Journal of Agricultural and Resource Economics 44(1):179–193 ISSN 1068-5502

Estimating the Basis Risk of Rainfall Index Insurance for Pasture, Rangeland, and Forage

Jisang Yu, Monte Vandeveer, Jerry D. Volesky, and Keith Harmoney

Using historical yield and rainfall data from three university-managed ranches in Kansas and Nebraska, we measure basis risk of Rainfall Index Insurance for Pasture, Rangeland, and Forage (PRF-RI). We investigate the relationship between forage yield and monthly precipitation and estimate the relationship between forage yield and PRF-RI indices. Finally, we estimate basis risk of PRF-RI. Our estimates suggest that using actual site-level precipitation values would reduce basis risk by only 5%–9%, indicating that basis risk stems mostly from nonprecipitation factors. Using more flexible contract forms with site-level precipitation would have little impact on decreasing the degree of basis risk.

Key words: Elastic net penalty, PRF-RI program

Introduction

The availability of index insurance products has expanded globally in the last few decades in both developed and developing countries (Miranda and Farrin, 2012; Smith and Glauber, 2012). Compared to traditional insurance products, index insurance costs less to implement since it has lower costs from information asymmetry problems (Barnett and Mahul, 2007), which can be minimized if farms and insurers have similar information about the indices used in index insurance.

Index insurance products are particularly useful when gathering historical individual data would be difficult or expensive. In an environment of expensive data collection, index insurance based on easily observable data such as rainfall or satellite indices offers an alternative to conventional, individual-based insurance (Jensen and Barrett, 2016; Barnett and Mahul, 2007).

However, imperfect correlations between individual outcomes and insurance indices introduce an uninsured risk called "basis risk," which reduces the effectiveness of index insurance as a risk management tool. Previous studies suggest that basis risk reduces demand for index insurance (e.g., Giné, Townsend, and Vickery, 2008; Binswanger-Mkhize, 2012; Elabed et al., 2013; Clarke, 2016). Despite the global expansion of index insurance programs and recognition of the negative impacts of basis risk on index insurance demand, only a few studies have investigated the degree of basis

---

Jisang Yu (corresponding author) is an assistant professor and Monte Vandeveer is an extension agricultural economist in the Department of Agricultural Economics at Kansas State University. Jerry D. Volesky is a professor and range and forage specialist in the Department of Agronomy and Horticulture at the University of Nebraska-Lincoln. Keith Harmoney is an associate professor in the Department of Agronomy at Kansas State University.
We thank participants at the SCC-76 2017 annual meeting, Gyuhyeong Goh, and two anonymous referees. We also acknowledge the generous editorial assistance provided by *JARE* editor Darren Hudson.
Review coordinated by Darren Hudson.

risk among existing index insurance programs (e.g., Giné, Townsend, and Vickery, 2008; Jensen, Barrett, and Mude, 2016).[1]

We investigate basis risk of Rainfall Index Insurance for Pasture, Rangeland, and Forage (PRF-RI), an index insurance product in the U.S. federal crop insurance program that first became available in 2007. The 2012 Census of Agriculture indicates that over 415 million acres were devoted to "permanent pasture and rangeland" in 2012 (U.S. Department of Agriculture, 2014), representing approximately 45% of all land in U.S. farms in that year. PRF-RI also offers coverage to established perennial forages, including alfalfa and grass meadows baled for hay, representing another approximately 50 million acres nationally (U.S. Department of Agriculture). Despite the potential importance of the PRF-RI program, only a few studies have examined the program (e.g., Nadolnyak and Vedenov, 2013; Ifft et al., 2014; Diersen et al., 2015; Westerhold et al., 2018).

Data on individual forage yields for grazing and haying lands are scarce, since few producers bother to carefully clip, dry, and weigh forage samples across their grazing areas (Smith, Panciera, and Probst, 2010). Since producers had no forage yield histories, there were no historical data for estimating premium rates and no means of determining whether a loss had occurred. Alternatively, weather-based or satellite-based index insurance products have been considered by policy makers. As of 2017, rainfall index insurance programs were available to U.S. ranchers.

Like any other index insurance products, PRF-RI participants are exposed to basis risk because the rainfall index (used as a proxy for crop output) and crop yield may not track perfectly. Basis risk may deter participation in the program and thus may affect the cost-effectiveness of the subsidy in the index insurance program. Understanding the basis risk is important to be able to assess the impact of the PRF-RI program.

We measure the basis risk of PRF-RI using historical yield and rainfall data from three university-managed ranches and use the estimated results to evaluate the effectiveness of PRF-RI. We contribute to the literature in two ways: First, we estimate the degree of basis risk of PRF-RI that participants in the U.S. Midwest may face. We also decompose the basis risk into i) index-related risk and ii) nonprecipitation risk and discuss how much of the basis risk can be potentially reduced. To our knowledge, there are very few estimates on the basis risk of the rainfall index insurance programs. Important exceptions are Westerhold et al. (2018), who investigate the risk-reducing interval selection of the PRF-RI program, and Maples, Brorsen, and Biermacher (2016), who investigate the effectiveness of the Annual Forage Program.[2] Second, we use a regularization method to estimate the relationships between yields and precipitation as well as between yields and rainfall indices. This empirical approach is appropriate when there are a relatively large number of explanatory variables compared to the number of observations.

We find that the rainfalls in May–July play the most important role on forage yield growth. Consistent with these results from the yield regression with actual monthly rainfall amounts as explanatory variables, the results of the yield regression with the PRF-RI indices as explanatory variables indicate that the May–June index is the most important variable explaining the forage yield. Our models also provide estimates on the basis risk of the PRF-RI program, measured by false negative probabilities. We find that the overall basis risk of the PRF-RI program ranges from 26% to 43%. Using actual precipitation data and more flexible contract forms has the potential to reduce basis risk by 5%–9%.

---

[1] Several studies have investigated basis risk in the context of designing index insurance programs or assessing the feasibility of weather derivatives (e.g., Vedenov and Barnett, 2004; Woodard and Garcia, 2008; Norton, Turvey, and Osgood, 2012; Chantarat et al., 2013). Estimating basis risk requires sufficient amounts of data on individual outcomes; many index insurance programs, especially those in developing countries, lack data on individual outcomes.

[2] Maples, Brorsen, and Biermacher (2016) examine the statistical relationships between annual forage production in southern Oklahoma, local rainfall, and the precipitation indices used in the Annual Forage Program, which uses the same indices as the PRF-RI program.

Index Insurance and Basis Risk

Index insurance pays indemnities based on the outcomes of indices such as area yields, rainfall, or vegetation indices. Index insurance is not subject to asymmetric information problems such as moral hazard or adverse selection because the indices are exogenous to the insured's behavior or characteristics (Barnett, Barrett, and Skees, 2008). Also, the cost of collecting data to estimate premiums is often lower than for traditional, individual-based insurance when the indices are based on easily observable data.

Poor actuarial performance of individual yield insurance in the U.S. federal crop insurance program in the 1980s prompted several studies investigating the feasibility of area-yield insurance (e.g., Miranda, 1991; Smith, Chouinard, and Baquet, 1994; Mahul, 1999). Insurance products based on area-yield or area-level revenue have been available since the early 1990s, but participation rates have been low (Glauber, 2013).

Because the indices are exogenous to individual outcomes, indemnities are not perfectly correlated with individual losses. Basis risk is defined as uninsured risk that results from imperfect correlations between individual outcomes and insurance indices. For example, Miranda (1991) describes how higher correlation between individual and area yields lowers the basis risk of area-yield index insurance. Rainfall index insurance has a higher level of basis risk if rainfall indices are not highly correlated with individual outcomes.

Several studies show that basis risk reduces demand for index insurance (e.g., Giné, Townsend, and Vickery, 2008; Binswanger-Mkhize, 2012; Elabed et al., 2013; Elabed and Carter, 2015; Clarke, 2016). Using standard expected utility theory, Clarke (2016) shows that risk-averse agents would not fully insure when there is basis risk. The negative impact of basis risk on the demand for index insurance is even greater when agents are compound-risk averse (Elabed and Carter, 2015). Thus, minimizing basis risk in index insurance products should lead to greater demand for index insurance. Despite the importance of minimizing basis risk in index insurance, few studies have examined basis risk in existing index insurance programs: Giné, Townsend, and Vickery (2008) and Jensen, Barrett, and Mude (2016). We contribute to the literature by measuring the magnitude of basis risk in the PRF-RI program that participants in the U.S. Midwest may face.

We also decompose basis risk into two sources: nonprecipitation risk and index-related risk.[3] We define nonprecipitation risk as basis risk from factors (other than rainfall) that can influence forage yields, such as temperature, disease, pests, hail, and poor plant vigor carried over from the previous growing season. Another source of basis risk is the fact that PRF-RI bases its index values on rainfall amounts at multiple locations that may be distant from the producer.

We first estimate the relationships between yields and precipitation and between yields and rainfall indices. We then obtain the false negative probabilities from the predictions of each estimation. Comparing the false negative probabilities of the two relationships allows us to infer the degree of design risk, which can potentially be reduced.[4] Our decomposition provides some useful information on potential improvements to the PRF-RI program.

Precipitation and Forage Yields

Range scientists have long studied the relationship between forage production and weather-related variables such as rainfall. Some studies also explore the interactions between weather and other environmental conditions (e.g., soil properties) and management decisions (e.g., grazing intensity).

The study with the widest geographic scope seems to be that of Sala et al. (1988), who examine total forage production as a function of annual precipitation for more than 9,000 sites in the central

[3] Elabed et al. (2013) decompose basis risk into idiosyncratic risk and design risk. Our decomposition has a slightly different nuance since we use individual-level precipitation data.

[4] Again, our decomposition is different from that of Elabed et al. (2013), but index-related risk is part of design-risk, which can potentially be reduced.

United States using data collected by the USDA's Soil Conservation Service.[5] They estimate yield response functions for average, favorable, and unfavorable years and find that roughly 90% of variability in forage production in this cross-sectional dataset could be explained using annual precipitation and soil water-holding capacity.

Several other studies perform time-series analyses for individual sites. Smoliak (1986) analyzes range forage yields over a 50-year period for southeast Alberta using different combinations of precipitation and temperature variables. The study shows that June and July precipitation plus May and June mean temperatures explained 63% of variation in forage yield production. Various measures of precipitation from the previous year were also considered, as they could affect plant vigor in the previous year or soil moisture at the beginning of the current year. However, they added little statistically in explanatory value.

Using a 52-year dataset, Lauenroth and Sala (1992) investigate the relationship between shortgrass production in north-central Colorado and a variety of weather variables, including annual precipitation, growing season (April to September) precipitation, precipitation by size of rainfall event, and temperature. A model using precipitation by size of rainfall event performed best in terms of explaining forage production, accounting for 45% of variation in forage production over this time. They also note that total annual precipitation often failed to account for variability in forage production.

Oesterheld et al. (2001) apply other methods to Lauenroth and Sala's data and find that the previous year's production (or, alternatively, the previous year's precipitation) added to the explanatory power of the forage production models. For this more arid location and particular plant community, previous-year effects were important in accounting for the buffering or amplifying effects of much higher or lower precipitation in the current year.

Smart et al. (2007) examine weather variables affecting forage output for three plant communities in southwestern South Dakota for 1945–1960. Mid-grass, shortgrass, and mixed-grass production were all evaluated using separate models of yield as a function of rainfall and temperature-related variables. Current-year spring (April–June) precipitation, number of days since the last freeze, and previous-year spring precipitation were all significant variables. These models accounted for 52%–81% of variability in annual forage production.

Heitschmidt and Vermeire (2006) investigate the yield response of Montana warm-season rangelands to summer precipitation when springtime precipitation had been lacking. These pastures showed the capacity to respond in summer after dry springs, but total yields still lagged those from years with the usual wet spring/dry summer pattern.

An extension resource for producers, *Managing Drought Risk on the Ranch: A Planning Guide for Great Plains Ranchers*, summarizes much of the relevant research: (National Drought Mitigation Center, 2012, p. 4):

> Forage research shows that the most important months for precipitation are the months just prior to the rapid growth periods of your dominant plant species. For much of the Great Plains, those critical rain months occur in spring through early summer. Rainfall that occurs after the rapid growth period of dominant plant species does not result in as much useable [*sic*] forage.

Rainfall Index Insurance in the United States

As of 2017, apiculture; annual forage; and pasture, rangeland, and forage were the only commodities eligible to purchase the rainfall index insurance governed by the Risk Management Agency (RMA) and the Federal Crop Insurance Corporation (FCIC).[6] This article focuses on the PRF-RI program.

---

[5] Annual precipitation in Sala et al. (1988) is defined as precipitation during the growing season, usually designated as October 1 through September 30 of the following year.

[6] A recent study by Maples, Brorsen, and Biermacher (2016) investigates the effectiveness of the Annual Forage Program.

The PRF-RI program's rainfall index is based on precipitation measured over 2-month intervals, selected by the producer at sign-up. Producers must decide which periods (called "index intervals") to insure and how to allocate their coverage across the various intervals selected. As described previously, significant agronomic research demonstrates the positive relationship between rainfall and forage output (e.g., Sala et al., 1988; Sala, Lauenroth, and Parton, 1992; Lauenroth and Sala, 1992; Heitschmidt and Vermeire, 2006; Smart et al., 2007).

The PRF-RI rainfall index is calculated using grids defined by the National Oceanic and Atmospheric Administration Climate Prediction Center (NOAA CPC). These grid areas are 0.25° of longitude by 0.25° of latitude (U.S. Department of Agriculture, 2015), resulting in areas approximately 17 by 13 miles at latitudes in the central United States. A producer's PRF-RI insurance is based on the rainfall index values of the grid in which their operation is located. Premium rates are based on each grid's precipitation history since 1948.

PRF-RI coverage is *not* based on observed precipitation at the producer's own location. The rainfall index for each grid area expresses observed rainfall as a percentage of normal based on observed precipitation at the four nearest weather stations used by the NOAA CPC system. The readings from these four stations are combined to create a single rainfall index value for the grid area, with the weights for each station based on their distance from the center point of the grid.

PRF-RI coverage is based on rainfall observed over 2-month intervals. Producers can insure 70%–90% of normal rainfall for their grid area. Actual rainfall during an insured period that is less than the guarantee level selected by the producer triggers an indemnity based on the size of the shortfall. A producer seeking to maximize the risk-reducing benefits from PRF-RI would seek to allocate his/her coverage to the index intervals whose rainfall had the greatest effect on forage production.

Because PRF-RI relies on the rainfall index as a proxy for forage production, the correspondence between rainfall and forage output determines how much of producers' risk the PRF-RI program can cover. Thus, low correlation between rainfall and forage output is one source of basis risk (nonprecipitation risk). Another source of basis risk is the difference between rainfall amounts at one's location and at those measured at the stations used by the insurance grid area (index-related risk). Anecdotally, some producers express concern about the distance between their locations and the reporting stations; our estimates on index-related risk could address this concern.

Data

We use both rainfall and pasture yield data from three university-managed ranches: The Barta Brothers Ranch (BBR) near Rose, in north-central Nebraska, and the Gudmundsen Sandhills Laboratory (GSL) in northwest Nebraska, near Whitman, were donated to the University of Nebraska-Lincoln and are operated in part for university research.[7] Kansas State University operates the third ranch, which is near Hays, in north-central Kansas.

Rainfall data come from on-site weather stations at each ranch. Forage yield data come from long-run forage yield studies at each location. The Nebraska ranches collect forage data using exclusion cages, which prevent cattle from grazing in a particular spot. These cages are scattered across the ranches; in August each year, forage growth from the cages is clipped and collected then dried and weighed to calculate yield. The BBR yield data series begins in 1999, and the GSL data series begins in 2004. Kansas State University range scientists' estimates of forage data from the Hays ranch begin in 2001.[8] Figure 1 shows the trends for total forage yield from the three

[7] Westerhold et al. (2018) explain details of BBR and GSL data.

[8] Plot-level yields have been observed at BBR. We use ranch-level average annual yield information for our main analysis, but we also pool BBR plot-level yields and GSL and Hays ranch-level for the sensitivity analysis. For the PRF-RI indices, we select Grid 26515 for GSL, Grid 26822 for BBR, and Grid 22623 for Hays. The majority of GSL upland pasture is in Grid 26515, and the BBR plot with the longest periods of yield data is in Grid 26822. The KSU ranch at Hays is located in the center of Grid 22623.

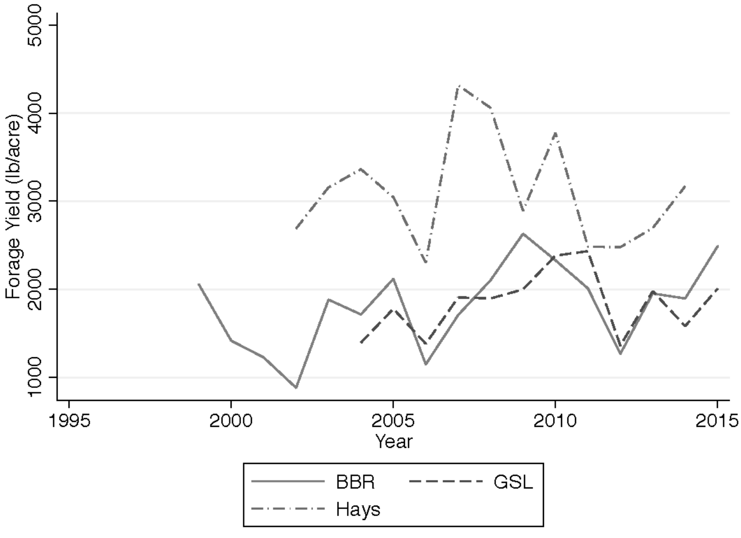

Figure 1. Forage Yield Trends of BBR, GSL, and Hays Ranches

ranches and indicates how widely forage yields can vary across years. Notable are droughts in 2002, 2006, and 2012, with corresponding declines in yields. Hays ranch has overall higher yields than the Nebraska ranches.

As described previously, the PRF-RI indices are based on the NOAA CPC grid system. We obtained the indices of the grids where the ranches are located for the years of our sample. PRF-RI indices for the corresponding years were obtained from the RMA.

We present means and standard deviations of precipitation and PRF-RI index values for each ranch in Tables 1 and 2. In Table 1, we observe the continental rainfall patterns for northern Nebraska: higher rainfall in the spring and summer months. The April-to-August period accounts for 67.2% of the 21.1 inches of total annual rainfall at BBR and for 75.7% of GSL's annual average of 20.8 inches. Climate and production conditions at Hays follow similar patterns, with slightly warmer temperatures. The April-to-August period accounts for 62.5% of Hays's 23 inches of annual rainfall.

Table 2 shows the statistics for the PRF-RI index values. Since an index of 100 indicates a value equal to the long-term mean (index values are calculated using data going back to 1948), we see that the means for the years in question do not exactly match the long-term averages for each time interval. Note that there are differences in index variability across intervals: At BBR, index values in January–February and November–December are most variable, while those in summer months tend to be less so. Variability in index values at GSL and Hays is more mixed across periods.

Estimation Methods

We first specify the yield–precipitation equation as

$$
\begin{aligned}
Yield_{it} = {} & \beta_0 + \sum_{k=1}^{12} \beta_{1,k} Precipitation_{ikt-1} + \sum_{k=1}^{7} \beta_{2,k} Precipitation_{ikt} \\
& + \sum_{k=1}^{12} \beta_{3,k} Precipitation^2_{ikt-1} + \sum_{k=1}^{7} \beta_{4,k} Precipitation^2_{ikt} + \sum_{i=1}^{3} \gamma_i + \delta_1 Time_t \\
& + \delta_2 {Time_t}^2 + \varepsilon_{it}
\end{aligned} \tag{1}
$$

Table 1. Summary Statistics: Precipitation (inches)

	BBR		GSL		Hays	
Variable	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
January	0.369	0.348	0.340	0.349	0.400	0.383
February	0.704	0.440	0.662	0.432	0.885	0.635
March	1.111	0.903	0.721	0.622	1.457	1.302
April	2.625	1.387	2.491	0.907	1.940	0.866
May	3.076	1.425	3.257	1.775	2.881	2.003
June	4.205	2.523	4.374	2.375	3.208	2.091
July	2.084	1.296	3.107	2.293	3.219	2.361
August	2.159	1.019	2.590	1.992	3.146	1.381
September	1.932	1.153	1.616	1.010	2.660	1.866
October	1.721	1.514	1.143	0.612	1.902	1.451
November	0.554	0.557	0.314	0.186	0.556	0.424
December	0.517	0.474	0.269	0.222	0.792	0.937
No. of years	17		12		14	

Table 2. Summary Statistics: PRF-RI Index Values

	BBR		GSL		Hays	
Variable	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
January–February	88.529	70.438	103.650	57.450	105.936	87.639
February–March	82.588	36.477	103.900	47.870	91.836	62.081
March–April	100.512	42.953	109.908	34.459	89.293	44.020
April–May	103.859	42.309	113.600	36.526	91.257	42.308
May–June	101.012	40.018	114.825	39.774	93.229	32.058
June–July	93.888	40.501	115.133	50.148	88.279	35.513
July–August	88.153	36.128	116.958	56.608	92.536	27.344
August–September	94.641	31.999	122.033	48.494	111.379	26.554
September–October	100.065	42.462	119.017	50.615	118.521	45.982
October–November	94.653	47.470	99.292	39.915	100.057	67.928
November–December	79.112	50.571	78.392	42.211	81.093	51.805
No. of years	17		12		14	

and the yield–PRF-RI indices equation as

$$\begin{aligned} Yield_{it} = \beta_0 + \sum_{k=1}^{11} \beta_{1,k} PRF_{ikt-1} + \sum_{k=1}^{7} \beta_{2,k} PRF_{ikt} + \sum_{k=1}^{11} \beta_{3,k} PRF_{ikt-1}^2 \\ + \sum_{k=1}^{7} \beta_{4,k} PRF_{ikt}^2 + \sum_{i=1}^{3} \gamma_i + \delta_1 Time_t + \delta_2 {Time_t}^2 + \varepsilon_{it} \end{aligned} \tag{2}$$

where $Yield_{it}$ is the forage yield at ranch i in year t; $Precipitation_{ikt}$ is the actual precipitation at ranch i in month k, year t; and PRF_{ikt} is the PRF-RI indices of ranch i's grid in month k, year t.[9]

We estimate equations (1) and (2) by using an elastic net estimator.[10] OLS does not perform well in models with a relatively large number of explanatory variables, particularly for the out-of-sample

[9] In order to consider the potential impacts of the previous year's precipitation on forage yield through soil moisture carry-over, we included the previous year's monthly precipitation. Note that Oesterheld et al. (2001) find that the previous year's precipitation adds additional explanatory power to the forage production model. The variables would not be selected if they did not have any significant impacts on forage yield.

[10] For the statistical inferences, we estimate the models with selected variables by the elastic net regression via OLS.

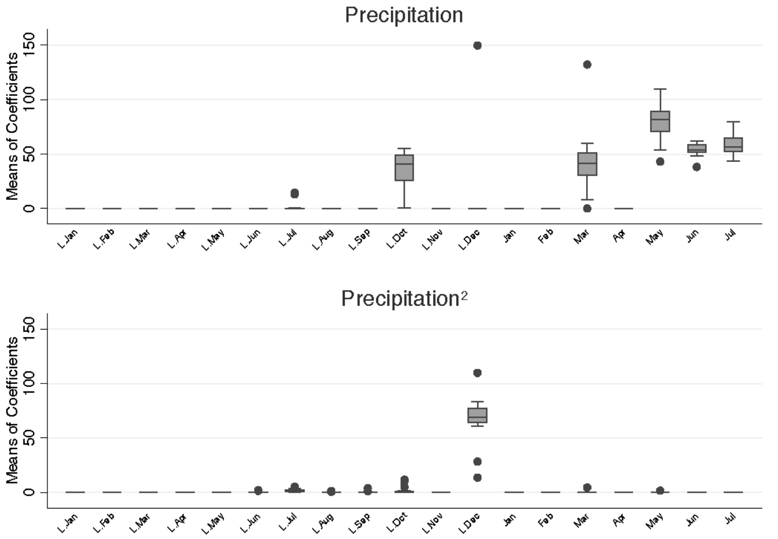

Figure 2. Effects of Monthly Precipitation on Forage Yields: Elastic Net Estimation

predictions (Belloni, Chernozhukov, and Hansen, 2014). Consistency of OLS estimation requires the number of observations to be greater than $p\ln(p)$, where p is the number of explanatory variables (Portnoy, 1984). Equation (1) has $p = 42$, and equation (2) has $p = 40$; both yield a number of observations less than $p\ln(p)$.

As Tibshirani (1996) document, if predictors are highly correlated, then the ridge regression outperforms the lasso regression in terms of minimizing out-of-sample prediction errors. However, the ridge regression is less appropriate for variable selection purposes since the ridge regression always keeps all predictors in the model. Our explanatory variables are highly correlated with one another: the means of variance inflation factors are 142 for equation (1) and 363 for equation (2). Thus, we use the elastic net regularization method developed by Zou and Hastie (2005), which has a better out-of-sample prediction power than the lasso when the predictors are correlated with one another.

Let **Y** be the vector of centered dependent variable and **X** be the vector of standardized explanatory variables after a location and scale transformation. The vector of coefficients is **B**, and p is the number of regressors. Then, the elastic net estimator is

$$\hat{B} = arg\ min_B\{|Y - XB|^2\}$$

(3)

$$\text{subject to } (1-\alpha)\sum_{j=1}^{p}|\beta_j| + \alpha\sum_{j=1}^{p}{\beta_j}^2 \leq s,$$

where α and s are tuning parameters. Note that if $\alpha = 1$, then the estimator is a lasso estimator and if $\alpha = 0$, then the estimator is a ridge estimator.

To find optimal α, we compute the root mean squared errors (RMSEs) using cross-validations. We first split our dataset into i) a training dataset and ii) a test dataset. In order to preserve the within-year correlation of observations, we leave 1 year of the data out and use the remaining years as the training dataset. We then estimate our regression equations using the training dataset. With the estimated coefficients, we obtain out-of-sample predicted values using the test dataset. The out-of-sample RMSE is computed from the test dataset. Finally, we repeat this procedure for each year so that every year serves as a test dataset. We conduct the cross-validation procedure for αs from 0

Table 3. Post-Selection Linear Regression Results: Effects of Monthly Precipitation on Forage Yields

	OLS with Selected Variables: Forage Yield			
Variables	**Ranch-Level Data**		**Field-Level BBR Data, Ranch-Level GSL and Hays Data**	
L. July	−17.8	(81.8)	−30.1	(57.3)
L. October	−1.1	(92.3)	41.5	(64.1)
March	101.9*	(59.3)	83.5*	(42.8)
May	125.9***	(29.6)	148.0***	(22.8)
June	80.0***	(11.1)	98.6***	(10.9)
July	60.4***	(21.4)	55.7***	(19.5)
L. June^2	3.6	(2.6)	4.2**	(1.8)
L. July^2	6.8	(10.9)	7.6	(8.5)
L. October^2	15.9	(16.4)	13.2	(12.9)
L. December^2	69.3*	(36.3)	9.7	(39.5)
Hays	1,265.0***	(123.5)	1,405.1***	(112.0)
Constant	650.2***	(198.2)	482.4***	(140.5)
R^2	0.905		0.824	
No. of obs.	42		118	

Notes: Robust standard errors are in parentheses. Single, double, and triple asterisks (*, **, ***) indicate significance at the 10%, 5%, and 1% level.

to 1 with 0.01 increments. For each α, we estimate the regression equation 17 times. We choose the α that minimizes average RMSE. For both equations (1) and (2), our optimal α are 0.92.

Results

In this section, we discuss the estimation results of equations (1) and (2) using the elastic net.[11] We discuss how our results compare with those found in the range science literature. Then, we decompose basis risk of PRF-RI into i) nonprecipitation risk and ii) index-related risk.

Effects of Precipitation on Forage Yields

We present the estimated results of equation (1) using the elastic net estimator. The optimal α is equal to 0.92, with average RMSE of 352.74. Figure 2 reports the estimated results of equation (1) from the elastic net estimation.

The estimated results are also consistent with the agronomy literature. In Figure 2, the variables selected by the elastic net estimation indicate that rainfall in May, June, and July is positively correlated with forage yields. This is consistent with findings in the range science literature, including Smoliak (1986), Sala et al. (1988), Lauenroth and Sala (1992), and Smart et al. (2007), who note the importance of spring and summer precipitation for forage output.

Table 3 reports post-selection linear regression results of equation (1) with the full sample. We estimate equation (1) only with selected variables via elastic net using OLS. We include the variables

[11] Using the last 3 years as the test dataset and the remaining part as the training dataset, the out-of-sample RMSE of OLS estimation of (1) without variable selection is 1,186 and the out-of-sample RMSE of Elastic Net estimation is 487. We also find that the elastic net estimator provides more reasonable estimates than OLS. We provide the result and the discussion in the appendix.

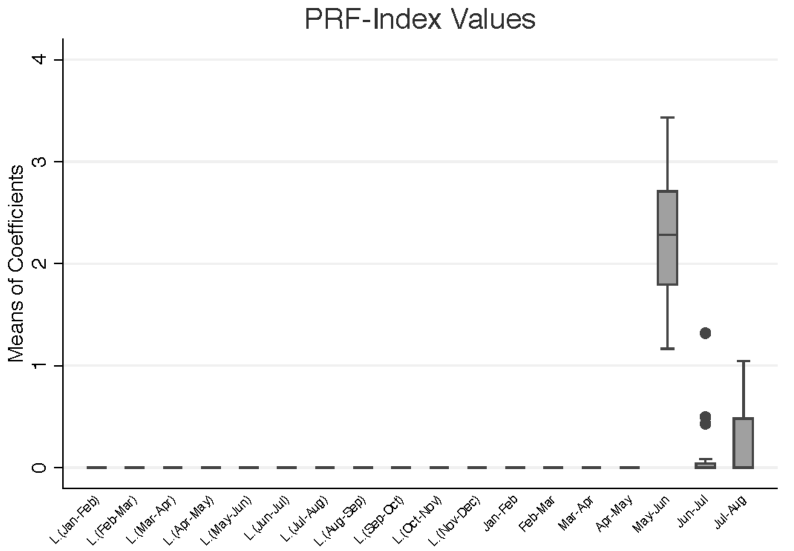

Figure 3. Effects of 2-Month PRF-RI Indices on Forage Yields: Elastic Net Estimation

Notes: None of the squared terms are selected by the elastic net estimation.

Table 4. Post-Selection Linear Regression Results: Effects of 2-Month PRF-RI Indices on Forage Yields

	OLS with Selected Variables: Forage Yield			
Variables	**Ranch-Level Data**		**Field-Level BBR Data, Ranch-Level GSL and Hays Data**	
May–June	7.35***	(1.95)	7.04***	(1.11)
June–July	−0.12	(2.14)	0.54	(1.47)
July–August	3.06*	(1.81)	1.62	(1.31)
Hays	1,444.09***	(145.36)	1,444.95***	(131.03)
Constant	747.70***	(178.60)	850.65***	(88.96)
R^2	0.778		0.672	
No. of obs.	42		118	

Notes: Robust standard errors are in parentheses. Single, double, and triple asterisks (*, **, ***) indicate significance at the 10%, 5%, and 1% level.

that are selected more than twice by elastic net regressions in the cross-validation process described above. Consistent with Figure 2 and the range science literature, we find significant and positive effects of rainfall in March, May, June, and July.[12] Results using both sets of data indicate that precipitation during the May–June period is most important.

Effects of PRF-RI Indices on Forage Yields

We now examine how well the 2-month interval PRF-RI indices explain forage yield. Similar to the relationship between precipitation and forage yields, we estimate equation (2) with the elastic

[12] Consistent with the range science literature, rainfall in the previous fall or winter has very little effect. The results show little evidence on the effect of soil moisture carry-over.

Table 5. Estimation of False Negative Probability

	Ranch-Level Data	Field-Level BBR Data, Ranch-Level GSL and Hays Data
Precipitation model	21%	33.9%
PRF-RI index value model	26%	42.7%
Difference	5%	8.8%
No. of obs.	42	118

net estimator. The optimal α is equal to 0.92 with average RMSE of 481.58. For cross-validation, we estimate the regression equation 17 times and report the average of the estimated coefficients. Figure 3 presents the results from the elastic net estimator. Similar to the estimation results of equation (1), the results are consistent with the agronomy literature: The PRF-RI index for the May–June interval has a positive impact on forage yield.

Table 4 presents the results from post-selection linear regressions using the full sample. The results are consistent with Figure 3. Again, we include the covariates that are selected more than twice in the cross-validation process we described above. Table 4 indicates that the May–June interval's index has a significant and positive effect.[13]

Decomposition of Basis Risk

We first compute false negative probability (*FNP*), which we define (similar to Elabed et al., 2013) as a measure of basis risk:[14]

$$FNP = Prob(\hat{y}_{it} > \bar{y}_i | y_{it} < \bar{y}_i) = \frac{Prob(\hat{y}_{it} > \bar{y}_i,\, y_{it} < \bar{y}_i)}{Prob(y_{it} < \bar{y}_i)} \tag{4}$$

where $\hat{y}_{it}$ is the predicted yield, $\bar{y}_i$ is the historical average yield, and y_{it} is the realized yield of ranch i in year t.

We estimate FNP with the following steps: First, we exclude 1 year of data, year t, and estimate each model with selected explanatory variables via elastic net. We then obtain the predicted yield for the excluded year, $\hat{y}_{it}$. We also compute the average yield for each ranch, $\bar{y}_i$ using all years except year t. We repeat these steps for all years and empirically compute the joint and marginal probabilities to obtain equation (4). We estimate FNPs for the yield–precipitation model and the yield–PRF-RI index value model.

Table 5 reports the estimated false negative probabilities. We find that there is 21% probability of the predicted yield being greater than the historical average when the true yield is actually less than the historical average, which is our measure of nonprecipitation risk. In other words, if there exists an index insurance product that uses the selected variables of table 3, the probability of the insurance payment not being triggered when the forage yield is lower would be approximately 21%.

If the yield is predicted by using the selected variables of table 4, the probability of predicted yield being greater than the historical average when the true yield is less than the historical average is higher, approximately 26%. This is the conservative measure of overall basis risk. The estimates

[13] Similar to Table 3, we also estimated same model with field-level data of BBR ranch. The results remain robust.

[14] Elabed et al. (2013) define false negative probability as

$$Prob(I_c > I_T | y_i < y_{iT}),$$

where I_c is the realized index value of an index insurance for area c, I_T is the trigger value of the index insurance, y_i is the realized individual outcome of individual i, and y_{iT} is the desired level of protection for individual i.

indicate that using monthly and site-level precipitation would decrease false negative probability by 5%–9%, which is our measure of index-related risk.

The results of Table 5 indicate that the overall basis risk is moderate. More importantly, index-related risk, which stems from the use of 2-month intervals and grid-level indices, is relatively small. At least for producers in the Midwest, the basis risk of the PRF-RI program is mostly due to nonprecipitation factors rather than the structure of the PRF-RI indices.

Conclusions and Policy Implications

We estimate the basis risk of the PRF-RI program, which is one of two crop insurance programs available to forage producers. Since we have a relatively large number of correlated explanatory variables compared to the number of observations, we utilize a regularization method, the elastic net penalty developed by Zou and Hastie (2005), to avoid overfitting and properly select explanatory variables that minimize out-of-sample prediction errors compared to lasso or ridge regressions. The estimated overall basis risk is moderate, approximately 26%. We find that most of the basis risk in the PRF-RI program is from nonprecipitation factors that affect forage yields.

The estimated index-related basis risk is small, 5%–9%. That is, using more flexible contract forms with site-level precipitation would have little impact on decreasing the degree of basis risk. This addresses a significant concern for producers: For the locations we examined, there was only a small amount of additional risk from using PRF-RI indices, compared to using rainfall on one's own ranch. Given that reporting stations in north-central and northwest Nebraska are more sparse or distant than in many other areas covered by the PRF-RI program, our estimates of index-related basis risk serve as upper bounds, suggesting that areas with more densely located stations face lower basis risk. However, studies in other locations would be useful to compare with this particular finding.

Our estimation results also suggest that monthly rainfall in May and June has significant and positive effects on forage yields. Consistent with this yield and actual monthly precipitation relationship, we find that the May–June PRF-RI index value is the only index value with a significant and positive effect on forage yield. The results are consistent with the agronomy literature (e.g., Smoliak, 1986; Sala et al., 1988; Lauenroth and Sala, 1992; Smart et al., 2007).

Understanding the basis risk of the PRF-RI program is crucial for policy makers to evaluate the impact of the program on production risk reduction and, thus, the cost-effectiveness of the program. Further research into geographical differences in the relationship between forage yield and monthly precipitation and the degree of basis risk would elaborate on how much forage production risk the PRF-RI program can reduce. Also, exploring how participation patterns across geographical regions depend on the degree of basis risk would help policy makers evaluate the PRF-RI program.

[Received March 2018; final revision received July 2018.]

References

Barnett, B. J., C. B. Barrett, and J. R. Skees. "Poverty Traps and Index-Based Risk Transfer Products." *World Development* 36(2008):1766–1785.

Barnett, B. J., and O. Mahul. "Weather Index Insurance for Agriculture and Rural Areas in Lower-Income Countries." *American Journal of Agricultural Economics* 89(2007):1241–1247.

Belloni, A., V. Chernozhukov, and C. Hansen. "High-Dimensional Methods and Inference on Structural and Treatment Effects." *Journal of Economic Perspectives* 28(2014):29–50.

Binswanger-Mkhize, H. P. "Is There Too Much Hype about Index-Based Agricultural Insurance?" *Journal of Development Studies* 48(2012):187–200.

Chantarat, S., A. G. Mude, C. B. Barrett, and M. R. Carter. "Designing Index-Based Livestock Insurance for Managing Asset Risk in Northern Kenya." *Journal of Risk and Insurance* 80(2013):205–237.

Clarke, D. J. "A Theory of Rational Demand for Index Insurance." *American Economic Journal: Microeconomics* 8(2016):283–306.

Diersen, M., P. Gurung, S. Fausti et al. "Optimal Allocation of Index Insurance Intervals for Commodities." 2015. Paper presented at the annual meeting of the Southern Agricultural Economics Association, Atlanta, Georgia, January 31–February 3.

Elabed, G., M. F. Bellemare, M. R. Carter, and C. Guirkinger. "Managing Basis Risk with Multiscale Index Insurance." *Agricultural Economics* 44(2013):419–431.

Elabed, G., and M. R. Carter. "Compound-Risk Aversion, Ambiguity and the Willingness to Pay for Microinsurance." *Journal of Economic Behavior & Organization* 118(2015):150–166.

Giné, X., R. Townsend, and J. Vickery. "Patterns of Rainfall Insurance Participation in Rural India." *World Bank Economic Review* 22(2008):539–566.

Glauber, J. W. "The Growth of the Federal Crop Insurance Program, 1990–2011." *American Journal of Agricultural Economics* 95(2013):482–488.

Heitschmidt, R., and L. Vermeire. "Can Abundant Summer Precipitation Counter Losses in Herbage Production Caused by Spring Drought?" *Rangeland Ecology & Management* 59(2006):392–399.

Ifft, J., S. Wu, T. Kuethe et al. "The Impact of Pasture Insurance on Farmland Values." *Agricultural and Resource Economics Review* 43(2014):390–405.

Jensen, N., and C. Barrett. "Agricultural Index Insurance for Development." *Applied Economic Perspectives and Policy* (2016):199–219.

Jensen, N. D., C. B. Barrett, and A. G. Mude. "Index Insurance Quality and Basis Risk: Evidence from Northern Kenya." *American Journal of Agricultural Economics* 98(2016):1450âĂŞ–1469. doi: 10.1093/ajae/aaw046.

Lauenroth, W., and O. E. Sala. "Long-Term Forage Production of North American Shortgrass Steppe." *Ecological Applications* 2(1992):397–403.

Mahul, O. "Optimum Area Yield Crop Insurance." *American Journal of Agricultural Economics* 81(1999):75–82.

Maples, J. G., B. W. Brorsen, and J. T. Biermacher. "The Rainfall Index Annual Forage Pilot Program as a Risk Management Tool for Cool-Season Forage." *Journal of Agricultural and Applied Economics* 48(2016):29–51.

Miranda, M. J. "Area-Yield Crop Insurance Reconsidered." *American Journal of Agricultural Economics* 73(1991):233–242.

Miranda, M. J., and K. Farrin. "Index Insurance for Developing Countries." *Applied Economic Perspectives & Policy* 34(2012):391–427.

Nadolnyak, D., and D. Vedenov. "Information Value of Climate Forecasts for Rainfall Index Insurance for Pasture, Rangeland, and Forage in the Southeast United States." *Journal of Agricultural and Applied Economics* 45(2013):109–124.

National Drought Mitigation Center. *Managing Drought Risk on the Ranch: A Planning Guide for Great Plains Ranchers*. Lincoln, Nebraska: University of Nebraska-Lincoln, National Drought Mitigation Center, 2012. Available online at https://drought.unl.edu/ranchplan/.

Norton, M. T., C. Turvey, and D. Osgood. "Quantifying Spatial Basis Risk for Weather Index Insurance." *Journal of Risk Finance* 14(2012):20–34.

Oesterheld, M., J. Loreti, M. Semmartin, and O. E. Sala. "Inter-Annual Variation in Primary Production of a Semi-Arid Grassland Related to Previous-Year Production." *Journal of Vegetation Science* 12(2001):137–142.

Portnoy, S. "Asymptotic Behavior of M-Estimators of p Regression Parameters When p2/n is Large. I. Consistency." *Annals of Statistics* (1984):1298–1309.

Sala, O., W. Lauenroth, and W. Parton. "Long-Term Soil Water Dynamics in the Shortgrass Steppe." *Ecology* 73(1992):1175–1181.

Sala, O. E., W. J. Parton, L. Joyce, and W. Lauenroth. "Primary Production of the Central Grassland Region of the United States." *Ecology* 69(1988):40–45.

Smart, A. J., B. H. Dunn, P. S. Johnson, L. Xu, and R. N. Gates. "Using Weather Data to Explain Herbage Yield on Three Great Plains Plant Communities." *Rangeland Ecology & Management* 60(2007):146–153.

Smith, R., M. Panciera, and A. Probst. "Using a Grazing Stick for Pasture Management." Extension Bulletin AGR-191, Lexington, KY, 2010.

Smith, V. H., H. H. Chouinard, and A. E. Baquet. "Almost Ideal Area Yield Crop Insurance Contracts." *Agricultural and Resource Economics Review* 23(1994):75–83.

Smith, V. H., and J. W. Glauber. "Agricultural Insurance in Developed Countries: Where Have We Been and Where Are We Going?" *Applied Economic Perspectives and Policy* (2012):363âĂŞ–390.

Smoliak, S. "Influence of Climatic Conditions on Production of Stipa-Bouteloua Prairie over a 50-Year Period." *Journal of Range Management* (1986):100–103.

Tibshirani, R. "Regression Shrinkage and Selection via the Lasso." *Journal of the Royal Statistical Society. Series B (Methodological)* (1996):267–288.

U.S. Department of Agriculture. *Census of Agriculture 2012*. Washington, DC: U.S. Department of Agriculture, National Agricultural Statistics Service, 2014.

———. *Summary of Business*. Washington, DC: U.S. Department of Agriculture, Risk Management Agency, 2015. Available online at http://www.rma.usda.gov/data/sob.html.

Vedenov, D. V., and B. J. Barnett. "Efficiency of Weather Derivatives as Primary Crop Insurance Instruments." *Journal of Agricultural and Resource Economics* (2004):387–403.

Westerhold, A., C. Walters, K. Brooks, M. Vandeveer, J. Volesky, and W. Schacht. "Risk Implications from the Selection of Rainfall Index Insurance Intervals." *Agricultural Finance Review* (2018):514–531.

Woodard, J. D., and P. Garcia. "Basis Risk and Weather Hedging Effectiveness." *Agricultural Finance Review* 68(2008):99–117.

Zou, H., and T. Hastie. "Regularization and Variable Selection via the Elastic Net." *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 67(2005):301–320.

Appendix A: OLS Regression Result

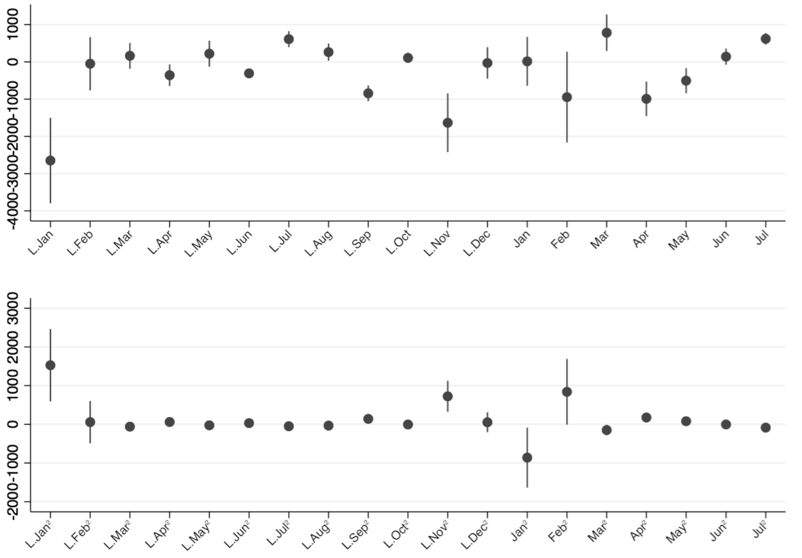

Figure A1. Estimation Result of Equation (1) via OLS Using Field-Level Data

Notes: The figure only reports the estimated coefficients and their standard errors of precipitation and squared precipitation variables ($N = 118$).

Figure A1 reports the estimated coefficients and their standard errors of precipitation and squared precipitation variables estimated by the OLS regression without variable selection. We find that the result from elastic net regression is more consistent with agronomic theory and previous studies compared to the result of Figure A1. For example, the negative coefficients of April and May are inconsistent with agronomic theory. Also, as we describe in the main text, we find that elastic net performs better in terms of out-of-sample predictions.