

Agricultural Data and Insurance

Innovations in agricultural data development for insurance

Background

Insurance requires high quality data. From an insurance perspective, data is of high quality if it is timely (so that claims can be paid quickly), relevant (so the product offers reliable protection), audited to international reinsurance standards, and available over a sufficiently long time horizon (time series). High quality data forms the basis for high quality, reliable insurance solutions. Without such data, insurance markets are unlikely to develop in a sustainable manner.

Insurance products for low-income farmers or herders are often built on indices that use agricultural or climatic data.

This includes yield data, livestock mortality data, remote sensing data, and weather data. Farm-level multiple peril crop insurance (MPCI) is generally not feasible for small farmers and herders, as the low sums insured and high cost of auditing data make the schemes uneconomic. By contrast, while index insurance can be cheaper to deliver, the quality of the protection that index insurance offers depends significantly on the quality of the index, which in turn depends on the quality of the underlying data. Only if the index reflects agricultural conditions experienced by the farmer is it likely to provide cost-effective, reliable protection.

Beyond insurance, high-quality agricultural data can support risk informed decisions made by policy makers and farmers.

Risk assessment based on agricultural data puts a price on agricultural production risk in a given area in the form of insurance premiums. It can thus help identify areas with high and low agricultural production risk and thus target input subsidies to areas that will benefit most. It can be used to develop cropping patterns for farmers to increase yields. It can also be used to select areas to focus expansion (or implementation) of irrigation programs. Moreover, by pricing and thus gauging production risk, agricultural insurance data has also been seen to crowd in better risk mitigation techniques. This was, for example, observed in a World Bank project with the Government of India (GoI): Initially, under the National Agriculture Insurance Scheme in India (NAIS), a cap of 3% was applied

Highlights

- High quality agricultural data is critical for sustainable agricultural insurance programs.
- For agriculture insurance in low or middle income countries, coordinated investments in agriculture data as a public good may be the best solution.
- Combining yield, weather, and satellite data in sensible ways can lead to higher quality, more cost-effective products than just using one type of data.
- Data sources are continually improving with technological and operational advancements. Agriculture insurance programs need to be responsive to these innovations.

to premiums charged regardless of the risk. This led to insurance being underpriced in some high-risk areas which may have exaggerated the production of certain crops in high-risk areas (for example groundnut in the state of Gujarat). Through an enhanced new scheme offering premiums that are closer to risk-based premiums, farmers have more information about production risk and are also incentivized to grow other crops that are more economically viable in the respective areas.

The Challenge

There are two frequent overarching challenges relating to the provision of high-quality data for agricultural insurance: 1) data availability and 2) data quality.

For 1), to make high quality data available for agricultural insurance purposes, governments have important roles to play in establishing a framework for data collection, auditing, financing and management. Given that data for agriculture insurance is expensive and non-rivalrous¹, it is likely to be a natural monopoly in most low or middle income countries. It does not make economic sense for every insurance company to set up their own weather stations in the same area to capture the same data. Instead, investments in weather stations could be coordinated, for example by the government, and the data could be made available on standard, reasonable terms to all insurance

¹ Non-rivalrous goods may be consumed by many at the same time at no additional cost (e.g., national defense or a piece of scientific knowledge).

providers. This would enable market-driven insurance where no single insurer has a monopoly that could distort the market. The Motor Third Party Liability database in Turkey offers an example. The data is collected and stored in TRAMER, a centralized data system, which is managed by an executive committee consisting of both public and private sector institutions. The data is used by multiple public and private sector organizations and is funded through fees paid by the users of the data.

For 2), even if data for agricultural insurance purposes is available, its quality is often not high enough, which has critical consequences for farmers. Lack of quality (timely, relevant, audited, long time horizon) can be problematic because it defines the reliability with which an index captures major agricultural shocks and thus triggers payouts. Poorer, vulnerable farmers, being the least equipped to cope with inaccurate indices, have the most to gain from more reliable protection. In fact, relying on poor value products due to low data quality may increase the risk farmers are exposed to rather than protect them from it: firstly, there is the risk of suffering a loss but not receiving a payout as the index may be inaccurate; secondly, insurance encourages higher-yield agricultural production with higher exposure to production risk. With poor value products, the additional risk taken by the farmer remains uncovered and in case of a production loss event, he or she may be worse off than before. Thus, coordinated investments in the right type of data is a necessary precondition of agricultural insurance that serves vulnerable farmers.

In addition, data must be of high quality in order to ensure access of insurers to international reinsurance markets. Agricultural shocks can be very large and access to international reinsurance markets is important to off-load some of this risk outside the country. However, reinsurance companies have high standards for the data they are willing to use to develop and price insurance products, and will charge significantly higher premiums if they have concerns about how the data is audited. Therefore, it is important that agriculture insurance data is of high quality and is audited through a transparent process (Box 1).

Box 1. Improving the quality of yield data in India

For the past eight years, the World Bank has been providing technical assistance to improve the National Agriculture Insurance Scheme in India (NAIS), which provides agriculture insurance coverage for nearly 30 million farming

² A crop cutting experiment involves a trained individual visiting a designated farm, harvesting an area (for example, a 5x5 meter square) for a designated crop, waiting for the produce to dry, and then weighing it. Sample-based area yield indices are typically calculated as the average yield from a series of CCEs of randomly

households. A key challenge was a lack of standardization, trained personnel, and monitoring for crop cutting experiments (CCEs)² which exposed the NAIS to significant delays, inaccurate indices, and the risk of manipulation.

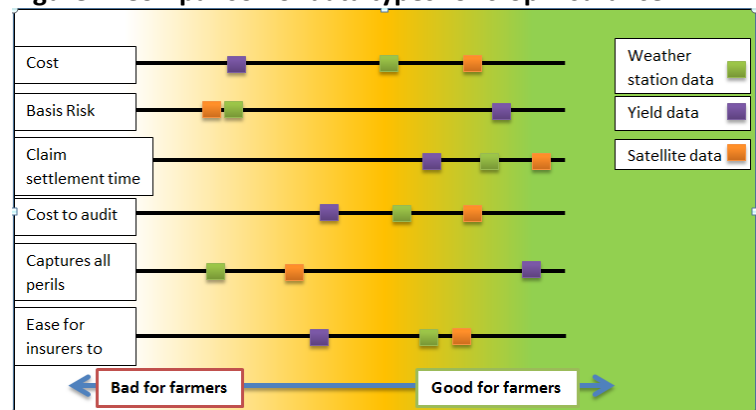
To address these challenges, a pilot was undertaken where CCEs were video recorded with GPS-tagged footage using mobile phones. The data was then provided to insurance companies by SMS at the time of the CCE to allow real-time monitoring. This innovative use of technology greatly improved the quality of data collected which could be verified through the video recordings. In this way, the trust of (re)insurers in the data was established.

It is also an example of how the speed of data collection can be significantly improved through using developments in technology. Whereas before, insurance companies could experience three-month delays in receiving the data under the paper-based CCE reporting system, under the new model, the data could be made available to insurance companies far quicker (potentially on the day the CCE is conducted through SMS), which would enable them to disburse claim payments sooner, ultimately benefiting effected farmers.

Types of Data for Agricultural Insurance

There are three key types of data that crop insurance can typically be based on: weather, yield and satellite data. These data types can be compared against several key metrics (Figure 1)³.

Figure 1. Comparison of data types for crop insurance



World Bank (2013)

While yield data comparatively tends to offer the most reliable indices, it can be the most expensive to collect and audit. Indices based on yield data (see box) typically provide the most comprehensive cover to farmers,

selected (using statistical methods) plots of land in the area the index is designed to cover.

³ Note these are indicative representations developed using expert opinion meant for illustrative purposes.

capturing agricultural production perils which other data sources cannot (for example pests and disease). However, yield data can be much more expensive to collect and audit due to the farm visits required, and as traditionally implemented has often led to long claim settlement time.

Insurance products based solely on weather or remote sensing indices can be less costly than products which require yield data. In recent years, there has been a lot of interest in pure weather index-based insurance (WII) or pure remote sensing index insurance (RSII) products to provide low cost insurance coverage for rural farmers and herders.

However, there is a growing body of statistical evidence that suggests that indices of WII and RSII can be too inaccurate for the product to reliably protect farmers and herders. An insured farmer could experience an event that destroys their crop yet the insurance does not pay. This may not be a big concern if the probability of this happening is very low but if the probability is too high it can significantly reduce the client value of the product. Whilst there is still a need for further research, the current evidence suggests that this risk is often higher for pure WII or pure RSII. For example, recent research by the World Bank and the Agricultural Insurance Company of India found that across one Indian state the correlation between weather indexed claim payments and farmer yield losses was only 14%. In addition, there was a 1-in-3 chance that a farmer would receive no insurance payout in a very bad year (World Bank, 2012).

Yet, pure WII and RSII are most attractive precisely in environments in which there is little yield data to fit and validate such products. This makes consumer protection challenging for the regulator.

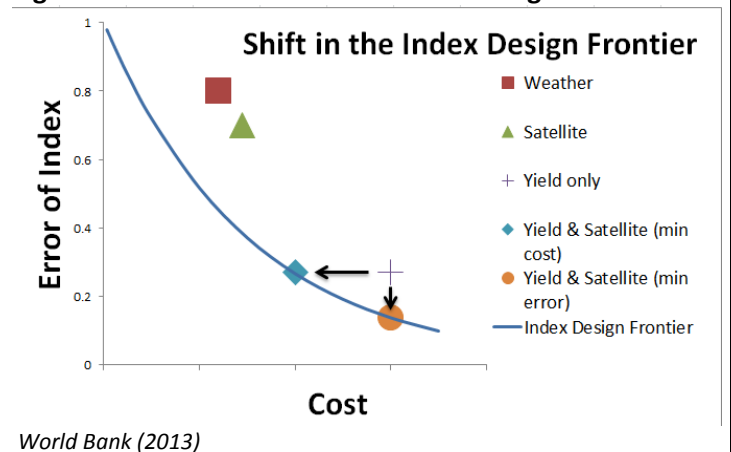
Towards a Solution

To tackle both lacking availability and quality of agricultural insurance data, various measures can be considered - firstly, sustainable development of agricultural insurance data can be achieved through coordinated investment in market data infrastructure. Governments have an important role to play in establishing and implementing a framework for data collection, auditing, financing and management. At the same time, the private sector can play various critical roles: for example, private sector may in some cases be better suited to collect

required data than the government, as they have the required expertise, can leverage innovation more readily, and are independent of short-term political cycles. In addition, by being involved through advisory and/or audit functions, private sector insurers can be key to ensuring that data investments are made into the most accurate and useful data.

Secondly, it is critical for data investment decisions to be based on credible Cost Benefit Analyses (CBA).⁴ A CBA tool like the Index Design Frontier captures two key aspects of indices⁵: the error of the index and the cost of creating indices representing agricultural production risk (Figure 2). The degree of accuracy of the index is captured by the y-axis, "Error of the Index", with high values associated with greater error as measured by the root mean squared error. The cost of the index is shown on the x-axis with high values representing higher costs. The frontier represents the most cost effective solutions available. Practitioners can investigate different index designs and decide where they wish to lie on the Index Design Frontier. Investments can then be made to develop the desired data market infrastructure based on this analysis. Investments could involve enhancing the weather station network by buying additional weather stations, developing yield data through crop cutting processes or investing in enhanced satellite data (see Box 2 for more details).

Figure 2. Illustrative Shift in the Index Design Frontier



Thirdly, sound investment decisions can be coupled with innovations in index design, for instance using more than one type of data, to increase index accuracy or reduce data costs (see Box 3). Practitioners can either have same quality data at a lower cost or same cost for data at a higher quality, by moving onto the frontier from behind it. This can be seen

⁴A CBA is an economic assessment that incorporates a statistical analysis of the implied index accuracy of different data sources, and the costs of investing in different data sources or the benefit of combining data sources to develop an index

⁵ Another key issue is timeliness of collection and claim payment. It has been set aside for this analysis as timeliness is becoming

less of an issue with technology. With new innovations data can be collected before harvest for certain technologies, and even audited CCE data can now be made available within a week of harvest.

in Figure 2. Using yield data on its own would place one at the purple cross. Combining yield data with satellite data one can (i) achieve the same accuracy for a lower cost – move to the blue diamond, or; (ii) improve the accuracy for the same cost – move to the orange circle; thus shifting along the Index Design Frontier.

Box 2. Shift in index design frontier: case study India

A recent analysis conducted under the World Bank project with NAIS in India was a study to investigate the potential benefits of combining data sources. In the study, satellite data was used to target the CCEs for generating a yield index. In areas where satellite data indicated crop yields may be unusually low, additional CCEs were carried out; in areas where satellite data indicated crop yields may be normal or good, fewer CCEs were carried out.

The results of the study were surprisingly strong: In areas where there is a high correlation between satellite data and crop yields (approximately 70%), using satellite data to target CCEs can either reduce the cost of conducting the CCEs by a factor of four, or improve the accuracy by a factor of two. It is also worth noting that under this scenario using satellite data alone, despite a fairly high correlation, would reduce the accuracy by at least a half when compared to using both satellite and yield data.

Fourthly, technological innovations can be harnessed to continually improve data quality. These improvements can involve: (i) building better data series such as more frequent, higher resolution satellite data (e.g., infrared technology to see through clouds); (ii) using different data sources in complementary ways (Box 2); (iii) innovations in satellite technology that drive down the cost of using maps and models to identify houses and roads and separate different types of crops in intercropped land; and (iv) using state of the art mobile phone technology in conducting CCEs, which can improve the accuracy and speed with which the data is collected for yield indices (see Box 1 for more details).

Box 3. Combining data sources to improve quality

The World Bank team working on the NAIS project in India investigated the benefits of combining data sources to improve the timeliness and accuracy of data. The total claim payment of the proposed product would be the maximum of the two indices, one based on yield data and the other based on weather station data. Given weather station data is quicker to obtain, under scenarios where a claim is payable, the claim payment due from the weather index would be paid at, or even before, harvest. Once the yield data becomes available, and if the area yield indexed claim payment exceeding the weather indexed claim payment, a “top-up” payment would be paid at the end of the season.

Other combinations of data are possible such as WII interpolation. Here remote sensing data is used to interpolate between the gaps in the ground station network of weather data, thus improving the accuracy of indices developed.

Bridging the Data Gap

During times of initial data scarcity, development of agricultural insurance can complement investments in the data market infrastructure. Agriculture insurance programs do not necessarily need to use only the existing agricultural data available. If current data quality is low, patchy or non-existent, investments in new data may be necessary. Under such scenarios, it is important that the other sections of the program are designed to use and support the new data that is being invested in. An example of this is developing a risk financing layer with the government acting as a reinsurer.

For example, in data sparse environments, the government could support risk financing for agriculture insurance programs in the short-medium term while high quality data series are being invested in. This can increase demand for insurance by reducing the costs for farmers, as it avoids the high additional premiums charges from reinsurance companies due to the uncertainty of the data. Data uncertainty is a key concern for private (re)insurance companies. It can lead to large increases in premiums charged to farmers through “data uncertainty” loadings or, in extreme cases where data is very poor, the absence of insurance cover. This means that high layers of risk (covering low probability events) and elements of coverage based on new datasets (designed to increase index reliability) are typically expensive to reinsure. In such scenarios, the government can establish a risk financing fund which provides risk financing on best estimate actuarially fair cost basis to programs investing in improved data sources.

The Government of Mongolia (GoM) implemented such an approach during the initial years of the Index Based Livestock Insurance project by financing providing a “social” layer of risk through a World Bank line of credit, to complement investments in the data market infrastructure. Over time the reinsurance markets became more comfortable with the new data collection and audit processes and now the GoM is able to access cost effective reinsurance even for high layers of risk.

Lessons Learnt

High quality agricultural insurance data is critical for agricultural insurance programs to develop sustainably. Without high quality data, insurance will not offer reliable

and cost-effective protection, which is particularly critical for poorer, more vulnerable farmers.

Governments have a key role in developing the data market infrastructure due to the fact that the collection of data can be a natural monopoly. This can be overcome by coordinated investment in data as a public good.

At the same time, the private sector can contribute to data collection and management, either via taking on specific tasks or by paying in order to be able to use the data. The latter can serve as an incentive for public data investments.

Indices can be based on a combination of yield, satellite and weather data. The data types have each have advantages and disadvantages, and combining different types of data can lead to products that offer both speed and reliability cost effectively.

As large agriculture shocks can affect an entire country, reinsurance is critical. Data quality needs to be high with robust audit mechanisms in place due to the high standards reinsurance companies have for data verifiability.

Through providing a layer of reinsurance, governments can support agriculture insurance programs in initial years, while investments are being made in market data infrastructure. The government can then offload the risk to reinsurance markets over time as data quality improves.

Further Information

World Bank (2012). *Improving Farmers' Access to Agriculture Insurance in India*. World Bank Policy Research Working Paper 5987. The World Bank. Washington, DC.

World Bank (2012). *Weather Based Crop Insurance in India*. World Bank Policy Research Working Paper 5985. The World Bank. Washington, DC.

Mahul, O. and Stutley (2010). *Government Support to Agricultural Insurance: Challenges and Options for Developing Countries*. The World Bank. Washington, DC.

DRFI Program website: worldbank.org/drfi

Contact

Olivier Mahul, Program Manager, Disaster Risk Financing & Insurance Program, GFMDR and GFDRR, The World Bank, omahul@worldbank.org

Daniel Clarke, Senior Disaster Risk Financing & Insurance Specialist, Disaster Risk Financing & Insurance Program, GFMDR and GFDRR, The World Bank, dclarke2@worldbank.org

Felix Lung, Disaster Risk Financing & Insurance Specialist, Disaster Risk Financing & Insurance Program, GFMDR and GFDRR, The World Bank, flung@worldbank.org