 하지 않는 것으로 나타났다.

포함된 통제변수를 측정하기 위해서 구체적인 측정방법을 설명하여야 한다. 또한 해당 변수 개념을 측정하기 위한 여러 가지 도구 중에서 연구 의도에 적합한 측정도구를 선택해야 한다. 앞의 설명에서 ‘조직의 규모’는 ‘종업원의 수’로 정의할 수도 있고 ‘매출액’으로도 나타낼 수 있다. 따라서 어떤 측정방법을 사용할 것인지 구체적인 언급이 있어야 독자의 이해를 증진시킬 수가 있다. 측정방법을 구체적으로 설명했는지 여부에 대해 분석이 이루어졌다. 모든 통제변수에 대해서 설명을 한 연구는 24%였고 모든 변수에 대해서 아무런 측정방법을 설명하지 않은 연구도 17.3%가 되어 문제점이 있었다.

### (3) 기술통계, 분석 결과 및 논의

측정된 MCV의 기술적 통계치의 제시여부를 확인하였다. 많은 연구에서 평균 및 표준편차, 상관관계의 크기와 유의수준을 제시하고 있었지만 12.3%의 연구에서 이러한 기술통계치를 발견할 수 없었다.

〈표 5〉 하위범주별 통제변수 활용 실태 비율

구 분	하 위 범 주	백분율(%)
통제변수 설정근거	통제변수를 포함하지 않음	46.4
	전 변수에 대해서 아무런 설명을 하지 않음	24
	1개 이상의 변수에 대해서 설명을 하지 않음	6.7
	1개 이상의 변수에 대해서 불명확하거나 단순 설명	11.7
	모든 변수에 대해서 인용 및 근거를 제시	11.2
측정방법 구 체 화	통제변수를 포함하지 않음	46.4
	전 변수에 대해서 아무런 설명을 하지 않음	17.3
	1개 이상의 변수에 대해서 설명을 하지 않음	12.3
	모든 변수에 대해서 구체화된 측정 방법을 제시	24
기술적 통계	통제변수를 포함하지 않음	46.4
	전 변수에 대해서 아무런 설명을 하지 않음	12.3
	1개 이상의 변수에 대해서 설명을 하지 않음	3.9
	모든 변수에 대해서 설명을 함	37.4
분석결과	통제변수를 포함하지 않음	46.4
	전 변수에 대해서 아무런 설명을 하지 않음	4.5
	1개 이상의 변수에 대해서 설명을 하지 않음	1.7
	모든 변수에 대해서 설명을 함	47.5
논의 포함	통제변수를 포함하지 않음	46.4
	전 변수에 대해서 아무런 설명을 하지 않음	40.8
	1개 이상의 변수에 대해서 설명을 실시하지 않음	1.7
	모든 변수에 대해서 설명을 실시	11.2

가설을 검증하기 위한 주요 분석 결과에서 통제변수에 대한 효과크기(effect size) 및 유의수준(significance)을 확인하였다. 47.5% 이상의 연구에서 제시되고 있었지만 분석결과에 대한 논의는 이루어지지 않고 있었다. 통제변수를 사용하여 분석한 논문 중에서 41.6%의 논문이 통제변수와 종속변수의 관계에 대한 아무런 논의 없이 지나치고 있었고, 모든 변수에 대해서 설명을 한 연구는 20편(21%) 이었다. 통제변수와 타 변수와의 관계에 대해 구체적으로 효과크기 및 유의수준을 언급하는 것은 미래의 연구자가 메타분석을 할 경우 혹은 선행연구를 검토할 경우 변수 간의 관계에 대한 구체적이고 정확한 정보를 전달함으로써 연구의 신뢰성을 높일 수 있을 것이다.

#### (4) 포함된 통제변수의 적절성 여부

포함된 통제변수가 과연 적절한지 아니면 불필요한 것인지를 확인하기 위해 추가적인 분석이 이루어졌다. 독립변수 및 종속변수와 유의미한 상관관계를 가지고 있다면 이는 통제변수로서 적절하다고 볼 수 있다. 이러한 통제변수들은 앞에서 논의된 것처럼 허위 변수, 편의, 그리고 왜곡변수의 기능을 가지고 있을 수 있다. 또한 통제변수가 억제변수로서의 역할 가능성을 고려하여 비록 종속변수와는 아무런 상관없이 독립변수와 유의미한 상관관계가 있을 경우(Kerlinger & Lee, 2000)도 통제되어야 한다. 반대로 독립변수와는 관련성이 없지만 종속변수와는 관련성이 있을 경우 독립외생변수로서 잡음의 역할을 하기 때문에 이러한 변수를 통제변수로 도입하면 보다 정확한 독립변수의 설명력을 측정할 수가 있다. 정리하면, 통제변수가 독립변수 및 종속변수와 상관관계가 나타나지 않고 포함하여야 할 아무런 이론적 근거가 없다면 부적절한 사용이 된다. 이런 사례를 확인하기 위해 이론적인 근거의 제시가 없으며 독립변수 및 종속변수와 아무런 관련성이 없는 상황에서 통제변수로 사용된 사례들을 분석하였다.

종속변수 및 독립변수와 아무런 관련성도 없고 이론적인 근거도 부족한 통제변수의 사용은 10편의 논문에서 발견되었는데, 이는 통제변수를 포함한 연구 96편의 10%에 해당하는 논문이다. 전술한 바와 같이, 무분별하게 통제변수로 포함하는 것은 II종 오류를 초래하므로 통제변수의 사용에 있어 논리적인 접근이 필요하다.

#### 3) 추가 분석 결과

통제변수의 활용에 영향을 미치는 원인을 탐색하기 위해서 저널의 종류, 페이지 크기, 데이터 특성, 사용된 변수의 개수, 분석방법의 종류 등과 사용된 통제변수 간의 상관관계를 확인하였다. 상관관계 분석 결과는 <표 6>과 같다.

상관관계표에서 포함되는 통제변수(CCV)의 개수는 저널의 종류 및 페이지 수와는 관련성이 없는 것으로 나타났다. 이를 고려해 볼 때 통제변수의 활용은 편집 면의 부족이나 특정 저널의 편집 방침에 의해서 영향을 받지 않는다는 점을 확인할 수 있다. 따라서 통제변수의 포함여부 및 적극적인 활용은 연구자의 재량에 따라서 달라짐을 확인할 수 있다. 반면에 데이터의 종류에서 현장 설문지 데이터와는 부정적 관계를 보였고 반면에 문헌 데이터일수록 통제변수의 수는 증가하였다. 이는 현장 설문지에 비해서 문헌을 통해 자료를 획득하는 것이 쉬울 수 있기 때문이거나 혹은 연구의 주제에서 차이가 날 수

〈표 6〉 상관관계 분석 결과

	평균	표준 편차	CCV수	저널종류	데이터	주요변수 수	페이지 수
CCV 수	2.32	2.96	1				
저널종류 <sup>a</sup>	68.7%	·	.094	1			
데이터 <sup>b</sup>	14%	·	.333**	.202**	1		
주요변수 수	7.77	3.71	-.173*	-.009	-.075	1	
페이지 수	30.78	8.51	.018	.527**	.094	.184*	1

주) a: 경영학연구 0, 인사조직연구 1 b: 설문자료 0, 문헌자료 1  
N = 179, \*. p < 0.05 ; \*\*. p < 0.01

있다. 상관분석에서 주요변수의 수(독립변수 및 종속변수)가 증가할수록 통제변수의 수가 줄어들음을 확인할 수 있었다.

위의 상관관계 분석 결과를 고려하여 데이터 특성에 따른 사용된 통제변수의 평균 차이를 분산분석한 결과, 유의미한 차이가 있음이 확인되었다. 문헌 데이터를 사용한 연구에서 보다 많은 통제변수를 도입하고 있었다.

〈표 7〉 데이터 특성별 CCV 수 분산분석 결과

	사례수	평균	표준편차	분산분석 결과
설문 데이터	154	1.929	2.4283	df:1
문헌 데이터	25	4.760	4.4933	F: 22.008
계	179	2.324	2.9597	p<.001

### 3. 연구 결과의 요약 및 이슈 논의

#### 1) 문제점 요약 및 제언

통제변수의 활용에 있어서 모범 사례와 문제점을 분석한 결과, Becker(2005)의 연구에서 제기된 문제들이 국내 연구에서도 동일하게 나타남을 확인할 수 있었다. 양적 분석 결과를 요약하고 연구에서 고려해야할 통제변수의 활용에 관한 제언과 논의의 필요성이 있는 이슈를 도출하였다.

첫째, 많은 연구들(46.4%)에서 통제변수를 포함하지 않았다는 점이다. 이는 외생변수

를 통제하지 않음으로 하여 연구 결과를 근거로 인과성을 추론하는 것에 대한 타당성의 문제를 야기 시킨다. 특히, 회귀분석을 제외한 구조방정식의 경우 1편을 제외한 모든 연구에서 통제변수를 사용하지 않았기 때문에 논의의 필요성이 제기된다.

제언 1: 연구에 통제변수를 활용하여 인과성을 제고하라. 특히 구조방정식모형에서 통제 변수의 활용을 강화하라.

둘째, 통제변수를 선정하고 포함하는 과정에 있어서 많은 연구(42.4%)에서 연구자의 임의적인 선택이 이루어지고 있다는 점이다. 독립변수 및 종속변수 간의 실제 관계를 왜곡할 수 있는 변수들은 연구의 엄밀성을 위해서 통제되어야 한다. 그러나 이러한 통제변수의 선택은 이론적인 배경이라든지 실증 근거가 있어야 적절하며 만약 임의적인 선택이 발생한다면 위의 분석 결과와 같이 아무런 관련성이 없는 통제변수가 선택되어지는 사례가 발생하게 된다. 이는 오히려 통제변수의 도입으로 인해서 독립변수와 종속변수와의 관계가 왜곡되어 나타나게 되는 문제점을 내포하게 된다. 이론적 기반이 없는 통제변수의 도입은 과대통제(overcontrol)의 문제가 발생하게 된다(Bandura, 1997; Breaugh, 2006; Gordon, 1968; Pedhazur, 1982).

제언 2: 연구에 통제변수를 포함한다면 '왜 그 변수를 선택하였는지' 이론적인 배경이나 근거를 인용하거나 구체적으로 제시하라.

제언 3: 통제변수를 선택하는 데 있어서 주의하라. 종속변수와의 관계를 뒷받침할 아무런 이론적 배경이 없거나 실제적으로 상관관계 분석 결과 종속변수 및 독립변수와 관련성이 없다면 통제변수로서 선택하는 것은 부적합하다. 오히려 통제변수에 포함함으로써 설명력만 감소시킨다.

통제변수의 선정 경향에서 발견된 또 하나의 문제점은 Breaugh(2006) 및 Rosenberg(1968)의 지적대로 상투적인 사용이 이루어진다는 점이다. Rosenberg는 사회과학연구에서 '나이', '성별', '계층' 등은 통제변수로서 포함시켜야 할 아무런 근거 없이 사용되고 있음을 지적하였는데, 이러한 경향은 조직연구에서도 유사한 현상이 벌어지고 있었다. 많은 연구에서 통제변수로 인구통계학적 변수와 조직특성 변수가 중점적으로 사용되었

다. 상투적이고 부적절한 통제변수의 쓰임에 대한 문제는 다음 절에서 추가적으로 논의하였다.

셋째, 통제변수를 측정하기 위한 방법에 대해서 구체적인 언급을 누락한다는 점이다. 이는 독자에 따라서 측정방법을 다르게 해석할 수 있는 문제점이 발생한다.

제언 4: 통제변수를 측정할 수 있는 구체적인 방법(MCV)을 제시하고 왜 그러한 MCV를 선택했는지 이론 및 실증 근거를 제시하라.

넷째, 기술적 통계 및 분석 결과의 보고는 비교적 정착되어 있지만 일부 연구에서 누락되어 있는 경우가 많았다. 누락된 경우, 연구에서 측정되어진 자료가 메타분석과 같은 2차 연구에서 활용될 수 없다는 단점이 있다.

제언 5: 통제변수에 대한 기술통계를 제시하라. 평균 및 표준편차 그리고 상관관계의 계수와 유의수준을 모두 제시하라(범주형 변수일 경우 비율을 제시하라). 다항목 메트릭 척도를 사용할 경우, 타당도 및 신뢰도를 제시하라.

다섯째, 통제변수 분석 결과에 대해서 연구자의 논의가 부족하다는 점이다. 전체의 11.2%만이 결과 분석 후 종속변수와의 관련성에 대한 결과와 해석을 덧붙이고 있다.

제언 6: 효과크기 및 유의수준을 기술하고, 분석결과에 통제변수에 대한 결과도 논의에 포함하라.

## 2) 주요 이슈 논의

본 절에서는 양적 분석 결과에서 제기된 주요 이슈에 대한 논의가 이루어졌다.

첫째, 상투적인 통제변수의 활용 문제다. 분석대상 논문 중에서 통제변수를 사용한 논문은 대부분 인구통계학적 변수들(성별, 나이, 학력)과 조직특성 변수들(직무연한, 근속년수, 규모, 산업의 종류 등)이 중점적으로 사용되었다. 개인의 태도나 행동과 관련된 연구에서 인구통계학적 변수들이 합리적인 근거 없이 사용되고 있는 점이 다른 연구에서도 지적되고 있다(Becker, 2004; Breugh, 2006). 특히, 아무런 이론적 배경이나

실증 근거 제시 없이 통제변수가 사용된다는 점에서 상투적 사용에 대한 의심을 배제할 수 없다.

통제변수로서 일반적으로 성별이나 나이 등의 변수가 사용되는 이유는 이들 변수를 대리변수(proxy variable)로 취급하는 경향이 있기 때문이다(Breaugh, 2006). 예를 들어, 직무만족에 대한 리더십의 영향을 검증하기 위해서 성별을 통제변수로 사용한다면 이는 여성과 남성의 경험 차이에서 오는 속성의 영향을 제거하기 위한 조치이다. 연구자가 여성은 대체적으로 남성에 비해 직장-가정 갈등이 심하다고 판단하고 직장-가정 갈등의 영향을 제거하기 위해서 이와 공변한다고 생각하는 성별을 통제 변수로 사용한 경우를 예로 들 수 있다. 이와 같이, 공변한다는 가정 때문에 인구통계학적 변수들이 대리변수로 사용되며, 실제 통제하고자 하는 변수에 비해 측정이 용이하다는 이점에서 더욱 쉽게 사용된다(Breaugh, 2006). 이러한 대리변수의 사용의 문제점은 실제 통제하고자 하는 장애변수 효과를 대리변수가 어느 정도 대체하느냐를 판단하기 어렵다는 점이다(Breaugh, 2006).

통제변수의 무분별한 사용은 오히려 모형의 적합성을 저하시키는 원인이 될 수 있다. 일례로, Cropanzano, Rupp, & Byrne(2003)의 연구에서 저자들은 직무성과평가(job performance ratings) 변수에 정서 소진(emotional exhaustion) 변수로 회귀분석을 했을 때, 회귀모형이 적합하지 않게 나온 것은 종속변수와 관련성이 약한 통제변수를 포함했기 때문이라고 설명하고 있다. 그러나 저자들은 통제변수의 사용에 대한 적절한 근거를 제시하지 않고, 상관관계에서도 아무런 관련성이 나타나지 않는 변수를 사용함으로써 발생한 결과임을 간과하고 있다. 그들의 모형에서 일부 통제변수가 잘못 사용되어졌고 결과적으로 전반적인 회귀모형의 적합성까지 떨어뜨리는 문제를 일으킨 것이다.

따라서 통제변수의 선택에서는 논리적인 절차가 필요하다. Rosenberg(1968)의 주장대로, 첫째, 이론적인 기반이 뒷받침 되어야 하며 논리적인 근거가 있어야 한다(Breaugh, 2006; Gordon, 1968). 둘째, 독립변수와 종속변수와의 상관관계를 확인하여 관련성 있는 변수를 포함하여야 한다.

두 번째로 제기되는 이슈는 구조방정식의 경우 거의 대부분의 연구가 통제변수를 사용하지 않았다는 점이다. 송보화(2006)의 연구에서 유일하게 구조방정식모형에 통제변수를 포함하여 분석을 하였다. 이런 현상은 구조방정식이 인과관계를 증명할 수 없다는 주장들(Bollen, 1989; Brannick, 1995; Breckler, 1990; Holland, 1988; Hoyle &

Smith, 1994; McDonald & Ho, 2002; Mittal, 1993; Sobel, 1996; Tomarken & Waller, 2003; Williams, 1995)을 고려해 볼 때 심각한 문제임을 알 수 있다. 많은 연구자가 구조방정식을 이용한 분석방법은 인과관계를 증명할 수 있다는 생각을 가지고 있었는데 이는 구조방정식모형을 초기에 '인과모형(causal modelling)'으로 칭한 것(예: Bentler, 1980)에서 비롯되었다(Brannick, 1995; Breckler, 1990; Williams, 1995). 이를 볼 때, 구조방정식을 이용한 연구에서 통제변수의 사용이 극히 저조한 원인을 구조방정식은 인과관계를 증명할 수 있다고 잘못 이해한 점에서 찾을 수 있다.

따라서 구조방정식모형 분석 시 통제변수 적용에 대한 논의가 필요하다. 구조방정식에서 통제변수의 사용에 대한 고찰은 다음 장에서 구체적으로 다루겠다.

#### IV. 구조방정식모형에서의 통제변수 적용

국내·외 연구에서 구조방정식모형을 사용한 연구가 점점 증가하고 있다. 구조방정식모형을 사용한 연구에서 가장 두드러진 특징은 앞에서 언급한 것처럼 통제변수를 잘 사용하지 않는다는 점이다. 그러면 구조방정식은 통제변수를 사용하지 않더라도 다른 외생변수를 모두 통제할 수 있고 그 결과에 대해 인과관계를 확신할 수 있는 방법인가? 그렇지 않다. 많은 학자들이 단순히 구조방정식모형을 사용함으로써 인과관계를 증명하는 것은 아니라고 주장한다. 이러한 이유로 근래에 들어 구조방정식을 사용한 논문에서 통제변수를 적용하여 분석한 사례가 증가하고 있다(예: Bakker, Demerouti, & Dollard, 2008; Baum & Locke, 2004; Deery, Iverson, & Walsh, 2006; Eddleston, Veiga, & Powell, 2006; Gong & Fan, 2006; Hammer et al., 2005; Hirschfeld et al., 2006; Lim & Sng, 2006; Rafferty & Griffin, 2006; Scott, Cloquitt, & Zapata-Phelan, 2007; Sprigg et al., 2007; Zhao, Seibert, & Hills, 2005).

전술한 바와 같이 인과관계는 무작위할당 실험연구에서만 추론 가능하므로 비실험연구에서 구조방정식모형의 분석결과로 인과관계를 설명하는 것은 불가능하다. 따라서 구조방정식에서도 독립변수 외에 종속변수에 영향을 주는 다른 외생변수의 영향을 통제하기 위해서는 통계적인 통제방법, 즉 통제변수를 활용하여야 한다(Bollen, 1989; Cliff, 1983; Hoyle & Smith, 1994).



## 1. 구조방정식모형에서 통제변수 사용이 저조한 원인

구조방정식모형에서 통제변수를 사용하지 않는 이유를 생각해 보고 바람직한 해결방안을 도출해 보자.

통제변수를 사용하지 않는 첫 번째 이유는 역시 구조방정식모형은 인과성을 증명할 수 있다는 오해에서 비롯되었다고 할 수 있다. 연구자의 이러한 오해는 구조방정식 모형은 초기에는 '인과모형'으로 부른 것에서 비롯되었고 많은 학자들이 '인과효과'를 검증할 수 있음을 전제하면서 강화된 것이다(Breckler, 1990; Williams, 1995).

두 번째 이유는 구조방정식을 이용한 분석과정이 대체적인 모형(alternative model)을 제거하면서 이루어진다는 점이다. 인과관계를 단정하는 것은 다른 대체적인 설명들이 배제되었을 때 가능한데 구조방정식에서는 다른 경쟁모형과 비교 과정을 거쳐 데이터에 가장 적합한 모형을 탐색하기 때문에 가장 사실에 가까운 모형을 구할 수가 있다. 따라서 구조방정식을 이용한 분석은 이러한 모형 선택과정을 거치기 때문에 보다 정확한 사실적 관계에 다가가는 이점이 있지만 이러한 과정을 거친다고 해서 인과관계를 주장할 수는 없다(Williams, 1995).

셋째, 통제변수로서 사용되는 대부분의 변수가 인구통계학적 변수 혹은 조직특성 등의 변수로서 다항목적도가 아니기 때문이다. 전술한 바와 같이, 통제변수의 상투적인 사용(Becker, 2004; Breugh, 2006; Rosenberg, 1968)은 구조방정식 분석방법이 다항목 측정치(multiple indicator)로 이루어진 잠재변수를 사용한다는 점에서 서로 상충되고 있다. 따라서 이제까지 주로 사용되어왔던 통제변수가 단일항목 측정치(single indicator)라는 점 때문에 구조방정식 분석에서 배제되는 경향이 있어 왔음을 짐작할 수 있다.

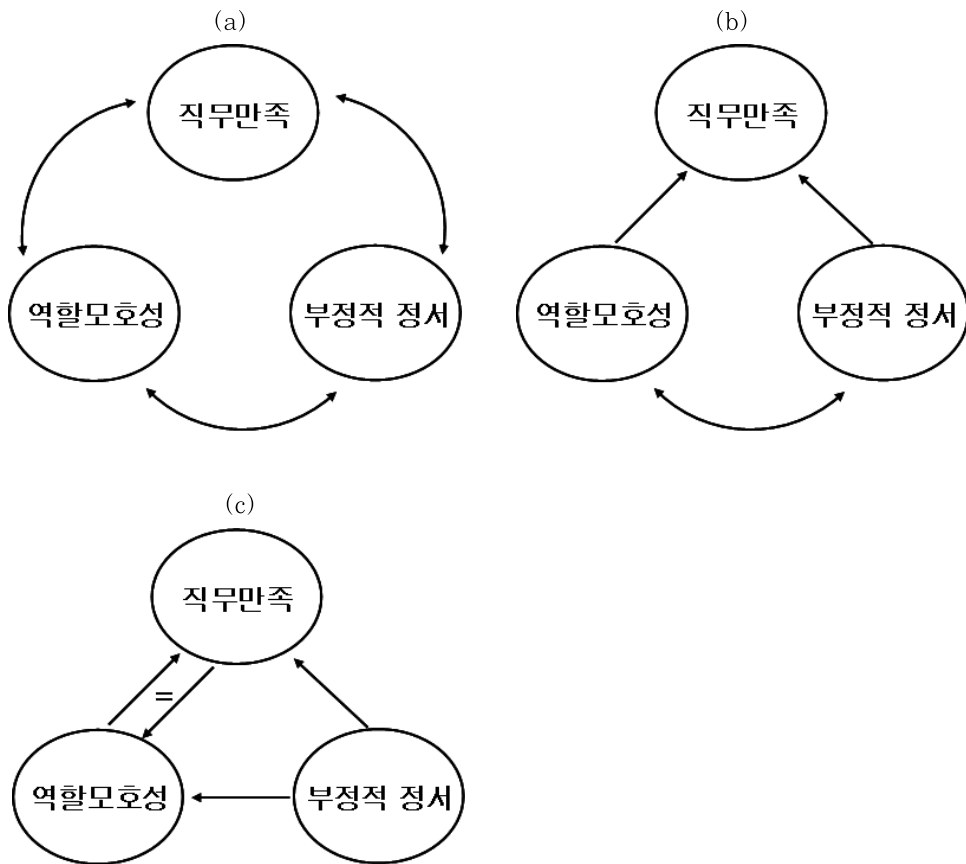
## 2. 구조방정식모형에서 통제변수 사용의 필요성

### 1) 동치모형의 존재

구조방정식모형의 분석 결과를 가지고 인과적 추론을 할 수 없다는 점은 동치모형(equivalent model)의 존재를 통해서 이해할 수 있다. 동치모형은 통계적으로 동일한 적합도를 가진 다른 대안적인 모형을 의미하는데,<sup>9)</sup> 동치모형이 존재한다는 것은 연구자

9) 적합도뿐만 아니라 변수 간 관계의 모수추정치도 완벽히 일치하는 사례도 존재함(배병렬, 2006;

의 가설 외에 다른 가설도 충분히 해당 데이터로 입증 가능함을 뜻한다(MacCallum et al., 1993). <그림 5>에서 3개의 모형은 3개의 잠재변수 간에 상관관계 및 인과관계를 가리키는 화살표로 구성되어 있는데 그 방향성은 차이가 있다. 그림 3a는 3개의 잠재변수가 단순 상관관계만을 가지고 있고 인과관계는 없다. 그림 3b는 역할모호성(role ambiguity)과 부정적 정서성향(negative affectivity)은 관련성이 있고 각각 직무만족(job satisfaction)에 영향을 주는 것으로 되어 있다. 그림 3c는 부정적 정서성향이 직무만족 및 역할 모호성에 영향을 미치고 직무만족과 역할모호성은 상호 영향을 미치는 모형이다.



<그림 5> 동치모형 사례

〈그림 5〉의 3개의 모형에서 모형의 적합도는 동일하게 나타난다. 위의 모형들은 구조 모형(structural model) 부분만을 다루고 있지만 측정모형(measurement model)까지 고려한다면 변수의 수가 증가함에 따라 동치모형은 급격하게 늘어난다(Brannick, 1995; Cliff, 1983; Stelzl, 1986). 따라서 구조방정식모형 분석을 통해 변수 간의 관련성을 통계적으로 입증하였다고 해서 인과관계를 주장할 수는 없으며 반드시 논리나 이론적인 근거를 기반으로 해서 주장하여야 한다(배병렬, 2006; Hoyle & Smith, 1994; Williams, 1995). 한편, 동치모형 중에서 이론적으로 납득 가능한 모형이 존재할 수 있는데 이런 경우 다른 대안적인 가설도 충분히 타당성을 가지고 있기 때문에 연구자의 주장을 정당화할 수 없다. 동치모형뿐만 아니라 충분히 수용할 수 있는 적합도를 가진 대안모형(alternative nonequivalent models)도 다수 존재하기 때문에 이러한 문제는 더욱 심각해진다(Tomarken & Waller, 2003).

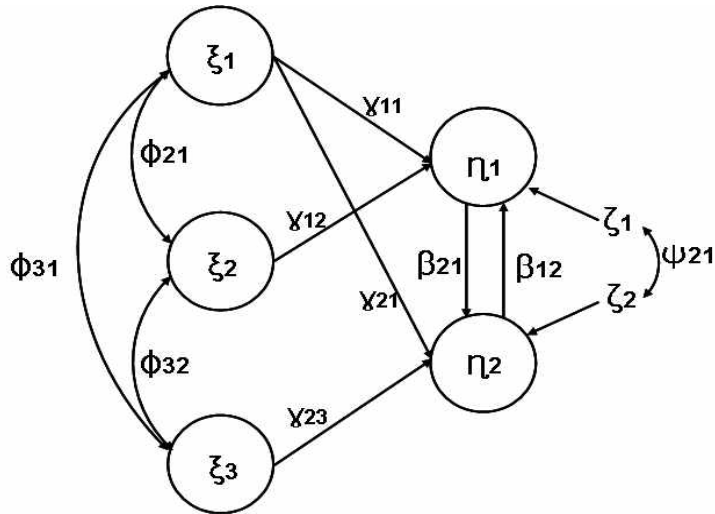
MacCallum et al.(1993)은 동치모형의 존재를 해결할 수 있는 방법으로 인구통계학적 변수의 사용을 강조하고 있다. 나이나 성별은 다른 변수에 영향을 주지만 영향을 받지 않는다. 따라서 어떤 동치모형에서 특정의 변수들이 나이나 성별에 영향을 주는 것으로 묘사 된다면 그 동치모형은 이론적으로 부적절하다. 따라서 인구통계학적 변수를 모형에 포함한다면 많은 동치모형들이 논리적으로 무의미하게 될 것이다. 만약, 해당 연구에서 인구통계학적 변수가 통제변수로서 타당성이 있다면 이러한 변수의 도입은 동치모형들을 억제하는 효과를 가져 올 것이다.

구조방정식모형 분석 결과가 인과성을 증명하는 것이 아니기 때문에 다른 외생변수의 통제가 불가피하고 또한 인구통계학적 변수의 도입이 동치모형을 억제하는 효과가 있기 때문에 구조방정식에서도 통제변수의 활용이 필요하다.

## 2) 누락변수(omitted variable)의 문제

구조방정식모형에서도 인과성을 입증하기 위해서는 2장에서 전술한 인과조건이 충족되어야 한다. 독립변수의 영향 외에 다른 대체적인 영향을 배제하는 분리(isolation), 독립변수와 종속변수의 유의미한 관련성(association) 그리고 원인이 결과보다 선행되어야 하는 인과의 방향성(direction of causation) 요건이 필요하다(Bollen, 1989). 따라서 분석하는 구조방정식모형의 자료가 실험 자료가 아닌 비실험연구 자료라면 다른 외생변수의 효과를 제거하기 위해 즉, 분리 요건을 만족시키기 위해 다른 변수의 영향을

통계적으로 통제하여야 한다.



주) 도식을 간략하게 하기 위해서 측정모형은 생략함

<그림 6> 구조방정식 구조모형

위 <그림 6>에서 교란(구조오차 혹은 잔차, disturbance:  $\zeta$ )은 독립변수에 의해 설명이 되지 않은 내생잠재변수(endogenous latent variable:  $\eta$ )의 변량(variation)을 의미하는 것으로 연구 모형에 생략된 변수(omitted variable)의 복합체라고 할 수 있다. 완전한 분리는 외생잠재변수(exogenous latent variab:  $\xi$ )와 교란 간의 상관관계가 없을 때 달성되어진다(Bollen, 1989). 즉, 'COV( $\xi, \zeta$ ) = 0'일 경우, 외생변수가 통제되어 분리되었다고 할 수 있는데, 이를 위해서는 외생잠재변수와 상관이 있는 생략된 변수를 모형 내의 독립변수로 많이 포함시켜야 한다.

구조방정식모형에서도 회귀분석과 마찬가지로 종속변수에 영향을 주는 중요한 변수들의 누락은 인과관계를 규명하는데 심각한 위험이 된다(Bollen, 1989; Cliff, 1983; Tomarken & Waller, 2003). 누락 변수로 야기되는 문제점은 먼저, 인과 모수(causal parameters) 추정치에서 편이가 발생하며, 둘째, 모수 추정치의 표준오차(standard error)가 부정확하게 된다. 또한 상술오류(misspecification)를 범하게 되며 많은 경우에 적합도도 떨어지게 된다(Tomarken & Waller, 2003). 따라서 이러한 상술오류를 극복하기 위해서는 모형과 관련된 외생변수들을 통제변수로 도입하여 포함하여야 한다.

### 3. 구조방정식모형에서 통제변수 적용 방법

구조방정식모형에서 외생변수의 효과를 제거하기 위해서는 통제변수의 도입이 필요하다는 점은 공감하고 있지만 그 구체적인 방법에 대한 관심은 소홀한 것 같다.

다수의 연구자들이 아무런 논리적 설명 없이 통제변수와 내생잠재변수 간에는 자유 경로를 부여하고 다른 외생잠재변수 간에는 경로를 고정(fixed=0)하는 경향이 있다(Becker, 2005). Becker(2005)는 이러한 문제점을 해결하는 방법으로 회귀분석과 동일하게 모든 예측변수와 상관관계를 설정하는 것(모든 외생잠재변수 간의 자유 경로를 허용)을 주장하고 있다. 하지만 Becker의 이러한 주장은 구조방정식의 이점을 활용하지 못할 뿐 아니라 식별(identification)의 문제가 발생한다. 이론적으로 관련이 없는 관계를 설정함으로써 설정 오차(specification error)가 발생할 수 있으며 추정해야 할 모수가 증가함에 따라 자유도가 감소하여 간명도가 낮아지며 과대식별(over-identified)의 요건을 충족시키지 못할 수도 있다.

이러한 문제는 저자들이 앞에서 제기한 외생변수의 구분을 가지고 설명을 할 수 있을 것이다. 저자들은 외생변수를 독립변수와 독립적일 경우 '독립적 외생변수'로 독립적이지 않을 경우 '비독립적 외생변수'라고 분류하였다. 이 두 구분을 기준으로 의문에 대한 답을 구할 수 있을 것이다.

만약, 연구자가 구조방정식모형에 포함하여서 통제하고자 하는 외생변수가 독립변수 즉 외생잠재변수와 비독립적일 경우 Becker의 주장처럼 외생잠재변수 간 자유모수를 부여하는 것이 바람직하다.(<그림 7a> 참조)

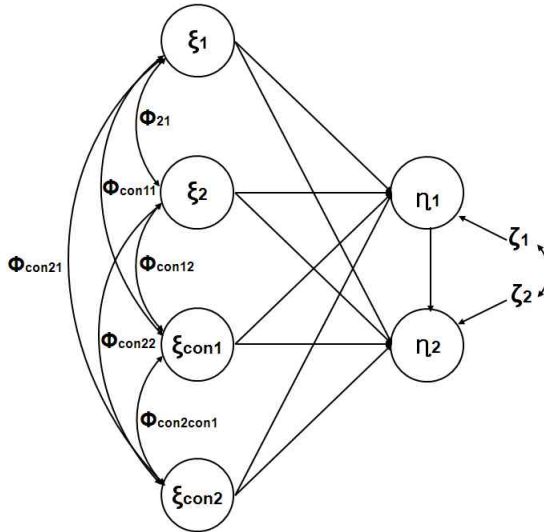
도입된 통제변수가 외생잠재변수와 상관관계가 없을 경우, 즉 독립적 외생변수일 경우에는 해당 통제변수와 외생잠재변수 간의 경로는 고정하여 분석하여야 할 것이다.(<그림 7b 참조>)

통제변수가 독립변수와 독립 혹은 비독립일 경우가 있을 수 있다. 이러한 경우는 각 변수 간의 관계 유무에 따라 경로를 부여하여야 할 것이다.(<그림 7c> 참조)

Rosenberg(1968)의 주장대로, 통제변수의 적용은 이론적인 기반과 독립변수 및 종속 변수와의 관련성을 고려하여 경로를 부여하는 것이 타당하다. 통제변수도 독립 및 종속 변수와 마찬가지로 이론을 근거로 해서 경로를 설정하여 모형을 구체화하여야 할 것이다(Cohen et al., 2003). 이를 통해 설정 오차의 발생을 방지할 수 있을 것이고 식별의 문제도 해결할 수 있을 것이다. 특히, 많은 연구자가 통제변수의 경로 설정에 대한 설명

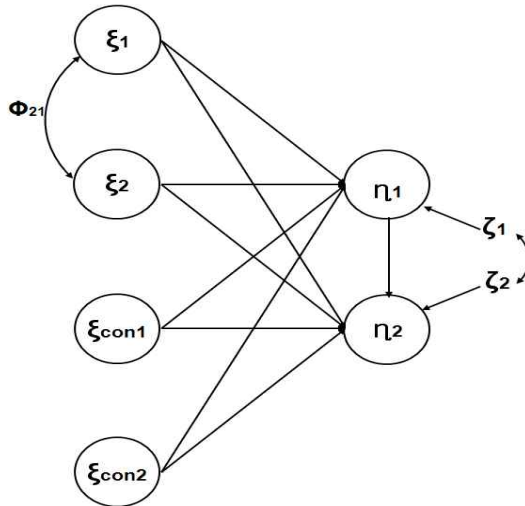
을 생략하는데 향후에는 ‘왜 그러한 통제변수의 경로를 설정했는가?’에 대한 설명이 뒤따라야 할 것이다.

(a)



주) 통제변수( $\xi_{con1}$ ,  $\xi_{con2}$ )는 독립변수( $\xi_1$ ,  $\xi_2$ )에 대해서 각각 비독립적임

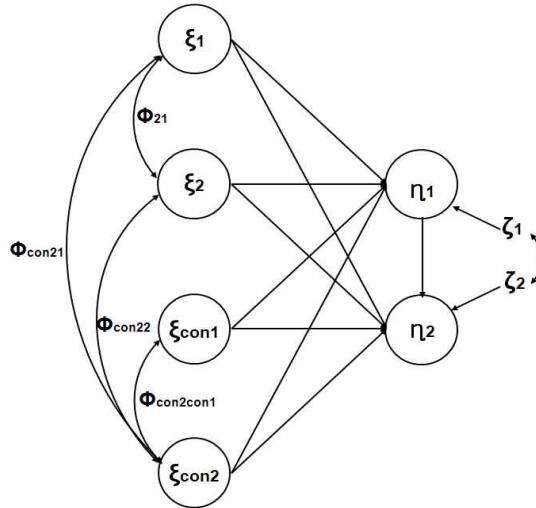
(b)



주) 통제변수( $\xi_{con1}$ ,  $\xi_{con2}$ )는 독립변수( $\xi_1$ ,  $\xi_2$ )에 대해서 각각 독립적임

<그림 7> 통제변수의 경로 설정 방법

(c)



주)  $\text{Corr}(\xi_{con1}, \xi_1) = \text{Corr}(\xi_{con1}, \xi_2) = 0$  경우

〈그림 7〉 통제변수의 경로 설정 방법 (계속)

한편, 통제변수가 단일항목 측정치일 경우에는 잠재변수로 전환하여 모형을 설정하여야 한다. 특히 인구통계학적 변수나 조직 특성변수들은 단일항목 측정치가 대부분이므로 이러한 변수들이 통제변수로 사용될 때 잠재변수로의 전환과정이 필요하다. 구체적인 전환방법은 이미 많은 연구들에서 제시되었고(예: 배병렬, 2006) 이 글의 범위를 벗어나기 때문에 논외로 한다.

## V. 결론

본 연구는 인과성에 대한 논의로부터 인과관계 검증을 충족시키기 위한 통제 수단과 통제변수 개념 및 의의와 필요성 등 전반적인 고찰을 실시하였다. 국내 인사조직 관련 주요 저널의 논문들에 대한 질적 분석은 통제변수의 활용에 대한 구체적인 강점과 문제점을 도출할 수 있었다. 이를 바탕으로 통제변수 적용에 대한 적절한 제언과 주요 이슈를 도출하고 논의가 이루어졌다.

본 연구는 그 동안 간과되어 왔던 통제변수의 중요성을 환기하였다는 점에서 그 첫 번

제 의의를 찾아 볼 수가 있다. 사회과학에서 인과관계는 다변수적 특성(채서일, 2002; Kerlinger & Lee, 2000; Marini & Singer, 1988; Schwab, 2005)을 가지고 있기 때문에 두 현상 간의 정확한 관계를 이해하기 위해서는 엄격한 외생변수의 통제가 중요하다. 특히, 비실험연구가 주를 이루는 조직 연구에서는 외생변수의 통제는 통계적인 방법으로 외생변수를 통제변수에 포함함으로써 달성된다. 따라서 통제변수에 대한 보다 많은 관심이 필요하다.

둘째, 결과변수에 영향을 주는 원인변수를 독립변수와 외생변수로 구분하고 다시 외생변수를 2차원(독립 및 통제가능)으로 구분하였다는 것이다. 이제까지 외생변수에 대한 용어의 구분이 정착되지 않아 많은 연구들에서 용어를 혼용하여 사용하고 있어 독자들의 혼란을 초래하였다. 본 연구에서 외생변수의 속성을 고려 구분함으로써, 연구자는 연구 단계별로 통제하여야 할 외생변수를 구분할 수 있을 것이다. 연구 설계와 데이터 수집의 과정에서는 통계적으로 통제 불가능한 외생변수의 통제에 관심을 기울이고 통계적인 분석 단계에서는 통제가능 변수를 통제를 하되, 독립변수와 독립성을 고려하여 포함하여야 할 것이다.

셋째, 주요 저널에 실린 연구의 분석을 통해서 통제변수 활용의 모범 사례와 문제점을 도출하여 바람직한 통제변수의 활용에 대한 제언을 하였다는 점이다. 특히, 통제변수를 전혀 사용하지 않거나 아무런 근거 없이 선정되는 현상은 시급히 개선되어야 할 문제점이었다. 본 연구에서는 통제변수의 활용에 있어 구체적인 기준과 절차를 제시함으로써, 향후 연구자들이 보다 신뢰할 수 있는 연구결과를 달성하는데 이바지할 것이다.

넷째, 간과되어 왔던 구조방정식에서의 통제변수의 사용을 강조하고 의의와 방법을 구체적으로 살펴보았다는 점이다. 구조방정식 분석기법으로 적용한 국내 연구 대부분이 통제변수를 사용하지 않은 것은 우려할 만한 수준이었다. 본 연구에서 이러한 문제에 대해 지적하고 적절한 방안을 제시함으로써, 향후 구조방정식 분석기법을 사용하는 연구자들이 인과성을 확신하는데 많은 기여를 할 것이다. 특히, 국내외를 막론하고 아직까지 정립되지 않은 '통제변수와 독립변수 간의 고정 혹은 자유 경로의 적절성'에 대한 이슈에 대해서 해결방안을 제시하였다.

결론적으로 본 연구에서 중점을 두어 강조한 것은 인과관계를 보다 확신하기 위해서는 적절한 통제변수를 포함하는 것이 중요하며 통제변수도 독립변수나 종속변수와 마찬가지로 동일하게 취급되어야 한다는 것이다. 통제변수의 포함과 신중한 선택, 측정방법의 구



체화, 기술통계 및 분석 결과의 기술, 결과에 대한 논의 등은 연구자의 연구 결과를 보다 가치 있게 하는데 기여할 것이다.

## 참 고 문 헌

- 김양민. 2006. 국제 인수합병의 주식시장 반응에 대한 실증분석: 내부화이론과 대리인 이론을 중심으로. 「인사·조직연구」, 14(1): 117-142.
- 김해룡. 2003. 기업집단 본부의 통제가 계열사 연구개발투자에 미치는 영향에 관한 연구. 「인사·조직연구」, 11(1): 35-71.
- 배병렬. 2006. 「LISREL 구조방정식 모델: 이해, 활용 및 프로그래밍」, 서울: 청람.
- 송보화. 2006. 인적자원관리시스템과 조직성과의 연계에 대한 실증적 고찰. 「경영학연구」, 35(1): 331-363.
- 양동훈. 2005. 보상제도가 협동행동에 미치는 영향에 관한 연구. 「인사·조직연구」, 13(3): 1-41.
- 이근희. 2005. 「사회과학 연구방법론」(수정판), 서울: 법문사.
- 임준철 & 윤정구. 1998. 분배공정성과 절차공정성이 직무만족과 조직몰입에 미치는 차별적 영향에 관한 연구: 문화적 맥락이 조직성원의 행위성향에 미치는 영향을 중심으로. 「경영학연구」, 27(1): 93-111.
- 채서일. 2002. 「사회과학 조사방법론」(2판), 서울: 학현사.
- Bakker, A. B., Demerouti, E., & Dollard, M. F. 2008. How job demands affect partners' experience of exhaustion: Integrating work-family conflict and crossover theory. *Journal of Applied Psychology*, 93: 901-911.
- Bandura, A. 1997. *Self-efficacy: The exercise of control*. New York: W. H. Freeman & Company.
- Baum, J. R., & Locke, E. A. 2004. The relationship of entrepreneurial traits, skill, and motivation to subsequent venture growth. *Journal of Applied Psychology*, 89: 587-598.
- Becker, T. E. 2005. Potential problems in the statistical control of variables

- in organizational research: A qualitative analysis with recommendations. *Organizational Research Methods*, 8(3): 274-289.
- Bentler, P. M. 1980. Multivariate analysis with latent variables: Causal modeling. *Annual Review of Psychology*, 31: 419-456.
- Blalock, H. M., Jr. 1961a. *Causal inferences in nonexperimental research*. Chapel Hill, NC: The University of North Carolina Press.
- Blalock, H. M., Jr. 1961b. Correlation and causality: The multivariate case. *Social Forces*, 39(3): 246-251.
- Blalock, H. M., Jr. 1962. Spuriousness versus intervening variables: The problem of temporal sequences. *Social Forces*, 40(4): 330-336.
- Boring, E. G. 1954. The nature and history of experimental control. *The American Journal of Psychology*, 67(4): 573-589.
- Bollen, K. A. 1989. *Structural equations with latent variables*. New York: John Wiley & Sons.
- Brannick, M. T. 1995. Critical comments on applying covariance structure modeling. *Journal of Organizational Behavior*, 16(3): 201-213.
- Breaugh, J. A. 2006. Rethinking the control of nuisance variables in theory testing. *Journal of Business and Psychology*, 20(3): 429-443.
- Breaugh, J. A. & Arnold, J. 2007. Controlling nuisance variables by using a matched-groups design. *Organizational Research Methods*, 10(3): 523-541.
- Breckler, S. J. 1990. Applications of covariance structure modeling in psychology: Cause for concern? *Psychological Bulletin*, 107, 260-273.
- Campbell, D. T., & Stanley, J. C. 1966. *Experimental and quasi-experimental designs for research*. Chicago: Rand McNally.
- Cliff, N. 1983. Some cautions concerning the application of causal modeling methods. *Multivariate Behavioral Research*, 18: 115-126.
- Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. 2003. *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd

- ed.). Mahwah, NJ: Lawrence Erlbaum Associates.
- Cook, D. T., & Campbell, D. T. 1979. *Quasi-experimentation: Design & analysis issues for field settings*. Boston: Houghton Mifflin Company.
- Cropanzano, R., Rupp, D. E., & Byrne, Z. S. 2003. The relationship of emotional exhaustion to work attitudes, job performance, and organizational citizenship behaviors. *Journal of Applied Psychology*, 88(1): 160-169.
- Deery, S. J., Iverson, R. D., Walsh, J. T. 2006. Toward a better understanding of psychological contract breach: A study of customer service employees. *Journal of Applied Psychology*, 91:166-175.
- Eddleston, K. A., Veiga, J. F., & Powell, G. N. 2006. Explaining sex difference in managerial career satisfier preferences: The role of gender self-schema. *Journal of Applied Psychology*, 91: 437-445.
- Frank, K. A. 2000. Impact of confounding variable on a regression coefficient. *Sociological Methods Research*, 29(2): 147-194.
- Gong, Y., & Fan, J. 2006. Longitudinal examination of the role of goal orientation in cross-cultural adjustment. *Journal of Applied Psychology*, 91: 176-184.
- Gordon, R. A. 1968. Issues in multiple regression. *The American Journal of Sociology*, 73(5): 592-616.
- Hammer, L. B., Neal, M. B., Newsom, J. T., Brockwood, K. J., & Colton, C. L. 2005. A longitudinal study of the effects of dual-earner couples' utilization of family-friendly workplace supports on work and family outcomes. *Journal of Applied Psychology*, 90: 799-810.
- Hirschfeld, R. R., Jordan, M. H., Feild, H. S., Giles, W. F., & Armenakis, A. A. 2006. Becoming team players: Team members' mastery of teamwork knowledge as a predictor of team task proficiency and observed teamwork effectiveness. *Journal of Applied Psychology*, 91: 467-474.
- Holland, P. W. 1986. Statistics and causal inference. *Journal of the American Statistical Association*, 81: 945-960.

- Holland, P. W. 1988. Causal inference, path analysis, and recursive structural equations models. *Sociological Methodology*, 18: 449-484.
- Hoyle, R. H., & Smith, G. T. 1994. Formulating clinical research hypotheses as structural equation models: A conceptual overview. *Journal of Consulting and Clinical Psychology*, 62: 429-440.
- Kerlinger, F. N., & Lee, H. B. 2000. *Foundations of Behavioral Research* (4th ed.). Orlando, FL: Harcourt College Publishers.
- Kish, L. 1959. Some statistical problems in research design. *American Sociological Review*, 24(3): 328-338.
- Lim, V. K. G., & Sng, Q. S. 2006. Does parental job insecurity matter? Money anxiety, money motives, and work motivation. *Journal of Applied Psychology*, 91: 1078-1087.
- Marini, M. M. & Singer, B. 1988. Causality in the social science. *Sociological Methodology*, 18: 347-409.
- Martin, M. W. & Sell, J. 1979. The role of the experiment in the social sciences. *The Sociological Quarterly*, 20: 581-590.
- Maassen, G. H. & Bakker, A. B. 2001. Suppressor variables in path models: Definitions and interpretations. *Sociological Methods & Research*, 30(2): 241-270.
- MacCallum, R. C., Wegener, D. T., Uchino, B. N., & Fabrigar, L. R. 1993. The problem of equivalent models in applications of covariance structure analysis. *Psychological Bulletin*, 114: 185-199.
- McDonald, R. P., & Ho, M-H. R. 2002. Principles and practice in reporting structural equation analysis. *Psychological Methods*, 7: 64-82.
- McNemar, Q. 1945. The mode of operation of suppressant variable. *American Journal of Psychology*, 58: 554-555.
- Mittal, B. 1993. Testing consumer behavior theories: LISREL is not a panacea. *Advances in Consumer Research*, 20: 647-653.
- Pan, W., & Frank, K. A. 2003. A probability index of the robustness of a

- causal inference. *Journal of Educational and Behavioral Statistics*, 28(4): 315-337.
- Pedhazur, E. J. 1982. *Multiple regression in behavioral research* (2nd ed.). New York: Holt, Rinehart and Winston.
- Rafferty, A. E., & Griffin, M. A. 2006. Perceptions of organizational change: A stress and coping perspective. *Journal of Applied Psychology*, 91: 1154-1162.
- Rosenberg, M. 1968. *The Logic of Survey Analysis*. New York: Basic Books.
- Rubin, D. B. 1974. Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, 66 (5): 688-701.
- Schwab, D. P. 2005. *Research methods for organizational studies* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum.
- Scott, B. A., Cloquitt, J. A., & Zapata-Phelan, C. P. 2007. Justice as a dependent variable: Subordinate charisma as a predictor of interpersonal and informational justice perceptions. *Journal of Applied Psychology*, 92: 1597-1609.
- Simon, H. A. 1954. Spurious correlation: A causal interpretation. *Journal of the American Statistical Association*, 49: 467-479.
- Singer, B., & Marini, M. M. 1987. Advancing social research: An essay based on stanley lieberman's making in count. *Sociological Methodology*, 17: 373-391.
- Smith, H. L. 1990. Specification problems in experimental and nonexperimental social research. *Sociological Methodology*, 20: 59-91.
- Sobel, M. E. 1996. An introduction to causal inference. *Sociological Methods & Research*, 24(3): 353-379.
- Solomon, R. L. 1949. An extension of control group design. *Psychological Bulletin*, 46(2): 137-150.
- Sprigg, C. A., Stride, C. B., Smith, P. R., Wall, T. D., & Holman, D. J.

2007. Work characteristics, musculoskeletal disorders, and the mediating role of psychological strain: A study of call center employees. *Journal of Applied Psychology*, 92: 1456-1466.
- Stelzl, I. 1986. Changing a causal hypothesis without changing the fit: Some rules for generating equivalent path models. *Multivariate Behavioral Research*, 21: 309-331.
- Tomarken, A. J., & Waller, N. G. 2003. Potential problems with “well fitting” models. *Journal of Abnormal Psychology*, 112: 578-598.
- Williams, L. J. 1995. Covariance structure modeling in organizational research: Problems with method versus applications of the method. *Journal of Organizational Behavior*, 16(3): 225-233.
- Winship, C., & Morgan, S. L. 1999. The estimation of causal effects from observed data. *Annual Review of Sociology*, 25: 659-706.
- Zhao, H., Seibert, S. E., & Hills, G. E. 2005. The mediating role of self-efficacy in the development of entrepreneurial intentions. *Journal of Applied Psychology*, 90: 1265-1272.

## 〈부록 1〉 질적 분석 결과 코딩 방법

저널명	1: 경영학연구 2: 인사조직연구	연도	1997-2006
권	해당 권수	호	해당 호수
시작페이지	논문의 첫 페이지	종료페이지	논문의 종료페이지
데이터 특성	1: 현장            2: archival		
분석방법	1: 회귀분석    2: 구조방정식    3: 분산분석    4: 기타		
통제변수의 수	통제변수를 제외한 변수(독립, 종속, 조절, 매개변수 등)		
CCV 수	연구에서 제시된 CCV의 수		
통제변수 설정 이유 설명	0: 통제변수를 포함하지 않음 1: 전 변수에 대해서 아무런 설명을 하지 않음 2: 1개 이상의 변수에 대해서 설명을 하지 않음 3: 1개 이상의 변수에 대해서 불명확하거나 단순 설명 4: 모든 변수에 대해서 인용 및 근거를 제시		
통제변수의 측정방법 구체화 여부	0: 통제변수를 포함하지 않음 1: 전 변수에 대해서 아무런 설명을 하지 않음 2: 1개 이상의 변수에 대해서 설명을 하지 않음 3: 모든 변수에 대해서 구체화된 측정 방법을 제시		
MCV 수	연구에서 제시된 MCV의 수		
기술적 통계 (평균/표준편차)	0: 통제변수를 포함하지 않음 1: 전 변수에 대해서 아무런 설명을 하지 않음 2: 1개 이상의 변수에 대해서 설명을 하지 않음 3: 모든 변수에 대해서 설명을 함		
기술적 통계 (상관관계 크기)	0: 통제변수를 포함하지 않음 1: 전 변수에 대해서 아무런 설명을 하지 않음 2: 1개 이상의 변수에 대해서 설명을 하지 않음 3: 모든 변수에 대해서 설명을 함		
기술적 통계 (상관관계 유의수준)	0: 통제변수를 포함하지 않음 1: 전 변수에 대해서 아무런 설명을 하지 않음 2: 1개 이상의 변수에 대해서 설명을 하지 않음 3: 모든 변수에 대해서 설명을 함		
분석결과 (효과크기)	0: 통제변수를 포함하지 않음 1: 전 변수에 대해서 아무런 설명을 하지 않음 2: 1개 이상의 변수에 대해서 설명을 하지 않음 3: 모든 변수에 대해서 설명을 함		

## 〈부록 1〉 질적 분석 결과 코딩 방법 (계속)

<b>분석결과 (유의수준)</b>	0: 통제변수를 포함하지 않음 1: 전 변수에 대해서 아무런 설명을 하지 않음 2: 1개 이상의 변수에 대해서 설명을 하지 않음 3: 모든 변수에 대해서 설명을 함
<b>다중공선성의 문제 발생 여부</b>	0: 통제변수를 포함하지 않음 1: 모든 변수에서 다중공선성 발생 우려 2: 1개 이상의 변수에서 다중공선성 발생 우려 3: 모든 변수에서 다중공선성 발생 우려 없음
<b>독립변수와의 상관관계</b>	0: 통제변수를 포함하지 않음 1: 모든 변수에서 독립변수와 무관한 통제변수 2: 1개 이상의 변수에서 독립변수와 무관한 통제 변수 3: 모든 변수에서 독립변수와 관련 있는 변수
<b>종속변수와의 상관관계</b>	0: 통제변수를 포함하지 않음 1: 모든 변수에서 종속변수와 무관한 통제변수 2: 1개 이상의 변수에서 종속변수와 무관한 통제 변수 3: 전 변수에서 종속변수와 관련 있는 변수
<b>논의 포함 여부</b>	0: 통제변수를 포함하지 않음 1: 전 변수에 대해서 아무런 설명을 하지 않음 2: 1개 이상의 변수에 대해서 설명을 실시하지 않음 3: 모든 변수에 대해서 설명을 실시



## Improving causality: The significance of control variable, phenomena, and recommendations

Won-Woo Park\*

Dong Woon Go\*\*

Eunsung Yun\*\*\*

### ABSTRACT

The purpose of this study is to offer some recommendations for use of control variables. For this purpose, the authors investigate the meaning and significance of control variable. And then they examine statistical control in the sample of 179 articles about management published in *Korean Management Review* and *Korean Journal of Management* during 1997 to 2006. Authors' bases for including control variables, clarity regarding measures and methods, and reporting of results were analyzed. Potential problems includes a lack of explanations for inclusion, unclear descriptions of measures and methods, incomplete reports of results, and specifically, excluding control variable in the analyses of structural equation modeling(SEM). Based on this analysis, the authors suggest six recommendations to organizational researchers for how to use the control variables. In conclusion, this paper emphasizes that if researchers want to assure the causalities of their studies, they should include some appropriate control variables, and control variables should be treated like independent variables and dependent variables.

Key word: causal relationship, control, extraneous variable, control variable

---

\* Professor, College of Business Administration, Seoul National University

\*\* Doctoral student, College of Business Administration, Seoul National University

\*\*\* Master student, College of Business Administration, Seoul National University