

빅데이터를 활용한 양파 관측의 사회적 후생효과 분석

주재창* · 문지혜**

Analysis of Social Welfare Effects of Onion
Observation Using Big Data

Joo, Jae-Chang · Moon, Ji-Hye

This study estimated the predictive onion yield through Stepwise regression of big data and weather variables by onion growing season. The economic feasibility of onion observations using big data was analyzed using estimated predictive data. The social welfare effect was estimated through the model of Harberger's triangle using onion yield prediction with big data and it without big data. Predicted yield using big data showed a deviation of -9.0% to 4.2%. As a result of estimating the social welfare effect, the average annual value was 23.3 billion won. The average annual value of social welfare effects if big data was not used was measured at 22.4 billion won. Therefore, it was estimated that the difference between the social welfare effect when the prediction using big data was used and when it was not was about 950 million won. When these results are applied to items other than onion items, the effect will be greater. It is judged that it can be used as basic data to prove the justification of the agricultural observation project. However, since the simple Harberger's triangle theory has the limitation of oversimplifying reality, it is necessary to evaluate the economic value through various methods such as measuring the effect of agricultural observation under a more realistic rational expectation hypothesis in future studies.

Key words : *big data, observation, onion yield, social welfare effects*

* First author, 한국농촌경제연구원 위촉전문연구원

** Corresponding author, 한국농촌경제연구원 위촉연구원(mjh91@krei.re.kr)

I. 서 론

1999년부터 시작된 농업관측사업을 통해 농산물의 수급 안정과 농업인의 농업소득 안정화를 위해서 주요 농산물에 대하여 관측정보를 제공하고 있다. 제공하는 주요 관측정보로는 기상정보, 재배면적, 작황, 예상생산량, 재고물량, 소비동향, 해외시장 정보 등이 있으며, 이러한 정보를 종합적으로 조사분석하여 예측한 미래정보를 제공하고 있다(Kim et al., 2009).

정부는 농업관측사업을 포함하여 수매비축, 수급 안정 사업 등을 시행하고 있지만, 양파와 같은 채소류는 재배면적 변동이 심하고 기상조건에 영향을 받아 여전히 가격의 불안정성의 문제를 해결하는 데 어려움을 겪고 있다. 농산물 가격의 불안정성은 농산물 수요와 공급의 비탄력성, 농업생산에 내재한 불확실성 등의 요인이 존재한다. 일반적으로 이러한 가격 위험은 과잉 생산에 의해 가격이 하락할 경우, 생산자인 농가의 소득을 감소시키며, 과소생산에 의해 가격이 상승할 경우 소비자의 후생을 감소시키는 것으로 알려져 있다(Hayami, 1972; Freebairn, 1976; Antonovitz and Roe, 1984; Nieuwoudt et al., 1988; Lee et al., 2002; Kim et al., 2009). 이러한 영향을 최소화하기 위해서는 보다 정교한 관측이 필요하다.

농업관측의 정교화를 위한 방안 중 하나는 4차 산업혁명의 핵심기술인 빅데이터 분석을 활용하여 수급예측 시스템을 고도화하는 것일 수 있다. 실제로 농식품신유통연구원은 빅데이터 기반 채소류 수급예측시스템(BIGFOS)를 개발하였다. BIGFOS는 기존 가격, 식부면적, 작황과 같은 데이터는 물론 대형유통업체의 POS데이터, 농업기상데이터, 인터넷상의 비정형데이터, 해외시장 및 수급 동향 등을 광범위하게 수집하여 제공하고 있으며, 인공지능망 알고리즘에 의해 농산물 5개 품목(배추, 무, 고추, 양파, 마늘)의 수급 상황을 예측한다. 또한 경상남도는 자체적으로 ‘주요 농산물 가격예측 시스템’을 개발을 완료하여 2020년부터 홈페이지를 통해 서비스를 시작하였다. 경남의 주요 농산물 10개 품목(마늘, 양파, 풋고추, 깻잎, 딸기, 양상추, 부추, 시금치, 호박, 파프리카)에 대해 인공지능 빅데이터를 기반으로 실시간 가격예측 정보를 제공하고 있다.

이처럼 빅데이터를 활용한 농산물 수급 예측사업에 대한 사회적 관심이 증폭되고 있는 만큼 그에 상응하는 사회적 요구도 커지게 되었다. 아울러 빅데이터를 활용한 관측의 경제적 효과는 무엇이며, 생산자 및 소비자의 후생 증대에 얼마만큼 기여할 것인가에 대한 관심 또한 높아지고 있는 상황이다. 이러한 상황을 반영하여 본 논문에서는 재배면적 변동이 심하고 기상조건 영향을 받아 가격의 폭등락 현상이 반복되고 있는 양파를 대상 품목으로 정하고 빅데이터를 활용하여 관측을 시행했을 때와 시행하지 않았을 때를 가정하여 사회적 후생변화 분석을 통해 빅데이터를 활용한 양파 관측의 경제적 가치를 평가하고자 한다.

일반 농업관측의 경제적 효과에 대해 분석한 선행 연구들은 지속적으로 진행되고 있다. Lee 등(2002)은 농업관측의 경제적 효과를 가격안정효과, 사회후생효과, 학습효과 등으로 구분하여 분석하였다. 실제 고추와 양파를 예를 들어 가격, 재배면적 자료 등을 이용하여

농업관측정보의 실제 사회적 후생효과를 계측하였다. Kim 등(2009)은 10년간의 농산물 14개의 품목에 대한 실증자료를 바탕으로 수요/공급 탄성치를 이용하여 생산자와 소비자 후생 변화 계측을 통해 농업관측사업의 경제적 효과를 추정하였다. 특히 사전적(ex-ante) 후생효과뿐만 아니라 실제 관측정보 이용률을 감안한 사후적(ex-post) 후생효과를 계측하였다. Kim (2014)은 농업관측의 과일류 7개 품목에 대한 최대 사회후생효과를 측정하였으나, 모든 시장 참여자가 관측정보를 채택한다는 가정에 대해 실제 효과와 다소 차이가 날 수 있는 한계를 지적하였다.

한편, 빅데이터를 활용한 양파 수급 및 가격 예측 등과 관련된 선행 연구들을 살펴보면, Na 등(2017)은 마늘과 양파를 대상으로 시계열 육지부문 위성데이터와 기상데이터를 이용하여 시군단위 단수예측모형을 개발하고, 예측시기별 정확도를 산정한 후 적용성을 검토하였다. Rah 등(2018)은 양파 소비 촉진과 관련된 정형 및 비정형 빅데이터를 수집하여, 양파 소비 촉진과 관련된 예측 변수를 찾았으며, Jung (2019)은 Stepwise, LASSO (Least Absolute Shrinkage and Selection Operator), 주성분 분석(Principal Components Analysis), 부분최소제곱(Partial Least Squares) 등의 회귀모델)과 1개년도의 재배면적, 양파 도매시장 가격, 기상정보 등의 빅데이터를 이용하여 양파의 재배면적과 생산단수를 각각 추정하였다.

기존의 연구에서는 농업관측사업에 대한 사회적 후생효과를 분석하거나, 빅데이터를 이용하여 양파의 도매가격 또는 생산성을 예측하였으나, 본 연구에서는 빅데이터를 활용하여 양파 관측을 하였을 때의 경제적 효과를 계측하고자 한다.

본 연구의 목적인 빅데이터를 활용한 양파 관측의 경제성을 분석하기 위해서는 빅데이터를 활용한 양파의 예측자료가 필요하다. 예측자료는 2002년부터 2020년 양파 생육 시기별 빅데이터와 기상정보 요인의 상관분석을 통해 양파 생산단수에 영향을 주는 변수를 선택한 후, 선택된 변수를 이용하여 추정된 단위면적당 예측단수 데이터를 활용하였다. 또한 2002년부터 2020년 동안의 양파의 실측자료인 재배면적, 실측단수, 가격 정보와 앞서 설명하였던 빅데이터를 활용한 예측단수 자료를 하버거삼각형(Harberger's triangle)의 모형을 이용하여 후생효과를 계측한다. 아울러, 빅데이터 활용 유무에 따른 사회적 후생효과와의 차이를 측정하기 위하여 빅데이터를 활용하지 않은 가상의 양파의 예측단수 데이터를 고려한 사회후생효과를 계측한다. 이러한 단계별 추정 기법은 빅데이터를 활용한 양파 관측의 경제적 파급 효과 추정에 필요한 자료의 제약을 조금이나마 해소할 수 있다고 판단한다.

1) Jung (2019)에 따르면, Stepwise은 반응변수에 가장 영향이 큰 설명변수부터 하나씩 증가시키면서 적절한 모형을 찾아가는 전진선택법이다. LASSO은 모든 설명변수를 포함하는 모델에 적용, 회귀계수들에 제약조건을 주어 계수의 추정치들의 크기를 축소시키는 방법이다. 주성분 분석(PCA)은 여러 개($p \geq 2$)의 양적변수들 사이의 분산-공분산 관계를 이용하여 변수들 간의 선형결합으로 나타나는 주성분을 찾아 차원을 축소하는 방법이다. 부분최소제곱(PLS)은 반응변수를 이용하여 주성분을 만들 때 지도식 방식으로 찾아 주성분분석과 같은 차원 축소의 방법이다.

본 논문의 구성은 다음과 같다. 제2장에서는 빅데이터 활용한 예측단수 추정 모형과 하버거삼각형(Harberger's triangle)의 분석모형에 대해 설명한 후, 제3장에서 분석 결과를 제시한다. 제4장에서는 추정결과를 요약하고 결론 및 한계 등을 논의한다.

II. 분석 모형

먼저 빅데이터 분석과 관련해서는 양파의 생육 시기(경엽신장기부터 구비대기)와 기상정보(전년도 10월~금년 4월, 시점: 월) 요인을 바탕으로 양파 생산단수에 영향을 주는 변수를 선택하였다. Jung(2019)에서 제시한 빅데이터 분석방법(Stepwise 회귀분석, LASSO, PCA, PLS) 중 Stepwise 회귀분석이 경제학적으로 설명력이 높아 본 연구의 분석 모형으로 채택하였다.

Stepwise는 변수선택법 중 하나의 방법으로 반응변수에 가장 영향이 큰 설명변수부터 하나씩 증가시키면서 적절한 모형을 찾아가며, 전진선택법의 단점인 한번 들어간 변수는 나오지 않는 단점을 보완한 방법이다. Stepwise의 알고리즘은 Table 1과 같다.

Table 1. Analysis process of stepwise algorithm

<p>1) Choose the descriptive variable X_k if the explanatory power($SSR(X_k)$) is the highest and the $F(= \frac{MSR(X_k)}{MSE(X_k)})$ value is significant among the considered descriptive variables</p>
<p>2) Choose the descriptive variable X_l if the $SSR(X_l X_k)$ is the largest and the $F(= \frac{MSR(X_l X_k)}{MSE(X_l X_k)})$ value is significant except for the portion of the descriptive variable already described</p>
<p>3) $SSR(X_k X_l)$ test for significance of already existing descriptive variables except for the descriptive portion of the newly selected descriptive variable X_l and exclude descriptive variable X_k if not significant</p>
<p>4) Repeat the preceding process until no significant explanatory variable exists</p>

이러한 Stepwise 회귀분석을 이용하여 양파 생산단수에 영향을 주는 변수를 파악하기 위해 양파 생육기간을 유묘기(처서~한로), 활착기(한로~대설), 월동기(대설~소한), 경엽신장기(소한~춘분), 구비대기(춘분~소만)로 나누고(Table 2), 각 기간별 일교차(daily cross), 평균기온(average temperature), 누적강수량(accumulated precipitation), 강수일수(the number of days of precipitation), 평균상대습도(average steam pressure), 가조시간(possible duration of sunshine)의 상호작용을 고려해 기상변수는 다음과 같다(Table 3).

Table 2. Classification of onions by season and period

Period	Season
1 August ~ 23 August	Sowing
24 August ~ 8 October	Seedling
9 October ~ 7 December	Rooting
8 December ~ 5 January	Wintering
6 January ~ 20 March	Stem and leaf growing
21 March ~ 31 May	Bulb hypertrophying

Table 3. Meteorological factor

Variable name	Unit
Daily temperature range (DTR)	℃
Average temperature (AT)	℃
Daily precipitation (DP)	mm
The number of days of precipitation (NDP)	day
Average steam pressure (ASP)	Pa
Average relative humidity (ARH)	%
Possible duration of sunshine (PDS)	hr

위의 변수들을 이용하여 Stepwise 회귀분석을 실시한 결과 추정된 결과는 Table 4와 같다. Stepwise 회귀분석을 통해 선택된 변수는 평균기온과 누적강수량에 대한 변수가 많은 것으로 나타났으며, 강수량, 평균기온과 관련된 절기별 변수가 유의성이 높은 것으로 분석되었다. 결과 회귀식의 수정된 R^2 값은 0.64이다.

Table 4. Stepwise regression result

Variable	Coefficient	Variable	Coefficient
Intercept	-30184.574	DTR_Wintering	345.574*
AT_Sowing	125.281*	ASP_Wintering	460.411
DP_Sowing	0.475**	DP_Wintering	-11.159*
AT_Seedling	-532.903*	NDP_Stem and leaf growing	-83.914
DP_Seedling	-1.197*	DTR_Stem and leaf growing	-470.756
AT_Rooting	-1157.530*	AT_Stem and leaf growing	602.561**

Variable	Coefficient	Variable	Coefficient
AT_Wintering	177.059*	AT_Bulb hypertrophying	20.470*
ARH_Wintering	-49.699*	DP_Bulb hypertrophying	5.418*
PDS_Wintering	5034.828	NDP_Bulb hypertrophying	56.618
Variable for intersection term	AT_Seedling × AT_Rooting		43.100*
	DP_Seedling × AT_Wintering		-1.193*
	DP_Sowing × AT_Stem and leaf growing		-0.263*
	AT_Rooting × AT_Stem and leaf growing		-39.370
	DP_Wintering × AT_Stem and leaf growing		-2.108*
	DP_Seedling × AT_Sowing		0.097*
	DP_Bulb hypertrophying × AT_Bulb hypertrophying		-0.360*
F-ratio	9.350***	Adjusted R ²	0.638

Note: Single, double, and triple asterisks (*, **, ***) denotes that coefficient is significant at the 10%, 5% and 1% significance levels, respectively.

다음으로 사회적 후생효과는 빅데이터를 활용한 예측정보에 의해 과잉생산 또는 과소생산이 모두 해소되어 균형 생산이 달성되었을 때의 사회적 후생 증대 효과를 말한다. Fig. 1은 과소생산이 예상될 때 빅데이터를 이용한 예측치의 소비자 및 생산자 후생효과를 도식화하고 있다.

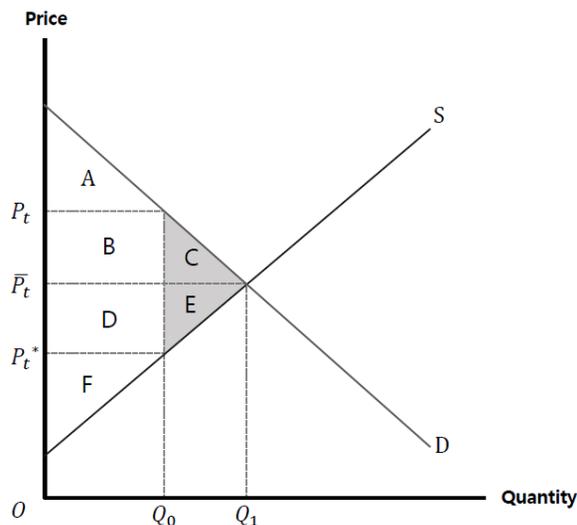


Fig. 1. Effect of resource allocation of predictions using big data.

수요곡선과 공급곡선을 각각 D 와 S 이다. 빅데이터로 예측한 가격은 P_t^* 이며, 생산량은 Q_0 이고, 그때의 시장가격은 P_t 가 된다. 만약 정확한 정보를 제공하여 예측가격이 \overline{P}_t 라면 시장가격은 \overline{P}_t 이고 생산량은 Q_1 이 될 것이다.

소비자잉여는 균형가격일 때 $A+B+C$ 에서 예측가격이 균형가격보다 낮을 경우 A 로 줄어든다. 생산자잉여는 균형가격일 때 $D+E+F$ 이며, 예측가격이 균형가격보다 낮을 경우 $B+D+F$ 가 된다.

소비자와 생산자를 합친 총 잉여는 $A+B+C+D+E+F$ 에서 $A+B+D+F$ 로 나타난다. 따라서 균형점에 도달할 수 있는 정보의 최대가치는 $C+E$ 에 해당한다.

$$\begin{aligned}
 \text{최대가치} &= \text{균형가격 사회후생} - \text{예측가격이 균형가격보다 낮을 경우의 사회후생} \\
 &= (\text{균형가격 소비자잉여} + \text{균형가격 생산자잉여}) - (\text{예측가격이 균형가격} \\
 &\quad \text{보다 낮을 경우의 소비자잉여} + \text{예측가격이 균형가격보다 낮을 경우} \\
 &\quad \text{의 생산자잉여}) \\
 &= (A+B+C+D+E+F) - (A+B+D+F) \\
 &= C+E
 \end{aligned} \tag{1}$$

최대가치는 하버거삼각형(Harberger's triangle)이라고 불리기도 하는 사회후생효과는 삼각형 $C+E$ 의 면적인데 이를 수식으로 나타내면 다음과 같다.

$$\begin{aligned}
 \text{사회후생효과} &= \frac{1}{2} |\Delta P| |\Delta Q| \\
 &= \frac{1}{2} [(P_t - \overline{P}_t) + (\overline{P}_t - P_t^*)] \frac{\Delta Q}{Q} Q
 \end{aligned} \tag{2}$$

수요의 가격탄력성을 ε_d , 공급의 가격탄력성을 ε_s 라 놓으면, 예측가격이 균형가격보다 낮을 경우 수요가격 변동분과 공급가격 변동분이 각각 식 (3a)와 식 (3b)로 도출된다.

$$P_t - \overline{P}_t = \Delta Q \times |\text{수요함수의 기울기}| = \frac{1}{\varepsilon_d} \frac{\Delta Q}{Q} P \tag{3a}$$

$$\overline{P}_t - P_t^* = \Delta Q \times |\text{공급함수의 기울기}| = \frac{1}{\varepsilon_s} \frac{\Delta Q}{Q} P \tag{3b}$$

식 (3a)의 좌변과 식 (3b)의 좌변을 더하면 $P_t - P_t^* = (P_t - \overline{P}_t) + (\overline{P}_t - P_t^*)$ 가 되는데,

이는 예측가격이 균형가격보다 낮을 경우의 가격상승분(f)이므로 개념상 fP 와 똑같아야 한다.

$$P_t - P_t^* = fP = \left(\frac{1}{\varepsilon_d} + \frac{1}{\varepsilon_s}\right) \frac{\Delta Q}{Q} P$$

$$\frac{\Delta Q}{Q} = \frac{f}{\frac{1}{\varepsilon_d} + \frac{1}{\varepsilon_s}} \quad (4)$$

식 (4)를 식 (3a)와 식 (3b)에 각각 대입하면 다음 두 식을 얻는다.

$$P_t - \bar{P}_t = \frac{\frac{1}{\varepsilon_d}}{\frac{1}{\varepsilon_d} + \frac{1}{\varepsilon_s}} fP \quad (5a)$$

$$\bar{P}_t - P_t^* = \frac{\frac{1}{\varepsilon_s}}{\frac{1}{\varepsilon_d} + \frac{1}{\varepsilon_s}} fP \quad (5b)$$

이제 식 (5a)와 식 (5b)를 식 (2)에 대입하면, 사회후생효과를 아래 식 (6)과 같이 가격변동률 f , 수요의 가격탄력성 ε_d , 공급의 가격탄력성 ε_s , 그리고 거래규모 PQ 로 나타낼 수 있다.

$$\begin{aligned} \text{사회후생효과} &= \frac{1}{2} f^2 \frac{1}{\frac{1}{\varepsilon_d} + \frac{1}{\varepsilon_s}} PQ \\ &= \frac{1}{2} f^2 \frac{\varepsilon_d \varepsilon_s}{\varepsilon_d + \varepsilon_s} PQ \end{aligned} \quad (6)$$

Fig. 1에서 논의된 빅데이터를 활용한 관측정보의 사회적 후생효과는 과소생산이 관측정보에 의하여 균형점에 도달했을 때 발생할 수 있는 사회적 후생효과의 최대치를 의미한다 (Kim et al., 2009). 그러나 빅데이터를 활용하지 않아 관측정보의 정확도가 떨어진다면 빅데이터 활용한 예측가격보다 낮아져, 사회적 후생효과의 최대치 역시 영향을 받게 될 것이다.

Ⅲ. 분석 결과

전국을 대상으로 2002년부터 2020년의 양파 생육 시기 빅데이터를 절기별(파종기, 유포기, 활착기, 월동기, 경엽신장기, 구비대기)로 분석하여 상관분석을 통해 재배 시기별로 양파 생산량에 요인을 주는 환경변수(일교차, 평균기온, 누적강수량, 강우일수, 평균증기압, 평균상대습도, 가조시간)를 추정하였다.

양파 수확량이 파종기에는 기온이 예년보다 높고 강수량은 적을수록, 유포기에는 기온이 낮고 강수량이 적을수록, 활착기에는 기온은 낮을수록, 월동기에는 기온이 높고 상대습도는 적을수록, 가조시간은 길고 일교차는 클수록, 증기압은 크고 강수량은 적을수록, 경엽신장기에는 일교차와 비 온 날은 적고 기온이 높을수록, 구비대기에는 기온이 크고 강수량이 많고 비 온 날이 많을수록 각각 많아지는 것으로 나타났다. 특히 상호작용으로 유포기의 강수량이 148.4 mm보다 작을 때는 월동기 평균기온이 평년보다 높을수록 생산량이 많아지고, 구비대기의 기온이 15.05℃보다 낮은 경우는 강수량은 예년보다 많을수록 생산량이 늘어난 것으로 분석되었다.

추정된 변수는 단계적 회귀분석을 통해 단위면적당 예측단수를 계측하였다(Table 4). 빅데이터를 활용한 양파의 예측단수는 실측단수와 비교하였을 때 편차는 -9.0%~4.2%의 차이를 보이고 있으며, 특히 2020년을 제외하면 -3.3%~4.2%의 근소한 편차라는 것을 알 수 있다. 2002년부터 2020년의 예측단수 데이터와 전국 양파 재배면적, 실제 양파가격을 이용하여 예측한 생산액을 기준으로 빅데이터를 활용한 양파 관측의 사회적 후생효과 분석 결과를 제시한다.

Table 4. Deviation of the actual onion yield and the predicted onion yield

Year	The actual onion yield (kg/10a)	The predicted onion yield (kg/10a)	Deviation
2002	6,093	6,068	-0.4%
2003	6,033	5,965	-1.1%
2004	6,090	6,160	1.1%
2005	6,114	6,032	-1.3%
2006	5,809	5,881	1.2%
2007	6,836	6,611	-3.3%
2008	6,725	6,865	2.1%
2009	7,412	7,465	0.7%

Year	The actual onion yield (kg/10a)	The predicted onion yield (kg/10a)	Deviation
2010	6,384	6,501	1.8%
2011	6,616	6,483	-2.0%
2012	5,703	5,925	3.9%
2013	6,458	6,651	3.0%
2014	6,649	6,865	3.2%
2015	6,072	6,133	1.0%
2016	6,528	6,760	3.6%
2017	5,858	5,965	1.8%
2018	5,756	6,000	4.2%
2019	7,322	7,618	4.0%
2020	7,962	7,247	-9.0%

Source: Statistics Korea, Crop Production Survey (2002-2020).

Table 5는 빅데이터를 활용한 양파 예측치를 통해 2002년부터 2020년까지의 연도별 사회 후생효과를 제시하고 있다. 본 연구에서는 빅데이터를 활용하지 않았을 때의 최대 사회 후생효과를 추정하기 위해서 양파 단수 2002년부터 2020년까지의 과거 3개년 산술평균을 예측단수로 설정하여 사회후생효과를 계측하였다. 이는 빅데이터를 활용하지 않고 과거의 단수를 바탕으로 산술평균한 값으로 빅데이터를 활용하였을 때와 그렇지 않았을 때의 경제적 파급 효과를 계측하기 적당하다고 판단되기 때문이다.

Table 5. Estimation of maximum social welfare effect of predicted onion yield using big data

Year	The actual onion yield (kg/10a)	The amount of production (Thousand won)	Price elasticity of demand	Price elasticity of supply	Volatility (%)	Social welfare effects (Thousand won)
2002	6,068	371,340,031	0.45	0.36	-0.006	1,199
2003	5,965	718,966,861	0.45	0.36	1.442	149,473,648
2004	6,160	603,289,838	0.45	0.36	-0.355	7,607,342
2005	6,032	611,177,096	0.45	0.36	-0.038	88,2423
2006	5,881	753,461,857	0.45	0.36	0.382	10,986,928

Year	The actual onion yield (kg/10a)	The amount of production (Thousand won)	Price elasticity of demand	Price elasticity of supply	Volatility (%)	Social welfare effects (Thousand won)
2007	6,611	630,667,491	0.45	0.36	-0.358	8,064,045
2008	6,865	806,751,802	0.45	0.36	0.421	14,276,692
2009	7,465	1,133,233,413	0.45	0.36	0.074	619,757
2010	6,501	1,360,122,592	0.45	0.36	0.154	3,220,673
2011	6,483	1,185,785,954	0.45	0.36	-0.159	2,982,493
2012	5,925	1,124,970,649	0.45	0.36	0.138	2,131,009
2013	6,651	1,725,078,238	0.45	0.36	0.429	31,807,295
2014	6,865	929,534,030	0.45	0.36	-0.563	29,417,555
2015	6,133	1,219,401,897	0.45	0.36	0.949	109,820,270
2016	6,760	1,391,620,201	0.45	0.36	-0.063	543,701
2017	5,965	1,382,213,856	0.45	0.36	0.146	2,956,115
2018	6,000	1,236,731,587	0.45	0.36	-0.342	14,491,122
2019	7,618	984,999,584	0.45	0.36	-0.239	5,617,907
2020	7,247	1,061,361,933	0.45	0.36	0.681	49,233,907
Average	-	1,012,142,574	-	-	-	23,333,679

Note: 1. The price change rate is the average value of the annual price change rate for the past three years including the current year. The price fluctuation rate was based on the wholesale price of onions of the Seoul Agricultural and Fisheries Corporation.

2. For the price elasticity of demand and the price elasticity of supply, the price elasticity of demand and the price elasticity of cultivation area presented in Lee et al. (2002), 「Analysis of Economic Effects of Agricultural Observation」 were used.

Source: Statistics Korea, Crop Production Survey (2002-2020).

Table 6는 빅데이터를 활용하지 않았을 경우의 2002년부터 2020년까지 연도별 사회후생효과를 제시하고 있다. 빅데이터를 활용하지 않았을 경우의 사회후생효과의 평균값은 224억 원으로 계측되었다. 총 사회후생효과는 2003년의 경우 약 1,418억 원으로 가장 높게 추정되었으며, 2002년도는 122만 원으로 가장 낮게 추정되었다. 이러한 사회적 후생효과의 격차는 주로 가격변동률에 기인하는 것으로 보인다. 가격 변동 폭이 높을수록 사회적 후생효과를 증대시키는 것으로 분석되며, 이는 가격 불안정성이 높을수록 관측정보의 경제적 가치가 높게 평가됨을 의미한다(Kim et al., 2009).

Table 6. Estimating the maximum social welfare effect when big data is not used

Year	The actual onion yield (kg/10a)	The amount of production (Thousand won)	Price elasticity of demand	Price elasticity of supply	Volatility (%)	Social welfare effects (Thousand won)
2002	5,562	340,374,630	0.45	0.36	-0.006	1,225
2003	5,659	682,124,581	0.45	0.36	1.442	141,838,530
2004	5,926	580,405,305	0.45	0.36	-0.355	7,314,558
2005	6,072	615,229,995	0.45	0.36	-0.038	88,839
2006	6,079	778,829,218	0.45	0.36	0.382	11,364,987
2007	6,004	572,793,502	0.45	0.36	-0.358	7,341,151
2008	6,253	734,831,613	0.45	0.36	0.421	13,024,229
2009	6,457	980,162,143	0.45	0.36	0.074	536,737
2010	6,991	1,462,639,139	0.45	0.36	0.154	3,468,795
2011	6,840	1,251,144,715	0.45	0.36	-0.159	3,163,019
2012	6,804	1,291,865,029	0.45	0.36	0.138	2,460,228
2013	6,234	1,617,006,881	0.45	0.36	0.429	29,759,556
2014	6,259	847,480,480	0.45	0.36	-0.563	26,862,504
2015	6,270	1,246,641,104	0.45	0.36	0.949	112,272,623
2016	6,393	1,316,069,223	0.45	0.36	-0.063	522,348
2017	6,416	1,486,797,123	0.45	0.36	0.146	3,169,257
2018	6,153	1,268,199,535	0.45	0.36	-0.342	14,833,369
2019	6,047	781,913,995	0.45	0.36	-0.239	4,466,371
2020	6,312	924,426,179	0.45	0.36	0.681	42,871,281
Average	-	988,364,968	-	-	-	22,387,348

Note: 1. The price change rate is the average value of the annual price change rate for the past three years including the current year. The price fluctuation rate was based on the wholesale price of onions of the Seoul Agricultural and Fisheries Corporation.

2. For the price elasticity of demand and the price elasticity of supply, the price elasticity of demand and the price elasticity of cultivation area presented in Lee et al.(2002), 「Analysis of Economic Effects of Agricultural Observation」 were used.

Source: Statistics Korea, Crop Production Survey (2002-2020).

마지막으로 빅데이터를 활용한 예측치를 이용했을 때와 그렇지 않았을 때의 사회후생효과를 비교해보았다(Table 7). 가장 큰 차이가 나타난 연도는 2003년으로 차이 값은 약 76억 원이며, 가장 적게 나타난 해 2002년은 25,692원으로 근소한 차이를 보였다. 결론적으로 2002년부터 2020년까지의 평균 사회후생효과 차이를 살펴보면 약 9억 5천만 원 수준인 것으로 추정되었다.

Table 7. Comparison of social welfare effects according to the use of big data

unit: Thousand won

Year	Social welfare effects		Gap (A-B)
	With big data (A)	Without big data (B)	
2002	1,199	1,225,	-26
2003	149,473,648	141,838,530	7,635,118
2004	7,607,342	7,314,558	292,784
2005	88,243	88,839	-596
2006	10,986,928	11,364,987	-378,059
2007	8,064,045	7,341,151	722,894
2008	14,276,692	13,024,229	1,252,463
2009	619,757	536,737	83,020
2010	3,220,673	3,468,795	-248,122
2011	2,982,493	3,163,019	-180,526
2012	2,131,009	2,460,228	-329,219
2013	31,807,295	29,759,556	2,047,739
2014	29,417,555	26,862,504	2,555,051
2015	109,820,270	112,272,623	-2,452,352
2016	543,701	522,348	21,353
2017	2,956,115	3,169,257	-213,142
2018	14,491,122	14,833,369	-342,247
2019	5,617,907	4,466,371	1,151,536
2020	49,233,907	42,871,281	6,362,626
Average	23,333,679	22,387,348	946,331

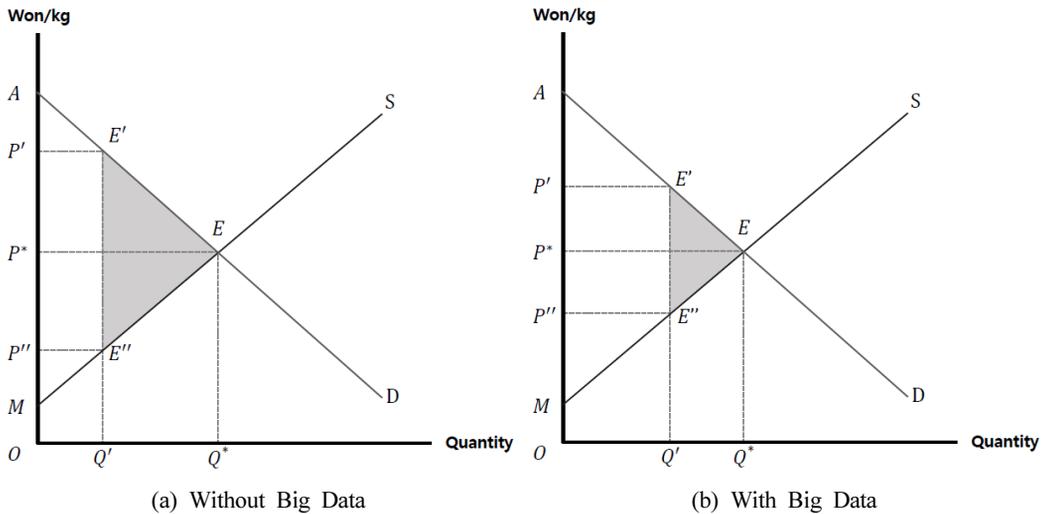


Fig. 2. Comparison of social welfare effects according to the use of big data.

IV. 요약 및 결론

정부는 농업관측사업을 통해 농산물의 수급 안정과 농업인의 농업소득 안정화를 위해서 주요 농산물에 대하여 관측정보를 제공함에도 불구하고 재배면적 변동이 심하고 이상기후로 인한 기상 조건의 영향으로 여전히 가격의 불안정성의 문제를 해결하는 데 어려움을 겪고 있다. 가격 불안정성을 최소화하기 위해서는 보다 정교한 관측이 필요하며, 이를 위한 방안 중 하나는 빅데이터를 활용하여 수급예측 시스템을 고도화하는 것이다. 농식품신용연구원에서 개발한 ‘BIGFOS’나 경상북도에서 개발한 ‘주요 농산물 가격예측 시스템’ 같이 이미 빅데이터 기술을 이용한 관측 시스템이 운영 중이거나 시범 운영을 하고 있다. 하지만 빅데이터를 활용한 농산물 수급 예측사업에 대한 사회적 관심만큼 요구도 커지게 되었다. 이에 따라 빅데이터를 활용한 관측의 경제적 파급 효과에 대한 객관적 평가가 필요함에도 불구하고 빅데이터를 활용한 관측의 경제적 가치 평가에 대한 연구는 이루어지지 못하였다.

이에 본 논문에서는 가격의 폭등락 반복이 심한 양파를 대상 품목으로 정해 빅데이터를 활용하여 관측을 시행했을 때와 그렇지 않았을 때를 가정하여 사회적 후생효과를 계측한 후 비교하였다. 사회적 후생효과를 계측하기 위해서 2002년부터 2020년 동안의 양파의 실측자료인 재배면적, 실측단수, 가격 정보, 빅데이터를 활용한 예측단수, 빅데이터를 활용하지 않았을 경우의 예측단수 등을 이용하여 하버거삼각형(Harberger's triangle)의 모형을 통해 사회후생효과를 추정하였다. 하버거삼각형(Harberger's triangle)의 모형은 수요와 공급 탄

성치를 이용하여 소비자와 생산자의 후생 변화를 계측하는 방법으로 빅데이터를 활용한 관측을 통해 최적 균형에 도달했을 때의 사회적 후생효과를 나타낸다.

빅데이터를 활용한 2002년부터 2020년 동안의 양과 예측단수는 실측단수와 비교하여 편차 -9.0%~4.2%를 보였다. 이를 이용하여 사회후생효과를 추정한 결과, 사회적 후생효과와 연도별 평균값은 233억 원으로 평균 생산액의 약 2% 수준에 해당하는 것으로 나타났다. 또한 빅데이터를 활용하였을 때와 그렇지 않았을 때의 경제적 파급 효과를 계측하기 위해, 2002년부터 2020년까지의 과거 3개년 양과 단수 산술평균을 예측단수 데이터로 이용하여 사회후생효과를 추정하였다. 빅데이터를 사용하지 않았을 경우의 2002년부터 2020년까지 사회후생효과와 연도별 평균값은 224억 원으로 계측되었다. 특히 관측정보의 연도별 사회적 후생효과는 가격 변동률의 증감 폭이 많은 영향을 받는 것으로 분석되었다. 마지막으로, 빅데이터를 활용한 예측치를 이용했을 때와 그렇지 않았을 때의 사회후생효과와 차이는 약 9억 5천만 원 수준인 것으로 추정되었다.

이처럼 본 연구는 빅데이터를 활용한 예측단수 데이터를 바탕으로 양과 관측의 사회적 후생효과를 분석하였다는 점에서 기존 연구와 차별성이 있다. 분석 결과 빅데이터를 활용하였을 때와 활용하지 않았을 때의 사회후생효과 차이는 약 9억 5천만 원 수준으로 나타난 만큼 양과 품목만이 아닌 다른 품목에도 적용하였을 때 그 효과는 크다고 할 수 있다. 이러한 결과는 농업관측정보 제공이 농업에만 국한되어 있는 것이 아니라 사회 전체적으로 후생효과가 있음을 알 수 있으며, 결과적으로 농업관측사업의 당위성 입증에 위한 기초 자료로도 활용이 가능할 것으로 판단된다.

다만 기존 연구와 본 연구에서 사용한 단순 거미집 이론은 현실을 지나치게 단순화하여 생산자와 최종 소비자 사이에 유통, 가공업자 등이 개입하는 보다 현실적인 상황을 고려하지 못했다는 한계점을 지니고 있다. 이에 추후 연구에서는 보다 현실적인 합리적 기대 가설 하에서의 농업관측효과 계측 등 다양한 방법을 통한 경제적 가치 평가가 이루어질 수 있도록 후속 연구가 지속적으로 필요할 것으로 생각된다.

[Submitted, February. 23, 2021 ; Revised, April. 9, 2021 ; Accepted, June. 3, 2021]

References

1. Antonovitz, F. and T. Roe. 1984. The Value of Rational Expectations Forecast in a Risky Market: Theoretical and Empirical Approach. *American J. Agricultural Economics* 66: 717-723.

2. Freebairn, J. W. 1976. The Value and Distribution of the Benefits of Commodity Price Outlook Information. *The Economic Record*. 52: 199-212.
3. Hayami, Y. and W. Peterson. 1972. Social Returns to Public Information Services: Statistical Reporting of US Farm Commodities. *The American Economic Review* 62: 119-130.
4. Jung, I. C. 2019. Comparison of the Prediction Model of the Production Unit and the Area of the Onion: Comparison of Stepwise, LASSO, PCA, and PLS methods, Chonnam National Univ. Graduate School.
5. Kim, B. S. 2014. Measuring the Maximum Social Welfare Effects of Outlook Information Service for Fruits Sector in Korea. *Korean J. Academia-Industrial Cooperation Society*, 15(11): 6646-6651.
6. Kim, K. S., D. W. An, and J. H. Sung. 2009. Estimating the Social Welfare Effects of Agricultural Outlook Information. *Korean J. Agricultural Economics*. 32(5): 1-16.
7. Lee, Y. S., Y. J. Kim, and Y. H. Kim. 2002. An Economic Analysis of Agricultural Outlook in Korea, *Korean J. Agricultural Economics*. 25(3): 1-15.
8. Na, S. I., C. W. Park, K. H. So, J. M. Park, and K. D. Lee. 2017. Development of Garlic & Onion Yield Prediction Model on Major Cultivation Regions Considering MODIS NDVI and Meteorological Elements. *Korean J. Remote Sensing*. 33(5_2): 647-659.
9. Nieuwoudt, W. L., A. W. Womack, and S. R. Johnson. 1988. Measurement of Importance of Risk on Supply Response of Corn and Soybeans. *North Central Journal of Agricultural Economics*. 10: 281-292.
10. Rah, H. C., E. H. Oh, D. I. Yoo, W. S. Cho, N. Aziz, S. H. Park, Y. B. Cho, and K. H. Yoo. 2018. Prediction of Onion Purchase Using Structured and Unstructured Big Data. *Korean J. Contents Association*. 18(11): 30-37.
11. Statistics Korea. Crop Production Survey.