

## 제조기업의 데이터 품질과 재무적 성과

김정철\* · 이춘열\*\* · 이상호\*\*\*

### Data Quality and Firm Financial Performance in the Manufacturing Industry

Jeong-Cheol Kim\* · Choon Yeul Lee\*\* · Sangho Lee\*\*\*

#### ■ Abstract ■

There is a belief that timely and precise data are important to decisions and the better decisions are related to better firm performance. However, empirical research investigating the effect of data quality on firm financial performance is still scarce up to recently. Current study empirically explores such an effect of data quality on firm accounting performance in the Korean manufacturing industry during 2008~2010 with secondary data. The results show that better data quality does not impact on sales and operating profit, but positively and significantly impacts on EVA(Economic Value Added). Raising the level of data quality management maturity by one level can increase EVA by about 34% in manufacturing firms.

Keyword : Data Quality, Firm Performance, Economic Value Added, Manufacturing Industry

## 1. 서 론

2011년 현재 한국 정부는 자동차, 조선, 기계, 의료, 섬유, 건설, 국방, 조명, 에너지, 로봇 등의 산업에 IT를 도입하거나 적극적으로 활용하여 산업의 생산성을 높이려는 산업-IT융합을 추진중이다[3]. 이러한 IT융합은 각 산업의 생산성뿐만 아니라 IT의 수요를 증가시켜 IT산업의 발전에도 중요한 역할을 할 수 있다. 산업이나 기업에 IT가 도입되거나 융합되면, 산업과 기업의 자료나 정보의 획득, 축적, 분석, 배포 등이 수월해진다. 산업-IT 융합은 기업이 보유하거나 활용할 수 있는 데이터량을 증가시킨다.

운영적이거나 전문적인, 그리고 전략적인 의사결정을 포함한 대부분의 기업 활동은 데이터를 필요로 한다[29]. 그리고 이러한 의사결정의 잠재적 가치는 의사결정에 사용되는 데이터의 품질에 따라 달라진다. 다양한 자료원으로부터 수집되는 원시 데이터들은 대부분 오류를 포함하게 되고, 이들 오류는 고객 불만을 야기시키거나 기업에 수집역원의 비용을 유발시킬 수 있으며, 새로운 전략의 구현을 어렵게 하거나 불가능하게 할 수 있다[27]. 데이터 웨어하우스 연구소(Data Warehouse Institute)는 데이터 품질 문제가 미국 기업에게 매년 600조원의 비용을 유발한다고 추정한다[18]. 이와 같이 데이터 품질이 기업 활동에 영향을 준 일화들은 자주 소개되고 있다. 즉, 높은 품질의 데이터는 기업 이익이나 수익 창출에 긍정적 효과를 주며, 이와 반대로 낮은 품질의 데이터는 기업 이익에 부정적 영향을 주는 것으로 받아들여지고 있다. 그러나 이들 데이터 품질과 기업의 재무성과에 대한 실증적 연구는 의외로 많지 않다.

데이터 품질과 관련된 대부분의 연구들(예 : [11, 12, 14, 19, 23, 31])은 데이터 품질 차원이나 데이터 품질을 측정하기 위한 절차를 주제로 하고 있다. 그리고 대부분의 실무자들은 데이터 품질이 개선되었을 때 조직이 얻을 수 있는 금전적인 혜택에 대해 잘 모를 수 있다. 본 연구는 이러한 기업

의 재무적 성과와 데이터 품질과의 연관성에 대한 답을 찾고자 한다.

데이터 품질과 기업의 재무적 성과에 대한 연구 부족은 다음과 같은 데이터 품질과 관련된 측정의 어려움 때문이다.

첫째, 조직의 데이터 품질 수준의 측정은 해당 조직의 협력을 필요로 한다. 즉, 기업의 데이터 품질 수준을 측정하기 위해서는 데이터 생산의 전체 과정을 추적할 수 있어야 하는데, 이들 과정은 대부분의 경우 보안 사항으로 분류되어 해당 기업의 허락을 얻는 것이 쉽지 않다.

둘째, 데이터 품질과 관련된 신뢰성 있는 이차 자료를 수집하는 것이 쉽지 않다. 데이터 품질 수준을 측정하기 위한 정보와 자료를 수집하는 것은 매우 노동집약적이며, 어려운 작업이다[24]. 따라서 연구자들이 직접 기업의 품질 수준을 객관적으로 평가하는 것은 쉬운 일이 아니다. 이차 자료 획득의 어려움으로 기존 연구들(예 : [32])은 리커트(Likert) 척도를 이용한 설문지를 사용하여 측정된 인지된 데이터 품질을 실제 데이터 품질의 측정치로 사용하곤 한다. 그러나 데이터 품질과 기업의 재무 성과와의 연관성을 객관적으로 검증하기 위해서는 인지된 품질 수준이 아닌 상호 객관적으로 비교 가능한 품질 수준이 필요하다.

본 연구는 이들 어려움을 극복하기 위한 방안으로서 한국데이터베이스진흥원(KDB : Korea Data Base agency, 이하 진흥원)에서 제공하는 데이터 품질 수준을 사용하여 데이터 품질과 기업 재무 성과와의 연관성을 검증한다. 또한 기업의 재무성과를 측정하는 지표로서 실제 기업들이 공시한 회계 자료를 사용하였으며, 산업 및 연도와 같이 기업 성과에 영향을 주는 요인들을 통제하여, 데이터 품질이 재무성과에 미치는 영향을 객관적으로 분석하고자 하였다.

진흥원의 데이터 품질관리 성숙 모형에서 데이터 품질은 유효성과 활용성의 차원으로 측정된다[10]. 유효성은 정확성과 일관성으로 세분되며, 활용성은 유용성, 접근성, 적시성, 보안성으로 세분된다. 이들

품질 요인들을 개략적으로 소개하면 다음과 같다.

- 정확성은 실제 세계에 존재하는 객체(이벤트, 사물, 개념 등)의 값이 오류 없이 저장되는 정도를 의미한다.
- 일관성은 데이터가 정보시스템 내에서 동일한 형식으로 제공되는 정도를 의미한다.
- 유용성은 조직이 요구하는 데이터의 범위와 상세화 정도를 충족시키는 정도를 의미한다.
- 접근성은 사용자가 원하는 데이터를 손쉽게 사용할 수 있는 정도를 의미한다.
- 적시성은 응답시간과 같은 비기능적 요구사항이나 데이터가 필요한 작업에 대하여 최신성을 유지하는 정도를 의미한다.
- 보안성은 데이터가 내부 및 외부 요인(예, 침입, 절도, 변경, 시스템 오류, 재해)으로부터 적절하게 보호되는 정도를 말한다[10, 26].

본 연구의 목적은 데이터 품질이 인지된 기업 성과가 아니라 기업의 재무 회계 성과에 주는 영향을 실증적으로 조사하는 것이다. 이러한 목적을 달성하기 위하여 본 연구는 다음의 내용을 수행하였다.

먼저 2008년부터 2010년까지 진흥원에서 발표하는 데이터 품질 수준 자료를 검토하여 연구 대상 기업을 선정하였다. 연구 대상 기업은 정보강도에 대한 통제를 위하여 제조업으로 한정하였다. 연구를 단일 산업에 집중하여 수행하는 것은 연구 결과의 일반화에 어려움이 있을 수 있지만, 통제하기 어려운 산업적 차이 요인을 통제할 수 있는 방법이다[33]. 이러한 점에서 본 연구는 금융업을 대상으로 한 기존 연구(예 : [4])를 제조업에 적용한 확장 연구라고 할 수 있다.

다음, 데이터 품질과 기업 성과 간의 관계를 조사하기 위해 계량경제학 모델을 사용하여 분석하였다. 데이터 품질과 기업 성과 간의 관계에 영향을 줄 수 있는 기업 크기, 하위산업 유형, 연도 등은 실증적 결과의 신뢰성을 높이기 위하여 통제변수로 취급하였다.

본 논문의 구성은 다음과 같다. 제 2장에서 관련

문헌을 검토하고, 제 3장에서 연구모형을 제시한다. 그리고 제 4장에서 실증 결과를 제시하고, 연구 결과의 의미를 설명한다. 마지막으로 제 5장에서 연구의 요약, 연구의 한계점, 미래 연구의 방향을 제시한다.

## 2. 선행연구 분석

### 2.1 데이터 품질과 데이터 품질관리 성숙모형

데이터 품질을 연구한 기존 연구들(예 : [2, 6-8, 23, 26, 30])은 주로 데이터 품질 요인들이나 데이터 품질 프레임워크를 구조화하여 제시하였다. 이들이 제시하고 있는 대표적인 품질 차원으로, 정확성, 적시성, 일관성, 완전성, 안정성 등이 이용되었다[19].

기존 연구(예 : [23, 26, 30])의 데이터 품질 차원과 명세를 기반으로 진흥원은 한국 기업이나 조직의 데이터 품질 요건을 만족하는 데이터 품질 프레임워크를 개발하였다[10]. 진흥원은 데이터 소비자를 만족시키는 데 필요한 데이터의 수준으로 데이터 품질을 정의하였고, 유효성과 활용성 차원에서 측정하고 있다. 본 연구는 진흥원의 데이터 품질의 정의와 차원을 따른다. 또한, 진흥원은 데이터 품질관리 성숙 수준을 측정하기 위한 데이터 품질관리 성숙모형을 개발하였다. 성숙도 모형은 도입 단계, 정형화 단계, 통합화 단계, 정량화 단계, 최적화 단계의 다섯 단계로 구성되어 있다. 제 1단계인 도입 단계에서는 조직이 데이터 품질관리에 대한 문제점과 필요성을 부분적으로 인식하고 있지만, 표준화와 정형화된 데이터 품질관리 프로세스가 미비하며, 데이터의 품질관리는 담당자 개개인의 능력에 의존해 수행된다. 정형화 단계는 데이터 품질관리를 위한 프로세스가 잘 정의되어 있는 단계이다. 통합화 단계에서는 데이터 품질관리가 일관성있고 전사적 관점에서 전개된다. 정량화 단계에서는 데이터 품질관리가 통계적 기법이나 정량적인 측정 방법을 통해 관리된다. 제 5단계인

최적화 단계에서는 조직이 데이터 품질관리 프로세스의 개선 사항을 지속적으로 도출하고 실행하며 평가를 통해 사후 관리를 수행한다[10, 13]. 진흥원은 데이터 품질 프레임워크를 이용하여 조직의 데이터 품질관리 성숙 수준을 측정하기 위하여 연례 설문 조사를 실시하고 있다.

## 2.2 데이터 품질과 기업 성과

데이터 품질은 정보의 품질과 밀접하게 관련되어 있기 때문에, 우리의 연구는 정보의 품질과 조직 성과와의 관계를 분석하는 정보시스템 성공 모형과 관련된 기존 문헌을 검토하였다. 문헌 검토가 끝난 후, 우리는 몇몇 성공 모형 연구(예 : [15, 17])가 정보의 품질과 조직성과 간의 긍정적인 관계를 제시하는 것을 알 수 있었다. 많은 연구들(예 : [16])은 설문지를 사용하여 정보 품질의 영향을 측정하였는데, 인지된 정보 품질이 인지된 효익에 긍정적이고 중요한 영향을 주는 것을 보이고 있다. 그리고 정보시스템 성공 모형의 메타분석을 수행한 연구(예 : [25])도 동일한 관계를 발견했다.

Wixom and Watson[32]은 111개 조직을 대상으로 데이터 웨어하우스 관리자 및 데이터 공급업체에 설문 조사를 실시하며 실증적으로 조사하였다. “데이터 웨어하우스는 나의 일을 크게 변경시켰다 [32, <표 2>, p.31]”와 같은 질문에 의해 측정되었는데, 결과는 데이터 웨어하우스의 우수한 데이터 품질은 인지된 더 많은 순효익과 관련된 것으로 나타났다.

본 연구는 또한 나쁜 품질의 데이터에 대한 문헌(예 : [19, 27, 28])과 관련되어 있다. 이러한 연구는 나쁜 품질의 데이터는 즉시적으로 재무적인 성과에 부정적인 영향을 주고, 간접적이고 장기적인 더 많은 악영향을 줄 수 있다고 주장했다. 전형적인 기업 사례를 분석한 연구(예 : [5, 27])는 기업의 나쁜 데이터 품질의 영향을 조사하고, 나쁜 품질의 데이터가 8~12%의 수익을 감소시키고 40~60%의 지출을 증가시킨 것으로 보고하였다. 이러

한 사례 연구는 기업 성과에 대한 데이터 품질의 효과의 이해를 위한 중요한 시작이었다. 그렇지만, 저자 자신들이 연구에서 서술한 것과 같이 그들의 사례 연구 결과는 몇 가지 사례에 기초하였다는 한계점이 있다.

기존 연구(예 : [4])는 한국 금융업에서 데이터 품질이 기업의 경제적인 성과에 주는 영향을 실증적으로 분석하였다. 연구자들이 예상했던 대로 데이터 품질은 기업의 성과에 통계적으로 유의하게 긍정적인 영향을 주는 결과를 발견하였다. 그러나 이들의 연구는 일반적으로 정보집약적 산업으로 분류되는 산업인 금융업에 대한 분석이었다. 정보 기술 투자가 기업의 성과에 주는 영향이 정보집약적 산업과 정보집약적이지 않은 산업에서 차이가 있기 때문에[21], 데이터 품질의 영향도 산업의 정보집약도에 따라 차이가 있을 수 있다.

데이터 품질 개선을 위한 실천 전략이나 좋지 않은 데이터 품질의 영향에 대한 상당히 실천적인 지혜를 제시하는 연구가 있었지만, 그것은 제한된 수의 일화적인 사례에 근거하는 경우가 많았다. 또한 데이터의 품질 인식과 혜택 인식의 관계를 분석하기 위한 노력도 있었지만, 데이터 품질의 효과에서 인지된 혜택(recognized benefit)과 기업의 재무적인 성과와는 차이가 있다. 그 연구들은 데이터 품질의 재무적인 영향을 체계적이고 엄격하게 조사하기 위해 횡단면 및 장기간의 데이터를 사용하지도 않았다. 또한 데이터 품질의 실증적인 증거를 제시한 기존 연구(예 : [4])는 정보집약적인 산업인 금융업에서의 분석이었기에, 분석 결과의 일반화를 위하여 정보집약적이지 않은 산업에서의 분석이 필요하다.

데이터 품질과 기업 성과에 대한 응답자의 인식을 이용한 설문 조사를 이용하였거나 몇 가지 사례를 조사했던 기존 연구와는 달리, 본 연구는 제조기업의 재무 회계 성과에 주는 데이터 품질의 영향을 실증적으로 조사하였다. 본 연구의 결과는 데이터 품질이 기업의 재무성과에 미치는 영향에 대한 우리의 현재 이해를 깊게 할 수 있으며, 기존에

수행된 연구 결과[4]의 일반화에 기여할 수 있다.

### 3. 연구방법

#### 3.1 연구모형

본 연구는 기업의 재무성과에 대한 데이터 품질의 효과를 측정하는 회귀 모델을 사용한다. 데이터 품질뿐만 아니라 다른 경영 활동(예 : 정보기술 투자, 연구 개발, 마케팅, 광고)도 기업 성과에 영향을 주고, 산업의 정보강도나 산업성장률과 같은 산업 특성, 연도별 상황도 기업 성과에 영향을 주는 요인이자[21]. 이 연구는 데이터 품질의 효과에 집중하며, 산업 특성, 기업의 크기, 연도와 같은 요인을 통제한다.

기존 연구(예 : [20])는 기업 크기가 독립 변수와 기업 성과 간의 관계에 영향을 미칠 수 있음을 보이고 있다. 따라서, 기업의 크기를 통제하며, 기업 성과에 대한 데이터 품질의 효과를 찾기 위한 연구 수식은 다음과 같다 :

$$FP(i, t) = \alpha + \beta_1 DQ(i, t) + \beta_2 Size(i, t) + \delta_1 D_1 + \delta_2 D_2 + \delta_3 D_3 + \epsilon(i, t)$$

여기서  $FP(i, t)$ 는  $i$ 기업의  $t$ 년도의 매출액의 로그값, 영업이익의 로그값, 부가가치의 로그값으로 표현되는 기업 성과;  $DQ(i, t)$ 는 데이터 품질;  $Size(i, t)$ 는 직원 수의 로그값; 2008년 자료이면  $D_1 = D_2 = 0$ , 2009년 자료이면  $D_1 = 1$ , 2010년 자료이면  $D_2 = 1$ , 전자부품, 컴퓨터, 전기장비 산업(정보집약적 산업)과 같이 두자리 산업 분류 코드가 26이상이면  $D_3 = 1$ , 그렇지 않으면  $D_3 = 0$ ;  $\alpha$ 는 상수항이고,  $\beta_1$ 은 데이터 품질의 계수이고,  $\beta_2$ 는 기업 크기의 계수이고,  $\delta$ 들은 통제변수의 계수이고,  $\epsilon(i, t)$ 는 에러항이다.

#### 3.2 독립, 종속, 통제 변수

데이터 품질 변수는 데이터 소비자의 요구사항

을 충족하는 데 필요한 데이터의 수준으로 정의되며, 진흥원에 의해 0부터 5(최적화 단계)의 규모로 평가된다. 성과 변수의 경우, 매출액, 영업이익, 부가가치의 로그 값을 이용한다. 매출액은 기업의 전체 판매액이고, 부가가치는 매출액에서 매출원가를 차감하고 판매 및 관리 비용을 추가하여 계산하고, 영업이익은 매출액에서 매출원가와 판매 및 관리 비용을 차감한 값이다. 본 연구는 패널 데이터를 사용하기 때문에, 성과 변수는 통계청의 국내총생산(GDP : Gross Domestic Product) 디플레이터[9]를 이용하여 2008년도 불변가치로 변환된다. 또한 기업의 크기(직원 수의 로그값), 연도, 그리고 하위산업의 정보집약도를 통제하였다.

#### 3.3 자료

본 연구의 데이터 품질 자료는 2008년부터 2010년까지 진흥원에서 조사한 결과를 이용하였다. 기업의 성과 자료는 한국금융감독원의 전자공시 시스템[1]을 이용하여 한국의 제조기업으로부터 수집하였다.

진흥원의 연도별 조사중 제조업에서 유효한 응답의 수는 각각 19, 25, 42로 표본기업의 수는 86이다(<표 1> 참조). 데이터 품질의 평균은 0.940으로, 한국 제조기업의 대부분은 데이터 품질관리의 1단계인 도입 단계에 있다는 것을 나타낸다. 매출액은 최소 284억 원, 최대 6조 7,034억 원, 평균 5,640억 원이다. 평균 영업이익은 452억 원, 평균 부가가치는 2,990억 원이다. 기업 성과의 로그값을 이용하였기 때문에 영업이익이나 부가가치를 창출하지 못한 기업은 표본에서 제외되었다. 평균 직원 수는 1,450명이며, 직원 수는 기초와 기말의 평균 값을 이용하여 계산되었다.

금융업을 대상으로 분석한 이상호, 박주석, 김재경[4, <표 1>, p.5]의 연구와 비교하면, 제조기업의 평균 매출액은 금융기업 평균의 8% 미만이지만, 부가가치는 약 50% 정도를 창출하여 매출액 대비 부가가치 창출율은 제조기업이 금융기업보다 5배

이상 더 높음을 알 수 있다.

〈표 1〉 기초통계량

항목	N	최소값	최대값	평균	표준편차
데이터 품질	86	0.0	2.67	0.940	0.604
매출액	82	284	67,034	5,640	9,780
영업이익	71	81	8,428	452	1,090
부가가치	84	412	93,308	2,990	10,700
직원 수	86	57	25,274	1,450	3,770

참고사항 : 기업 성과 변수의 단위는 억원(2008년도 불변가치). 직원 수는 기초와 기말의 평균값.

## 4. 결과와 토의

### 4.1 결과

피어슨 상관관계 분석의 결과는 〈표 2〉와 같다. 데이터 품질은 매출액의 로그값과 상관관계를 보이지 않고 있지만, 영업이익이나 부가가치의 로그값과 같은 기업 성과와는 긍정적인 상관관계를 보인다. 기업 성과 변수 간에는 통계적으로 유의한 긍정적인 상관관계를 보이고 있다. 또한, 분석 결과는 기업 크기가 높은 수준의 데이터 품질과 기업 성과와 연관되어 있음을 나타낸다. 따라서 기업 크기는 분석에서 통제될 필요가 있다.

〈표 2〉 상관관계 분석

	$\log S$	$\log OP$	$\log VA$	$DQ$
$\log S$	1.00			
$\log OP$	0.737***	1.00		
$\log VA$	0.770***	0.737***	1.00	
$DQ$	0.147	0.219*	0.319***	1.00
$Size$	0.717***	0.644***	0.805***	0.197*

주) \*\*\*  $p < 0.01$ ; \*\*  $p < 0.05$ ; \*  $p < 0.1$ .

$Size$  : 직원 수의 로그값.

회귀분석을 위해서 먼저 회귀분석의 기본 가정사항들(예, 정규성, 직선성, 극단치 없음 등)[21]의 유효성을 검사하였으며, 기본 가정사항들을 위해

하지 않아 회귀분석을 계속할 수 있었다. 매출액의 로그값( $\log S$ )과 영업이익의 로그값( $\log OP$ )을 종속변수로 하는 회귀분석 결과에서는 데이터 품질의 의미있는 효과를 찾지 못하였다(〈표 3〉, 〈표 4〉 참조).

〈표 3〉 매출액의 로그값( $\log S$ )에 대한 회귀분석

	모형 i	모형 ii	모형 iii	모형 iv
상수	5.36*** (0.093)	2.95*** (0.273)	2.83*** (0.279)	2.83*** (0.280)
$DQ$	0.116 (0.087)	0.066 (0.062)	0.038 (0.064)	0.034 (0.064)
$Size$		0.904*** (0.100)	0.914*** (0.101)	0.908*** (0.101)
$D_1$			0.096 (0.106)	0.095 (0.106)
$D_2$			0.176* (0.096)	0.177* (0.096)
$D_3$				0.054 (0.071)
관측 수	82	82	82	82
$R^2$	0.022	0.521	0.543	0.546
조정 $R^2$	0.009	0.509	0.519	0.517
$F$ 통계량	1.77	43.0***	22.9***	18.3***
더빈왓슨 $d$	1.95	1.73	1.80	1.82

주) \*\*\*  $p < 0.01$ ; \*\*  $p < 0.05$ ; \*  $p < 0.1$ .

괄호안의 값은 표준편차.

〈표 5〉는 부가가치의 로그값( $\log VA$ )에 대한 데이터 품질의 영향을 분석한 결과이다. 모형 ix는 관련된 변수의 통제없이 데이터 품질만의 영향을 보이는데, 약 10%의 설명력을 보이고 있다. 모형 x는 기업 크기를, 모형 xi는 기업 크기와 연도를, 모형 xii는 기업 크기, 연도, 하위산업의 정보강도를 통제한 결과이다. 기업의 성과에 영향을 주는 요인을 통제한 회귀모형들(모형 x, xi, xii)의  $F$ -통계값은 각각 0.01수준에서 유의하며, 모형의 설명력은 약 65% 정도이다. 데이터 품질의 회귀계수는 통계적으로 유의하여, 데이터 품질이 부가가치의 로그값으로 표현되는 기업 성과에 통계적으로 유의하게 긍정적인 영향을 주는 것을 알 수 있다. 데이터 품

질의 회귀계수가 0.126이므로(모형 xii), 제조기업의 데이터 품질관리 성숙도가 1단계 상승하면, 부가가치는 약 33.7% 증가한다.

〈표 4〉 영업이익의 로그값( $\log OP$ )에 대한 회귀분석

	모형 v	모형 vi	모형 vii	모형 viii
상수	4.07*** (0.121)	1.63*** (0.381)	1.51*** (0.402)	1.51*** (0.404)
$DQ$	0.200* (0.108)	0.073 (0.087)	0.056 (0.090)	0.060 (0.091)
$Size$		0.942*** (0.142)	0.953*** (0.144)	0.958*** (0.146)
$D_1$			0.111 (0.147)	0.109 (0.148)
$D_2$			0.129 (0.135)	0.127 (0.136)
$D_3$				-0.035 (0.103)
관측 수	71	71	71	71
$R^2$	0.048	0.421	0.429	0.430
조정 $R^2$	0.034	0.404	0.395	0.386
$F$ 통계량	3.48*	24.7***	12.4***	9.82***
더빈왓슨 $d$	1.64	1.53	1.55	1.55

주) \*\*\*  $p < 0.01$ ; \*\*  $p < 0.05$ ; \*  $p < 0.1$ .  
팔호안의 값은 표준편차.

금융업에서 데이터 품질의 영향을 분석한 기존 연구[4]의 결과와 제조업에서 데이터 품질의 영향을 분석한 본 연구의 결과를 비교하여, 제조업에서의 영향이 금융업에서의 영향보다 작다는 것을 알 수 있다. 본 연구에서는 매출액과 영업이익의 로그값을 종속변수로 하는 모형에서는 유의한 결과를 찾을 수 없었던 반면, 기존 연구[4]에서는 유의한 양의 영향관계를 찾을 수 있었다. 이것을 통하여 금융업에서 더 큰 효과를 보이는 것을 알 수 있다. 부가가치의 로그값을 종속변수로 하는 모형의 비교에서는 0.126(제조업, <표 5>의 4번째 열의 모형 xii)과 0.237(금융업, <표 6>의 4번째 열의 모형 12)로 제조업에서의 영향은 금융업에서의 영향의 약 1/2수준이다.

〈표 5〉 부가가치의 로그값( $\log VA$ )에 대한 회귀분석

	모형 ix	모형 x	모형 xi	모형 xii
상수	4.66*** (0.108)	1.98*** (0.241)	2.06*** (0.266)	2.08*** (0.261)
$DQ$	0.297*** (0.098)	0.111* (0.062)	0.120* (0.064)	0.126** (0.063)
$Size$		1.04*** (0.090)	1.02*** (0.093)	1.04*** (0.091)
$D_1$			-0.051 (0.105)	-0.045 (0.103)
$D_2$			-0.081 (0.097)	-0.080 (0.096)
$D_3$				-0.149** (0.073)
관측 수	84	84	84	84
$R^2$	0.102	0.661	0.664	0.681
조정 $R^2$	0.091	0.653	0.647	0.660
$F$ 통계량	9.27***	78.9***	39.0***	33.3***
더빈왓슨 $d$	2.08	1.69	1.71	1.79

주) \*\*\*  $p < 0.01$ ; \*\*  $p < 0.05$ ; \*  $p < 0.1$ .  
팔호안의 값은 표준편차.

〈표 6〉 제조업과 금융업의 비교 :  $\log VA$

	제조업		금융업[4, p.7]	
	모형 xiii	모형 xiv	모형 11	모형 12
상수	4.80*** (0.180)	1.94*** (0.252)	4.64*** (0.222)	2.62*** (0.323)
$DQ$	0.336** (0.149)	0.152** (0.071)	0.628*** (0.131)	0.237** (0.103)
$D_1$	-0.119 (0.191)	-0.060 (0.089)	-0.111 (0.196)	0.116 (0.133)
$Size$		1.06*** (0.088)		0.805*** (0.114)
관측 수	43	43	40	40
$R^2$	0.112	0.813	0.396	0.746
조정 $R^2$	0.068	0.798	0.363	0.724
$F$ 통계량	2.53*	56.4***	12.1***	35.2***
더빈왓슨 $d$	2.42	1.89	2.43	1.76

주) \*\*\*  $p < 0.01$ ; \*\*  $p < 0.05$ ; \*  $p < 0.1$ .  
팔호안의 값은 표준편차.

<표 6>은 본 연구가 연도를 통제하였음에도 불구하고, 본 연구와 기존 연구[4]와의 영향 차이가

표본기간이나 사용된 통제변수의 차이 때문일 수 있기에 기존 연구와 동일한 표본기간과 통제변수를 이용하여 비교한 결과이다. 기존 연구[4]의 연구와 동일하게 2008년과 2009년의 자료만을 이용하였고, 연도와 기업 크기만을 통제하였다. 제조업에서의 데이터 품질의 영향(0.152)은 금융업에서의 데이터 품질의 영향(0.237)에 비하여 약 65%에 그친다. 또한 데이터 품질 관리 성숙도가 1단계 상승하면, 금융업에서는 부가가치가 72.6% 증가하며, 제조업에서는 부가가치가 41.9% 증가한다.

## 4.2 토의

이 연구는 한국 제조기업의 데이터 품질과 재무성과 간의 관계를 조사하였다. 데이터 품질은 매출액과 영업이익에는 영향을 주지 못하였으나, 부가가치에는 통계적으로 유의한 긍정적인 영향을 주었다. 부가가치에 대한 데이터 품질의 탄력성( $\beta_1$ )은 0.126으로, 데이터 품질 관리 성숙도의 1단계 상승은 부가가치를 약 34% 증가시킬 수 있다.

본 연구의 표본기업의 평균 부가가치는 2,990억 원이다(<표 1> 참조). 데이터 품질이 0.94(도입 단계)에서 1.94(정형화 단계)로 개선되면, 기업의 평균 부가가치는 약 1,000억 원 정도가 증가하여 약 4,000억 원이 될 수 있다.

한국 제조업의 2008년~2010년 평균 부가가치는 204조 원이다[9]. 데이터 품질의 수준이 0.94(도입 단계)에서 1.94(정형화 단계)로 상승하면, 한국 제조업의 부가가치는 약 69조 원이 증가되어 273조 원으로 증가될 수 있다. 그리고 2008년~2010년 한국의 평균 국내총생산은 1,088조 원이다[9]. 데이터 품질의 수준이 0.94(도입 단계)에서 1.94(정형화 단계)로 상승하고, 부가가치에 주는 데이터 품질의 영향이 국내총생산에도 동일하게 영향을 줄 수 있다면, 한국의 국내 총생산은 약 366조 원이 증가되어 약 1,454조 원이 될 수 있다.

금융업을 대상으로 한 기존 연구[4, p.8]에서는 약 1,803조 원으로 증가될 수 있다고 하였지만, 제

조업을 대상으로 한 본 연구에서는 약 1,454조 원으로 증가될 수 있다. 약 350조 원의 큰 추정오차가 발생한다. 정보집약적인 산업인 금융업에서의 결과를 전체 산업에서의 결과로 일반화시키기에는 과추정될 수 있는 위험이 있음을 알 수 있다. 데이터 품질의 영향을 다양한 산업의 총생산을 합한 국내 총생산에 적용하는 것에는 제한사항이 있음을 알 수 있다.

## 5. 결 론

기존 연구들(예 : [26-28])은 데이터 품질이 기업 성과에 중요한 영향력을 가지고 있다고 주장하지만, 그들은 결론을 지원하는 명확한 실증적 증거를 제공하지 않았다. 반면 본 연구는 데이터 품질이 기업의 재무적인 성과에 통계적으로 유의하게 긍정적인 영향을 주고 있다는 실증적인 증거를 제시하였다.

기업의 재무성과에 대한 데이터 품질의 효과를 이해하는 것은 실무적이고 학문적으로 중요한 의미가 있다. 첫째, 우리의 결과는 한국 제조기업의 관리자들이 전략적으로 데이터 품질의 향상에 투자해야 하고 정보시스템을 개발할 때 데이터 품질에 충분한 주의를 기울여야 함을 보여준다. 또한, 고품질의 데이터를 보장하기 위해 조직의 비즈니스 프로세스, 정보시스템, 조직 구조를 생성하고 개선해야 함도 보인다. 데이터 품질관리 성숙 수준을 1단계 개선하면, 부가가치가 34% 정도 증가될 수 있다. 물론 관리자는 비용편익분석을 수행해야 한다. 자신의 기업에서 34%의 부가가치 증가의 가치가 데이터 품질관리 성숙 수준 1단계의 개선 비용보다 큰 경우에만, 관리자는 데이터 품질의 향상에 투자해야 한다. 둘째, 본 연구는 데이터베이스 산업에 대한 한국 정부의 현재 정책은 기업의 성장이나 성과 창출을 위하여 중요하고 의미가 있다는 것을 보여 준다. 한국 정부는 국가의 데이터베이스 산업을 지원하고 육성하여 정보 사회의 신속한 실현을 목적으로 1993년 진흥원을 설립하였다. 그



리고 진흥원은 한국 기업이나 기관의 데이터 품질을 향상시키기 위하여 데이터 품질관리 인증 센터를 개설하였고, 2006년부터 기업이나 기관의 데이터 품질을 평가하고 고품질 데이터에 대한 인증을 제공하고 있다. 마지막으로, 본 연구는 제조업을 대상으로 분석하여 금융업을 대상으로 분석한 기존 연구[4] 결과의 일반화를 추구하였다. 금융업은 정보집약적인 산업이기에 데이터 품질의 영향이 제조업보다 컸으며, 특정 산업의 결과를 전체 산업에 적용하는 것은 일반화의 오류가 발생할 수 있음을 보였다.

우리의 연구에는 몇 가지 제약 사항들이 있다. 첫째, 제조업에 한정되는 연구 결과이기에, 통신업이나 도소매업과 같은 서비스 산업으로의 일반화는 어려울 수 있다. 둘째, 신뢰할 수 있는 IT 투자에 대한 자료의 부족으로, 기존 연구에서 중요한 변수(예 : [20, 21])로 실증된 IT 투자를 통제하지 못하였다. 마지막으로, 장기간의 충분하지 못한 자료때문에 데이터 품질이 기업 성과에 주는 즉시적인 영향만을 측정하였다.

미래 연구의 첫 번째 주제는 전체 산업으로부터 표본을 수집하여 분석하는 연구이며, 데이터 품질의 비즈니스 가치가 전체 산업 분야에서 파생되는 방법에 대한 우리의 이해를 증진시키기 위하여 필요하다. 둘째, IT 투자는 기업 성과에 영향을 주는 중요한 요인이므로[22], 데이터 품질과 IT 투자를 모두 고려하는 연구가 데이터 품질의 효과에 대한 이론을 고안하는 데 필요하다. 마지막으로, 데이터 품질 자료의 축적으로 시차를 고려하는 연구모형은 데이터 품질의 실제 가치나 데이터 품질과 기업 성과 사이의 인과 관계를 이해하는 데 도움을 줄 수 있다.

## 참 고 문 헌

- [1] 금융감독원, 전자공시 시스템, <http://dart.fss.or.kr/>, 2012.
- [2] 박현지, 이춘열, “고객 DB 품질평가 및 품질원인에 관한 연구 : 고객포인트 관리 DB에 대한 실증적 평가”, 『한국정보기술응용학회 춘계학술대회』, (2004), pp.190-209.
- [3] 유천수, 심승배, “국방-IT융합을 위한 전략적 접근방안”, 『한국통신학회지(정보와 통신)』, 제28권, 제4호(2011), pp.3-10.
- [4] 이상호, 박주석, 김재경, “금융산업에서의 데이터 품질이 경제적인 성과에 주는 영향의 실증 분석”, 『Information Systems Review』, 제13권, 제1호(2011), pp.1-11.
- [5] 이춘열, “데이터베이스 품질평가에 관한 사례 연구”, 『Journal of Information Technology Applications and Management』, 제11권, 제4호 (2004), pp.209-225.
- [6] 이춘열, “비즈니스 인텔리전스 환경에서 변화관리를 이용한 데이터 품질 향상에 대한 연구”, 『Information Systems Review』, 제6권, 제2호 (2004), pp.65-77.
- [7] 이춘열, “정보구조 그래프를 이용한 통합 데이터 품질 관리방안 연구”, 『Journal of Information Technology Applications and Management』, 제10권, 제4호(2003), pp.103-118.
- [8] 이춘열, 박재현, 배광현, 안희진, “데이터베이스 품질평가의 실증연구 : 한국데이터베이스진흥센터 품질평가 프레임워크를 중심으로”, 『데이터베이스 그랜드 컨퍼런스』, (2003), pp. 455-461.
- [9] 통계청, 국가통계포털, <http://kosis.kr/index/index.jsp>, 2012.
- [10] 한국데이터베이스진흥원(구 한국데이터베이스진흥센터), 『데이터 품질관리 성숙 모형(Ver 1.0)』, [http://www.dbq.or.kr/pds/pds\\_read.php?DBS\\_IDX=123&page=1&field=&keyword=&DBM\\_IDX=3](http://www.dbq.or.kr/pds/pds_read.php?DBS_IDX=123&page=1&field=&keyword=&DBM_IDX=3), 2006.
- [11] Ballou, D. P. and H. L. Pazer, “Designing Information Systems to Optimize the Accuracy-timeliness Tradeoff”, *Information Systems Research*, Vol.6, No.1(1995), pp.51-72.

- [12] Batini, C., C. Cappiello, C. Francalanci and A. Maurino, "Methodologies for Data Quality Assessment and Improvement", *ACM Computing Surveys*, Vol.41, No.3(2009), pp.1-52.
- [13] Caballero, I., A. Caro, C. Calero and M. Piatini, "IQM3 : Information Quality Management Maturity Model", *Journal of Universal Computer Science*, Vol.41, No.22(2008), pp. 3658-3685.
- [14] Cappiello, C., C. Francalanci and B. Pernici, "Time-related Factors of Data Quality in Multichannel Information Systems", *Journal of Management Information Systems*, Vol. 20, No.3(2003), pp.71-92.
- [15] DeLone, W. H. and E. R. McLean, "Information Systems Success : The Quest for the Dependent Variable", *Information Systems Research*, Vol.3, No.1(1992), pp.60-95.
- [16] DeLone, W. H. and E. R. McLean, "Measuring E-commerce Success : Applying the DeLone and McLean Information Systems Success Model", *International Journal of Electronic Commerce*, Vol.9, No.1(2004), pp. 31-47.
- [17] DeLone, W. H. and E. R. McLean, "The DeLone and McLean Model of Information Systems Success : A Ten-year Update", *Journal of Management Information Systems*, Vol.19, No.4(2003), pp.9-30.
- [18] Eckerson, W. W., *Data Quality and the Bottom Line : Achieving Business Success through a Commitment to High Quality Data*, TDWI Report Series, The Data Warehousing Institute, <http://download.101com.com/pub/tdwi/Files/DQReport.pdf>, 2002.
- [19] Haug, A., F. Zachariassen, and D. van Liempd, "The Costs of Poor Data Quality", *Journal of Industrial Engineering and Management*, Vol.4, No.2(2011), pp.168-193.
- [20] Kim, J. K., J. Y. Xiang and S. Lee, "The Impact of IT Investment on Firm Performance in China : An Empirical Investigation of the Chinese Electronics Industry", *Technology Forecasting and Social Change*, Vol.76, No.5(2009), pp.678-687.
- [21] Lee, S. and S. H. Kim, "A Lag Effect of IT Investment on Firm Performance", *Information Resources Management Journal*, Vol. 19, No.1(2006), pp.43-69.
- [22] Lee, S., J. Y. Xiang and J. K. Kim, "Information Technology and Productivity : Empirical Evidence from the Chinese Electronics Industry", *Information and Management*, Vol.48, No.2/3(2011), pp.79-87.
- [23] Madnick, S. E., Y. W. Lee, R. Y. Wang, and H. Zhu, "Overview and Framework for Data and Information Quality Research", *ACM Journal of Data and Information Quality*, Vol.1, No.1(2009), pp.1-22.
- [24] Miles, M. B., "Qualitative Data as an Attractive Nuisance : The Problem of Analysis", *Administrative Science Quarterly*, Vol.24, No.4(1979), pp.590-601.
- [25] Petter, S. and E. R. McLean, "A Meta-analytic Assessment of the DeLone and McLean IS Success Model : An Examination of IS Success at the Individual Level", *Information and Management*, Vol.46, No.3 (2009), pp.159-166.
- [26] Pipino, L. L., Y. W. Lee, and R. Y. Wang, "Data Quality Assessment", *Communications of the ACM*, Vol.45, No.4(2002), pp.211-218.
- [27] Redman, T. C., "Improve Data Quality for Competitive Advantage", *Sloan Management Review*, Vol.36, No.2(1995), pp.99-107.
- [28] Redman, T. C., "The Impact of Poor Data

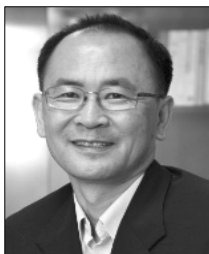
- Quality on the Typical Enterprise”, *Communications of the ACM*, Vol.41, No.2(1998), pp.79-82.
- [29] Tee, S. W., P. L. Bowen, P. Doyle, and F. H. Rohde, “Factors Influencing Organizations to Improve Data Quality in their Information Systems”, *Accounting and Finance*, Vol.47, No.2(2007), pp.335-355.
- [30] Wang, R. Y. and D. M. Strong, “Beyond Accuracy : What Data Quality Means to Data Consumers”, *Journal of Management Information Systems*, Vol.12, No.4(1996), pp.5-34.
- [31] Wang, R. Y., V. C. Storey and C. P. Firth, “A Framework for Analysis of Data Quality Research”, *IEEE Transactions on Knowledge and Data Engineering*, Vol.7, No.4(1995), pp.623-639.
- [32] Wixom, B. H. and H. J. Watson, “An Empirical Investigation of the Factors Affecting Data Warehousing Success”, *MIS Quarterly*, Vol.25, No.1(2001), pp.17-41.
- [33] Zhu, K., K. L. Kraemer, S. Xu and J. Dedrick, “Information Technology Payoff in E-business Environments : An International Perspective on Value Creation of E-business in the Financial Services Industry”, *Journal of Management Information Systems*, Vol.21, No.1(2004), pp.17-54.

## ◆ 저 자 소 개 ◆



**김 정 철** (kjcacc@mnd.go.kr)

성균관대학교 회계학과 졸업, 국민대학교 비즈니스IT전문대학원 석사를 마치고 동대학원 박사과정에 재학 중에 있다. 국방부 회계관리과장, 정보화정책과장, 시설기획과장 등 주요 보직을 역임하였으며 현재 시설기획환경과장으로 재직 중이다. 주요 관심분야는 회계 및 자산 관리, 데이터 자산 및 마스터 데이터 관리(Master Data Management) 등이다.



**이 춘 열** (cylee@kookmin.ac.kr)

서울대학교 산업공학과를 졸업하고, 동 대학원 경영학과에서 경영학석사 학위를 취득하였으며, 미시간대학교에서 경영정보학박사(Computer and Information Systems) 학위를 취득하였다. 또한, 국방정보체계연구원에서 연구원으로 근무하였으며, 한국통신 소프트웨어연구소에서 근무하였다. 국민대학교 비즈니스IT전문대학원 원장을 역임하였으며, 현재 국민대학교 경영정보학부 교수로 재직하고 있다. 한국경영정보학회, 한국경영학회, 대한산업공학회, 한국데이터베이스학회 등의 정회원이며, 주요 관심분야는 데이터베이스, 데이터웨어하우징, 정보 자원 계획 및 관리 등이다.



**이 상 호** (slee@sunmoon.ac.kr)

선문대학교 IT경영학과 조교수이며, 성균관대학교 경영학과에서 학사 및 석사를 취득하고 한국과학기술원(KAIST) 경영대학에서 경영공학박사를 취득하였다. Information Resources Management Journal, Technological Forecasting and Social Change, Information and Management, 한국IT서비스학회지, 경영정보학연구(현, Asia Pacific Journal of Information Systems) 등의 학술지에 논문을 게재하였다. 주요 관심분야는 정보기술 투자 평가, S/W개발 프로젝트관리, S/W 프로세스 개선 등이다.