Flexible weather index-based insurance design

Sarah Conradt a,⇑, Robert Finger b, Martina Spörri a

a ETH Zurich, Institute for Environmental Decisions, Agricultural Economics Group, Zurich, Switzerland
b University of Bonn, Institute for Food and Resource Economics, Production Economics Group, Germany

A R T I C L E   I N F O

Article history:
Available online 30 June 2015

Keywords:
Agricultural risk management
Weather index-based insurance
Index design
Downside risk
Quantile Regression

A B S T R A C T

This study investigates the performance of a flexible index design for weather index-based insurances using farm-level panel data on wheat production from Kazakhstan. The proposed flexible design is a generic framework that uses Growing Degree Days to determine annual variable start and end dates for the insured period. This approach reflects the progress of phenological plant growth phases more accurately than fixed periods and hence is expected to reduce the basis risk of the index insurance. In addition, we develop an economic framework that focuses on the role of downside risks and apply Quantile Regression to tailor optimal insurance specifications. This framework is then used to compare the downside risks associated with the use of flexible and fixed insurance periods. The results show that the introduction of flexibility in the index design leads to a reduction in farmers’ downside risk exposure and to a more efficient contract design.

© 2015 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Introduction

Agriculture production depends heavily on weather conditions. While idiosyncratic adverse weather impacts can be managed through informal measures, systemic weather events require formal risk management approaches such as insurances. Traditional indemnity based insurances however, are plagued by asymmetric information problems and high transaction costs (Chambers, 1989). Index-based insurance products overcome these challenges by conditioning the payout not on actual yield losses but on the realization of an independent and transparent index. For weather index-based insurances, the index is designed to adequately reflect weather conditions, which decisively influence crop yields or better crop yield reductions in a specific region 2. However, the discrepancy between index and crop yield loss leads to a residual risk borne by the insured farmer that is referred to as basis risk (Woodard and Garcia, 2008).

Along these lines, it has been shown that the basis risk decreases with a higher correlation between the chosen weather index and crop yield, which implies a higher effectiveness and higher potential up-take of these risk transfer products (Fuchs and Wolff, 2011). Improving the index insurance design is therefore a key challenge to achieve the potential benefits from weather index based insurances (Fuchs and Wolff, 2011; Woodard and Garcia, 2008; Norton et al., 2013). Most studies use a simple index based either on rainfall or temperature, summing up the weather information within the main vegetation period of a specific crop in a specific region (cf. Turvey, 2001; Martin et al., 2001; Barnett and Mahul, 2007; Berg and Schmitz, 2008; Kellner and Musshoff, 2011; Daron and Stainforth, 2014). However, these studies rely on indices that are based on

⇑ Corresponding author.
E-mail addresses: conradts@ethz.ch (S. Conradt), r.finger@ilr.uni-bonn.de (R. Finger).

1 The main challenges of asymmetric information problems are moral hazard and adverse selection. Both may lead to insurance market failure.

2 Note that alternative approaches focus on the use of other index sources such as remote sensing data (see e.g. Hochrainer-Stigler et al., 2014, for details).

http://dx.doi.org/10.1016/j.crm.2015.06.003
2212-0963/© 2015 The Authors. Published by Elsevier B.V.
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
fixed calendar dates. This means that each year the weather information is calculated in the exact same period, i.e. the yearly start and end days are identical (e.g. often chosen are start and end dates of months).

The fixed calendar date approach, however, neglects important information. First, it does not consider that the vulnerability of a specific crop changes across phenological phases (Nairizi and Rydzewski, 1977). Second these critical phases are in actual crop growth not constant with regard to calendar times but depend on external weather conditions. It has been acknowledged in previous research that the interaction between damaging weather event and the specific plant growth stage is important and that this information has to be incorporated in the index design (cf. Nieto et al., 2010). More specifically, Nieto et al. (2010) describe a procedure that divides the growing period into 10 days intervals and weights these decades by the sensitivity of the crop yield towards the index. Even though this approach integrates the plant sensitivity during different phenological phases, the specification of the considered decades is fix over the years, i.e. start and end dates remain unchanged.

The explicit consideration of plant growth stages in the index-based insurance literature is rare. For instance, Leblois et al. (2013) use observed and modeled sowing information to incorporate the timing of biological processes in an insurance application. Kapphan et al. (2012) extract information on phenological plant phases from a crop model and use Growing Degree Days3 (GDD’s) for index construction, however applying only simulated data. Not focusing on insurance applications, Meyer et al. (1993) developed a drought index for corn by relying on GDD’s to detect the progress of plant growth throughout the growing season. Their results show that using GDD’s indeed could contribute to a better explanation of the vulnerability of plant growth to drought events. Thus, a generic approach to describe crop growth stages flexibly across years based on GDDs tends to be superior to an approach using fixed dates to specify potentially critical phases.

Based on this background, we expect a stronger dependency between the weather index and the actual crop yield by using flexible time periods (defined through GDD’s) instead of fixed calendar dates to specify the index. Considering this element in the design of an index could result in transparent insurance solutions that reduce the basis risk and thus potentially improve the effectiveness and uptake of index-based insurance solutions. This link has, however, not been addressed in the literature so far.

The here presented study extends previous research in several directions. First, we develop a GDD concept to determine annual start and end dates of a cumulative index. Second, we compare and quantify potential benefits of this approach to the standard approach where fixed calendar dates are used to specify the index accumulation period. Third, we establish a consistent downside risk perspective, which aims to reduce the probability of exceeding certain loss thresholds from index design to risk evaluation of the index-based insurance contract. Our empirical application focuses on wheat farms in northern Kazakhstan.

The remainder of the article is structured as follows: In Section “Methodology and empirical procedure”, we present our methodology. More precisely, this section introduces the index insurance framework, the expected utility approach to evaluate our results and the procedure how to derive two indices: a flexible with varying yearly calendar dates and a fixed index with constant yearly dates. This section also introduces the case study on wheat production in Kazakhstan and presents the data. The results are shown in Section “Results”, followed by the discussion of the results and the conclusion in Section “Discussion and conclusion”.

Methodology and empirical procedure

In a first subsection, we explain the index insurance framework applied in our analysis and derive an expected utility approach based on mean and semi-variance. We adapt this approach to our objective function in the second subsection. In the third subsection, we describe the GDD concept and the index design approach. In the fourth subsection, we present our case study and explain the empirical procedure.

Conceptual framework

Index-insurance framework

Consider a farmer who is confronted with production risks due to varying weather conditions. Crop yield $y$ is expressed as a function of weather, represented by a weather index $WI$. Additional factors that affect crop yields but are uncorrelated with $WI$, are summarized within the term $\varepsilon$.

$$y = g(WI) + \varepsilon$$  \hspace{1cm} (1)

We assume that farmers are risk averse and purchase a weather index insurance to reduce their risk exposure. Following Deng et al. (2007), the farmer is assumed to hold only two assets: the production of one single crop and the index insurance contract, which is a valid assumption as our sample consists of Kazakh farmers solely focusing their production on wheat. In addition, wheat grown in northern Kazakhstan is characterized by low-cost, low-input and rain-fed agriculture, which justifies the simplifying assumption that production inputs are independent of weather realizations (cf. also Turvey, 2001).

3 GDD is a measure of heat accumulation and reflects the importance of temperature as a crucial driver of plant growth (McMaster and Wilhelm, 1997).
For index-based insurances, the payout \( PO \) is conditional on the weather realization and, designed as a put option, defined as:

\[
PO = \gamma \cdot \max \{0, (S - WI)\},
\]

where \( \gamma \) is the tick size, which is the incremental change of \( PO \) for a change of \( WI \) and is, following Woodard and Garcia (2008), normalized to unity. \( S \) is called the strike level, i.e. the threshold value at which the insurance payout is triggered. We define the strike as \( S = g^{-1}(0.7 \cdot E(y)) \), with \( E \) being the expectation operator. This approach allows expressing the strike level in terms of yield observations, while still avoiding asymmetric information problems as payouts are solely triggered by realizations of \( WI \), and not by direct yield observation as in the case of standard yield insurance. The income \( \pi \) in year \( t \) is then given by:

\[
\pi_t = y_t + PO_t - P,
\]

where \( P \) is the premium, which we assume to be actuarily fair, thus fulfilling \( P = E(PO) \). Note that we express \( \pi \) not in monetary terms but is standardized to be expressed in yield units (see also Leblois and Quirion, 2013).

Using a regression framework to deterministically link the influence of a random weather variable to yields, we aim to minimize the error term \( \varepsilon \) of Eq. (1). Fitting an appropriate relationship between yield and weather index is fundamental because \( \varepsilon \) reflects the unexplained variance and represents the basis risk of an index-based insurance contract.

Mean-(semi-)variance approaches

To evaluate farmer’s decisions under risk, we apply the expected utility (EU) approach. Accordingly, farmer’s preferences are described by a concave utility function \( U \), with \( U' (\pi) > 0 \) and \( U'' (\pi) < 0 \) where \( U' \) and \( U'' \) indicate the first and second derivative (Scott and Horvath, 1980). Using a mean–variance framework, the maximization of \( EU \) can be approximated as:

\[
EU(\pi) \approx U(\mu_\pi) + 0.5 \cdot \sigma^2_\pi \cdot U''(\mu_\pi),
\]

where \( \mu_\pi = E(\pi), \sigma^2_\pi = V[\pi - E(\pi)]^2 \). The subscript \( V \) indicates that we focus on the variance of yields. Farmers’ preferences for higher moments of the distribution could be depicted as well by adding further moments using a Taylor series expansion (Scott and Horvath, 1980; Levy and Markowitz, 1979). The focus on the first two moments of the profit distribution is however a valid assumption if and only if (a) these two moments are sufficient to describe the profit distribution, i.e. no higher moments (e.g. skewness or kurtosis) exist, and/or (b) the farmer has no preferences for these higher moments, i.e. derivatives of the utility function of higher order than two are zero. This is, for instance, the case for a quadratic utility function.²

In the context of agricultural insurance, however, these assumptions have been challenged for at least two reasons. First, crop yields and associated profits are often not normally distributed, but are characterized by non-normal and non-symmetrical distributions with negative skewness (e.g. Hennessy, 2009; Turvey, 2001). Second, farmers have been found to have preferences for higher moments of profit distributions. In particular, they are concerned about downside risks, i.e. risks of experiencing losses exceeding certain thresholds (e.g. Di Falco and Chavas, 2006; Groom et al., 2008). In terms of utility, this is expressed with a positive third derivative \( U'''(\pi) > 0 \) (Menezes et al., 1980).

Based on this background, the exclusive focus on the variance of profits has been seen as too restrictive and the use of semi-variance approaches has been suggested (Miranda, 1991; Miranda and Glauber, 1997; Markowitz, 1991; Vedenov and Barnett, 2004). The use of semi-variance implies that only losses that are below a specific benchmark are described by a concave utility function

\[
EU(\pi) \approx U(\mu_\pi) + 0.5 \cdot \sigma^2_\pi \cdot U''(\mu_\pi).
\]

The subscript \( SV \) now indicates that we focus on the semi-variance of yields. If the distribution would be symmetric, expected utility expressed in Eqs. (4) and (5) would be identical because \( 2 \cdot SV_\pi (B = E(\pi)) = \sigma^2_\pi \). In contrast, expected utility derived from both approaches differs if the underlying distribution is skewed (Estrada, 2004). More specifically, if the profit distribution is negatively skewed, i.e. \( E(\pi - E(\pi))^3 < 0 \) implying that \( 2 \cdot SV_\pi (B = E(\pi)) > \sigma^2_\pi \), the expected utility derived from the semi-variance approach will be smaller. In contrast, \( EU_{SV} < EU_V \). However, positive skewness results in a higher expected utility, because the larger variability above the expected yield is not perceived as a loss for the farmer. The concept of semi-variance thus allows representing downside risk in a simple two-parametric framework.³

---

² We choose the level of 70% of the expected profit – calculated as expected value based on past observations – as indicator for severe loss, for instance, following the proposal for an income stabilization tool of the European Union (Dell’Aquila and Cimino, 2012).

³ See Markowitz (1991) for an overview of utility functions that are well approximated with the mean–variance framework.

⁶ The semi-variance approach can be also closely related to stochastic dominance concept (Porter, 1974; Bawa, 1978; Nawrocki, 1991).
Quantile Regression and implementation

Next, the regression approach that is used to establish the yield-index dependency shown in Eq. (1) is presented. The chosen approach aims to derive a consistent downside risk perspective. Instead of using a mean-conditioning framework, such as Ordinary Least Squares, we derive the yield-index dependency by using Quantile Regression (QR), which leads to the following minimization problem:

$$\hat{\beta}(\tau) = \arg\min_{\beta \in \mathbb{R}} \left( \tau \cdot \sum_{y_t < x_T^* \beta} |y_t - x_T^* \beta| + (1 - \tau) \cdot \sum_{y_t > x_T^* \beta} |y_t - x_T^* \beta| \right),$$  \hspace{1cm} (6)

where $\tau \in (0, 1)$. According to Eq. (6), QR minimizes the sum of absolute residuals, which are asymmetrically weighted. The weighting factor depends on the sign of the residuals: positive residuals receive a weighting factor of $\tau$, negative residuals are weighted by $(1 - \tau)$ (Koenker, 2005). Besides being robust against outliers, this approach allows to specify $\tau \leq 0.5$ and thus to identify the differential impact of $WI$ on $y$ in the lower tails. This is in line with farmer’s concern on below-average outcomes outlined above. For our analysis, we specify $\tau = 0.3$, to adequately reflect low yield observations while still moderating the influence of single, very low yield values (Conradt et al., 2015; Conradt, 2014).

To be consistent with QR’s objective function (Eq. (6)), i.e. to minimize absolute residuals, we reformulate Eq. (5), by replacing the semi-variance with (absolute) semi-deviation $SDV$ (Konno and Yamazaki, 1991; Speranza, 1993):

$$EU(\pi)_{SDV} \approx U(\mu_n) + 0.5 \cdot SDV_n(B)^2 \cdot U'(\mu_n),$$  \hspace{1cm} (7)

where $SDV$ is here defined as $SDV_n(B) = E[|\min((\pi - B), 0)|]$, with $B$ being equal to the strike, i.e. $B = g^{-1}(0.7 \cdot E(y))$.

Following Deng et al. (2007), we base our analysis on an exponential utility function, $U = 1 - e^{-\pi}$, which allows to incorporate negative values.\(^7\) We specify the Arrow–Pratt risk aversion coefficient as $r_a = 0.1$. (Babcock et al., 1993; Raskin and Cochran, 1986).\(^8\)

Index design

The hypothesis underlying our analysis is that the incorporation of flexibility in terms of a yearly varying index accumulation period allows to increase the explanatory power of Eq. (1), and to reduce the term $\epsilon$. The flexibility arises due to GDD values, which are used to approximate the time of critical plant phases. We evaluate our approach by comparing results of Eq. (7) for (a) an index calculated over a varying time period for every year (referred to as FLEX) and (b) an index calculated over a fixed period for all years (referred to as FIX). In the following subsections, we provide the concept underlying the use of GDD’s and describe the general design of the FLEX and FIX indices.

General concept of GDD’s

A plant is expected to have varying sensitivity towards stress during the growing cycle (Nairizi and Rydzewski, 1977; Leblois and Quirion, 2013). Critical phases for wheat plants are for instance around tillering, shooting and ear emergence (Acevedo et al., 2002). These critical phases are not constant with regard to calendar dates but are shifted by varying weather conditions. Daily air temperature is an important factor of growth due to its effect on the enzymatic activity (Bonhomme, 2000). Hence, thermal units or so-called GDD’s can be used to approximate development rates of plants (Slafer and Savin, 1991). However, low and high temperature values limit plant growth and often lower and upper limits are assumed (“cut-off method”, Neild and Newman, 1987). Based on this background, GDDs are determined using $H_n^{up}$, $H_n^{av}$ and $H_n^{base}$:

$$GDD = \sum_{n=1}^{N} \max(\min\{H_n^{av}, H_n^{up}\} - H_n^{base}, 0)$$  \hspace{1cm} (8)

where $n = 1..N$ are the days in the growth period and $H_n^{up}$ is the maximal temperature, above which additional degrees in temperature do no longer accelerate plant development. $H_n^{av}$ is the average daily air temperature that is calculated as average of the observed daily minimum ($H_n^{min}$) and maximum temperature ($H_n^{max}$), $H_n^{av} = \frac{H_n^{max} + H_n^{min}}{2}$ and $H_n^{base}$ represents the base or threshold air temperature, assuming that below this value wheat stops its growth activity.

Note that we use the GDD concept not as an index but to more precisely determine the time of plant phases. This temporal information is applied to define the accumulation period of a cumulative rainfall index.

\(^7\) According to Babcock et al. (1993) an exponential utility is often used in the insurance context and, in addition, allows to incorporate negative end of season values (revenue minus insurance premium), which occurred in a few cases in our analysis.

\(^8\) We varied the values for $r_a$ in a sensitivity analysis, which has not affected the qualitative interpretation of our results.
Index specification

For each plant a specific amount of heat units (GDD’s) is required to advance in the plant growing cycle. The thermal requirements, which correspond to the start and end value of a critical plant phase, are referred to as GDD\textsubscript{start} and GDD\textsubscript{end} (estimated using Eq. (8)). These GDD\textsubscript{start} and GDD\textsubscript{end} values are applied to determine for each year \( t \) the flexible start dates \( D_{\text{FLEX}}(t) \) and end dates \( D_{\text{FLEX}}(t) \), which are used to specify the time period over which the daily rainfall \( R(n,t) \) is summed up. Accordingly, the FLEX index is defined as:

\[
FLEX(t) = \frac{1}{N(t)} \sum_{D_{\text{FLEX}}(t)} R(n,t). \tag{9}
\]

Depending on the temperature gradient, the length of the accumulation period varies between the years. Hence, we normalize the index by dividing by the effective number of days of the yearly period, \( N(t) = D_{\text{FLEX}}(t) - D_{\text{FLEX}}(t) \).

The FIX index is calculated by using the average of the flexible start and end dates, \( D_{\text{FIX}} = \frac{\sum_{t=1}^{T} D_{\text{FLEX}}(t)}{T} \) and \( D_{\text{FIX}} = \frac{\sum_{t=1}^{T} D_{\text{FLEX}}(t)}{T} \), where \( T \) is the total number of years. We determine:

\[
FIX(t) = \frac{1}{N(t)} \sum_{D_{\text{FIX}}(t)} R(n,t). \tag{10}
\]

Since FIX is accumulated over a fixed period, \( N \) has the same value for each year and is hence a constant. Comparing Eqs. (9) and (10) show that the only difference between the two indices FLEX and FIX is the timing of the accumulation period and that both indices are directly comparable to each other.

In the following section, we present our case study Kazakhstan and thereafter derive the empirical procedure to determine the indices FLEX and FIX.

Data and case study

Agriculture is an important sector for Kazakhstan, both, in terms of economic development and employment (World Bank, 2012). Moreover, as a large wheat exporter, the country also contributes to global grain production and ranks among the top 10 exporters worldwide (FAO, 2013). The vast steppe-like areas in northern Kazakhstan are the main production areas, and are based on a low-input and rain-fed agriculture. Droughts represent the main risk for agricultural production. Spring wheat is the major grain produced in the country and sown around May 25, confined by late spring and early autumn frost (World Bank, 2012; Doraiswamy et al., 2002). Our data set includes wheat yields of 47 single farms of 5 different countries (Rayons), producing solely wheat (Table 1)

9. Following a survey from Heidelbach (2007), a majority of farmers would like to purchase an insurance product due to the high degree of specialization (hence risky production) and due to the challenging climatic conditions.

An overview on the study region is shown in Fig. A.1 in the Appendix. Data are provided by the regional statistical offices and comprises the years 1980–2009. We detrend these farm yields using a robust regression technique (MM-estimator) to remove farm-specific technological trends (Finger, 2010).

Beside farm-level yield data, we use weather information, namely daily temperature and precipitation values, from 5 weather stations (one station per county), which were provided by the National Hydro-Meteorological Agency of Kazakhstan.

As can be seen in Table 1, wheat yields are subject to a strong annual variability. This is (mainly) due to a continental climate with cold winters, hot summers and widespread, extreme droughts, which is the major abiotic stress factor of wheat. Average temperature at sowing is around 15 °C and 19 °C during the growing season (mid May to the beginning of September). However, on single days, air temperatures may exceed 40 °C. Average yearly precipitation values in these regions are around 300 mm. Throughout the paper, our analysis is based on single farm estimations due to the high heterogeneity among farms11.

Empirical procedure

We use data from a Kazakh field experiment that provides the average time lengths of different wheat phases in days as well as the average sowing date (Fig. 1) (see also Conradt, 2014). Especially important for wheat plants are the phases tillering, shooting and ear emergence (i.e. phases 2–4) (Acevedo et al., 2002; Blum et al., 1990; Wollenweber et al., 2003) and hence, we account for this time period for the index design. Field experiments confirm that wheat plants are particularly sensitive and vulnerable towards water stress some days before ear emergence (Fischer, 1973). In addition, the phase from shooting to ear emergence is characterized by very high water requirements with an optimal value of 5.3 mm/day (see

11. Note that due to the lack of differences in the explanatory variables across individual farm-level regressions, there is no gain of using joint estimation approaches such as seemingly unrelated regression.
Fig. 1). This is also due to a continuous rise of air temperature from beginning of June, which increases evapotranspiration rates and drought stress. Often, GDD values for the different plant phases are known from regional field experiments or are provided by breeders. However, for Kazakhstan only the average calendar dates of wheat phases and not the GDD values are known. Hence, we have to approximate the GDD values for the wheat phases first. Our procedure is illustrated in Fig. 2. Based on the given average calendar dates, we estimate the corresponding yearly GDD values (step 1, Fig. 2). Computing the average over all years, results in GDD values per phase (step 2). For the seeding date (i.e. GDD_{start} for phase 1) we fix May, 25 for all farms. As the critical plant growth phases 2–4 are successive, we can define the overall period by using a GDD_{start} value, which corresponds to the GDD value at the beginning of phase 2, and a GDD_{end} value, which represents the GDD values at the end of phase 4 (Fig. 2).

Furthermore, the temperature boundaries $H_{\text{base}}$ and $H_{\text{up}}$ have to be determined in order to apply Eq. (7). For wheat, these boundaries can be set to 5 °C for $H_{\text{base}}$ and 30 °C for $H_{\text{up}}$ (Neild and Newman, 1987).\footnote{Similar GDD values are indicated by other studies (e.g. Neild and Newman, 1987; Slafer and Savin, 1991).} With this information, $D_{\text{FLEX}}(t)$, $D_{\text{FIX}}(t)$, $D_{\text{FLEX}}$, and $D_{\text{FIX}}$, and subsequently the indices FLEX and FIX are determined for every year and for each of the 5 counties (step 3 and step 4 in Fig. 2). These indices are then used to specify Eq. (6) for every single farm, and to evaluate the performance in terms of expected utility, $E_{\text{USDV}}$ (Eq. (7)).

\footnote{Temperature above 30 °C may not only decelerate growth, but may massively harm plants and even provoke plant sterility (Owen, 1971; Saini and Apinall, 1981).}

**Table 1**


<table>
<thead>
<tr>
<th>Region</th>
<th>County</th>
<th>Number of farms</th>
<th>Yields</th>
<th>Farm-level area under wheat</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Mean dt/ha\textsuperscript{a}</td>
<td>Min. dt/ha\textsuperscript{a}</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>12</td>
<td>9.1</td>
<td>0.7</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>11</td>
<td>8.9</td>
<td>1.0</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>7</td>
<td>8.5</td>
<td>1.9</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>10</td>
<td>10.8</td>
<td>0.9</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>7</td>
<td>9.2</td>
<td>0.3</td>
</tr>
<tr>
<td>All</td>
<td>All</td>
<td>47</td>
<td>9.4</td>
<td>0.3</td>
</tr>
</tbody>
</table>

Source: Regional statistical offices of Kazakhstan.

\textsuperscript{a} Unit: deci-tons per hectare.

\textsuperscript{b} sd: standard deviation.

\textsuperscript{c} Skew: skewness.

\textsuperscript{d} ha: hectares.

Fig. 1. Average calendar dates of wheat plant phases and their water availability for Kazakhstan. Note: The thick black bars represent the required water availability under optimal conditions during a specific plant phase in mm per day, whereas the boxplots illustrate the realized water availability (1980–2009) for the 5 counties (C1–C5). Source: Authors (data from field experiments in Akmola region).
Results

Shifts in critical phases

Fig. 3 shows the spread of the start and end calendar dates of the critical plant phase determined by $GDD_{\text{start}} = 148$ and $GDD_{\text{end}} = 881$ for the 5 counties. The starting date $D_{\text{start}}$ varies only slightly between June 2nd and June 14th with an average calendar date $D_{\text{fix}}$ of June 7th while the end date $D_{\text{end}}$ varies between July 12th and August 12th with an average date $D_{\text{fix}}$ of July 24th (slightly varying between counties). This means that in extremely warm years, the index is computed about 10 days earlier (depending on county) than in the coldest year. Similarly, the end date may vary for about 25 days (depending on county). The greater variability for $D_{\text{end}}$ compared to $D_{\text{start}}$ arises due to a longer GDD accumulation period. The variability of $D_{\text{start}}$ is solely based on the yearly temperature differences of phase 1 (constant sowing date), whereas the variability of $D_{\text{end}}$ represents all the yearly temperature variability from sowing to the end of phase 4.

Weather indices and regression parameters

We construct the two indices FLEX and FIX by determining the start and end date of the accumulation period once yearly variable, and once as an average over all years (as illustrated in Fig. 3 and defined in Eq. (9) and (10)). The summary statistics of these weather indices for the 5 different counties are provided in Table 2. The index values are on average around 1.5 mm per day, with maximal and minimal values being similar across counties. The presented average statistics do not reveal noticeable differences between the FLEX and the FIX approach. However, the shifts in the accumulation periods provoke major changes in single years and farms that can cause a halving or doubling of index values (single farm values are available upon request).

Table 3 shows a summary of the (farm-level) beta values of the QR analysis (based on Eq. (6)) for the 5 counties. Higher beta values represent a steeper slope of the regression curve and are thus an indication for a stronger yield-index dependency. We find that for all counties but county 4, the mean parameters based on FLEX are (significantly) higher (around +0.5) than those based on FIX.

The exceptional result for County 4 is expected to be caused by a particular weather pattern that is characterized by a concentrated and higher level of rainfall (rainfall is about 15% higher compared to the other counties). This concentrated rainfall period is captured within both indices, leading to the only marginal differences in FLEX and FIX. Thus, FLEX and FIX approaches result in similar index and regression values.

With regard to the minimal values (cp. the last column of Table 3), it has to be noted that there are negative values for two farms when using FIX and one farm when using FLEX. Negative values for beta indicate that an additional (rainfall) index unit decreases yields, which is inconsistent with the conceptual basis of our case study, where the index insurance contract is designed as put option (cp. Eq. (2)). The yield-index dependency of these farms seems to be poorly captured by our approach.

Expected utility estimations

Table 4 shows the EU values for the situation with an insurance insurance based on both specifications FLEX and FIX as well as for of the situation without insurance (“no”). Moreover, Table 4 presents the relative comparison of insurances (in terms of EU) based on FLEX and FIX (“rel”). The latter is determined as follows: $EU_{\text{rel}} = (EU_{\text{FLEX}} - EU_{\text{FIX}})/EU_{\text{FIX}}$. We find that the EU based on a FLEX insurance is always higher (except for one farm) than the EU without insurance. Note that the here
used approach to specify EU is based on fair insurance premiums and absolute semi-deviation (EUSDV). Gains in EU values are thus solely resulting from downside risk reductions. Comparing the EU values across the two considered specifications, we find that in about 90% of the cases the FLEX insurance leads to an increase in EU compared to the FIX insurance. The hypothesis of similar EUSDV values can be rejected using the non-parametric Wilcoxon-test based on pair-wise single farm comparisons over the entire sample (not shown).

Sensitivity analysis

Moreover, sensitivity analyses are conducted with respect to the assumptions made in our analysis. More specifically, we investigate the effects of changes in the GDD calculation as well as in the assumption on the sowing date. We find that slight de- and increases of the base and cut values (Hbase and Hcut) lead only to minor changes, whereas a change in the sowing date seems to be more influential (not shown). While shifts of a few days lead to minor changes only, shifts of e.g. 20 days have a greater impact on the result. For the sensitivity analyses, a clear pattern among the counties can be observed, i.e. almost all farms in a county behave similarly with regard to these changes.

Table 2
Summary statistics of weather index values (mm per day) based on single farms, aggregated to counties (C1–C5).

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>sd</th>
<th>Max</th>
<th>Min</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Fix</td>
<td>Flex</td>
<td>Fix</td>
<td>Flex</td>
</tr>
<tr>
<td>C1</td>
<td>1.54</td>
<td>1.40</td>
<td>0.66</td>
<td>0.63</td>
</tr>
<tr>
<td>C2</td>
<td>1.42</td>
<td>1.36</td>
<td>0.74</td>
<td>0.67</td>
</tr>
<tr>
<td>C3</td>
<td>1.31</td>
<td>1.26</td>
<td>0.72</td>
<td>0.68</td>
</tr>
<tr>
<td>C4</td>
<td>1.73</td>
<td>1.69</td>
<td>0.96</td>
<td>0.99</td>
</tr>
<tr>
<td>C5</td>
<td>1.38</td>
<td>1.37</td>
<td>0.80</td>
<td>0.71</td>
</tr>
</tbody>
</table>

Summary statistics based on a 30 years long time series. Source: Authors.

* sd: standard deviation.

Table 3
Summary statistics of beta value estimations of Quantile Regression based on single farms, aggregated to counties (C1–C5).

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Wilcoxon test</th>
<th>sd</th>
<th>Max</th>
<th>Min</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Fix</td>
<td>Flex</td>
<td>Fix</td>
<td>Flex</td>
<td>Fix</td>
</tr>
<tr>
<td>C1</td>
<td>2.42</td>
<td>2.98</td>
<td>***</td>
<td>0.76</td>
<td>0.79</td>
</tr>
<tr>
<td>C2</td>
<td>1.51</td>
<td>2.00</td>
<td>*</td>
<td>1.06</td>
<td>1.28</td>
</tr>
<tr>
<td>C3</td>
<td>1.32</td>
<td>1.71</td>
<td>**</td>
<td>0.96</td>
<td>1.05</td>
</tr>
<tr>
<td>C4</td>
<td>2.01</td>
<td>1.94</td>
<td>n.s.</td>
<td>0.75</td>
<td>0.72</td>
</tr>
<tr>
<td>C5</td>
<td>3.44</td>
<td>3.82</td>
<td>**</td>
<td>0.61</td>
<td>0.60</td>
</tr>
</tbody>
</table>

Source: Authors.

* Paired Wilcoxon-test based on single farm estimations, summarized on county level: n.s. denotes no significance, whereas *, ** and *** describe significance on the 10%, 5% and 1% level, respectively.

b sd: standard deviation. Beta value estimations for single farms are available upon request.

Comparing the EU values across the two considered specifications, we find that in about 90% of the cases the FLEX insurance leads to an increase in EU compared to the FIX insurance. The hypothesis of similar EUSDV values can be rejected using the non-parametric Wilcoxon-test based on pair-wise single farm comparisons over the entire sample (not shown).
Discussion and conclusion

Our analysis shows that a weather index insurance with flexible specifications to calculate indexes is, ceteris paribus, more effective in terms of reducing the downside risk exposure of farmers than the usually employed approach based on fixed start and end dates. More specifically, the generic approach developed in this study based on GDDs results in a significant increase of farmers’ expected utility (EU) levels. While the flexible index (FLEX) uses GDD values to specify the accumulation period per year, the fixed index (FIX) uses an average of these yearly varying start and end dates. Due to our...
assumption of fair premiums and focus on an expected utility framework considering absolute semi-deviation instead of variances, the increase of farmers’ expected utility levels under the FLEX specification is solely due to a reduction of downside risks. Thus, the proposed FLEX specification reduces basis risk, especially farmers’ risk of receiving no or a not sufficient insurance payout in case of a yield loss.

Further refinements of our approach could be achieved through a more comprehensive availability of data. For instance we are unable to consider shifts of sowing dates across time and space due to lack of information but rather rely on a single date for all counties. Although end of May is a typical sowing time in northern Kazakhstan, Muratova and Terekhov (2004) emphasize the long duration of sowing and its influence on crop productivity. Knowledge on the precise yearly sowing date for every farm (or at least per county) and GDD values estimated by experiments or breeders, would allow a more precise estimation of the critical growing phases. A consequence of our simplified approach is that, for a few farms, we do not precisely (enough) capture a critical phase, and thus limit the advantage of the FLEX index. However, even with this approach, we find the FLEX index design outperforms the FIX index design for the majority of the farms.

By using a cumulative precipitation index in our analysis, we follow the dominant approach of weather index insurance design and apply a simple but meaningful index for a (semi-) arid region. However, we did not capture disturbances beyond low rainfall levels such as hail and storm, or biotic stress factors such as fungal diseases. Since the impact of these disturbances depends on the vulnerability of the plant phase, similar to the effects of precipitation, our model can be easily adapted to incorporate such additional variables.

Further research should be focused on how to implement such a product and how it would be perceived by farmers. In general, the structure of the agricultural landscape in Kazakhstan with vast farms, a high degree of specialization and a rather high educational level (Heidelbach, 2007) might favor the introduction of these products. However, it needs to be assessed whether the reduction of basis risk that is realized with the here proposed approach outweighs the drawback of an additional layer of complexity implied by this specification.

For any index insurance solution it is a key challenge to reduce the amount of basis risk in order to develop a beneficial and viable instrument for both farmers and the insurance sector. With this study we address this key challenge by developing a consistent downside risk approach to consider critical plant growing phases, and show that accounting for the shifts of these phases decisively influences the hedging effectiveness of these index insurance instruments.

Appendix A.

**Fig. A.1.** Map of Kazakhstan and study region. Note: County 1–3 (R1–R3) are located in region 1 (Akmola oblast “O1” with capital Astana) and county 4–5 (R4–R5) are situated in region 2 (Kostanay oblast “O2” with administrative center Kostanay). R1–R3 stand for Zelinograd, Atbasar and Esil rayon and R4–R5 represent Denisovka and Kamisty rayon. Source: Authors.
References


Dell’Aquilla, C., Cimino, O., 2012. Stabilization of farm income in the new risk management policy of the EU: a preliminary assessment for Italy through FADN data. Paper prepared for European Association of Agricultural Economics, Capri, Italy.


