Title: "*Agricultural Economics* appendix for Weather Risk: How does it change the yield benefits of nitrogen fertilizer and improved maize varieties in sub-Saharan Africa?"

Authors: Terrance Hurley[1], Jawoo Koo[2] and Kindie Tesfaye[3]

Date: April 27, 2018

Note: The material contained herein is supplementary to the article named in the title and published in the *Agricultural Economics*.

[1]Department of Applied Economics, University of Minnesota, tmh@umn.edu, +1-612-625-1238. [2]Environment and Production Technology Division, International Food Policy Research Institute (IFPRI), j.koo@cgiar.org. [3]Socioeconomics Program, International Maize and Wheat Improvement Center (CIMMYT), K.TesfayeFantaye@cgiar.org.

This appendix and its associated article was prepared for the HarvestChoice project with support from the Bill and Melinda Gates Foundation, the University of Minnesota, the International Food Policy Research Institute, the International Maize and Wheat Improvement Center, and Agricultural Experimentation Station Project No. MIN-14-034.

Correspondence should be addressed to tmh@umn.edu.

Appendix A: Derivation of Willingness to Pay Bounds

The purpose of this appendix is to provide the theoretical motivation for the upper and lower willingness to pay (WTP) bounds used to compare traditional and improved seed varieties with and without nitrogen fertilizer. We begin with some notation. Let $\boldsymbol{\varepsilon}$ be a vector of random factors that may jointly influence a farmer's yield and maize price. The joint density of this random vector is $\phi(\boldsymbol{\varepsilon}) \geq 0$. Let $f^k(y|\boldsymbol{\varepsilon}) \geq 0$ for $\in [y^L, y^U]$ and $f^k(y|\boldsymbol{\varepsilon}) = 0$ otherwise be the yield probability density function for a farmer's kth production alternative conditional on the random vector $\boldsymbol{\varepsilon}$. Let $h(p|\boldsymbol{\varepsilon}) \geq 0$ for $p > 0$ and $h(p|\boldsymbol{\varepsilon}) = 0$ otherwise be the maize price probability density function conditional on the random vector $\boldsymbol{\varepsilon}$. This conditioning of yields and price on $\boldsymbol{\varepsilon}$ implies the two maybe indirectly correlated through some random exogenous shock such as market supply shocks.

Consider two production alternatives denoted by $k \in \{r, g\}$. Assume farmers exhibit risk averse, expected utility preferences such that alternative g is weakly preferred to r when

$$EU^g = \int_{\boldsymbol{\varepsilon}} \int_{p>0} \int_{y^L}^{y^U} U\big((py - c^g)A\big) f^g(y|\boldsymbol{\varepsilon})\, dy h(p|\boldsymbol{\varepsilon}) dp\, \phi(\boldsymbol{\varepsilon}) d\boldsymbol{\varepsilon} \geq \qquad \text{A1}$$

$$EU^r = \int_{\boldsymbol{\varepsilon}} \int_{p>0} \int_{y^L}^{y^U} U\big((py - c^r)A\big) f^r(y|\boldsymbol{\varepsilon})\, dy h(p|\boldsymbol{\varepsilon}) dp\, \phi(\boldsymbol{\varepsilon}) d\boldsymbol{\varepsilon}$$

where $A > 0$ represents a constant number of hectares planted to maize and $c^k \geq 0$ is a constant per hectare production cost for alternative k such that $(py - c^k)A$ is the the random net return to maize production for alternative k; and $U(\cdot)$ is a continuous and twice differentiable utility function such that $U'(\cdot) > 0$ and $U''(\cdot) < 0$, signifying risk averse preferences.

Building on the intuition presented in Figure 2 (b), consider the question: What is the most a farmer would be willing to pay to switch from alternative r to g? That is, how big can c^g get relative to c^f with equation A1 still being satisfied? Sufficiency conditions for the answer to this question based only on yield distributions can be obtained by evaluating

$$wtp = \max_w \Big\{w: \int_{\boldsymbol{\varepsilon}} \int_{p>0} \int_{y^L}^{y^U} U\big((p(y-w) - c^r)A\big) f^g(y|\boldsymbol{\varepsilon})\, dy h(p|\boldsymbol{\varepsilon}) dp\, \phi(\boldsymbol{\varepsilon}) d\boldsymbol{\varepsilon} \geq \qquad \text{A2}$$

$$\int_{\boldsymbol{\varepsilon}} \int_{p>0} \int_{y^L}^{y^U} U\big((py - c^r)A\big) f^r(y|\boldsymbol{\varepsilon})\, dy h(p|\boldsymbol{\varepsilon}) dp\, \phi(\boldsymbol{\varepsilon}) d\boldsymbol{\varepsilon}\Big\},$$

using second order stochastic dominance (Rothschild and Stiglitz, 1970 and 1971). To understand how, first transform the distribution of $f^g(y|\boldsymbol{\varepsilon})$ by defining $y = y^o + w$ and the distribution $f^r(y|\boldsymbol{\varepsilon})$ by defining $y = y^o$, which yields $f^{g^o}(y^o|\boldsymbol{\varepsilon}) = f^g(y^o + w|\boldsymbol{\varepsilon}) \geq 0$ for $y^o \in [y^L - w, y^U - w]$ and $f^{g^o}(y^o|\boldsymbol{\varepsilon}) = 0$ otherwise; and $f^{r^o}(y^o|\boldsymbol{\varepsilon}) = f^r(y^o|\boldsymbol{\varepsilon})$. Let

$y^{L^o} = \min\{y^L - w, y^L\}$, $y^{U^o} = \max\{y^U - w, y^U\}$, and rewrite the comparison in equation A2 based on these transformed distributions:

$$\int_{\boldsymbol{\varepsilon}} \int_{p>0} \int_{y^{L^o}}^{y^{U^o}} U\big((py^o - c^r)A\big) f^{g^o}(y^o|\boldsymbol{\varepsilon})\, dy^o h(p|\boldsymbol{\varepsilon})dp\, \phi(\boldsymbol{\varepsilon})d\boldsymbol{\varepsilon} \geq \qquad \text{A3}$$

$$\int_{\boldsymbol{\varepsilon}} \int_{p>0} \int_{y^{L^o}}^{y^{U^o}} U\big((py^o - c^r)A\big) f^{r^o}(y^o|\boldsymbol{\varepsilon})\, dy^o h(p|\boldsymbol{\varepsilon})dp\, \phi(\boldsymbol{\varepsilon})d\boldsymbol{\varepsilon}.$$

Integrating equation A3 by parts yields

$$\int_{\boldsymbol{\varepsilon}} \int_{p>0} U\left(\left(py^{U^o} - c^r\right)A\right) F^{g^o}\left(y^{U^o}|\boldsymbol{\varepsilon}\right) h(p|\boldsymbol{\varepsilon})dp\, \phi(\boldsymbol{\varepsilon})d\boldsymbol{\varepsilon} \qquad \text{A4}$$

$$-\int_{\boldsymbol{\varepsilon}} \int_{p>0} U\left(\left(py^{L^o} - c^r\right)A\right) F^{g^o}\left(y^{L^o}|\boldsymbol{\varepsilon}\right) h(p|\boldsymbol{\varepsilon})dp\, \phi(\boldsymbol{\varepsilon})d\boldsymbol{\varepsilon}$$

$$+\int_{\boldsymbol{\varepsilon}} \int_{p>0} \int_{y^{L^o}}^{y^{U^o}} pAU'\big((py^o - c^r)A\big) F^{g^o}(y^o|\boldsymbol{\varepsilon})\, dy^o h(p|\boldsymbol{\varepsilon})dp\, \phi(\boldsymbol{\varepsilon})d\boldsymbol{\varepsilon} \geq$$

$$\int_{\boldsymbol{\varepsilon}} \int_{p>0} U\left(\left(py^{U^o} - c^r\right)A\right) F^{r^o}\left(y^{U^o}|\boldsymbol{\varepsilon}\right) dp\, \phi(\boldsymbol{\varepsilon})d\boldsymbol{\varepsilon}$$

$$-\int_{\boldsymbol{\varepsilon}} \int_{p>0} U\left(\left(py^{L^o} - c^r\right)A\right) F^{r^o}\left(y^{L^o}|\boldsymbol{\varepsilon}\right) dp\, \phi(\boldsymbol{\varepsilon})d\boldsymbol{\varepsilon}$$

$$+\int_{\boldsymbol{\varepsilon}} \int_{p>0} \int_{y^{L^o}}^{y^{U^o}} pAU'\big((py^o - c^r)A\big) F^{r^o}(y^o|\boldsymbol{\varepsilon})\, dy^o h(p|\boldsymbol{\varepsilon})dp\, \phi(\boldsymbol{\varepsilon})d\boldsymbol{\varepsilon}$$

or

$$-\int_{\boldsymbol{\varepsilon}} \int_{p>0} \int_{y^{L^o}}^{y^{U^o}} pAU'\big((py^o - c^r)A\big) F^{g^o}(y^o|\boldsymbol{\varepsilon})\, dy^o h(p|\boldsymbol{\varepsilon})dp\, \phi(\boldsymbol{\varepsilon})d\boldsymbol{\varepsilon} \geq$$

$$-\int_{\boldsymbol{\varepsilon}} \int_{p>0} \int_{y^{L^o}}^{y^{U^o}} pAU'\big((py^o - c^r)A\big) F^{r^o}(y^o|\boldsymbol{\varepsilon})\, dy^o h(p|\boldsymbol{\varepsilon})dp\, \phi(\boldsymbol{\varepsilon})d\boldsymbol{\varepsilon}$$

where $F^{g^o}(y^o|\boldsymbol{\varepsilon})$ and $F^{r^o}(y^o|\boldsymbol{\varepsilon})$ are the cumulative distributions for $f^{g^o}(y^o|\boldsymbol{\varepsilon})$ and $f^{r^o}(y^o|\boldsymbol{\varepsilon})$. Integrating equation A4 by parts yields

$$-\int_{\boldsymbol{\varepsilon}} \int_{p>0} pAU'\left(\left(py^{U^o} - c^r\right)A\right) \int_{y^{L^o}}^{y^{U^o}} F^{g^o}(z|\boldsymbol{\varepsilon})dz\, h(p|\boldsymbol{\varepsilon})dp\, \phi(\boldsymbol{\varepsilon})d\boldsymbol{\varepsilon} \qquad \text{A5}$$

$$+\int_{\boldsymbol{\varepsilon}} \int_{p>0} pAU'\left(\left(py^{L^o} - c^r\right)A\right) \int_{y^{L^o}}^{y^{L^o}} F^{g^o}(z|\boldsymbol{\varepsilon})dz\, h(p|\boldsymbol{\varepsilon})dp\, \phi(\boldsymbol{\varepsilon})d\boldsymbol{\varepsilon}$$

$$+\int_{\boldsymbol{\varepsilon}} \int_{p>0} \int_{y^{L^o}}^{y^{U^o}} (pA)^2 U''\big((py^o - c^r)A\big) \int_{y^{L^o}}^{y^o} F^{g^o}(z|\boldsymbol{\varepsilon})dz\, dy^o h(p|\boldsymbol{\varepsilon})dp\, \phi(\boldsymbol{\varepsilon})d\boldsymbol{\varepsilon} \geq$$

$$-\int_{\boldsymbol{\varepsilon}} \int_{p>0} pAU'\left(\left(py^{U^o} - c^r\right)A\right) \int_{y^{L^o}}^{y^{U^o}} F^{r^o}(z|\boldsymbol{\varepsilon})\, dz h(p|\boldsymbol{\varepsilon})dp\, \phi(\boldsymbol{\varepsilon})d\boldsymbol{\varepsilon}$$

$$+\int_{\boldsymbol{\varepsilon}} \int_{p>0} pAU'\left(\left(py^{L^o} - c^r\right)A\right) \int_{y^{L^o}}^{y^{L^o}} F^{r^o}(z|\boldsymbol{\varepsilon})\, dz h(p|\boldsymbol{\varepsilon})dp\, \phi(\boldsymbol{\varepsilon})d\boldsymbol{\varepsilon}$$

$$+\int_{\boldsymbol{\varepsilon}} \int_{p>0} \int_{y^{L^o}}^{y^{U^o}} (pA)^2 U''\big((py^o - c^r)A\big) \int_{y^{L^o}}^{y^o} F^{r^o}(z|\boldsymbol{\varepsilon})\, dz\, dy^o h(p|\boldsymbol{\varepsilon})dp\, \phi(\boldsymbol{\varepsilon})d\boldsymbol{\varepsilon}$$

or

$$\int_{\boldsymbol{\varepsilon}} \left(\int_{p>0} pAU'\left(\left(py^{U^o} - c^r\right)A\right) h(p|\boldsymbol{\varepsilon})dp\right)\left(\int_{y^{L^o}}^{y^{U^o}} [F^{r^o}(z|\boldsymbol{\varepsilon}) - F^{g^o}(z|\boldsymbol{\varepsilon})]dz\right)\phi(\boldsymbol{\varepsilon})d\boldsymbol{\varepsilon}$$

$$-\int_{\boldsymbol{\varepsilon}} \int_{y^{L^o}}^{y^{U^o}} \left(\int_{p>0}(pA)^2U''\left((py^o - c^r)A\right)h(p|\boldsymbol{\varepsilon})dp\right) \times$$

$$\left(\int_{y^{L^o}}^{y^o}[F^{r^o}(z|\boldsymbol{\varepsilon}) - F^{g^o}(z|\boldsymbol{\varepsilon})]dz\right) dy^o\phi(\boldsymbol{\varepsilon})d\boldsymbol{\varepsilon} \geq 0$$

where *z* is a variable of integration. Sufficient conditions for equation A3 to be true are then $\int_{y^{L^o}}^{y^o} F^{r^o}(z|\boldsymbol{\varepsilon})dz \geq \int_{y^{L^o}}^{y^o} F^{g^o}(z|\boldsymbol{\varepsilon})dz$ for all $y^o \in \left[y^{L^o}, y^{U^o}\right]$ and $\boldsymbol{\varepsilon}$. Thus, the largest possible *w* that still satisfies equation A3 provides a bound for the WTP where the distribution *g*, adjusted by *w*, is preferred to the distribution *r*, yielding the definition

$$wtp^{LB} = \max_w \left\{w\colon \int_{y^L}^{y} F^r(z|\boldsymbol{\varepsilon})dz \geq \int_{y^L-w}^{y-w} F^g(z|\boldsymbol{\varepsilon})dz \text{ for all } y \geq \min\{y^L, y^L - w\} \text{ and } \boldsymbol{\varepsilon}\right\} \quad \text{A6}$$

Alternatively, consider

$$\int_{\boldsymbol{\varepsilon}} \int_{p>0} \int_{y^{L^o}}^{y^{U^o}} U\left((py^o - c^r)A\right)f^{r^o}(y^o|\boldsymbol{\varepsilon})\, dy^o h(p|\boldsymbol{\varepsilon})dp\, \phi(\boldsymbol{\varepsilon})d\boldsymbol{\varepsilon} \geq \quad \text{A3'}$$

$$\int_{\boldsymbol{\varepsilon}} \int_{p>0} \int_{y^{L^o}}^{y^{U^o}} U\left((py^o - c^r)A\right)f^{g^o}(y^o|\boldsymbol{\varepsilon})\, dy^o h(p|\boldsymbol{\varepsilon})dp\, \phi(\boldsymbol{\varepsilon})d\boldsymbol{\varepsilon},$$

which by analogous arguments has the sufficiency conditions $\int_{y^{L^o}}^{y^o} F^{g^o}(z|\boldsymbol{\varepsilon})dz \geq \int_{y^{L^o}}^{y^o} F^{r^o}(z|\boldsymbol{\varepsilon})dz$ for all $y^o \in \left[y^{L^o}, y^{U^o}\right]$ and $\boldsymbol{\varepsilon}$. Thus, the smallest possible *w* that satisfies equation A3' also provides a WTP bound. This WTP bound is where the distribution *r* is always preferred to the distribution *g*, adjusted by *w*, yielding the definition

$$wtp^{UB} = \min_w \left\{w\colon \int_{y^L-w}^{y-w} F^g(z|\boldsymbol{\varepsilon})dz \geq \int_{y^L}^{y} F^r(z|\boldsymbol{\varepsilon})dz \text{ for all } y \geq \min\{y^L, y^L - w\} \text{ and } \boldsymbol{\varepsilon}\right\} \quad \text{A6'}$$

Note the expression $\int_{\boldsymbol{\varepsilon}} \int_{p>0} \int_{y^L}^{y^U} U\left((p(y - w) - c^r)A\right)f^g(y|\boldsymbol{\varepsilon})\, dyh(p|\boldsymbol{\varepsilon})dp\, \phi(\boldsymbol{\varepsilon})d\boldsymbol{\varepsilon}$ is strictly decreasing in *w*, which immediately implies three additional results. For all continuous and twice differentiable $U(\cdot)$ such that $U'(\cdot) > 0$ and $U''(\cdot) < 0$:

- $wtp^{UB} \geq wtp^{LB}$,
- $wtp^{UB} \geq wtp^{LB} > 0$ if $EU^g > EU^r$, and
- $0 > wtp^{UB} \geq wtp^{LB}$ if $EU^r > EU^g$.

That is, wtp^{LB} is a lower bound, while wtp^{UB} is an upper bound on the WTP for alternative *g* instead of alternative *r*. Furthermore, if both WTP bounds are positive, then any risk averse individual will prefer alternative *g* to *r*. Conversely, if both WTP bounds are negative, then any risk averse individual will prefer alternative *r* to *g*.

References

Rothschild, M., Stiglitz, J., 1970. Increasing Risk I: A definition. J. Econ. Theory 2, 225-243.

Rothschild, M., Stiglitz, J., 1971. Increasing Risk II: Its economic consequences. J. Econ. Theory 3, 66-84.

Appendix B: Supplementary Tables and Figures

Table B1: Weather Year Used in DSSAT Yield Simulations by Simulation Year for Each of the 30 Replications Used to Construct the Yield Distributions

Replication	Simulation Year									
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
1	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990
2	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991
3	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992
4	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993
5	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994
6	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995
7	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996
8	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997
9	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998
10	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999
11	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000
12	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001
13	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002
14	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003
15	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004
16	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005
17	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006
18	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
19	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
20	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
21	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
22	2002	2003	2004	2005	2006	2007	2008	2009	2010	1981
23	2003	2004	2005	2006	2007	2008	2009	2010	1981	1982
24	2004	2005	2006	2007	2008	2009	2010	1981	1982	1983
25	2005	2006	2007	2008	2009	2010	1981	1982	1983	1984
26	2006	2007	2008	2009	2010	1981	1982	1983	1984	1985
27	2007	2008	2009	2010	1981	1982	1983	1984	1985	1986
28	2008	2009	2010	1981	1982	1983	1984	1985	1986	1987
29	2009	2010	1981	1982	1983	1984	1985	1986	1987	1988
30	2010	1981	1982	1983	1984	1985	1986	1987	1988	1989

Table B2: Willingness to Pay (WTP) Bound (maize ton/ha) Descriptive Statistics

	Year 1		**Year 10**	
	Upper Bound	**Lower Bound**	**Upper Bound**	**Lower Bound**
	From Traditional without to Traditional with 40 kg/ha N			
Mean	0.10	0.02	0.25	0.10
Standard Deviation	0.28	0.11	0.47	0.26
Median	0.00	0.00	0.00	0.00
Inter-Quartile Range	0.01	0.00	0.19	0.00
[Min, Max]	[-0.18, 1.69]	[-0.46, 1.17]	[-0.06, 2.41]	[-0.33, 1.54]
Clearly Better (%)	18.5		29.7	
Clearly Worse (%)	1.5		3.6	
	From Traditional without N to Improved without N			
Mean	2.58	1.33	1.45	0.53
Standard Deviation	2.15	1.94	1.56	1.35
Median	2.50	1.01	1.03	0.19
Inter-Quartile Range	3.52	2.89	2.18	1.49
[Min, Max]	[-1.63, 10.80]	[-3.38, 9.11]	[-1.63, 8.49]	[-3.36, 7.64]
Clearly Better (%)	69.6		58.5	
Clearly Worse (%)	10.5		13.0	
	From Traditional without N to Improved with N			
Mean	3.33	2.02	2.81	1.75
Standard Deviation	2.13	2.08	1.70	1.70
Median	3.18	2.06	2.76	1.90
Inter-Quartile Range	2.75	3.27	2.28	2.76
[Min, Max]	[-1.63, 11.32]	[-3.52, 8.70]	[-1.63, 8.97]	[-3.52, 7.68]
Clearly Better (%)	77.6		79.4	
Clearly Worse (%)	3.3		2.9	
Total Hectares	24.8			
Cells	3,854			

Note: Descriptive statistics are weighted by maize hectares.

Table B3: Genetic Coefficient Values of Traditional Varieties used in the CERES-Maize Model

	CM1509 (Long maturity)	**CM1510 (Short maturity)**
P1: Thermal time from seedling emergence to the end of juvenile phase (expressed in degree days above a base temperature of 8 °C) during which the plant is not responsive to changes in photoperiod.	238.6	125.0
P2: Extent to which development (expressed as days) is delayed for each hour increase in photoperiod above the longest photoperiod at which development proceeds at a maximum rate (which is considered to be 12.5 h)	0.5	0.5
P5: The thermal time from silking to physiological maturity (expressed in degree days above a base temperature of 8 °C)	654	560
G2: Maximum possible number of kernels per plant	450	450
G3: Kernel filling rate during the linear grain filling stage and under optimum conditions (mg/day)	8.5	9.5
PHINT: Phyllochron interval; the interval in thermal time (degree days) between successive leaf tip appearances.	75	75

Figure B3: Willingness to Pay Bounds (maize ton/ha) for Switching from the Traditional Variety Without Nitrogen Fertilizer to the Traditional Variety With 40 kg/ha of Nitrogen Fertilizer in Year 1 (Panels (a) and (b)) and 10 Years (Panels (c) and (d)) After Adoption

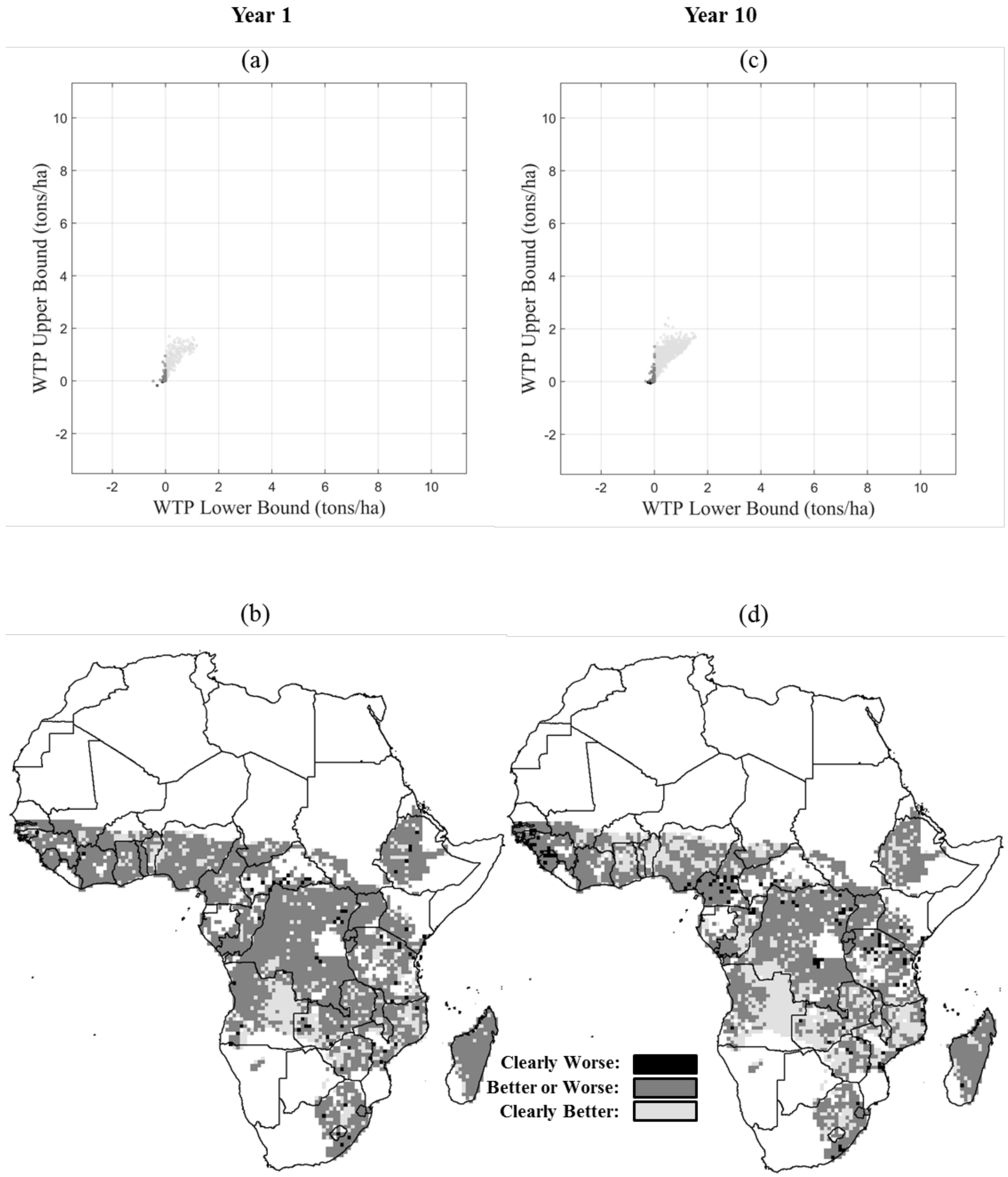


Figure B4: Willingness to Pay Bounds (maize ton/ha) for Switching from the Traditional Variety Without Nitrogen Fertilizer to the Improved Variety Without Nitrogen Fertilizer in Year 1 (Panels (a) and (b)) and 10 Years (Panels (c) and (d)) After Adoption

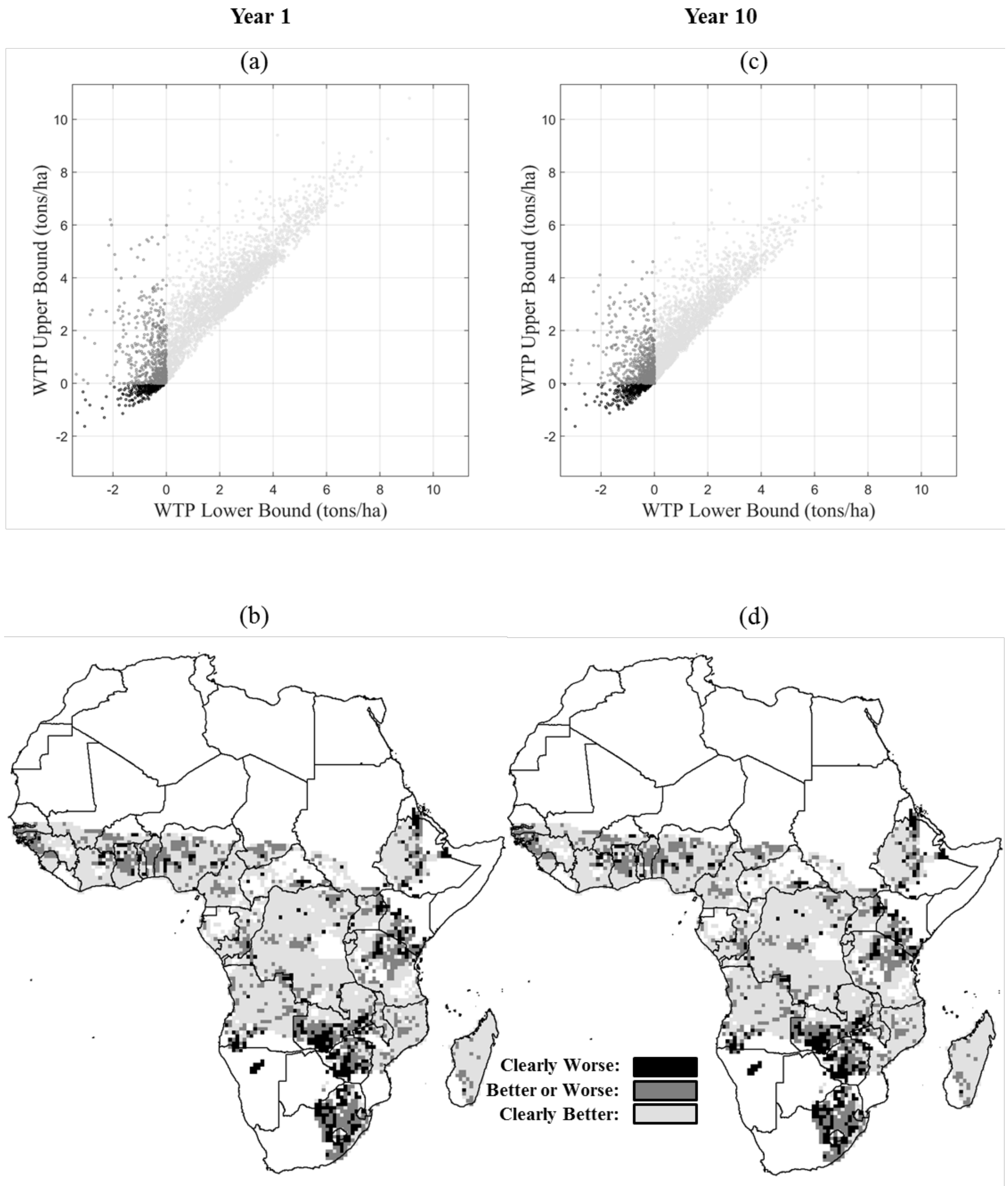


Figure B5: Willingness to Pay Bounds (maize ton/ha) for Switching from the Traditional Variety Without Nitrogen Fertilizer to the Improved Variety With 40 kg/ha of Nitrogen Fertilizer in Year 1 (Panels (a) and (b)) and 10 Years (Panels (c) and (d)) After Adoption

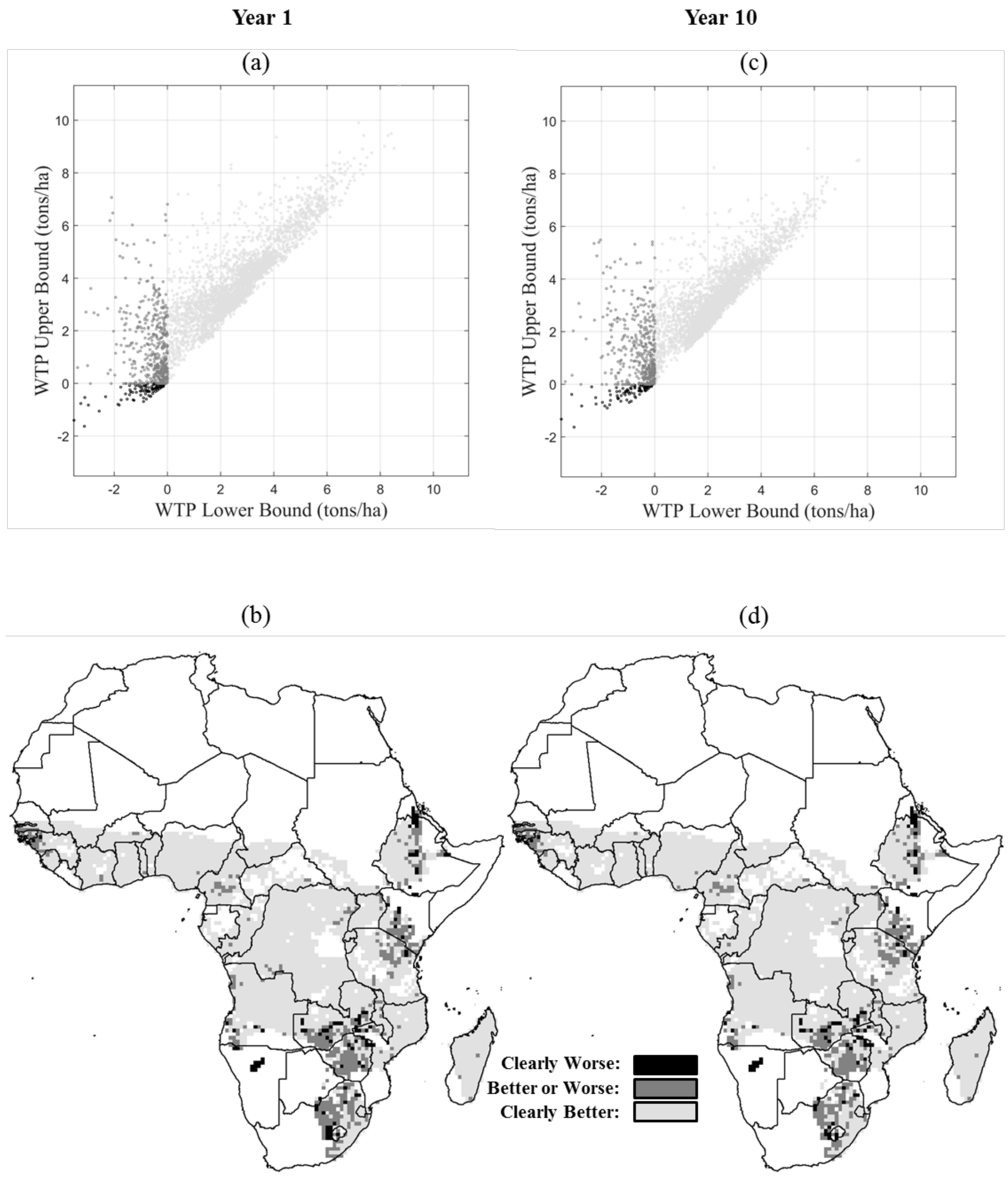


Figure B6: Sensitivity Analysis for Willingness to Pay Bounds (maize ton/ha) when Switching from the Traditional Variety Without Nitrogen Fertilizer to the Improved Variety With 40 kg/ha of Nitrogen Fertilizer at a Price Multiple of 122 kg/ha in Year 1 (Panel (a)) and 10 Years (Panel (b)) After Adoption

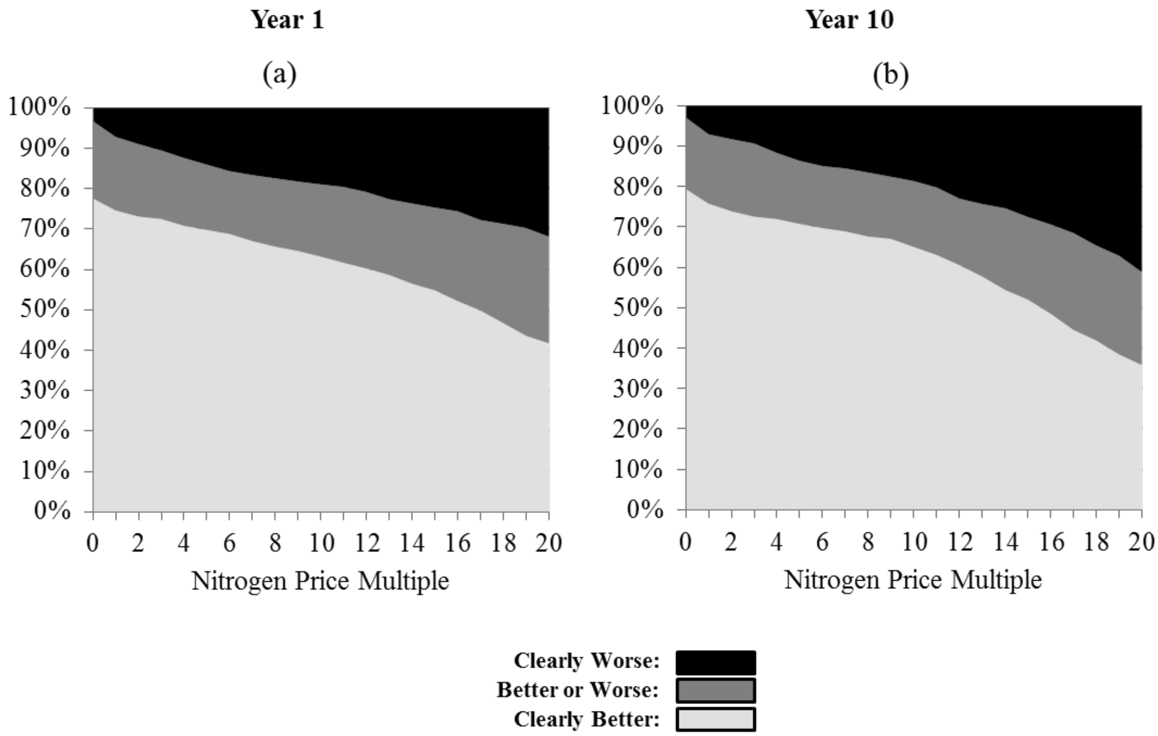
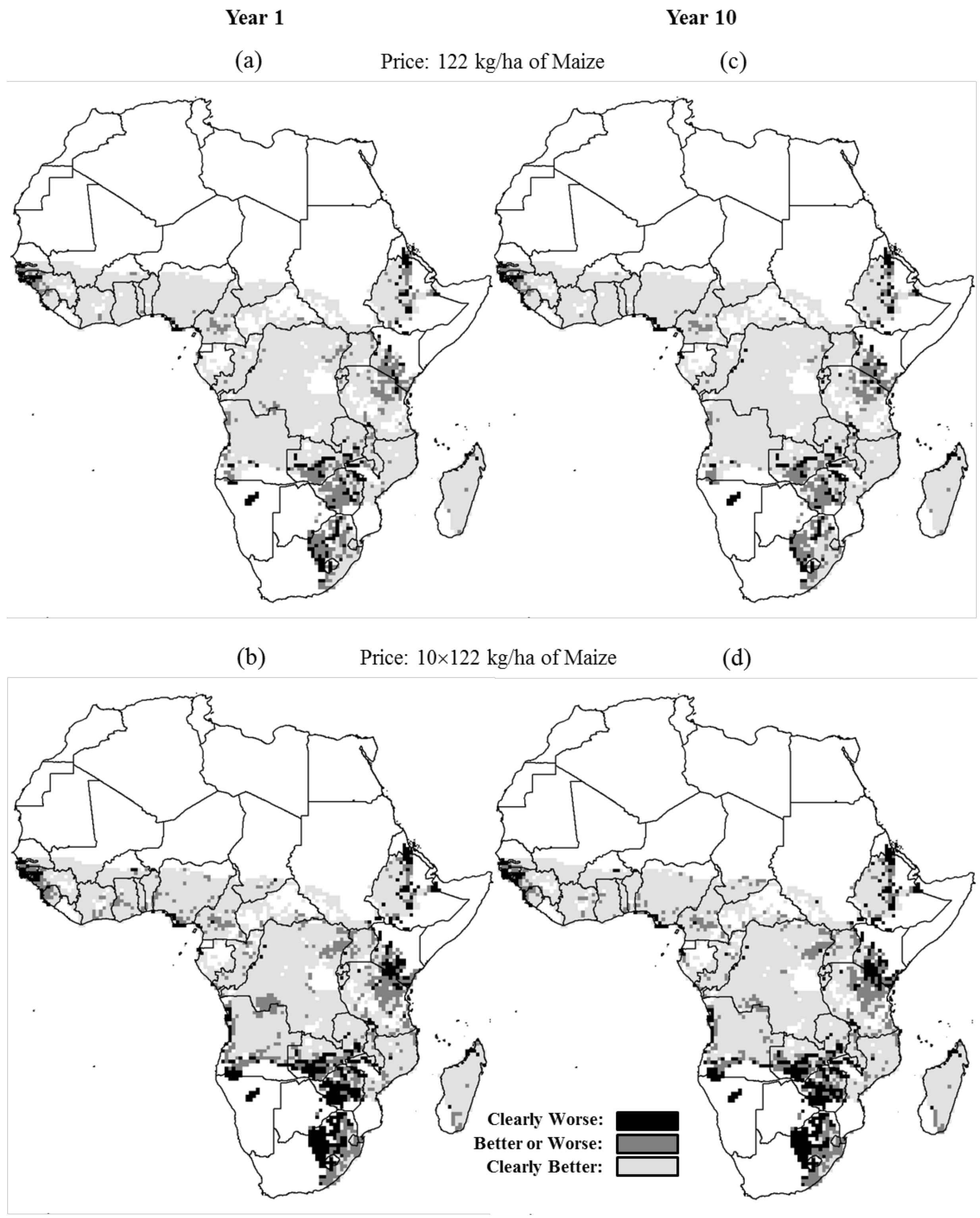


Figure B7: Sensitivity Analysis for Willingness to Pay Bounds (maize ton/ha) when Switching from the Traditional Variety Without Nitrogen Fertilizer to the Improved Variety With 40 kg/ha of Nitrogen Fertilizer at a Price of 122 kg/ha and 10×122 kg/ha of Maize in Year 1 (Panels (a) and (b)) and 10 Years (Panels (c) and (d)) After Adoption



Appendix C: Data Description and Analysis

The data used to produce the figures and tables in the article are in two matrices in the Matlab workspace <AGECON MS2017410> in the <MatLabFiles> directory: < ndataseq01060216> and < ndataseq10060216>. These two matrices contain the crop simulation output for the initial and 10th simulation years. The columns of these matrices are

1 = cell30m / Unique Cell ID
2 = Replication
3 = yearseq / Simulation Year
4 = Maize Area in Cell (ha)
5 = Maize Yield (kg/ha) for CM1510 with 40 kg N/ha
6 = Maize Yield (kg/ha) for CM1510 with 0 kg N/ha
7 = Maize Yield (kg/ha) for CM1509 with 40 kg N/ha
8 = Maize Yield (kg/ha) for CM1509 with 0 kg N/ha
9 = Maize Yield (kg/ha) for Improved with 40 kg N/ha
10 = Maize Yield (kg/ha) for Improved with 0 kg N/ha

Running the script <CreateWTPBoundsbyCell> uses the functions <ConstRiskThresh>, <SOSDConstBoundsv3> and <SOSDIntegralTestv3> to produce the matrix <RA>. The columns of this matrix are

1 = yearseq / simulation year
2 = cell30m / Unique Cell Identifier
3 = Comparison (Comp) Cultivar ID (CM1510 = 10, CM1509 = 9, Improved = 0)
4 = Comp Nitrogen ID
5 = Base Cultivar ID (CM1510 = 10, CM1509 = 9, Improved = 0)
6 = Base Nitrogen Management ID
7 = Mean Yield for Comp
8 = Standard Deviation of Yield for Comp
9 = CV of Yield for Comp
10 = Maximum Yield for Comp
11 = Minimum Yield for Comp

12 = Probability of Crop Failure for Comp

13 = Min Proportion for Comp to SOSD Base/ Minus the Upper Bound WTP

14 = Mean Yield for Base

15 = Standard Deviation of Yield for Base

16 = CV of Yield for Base

17 = Maximum Yield for Base

18 = Minimum Yield for Base

19 = Probability of Crop Failure for Base

20 = Min Proportion for Base to SOSD Comp/ Minus the Lower Bound WTP

21 = Difference in mean Comp - Base

22 = Difference in standard deviation Comp - Base

23 = Difference in CV Comp - Base

24 = Difference in Prob of Crop Failure Comp - Base

25 = Min Proportion for Comp to SOSD Base divided by Average Base Yield

26 = Min Proportion for Base to SOSD Comp divided by Average Base Yield

27 = Comp Clearly Worse (-1)/Clearly Better (1)/ Better or Worse (0) compared to Base

28 = Maize Area in Cell (ha)

For each cell, the <RA> matrix includes for the initial and 10th simulation years results for the comparison of the traditional variety (CM1509 and CM1510) with and without 40 kg/ha N, the improved and traditional varieties without N, and the improved variety with 40 kg/ha N and traditional variety without N. Note that <ConstRiskThresh> uses <SOSDConstBoundsv3>, which then uses <SOSDIntegralTestv3>. <ConstRiskThresh> organizes the results by cell and scenario comparison. <SOSDConstBoundsv3> calculates various statistics for each comparison yield distribution and the WTP bounds using a golden-section search methodology. <SOSDIntegralTestv3> performs a second-order stochastic dominance integral test following the algorithm found in Levy (2006, pages 180-182)

The script <CreateTableData> uses <RA> to calculate weighted statistics across cells for each of the three comparisons, both traditional varieties, and the initial and 10th simulation years. It produces the matrix <DescriptiveStat> with rows

1 = Weighted Mean UB WTP

2 = Weighted S.D. UB WTP

3 = Minimum UB WTP

4 = 10th Percentile UB WTP

5 = 25th Percentile UB WTP

6 = Median UB WTP

7 = 75th Percentile UB WTP

8 = 90th Percentile UB WTP

9 = Maximum UB WTP

10 = Weighted Mean LB WTP

11 = Weighted S.D. LB WTP

12 = Minimum LB WTP

13 = 10th Percentile LB WTP

14 = 25th Percentile LB WTP

15 = Median LB WTP

16 = 75th Percentile LB WTP

17 = 90th Percentile LB WTP

18 = Maximum LB WTP

19 = Proportion of Acres in Green

20 = Proportion of Acres in Yellow

21 = Proportion of Acres in Red

22 = Total Acres

23 = Number of Cells

and Columns

1 = RC_01_00_00_09_00

2 = RC_01_00_00_10_00

3 = RC_01_00_40_00_00

4 = RC_01_00_40_09_00

5 = RC_01_00_40_10_00

6 = RC_01_09_40_09_00

7 = RC_01_10_40_10_00

8 = RC_10_00_00_09_00

9 = RC_10_00_00_10_00

10 = RC_10_00_40_00_00

11 = RC_10_00_40_09_00

12 = RC_10_00_40_10_00

13 = RC_10_09_40_09_00

14 = RC_10_10_40_10_00

where RC_SY_CV_CN_BV_BN is for simulation year SY (01 for initial and 10 for 10^{th} year), comparison variety CV (00 for improved, 09 for CM1509, and 10 for CM1510), comparison N rate CN (00 for 0 N and 40 for 40kg/ha N), base variety BV (09 for CM1509 and 10 for CM1510), and base N rate BN (00 for 0 N). These results are reported in Tables 1 and B2.

The script <CreateFigureData> creates the matrix <CELL30MIDMAP> for GIS maps (Figures 3 – 5 and Figures B3 – B5). The geographical components of these figures were created in ArcMap 10.3.1 using the supporting files found in the < ArcViewShapeFiles> directory. <countries.shp> is the shape file for the African countries, while <cell30m_ssa_const.shp> is the shape file for the 30 arc minute cells used in the analysis. Note that the variable <cell30m> in <cell30m_ssa_const.shp> is used to link to the < cell30m> variable in the <CELL30MIDMAP> matrix. The scatter plots were done in Matlab using <RA> matrix data. Code for accomplishing this is not included because they can be easily reconstructed with the available data using any scatter plot tool. The script also creates the data for the price sensitivity analysis (Figures 6, 7, B6, and B7) in the matrix <PriceSensitivity> and matrix <CELL30MIDMAP>. Again, geographical maps were made using ArcMap 10.3.1. Figure 6 and B6 were constructed with the Microsoft Excel.

Note that the columns for <CELL30MIDMAP> are:

1 = cell30M

2 = Cell Agricultural Acreage

3 = RC_01_00_00_09_00

4 = RC_01_00_00_10_00

5 = RC_01_00_40_00_00

6 = RC_01_00_40_09_00

7 = RC_01_00_40_10_00

8 = RC_01_09_40_09_00

9 = RC_01_10_40_10_00

10 = RC_10_00_00_09_00

11 = RC_10_00_00_10_00

12 = RC_10_00_40_00_00

13 = RC_10_00_40_09_00

14 = RC_10_00_40_10_00

15 = RC_10_09_40_09_00

16 = RC_10_10_40_10_00

17 = PX01_RC_01_00_40_09_00

18 = PX01_RC_10_00_40_09_00

19 = PX01_RC_01_00_40_10_00

20 = PX01_RC_10_00_40_10_00

21 = PX10_RC_01_00_40_09_00

22 = PX10_RC_10_00_40_09_00

23 = PX10_RC_01_00_40_10_00

24 = PX10_RC_10_00_40_10_00

where RC_SY_CV_CN_BV_BN are for simulation year SY (01 for initial and 10 for 10^{th} year), comparison variety CV (00 for improved, 09 for CM1509, and 10 for CM1510), comparison N rate CN (00 for 0 N and 40 for 40kg/ha N), base variety BV (09 for CM1509 and 10 for CM1510), and base N rate BN (00 for 0 N). The PX01_ prefix is for the base nitrogen price (122 kg/ha) and the prefix PX10 is for ten times the nitrogen base price (10×122 kg/ha).

Finally, the rows of the <PriceSensitivity> matrix correspond to a multiple of the base nitrogen price (122 kg/ha), while the columns correspond to:

1 = Multiple of base nitrogen price

2 = proportion of the initial simulation year cells where the improved variety with 40 kg/ha N is clearly better than CM1509 without N

3 = proportion of the initial simulation year cells where the improved variety with 40 kg/ha N is clearly worse than CM1509 without N

4 = proportion of the 10th simulation year cells where the improved variety with 40 kg/ha N is clearly better than CM1509 without N

5 = proportion of the 10th simulation year cells where the improved variety with 40 kg/ha N is clearly worse than CM1509without N

6 = proportion of the initial simulation year cells where the improved variety with 40 kg/ha N is clearly better than CM1510 without N

7 = proportion of the initial simulation year cells where the improved variety with 40 kg/ha N is clearly worse than CM1510 without N

8 = proportion of the 10th simulation year cells where the improved variety with 40 kg/ha N is clearly better than CM1510 without N

9 = proportion of the 10th simulation year cells where the improved variety with 40 kg/ha N is clearly worse than CM1510without N

References

Levy, H. (2006). *Stochastic Dominance: Investment Decision Making Under Uncertainty* (Second Edition). Springer. New York, NY.