Inter-American Development Bank

Institutions for Development Sector

Capital Markets and Financial Institutions Division

DISCUSSION PAPER

No. IDB-DP-362

Default Risk in Agricultural Lending

The Effects of Commodity Price Volatility and Climate

Carlos Castro
Karen Garcia

September 2014

Default Risk in Agricultural Lending

The Effects of Commodity Price Volatility and Climate

Carlos Castro
Karen Garcia

IDB
Inter-American Development Bank
2014

http://www.iadb.org

Contact: Pablo Carrion Florez, pcarrion@iadb.org

Default Risk in Agricultural Lending

The Effects of Commodity Price Volatility and Climate

Carlos Castro and Karen Garcia

Faculty of Economics, Universidad del Rosario, Colombia

Abstract*

This paper proposes and estimates a default risk model for agricultural lenders that explicitly accounts for two risks that are endemic to agricultural activities: commodity price volatility and climate. The results indicate that both factors are relevant in explaining the occurrence of default in the portfolio of a rural bank. In addition, the paper illustrates how to integrate the default risk model into standard techniques of portfolio credit risk modeling. The portfolio credit risk model provides a quantitative tool to estimate the loss distribution and the economic capital for a rural bank. The estimated parameters of the default risk model, along with scenarios for the evolution of the risk factors, are used to construct stress tests on the portfolio of a rural bank. These stress tests indicate that climate factors have a larger effect on economic capital than commodity price volatility.

Keywords: Credit risk, rural banks, climate risk
JEL Classification: Q14, Q54, G17

* Corresponding author: carlos.castro@urosario.edu.co. The authors gratefully acknowledge the financial support of the Inter-American Development Bank (IDB).

1. Introduction

Quantitative risk management has become increasingly important for financial institutions ever since the publication of Basel II in 2004. Rural banks have not been exempt from this process, whether at their own initiative or required by the regulator. Some of these institutions receive considerable government support to enable them to provide loan opportunities to economic sectors considered too risky for most financial intermediaries. Given their significant specialization, rural banks have a smaller margin to diversify risk in their portfolios. They must also manage risks endemic to agricultural activities, including: (i) production risk (e.g., climate, disease, pests, technological change, and natural resource management); (ii) market risk (e.g., commodity price volatility, production factor price volatility, and uncertainty in the supply chain); and (iii) regulatory risk (e.g., changing policies, subsidies, and sanitary requirements).

Two underlying elements complicate the picture for rural banks. First, the agricultural sector is an important economic sector (especially in developing countries), a source of employment, and critical in guaranteeing food security for the population. Second, volatility in the agricultural sector is expected to increase in the coming decades (Gilbert and Morgan, 2010; Schaffnit-Chatterjee, 2010;). According to some analysts (Khan and Zaks, 2009), production risks are expected to increase due to climate change and globalization. Management of such risks, endemic to agricultural activities, will require multiple strategies and new participants apart from the government, which has been the primary source of risk coverage, especially in developing countries. From the financial standpoint, there is growing interest in the use of financial derivatives and insurance (Geman, 2005; Hess et al., 2004; Larson et al., 2004; OECD, 2011; Schaffnit-Chatterjee, 2010). However, in order to determine the most appropriate design of financial instruments, quantitative tools are desperately needed.

The techniques and the academic and empirical literature on portfolio credit risk modeling have evolved significantly in the last ten years. The implementation of internal models has given financial institutions the opportunity to calibrate their risk weightings in a more balanced and realistic matter. One important development is the mapping between prudential capital, the unexpected loss function, and elusive risk factors (de Servigny and Renault, 2004). In order to implement internal models and quantify the capital charge on their exposure, financial entities need to build portfolio credit risk models. For a rural bank, this is an important challenge. However, as distinct from international banks, the challenge is not so much mapping the complex interdependencies between their exposures, but rather the multiple risk factors that they face.

A parsimonious way of introducing risk factors into the portfolio model is to build a multi-factor model to explain default risk. Factor models are common in finance (Fama and French, 1993; Ross, 1976) and credit risk management (Demey and Roget, 2004; Gordy, 2000).

The literature on risk management for agricultural lending is scarce. Behrens and Pederson (2007), Escalante et al. (2004), and Gloy et al. (2005) analyze rating transition matrices on agricultural loan portfolios, and their results are not particularly different from what is observed in non-agricultural portfolios: macroeconomic factors are important in explaining credit migrations (farm-level factors are not significant), small and younger farms have more unstable ratings and downgrade momentum.[1] Katchova and Barry (2005) use farm association data to implement two commercial structural credit risk portfolio models: CreditMetrics and KMV. Both models are driven by a one-factor model representing the conditions of the average farm. Their results show that the probability of default (over one-year horizon) for agricultural lenders varies substantially depending on their ratings from $0\,percent$ the highest rating to $96\,percent$ for the lowest. Pederson and Zech (2009) propose a modified version of the CreditRisk+ approach for agricultural lenders. In their model they incorporate a proxy for the economic cycle of the agricultural sector.

The objective of this document is to build a portfolio credit risk model for a rural bank that is able to account for two specific risk factors that are relevant (and so far unexplored) for an agricultural loan portfolio: commodity price volatility and climate effects. In order to incorporate such effects, in a standard portfolio credit risk model, we specify and estimate a default risk model that incorporates price volatility and climate factors as determinants of the number of defaults in an agricultural portfolio. Information is used from the largest agricultural lender in Colombia, and the portfolio is organized into homogenous risk groups determined by the produce or crop and the zone of production.

We find significant heterogeneity in the occurrence of default across crops and zones of production within the same produce. This heterogeneity in part helps to find economically and statistically significant effects of both commodity price volatility and climate effects on the probability of default. We use the estimated default probability model to estimate conditional default probabilities. These probabilities can be used in a portfolio credit risk

[1] Downgrade momentum is the situation where recently downgraded obligors are at an increasing risk of experiencing further downgrades.

model to derive simulated loss distributions and quantify economic capital. Different versions of the estimated conditional default probabilities (estimated at different realizations of the risk factors) can be used to establish a baseline of economic capital as well as stressed estimates of economic capital. The results from simulating the evolution of the loan portfolio indicate that climate effects are more important than commodity price volatility. The results also indirectly underscore the importance of using counterparty risk mitigation schemes (such as insurance) to reduce losses in the event of default, which does not seem to be a rare event in agricultural portfolios.

The document is organized as follows: Section 2 describes the rural portfolio and discusses the relevant risk factors. Section 3 presents the structural credit risk model based on the mapping of the risk factors. Section 4 introduces a statistical model of default risk and discusses the estimation results. Section 5 explains the design of the credit risk portfolio model used to simulate the losses of the rural portfolio. Section 6 presents the baseline results on the estimation of the loss distribution and the economic capital. Section 7 discusses the results of a series of stress tests based on scenarios of changing macroeconomic, climate, and price conditions. Section 8 concludes.

2. Risk Factors in Agricultural Projects

The first step in the development of a portfolio credit risk model for rural banks is to understand how the exposures (i.e., loans) within the portfolio are organized, specifically, how these exposures are related to a set of risk factors that can be quantified and that will be the driving force behind the credit quality of such exposures. For a rural bank, the loan is tied to its purpose, that is, to provide financial capital to a rural project: a particular crop (produce), the acquisition of farm equipment, or land improvement. The loan is also related to the geographic area where the project will be carried out (zone of production).

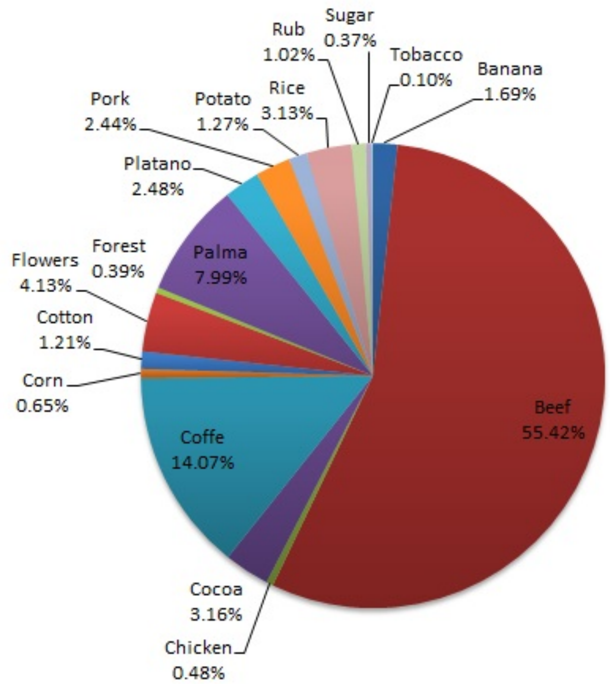
In risk management, an important concept that helps in envisaging the linkages between exposures and risk factors is the notion of homogenous risk groups. A homogenous risk group is a set of units (e.g., loans) that, because of their common and identifiable characteristics, are exposed to similar risk factors. Therefore, in order to map exposures to the set of risk factors, the loans must be grouped within these homogenous risk groups. Since the subject of interest is identifying how commodity price volatility and climate effects affect the quality of the agricultural loan portfolio, there will be two sets of homogenous risk groups: one based on produce and another based on zone of production. Table 1 presents the homogenous risk groups that we identified from the loan portfolio (Figures 1 and 2), with 17 crops and 8 geographical zones of production. The latter are organized according to the altitude (distance above sea level) of the region where the loan was granted. The grouping of loans according to homogenous risk groups represents around 60 percent of the loan book; the remaining part of the loan book cannot be classified into viable risk groups.

Table 1: Homogenous Risk Groups Based on Product and Geographical Zones of Production Classification of the Portfolio

Produce	Code	Zone
Cotton	Cot	0-50
Rice	Ric	51-100
Chicken	Chi	101-500
Bananas	Ban	501-1200
Forrest	For	1201-1800
Cocoa	Coc	1801-2200
Coffee	Cof	2201-3000
Sugar	Sug	>3000
Rubber	Rub	
Flowers	Flo	
Beef	Bee	
Corn	Cor	
Palm Oil	Pal	
Pottato	Pot	
Plantain	Pla	
Pork	Por	
Tabaco	Tab	

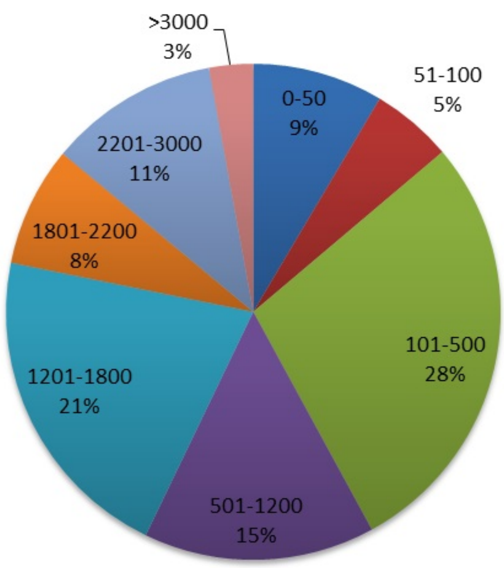
Source: Authors' elaboration.

Figure 1: Distribution of the Loan Portfolio across Produce and Based on the Average Historical Value of Exposures



Source: Authors' elaboration.

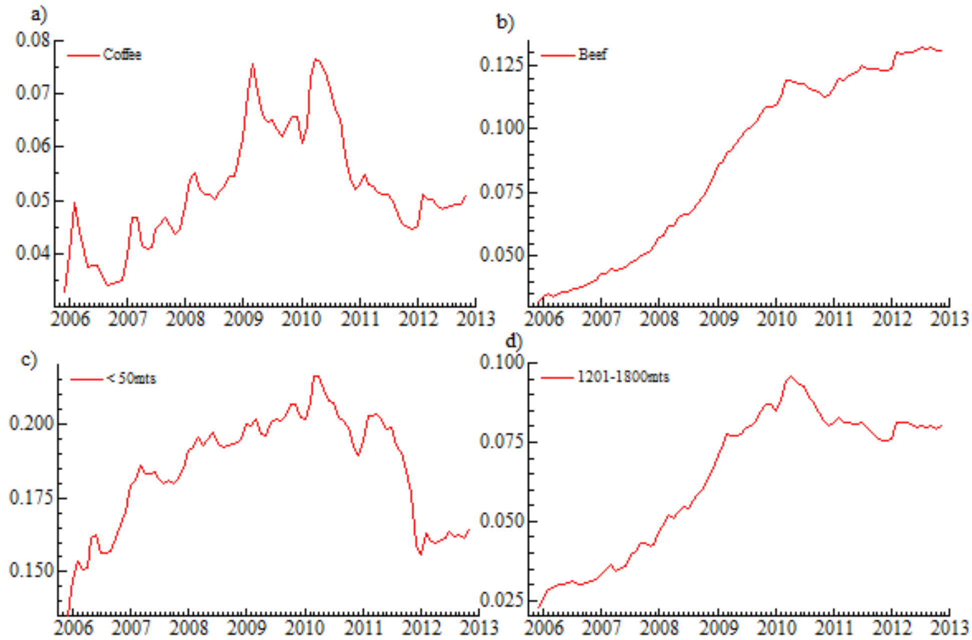
Figure 2: Distribution of the Loan Portfolio across Zones of Production (altitude) and Based on the Average Historical Value of Exposures



Source: Authors' elaboration.

Even though the notion homogenous risk groups implies different ex-ante risk behavior across the groups, shocks that are common to the different groups cannot be ruled out. This is precisely what we find when we look at the time-varying hazard rates across groups. Figure 3 shows the evolution of such rates for some products and zones of production: coffee and beef, less than 50 meters and 1201 to 1800 meters. Similar patterns are observed across the time series, peaking around 2009 and mid-2010, which implies that there is a common systemic factor. Similarities also surface across the two risk groups, meaning that there are some factors that might be local to some groups but not entirely idiosyncratic. The hazard rates are generally larger than zero; therefore, default is not such an uncommon event as one might expect (or as observed in other types of credit data (Koopman et al., 2008).

Figure 3: Hazard Rates for Selected Products and Zones in the Portfolio *(monthly data for 2005–12)*

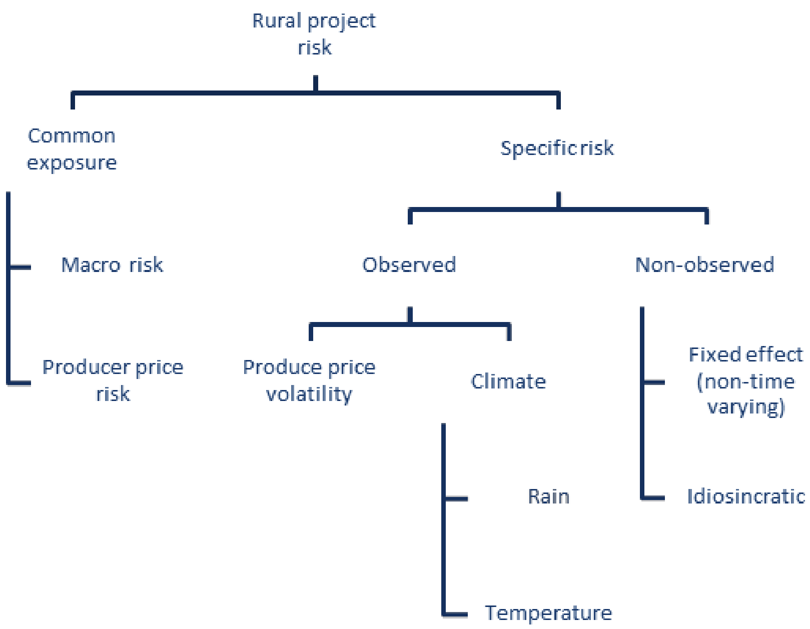


Source: Authors' elaboration.

After determining the risk groups, the risk factors of the rural projects need to be mapped to each of these paired exposures (produce and zone of production) within the portfolio. Figure 4 presents the main risk factors of rural projects and how they will be classified within the default risk model. This risk map is similar to the mapping of risk factors of commercial portfolio credit risk models (G-Corr of Moody's KMV, Portfolio

Tracker of Standard & Poor's and Creditmetrics). The purpose of such maps is not only to identify the main risk factors but also to depict the linkages between the elements of the portfolio. The map indicates a set of common risk factors, such as macroeconomic conditions or producer prices (captured by annual variations in the agricultural producer price index) and the produce- or zone-specific observed risk factors associated with commodity price volatility and climate effects (temperature and rainfall). The statistical model also allows unobserved non-time-varying risk factors (produce- and zone-specific intercepts in the linear model) to be estimated.

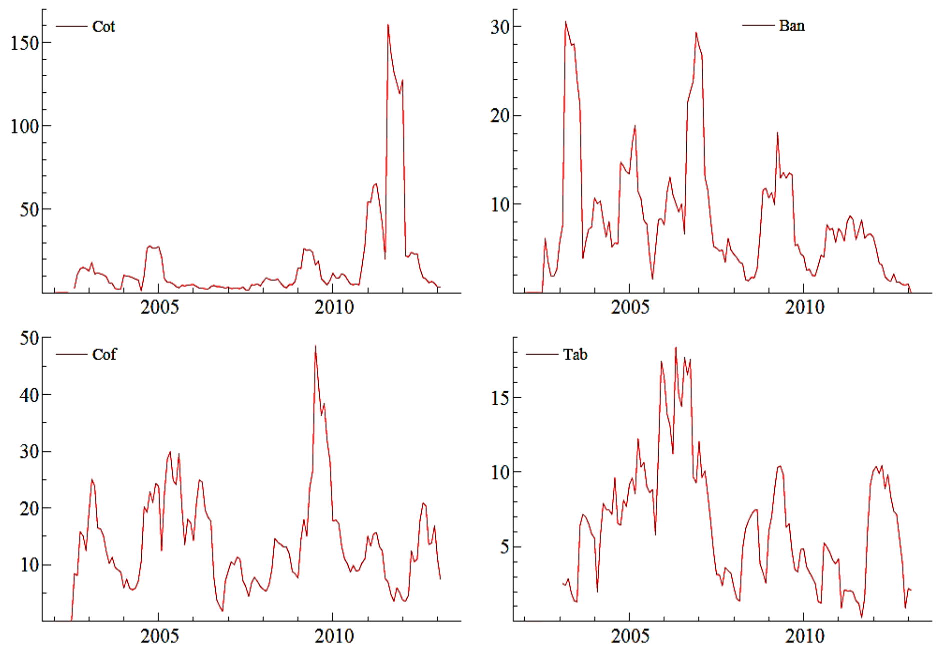
Figure 4: Rural Project: Hierarchy of Risk Factors



Source: Authors' elaboration.

Commodity price uncertainty is one of the greatest concerns for agricultural producers. The difference between the price at the start of the project and at the end is an important source of uncertainty; therefore, price volatility (in the form of market risk) is an issue. We introduce commodity prices as a factor that drives creditworthiness, but since we are interested in the uncertainty that they transmit to agricultural projects, we measure such uncertainty in the form of volatility of the return process. The challenge is that volatility is a latent variable; therefore, a statistical model is needed to measure (filter out) the instantaneous volatility of the commodity price process. We use the exponential weighted moving average model (EWMA) because it provides a simple method to measure the instantaneous volatility of the commodity price process. The moving average is estimated over a period of six months on the squared monthly returns, and the exponential weight, λ, is set to 0.94, which is standard in the literature. Figure 5 presents an example of the time series evolution of the estimated volatilities for some of the commodity price data for the crops in the portfolio. We observe that cotton prices are more volatile than coffee or tobacco. Cotton went through a period of important price variation at the end of 2011. Some other crops that show similar spikes are forest products (lumber) and potatoes (more than one spike in the sample).

Figure 5: Estimated Volatilities for Cotton, Bananas, Coffee, and Tobacco *(monthly data for 2002–12)*

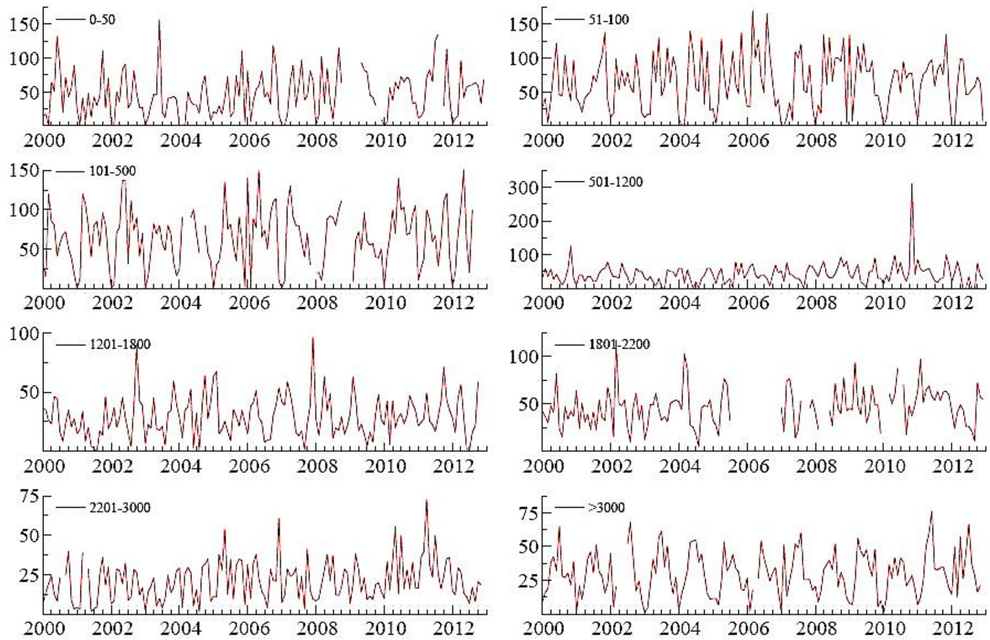


Source: Authors' elaboration.

Of similar importance in terms of agricultural production risk is climate risk. Therefore, we also consider it a determinant of creditworthiness. Small, unexpected changes in climate-related factors (temperature, rainfall, sunlight, and humidity, among others) can have important consequences for crop yields and, more importantly, for the success or failure of agricultural projects. Climate information is frequently collected from different climate stations around the country. We use two climate risk factors, based on the following climate information: average (over daily temperature readings) monthly temperature and precipitation, and average (over daily 24 hour readings) monthly precipitation.

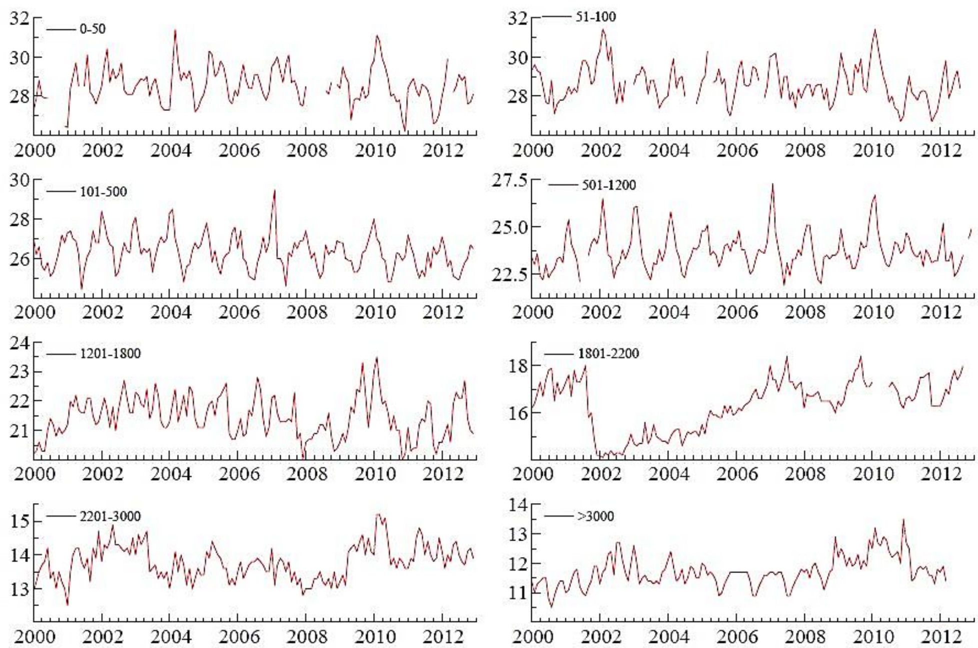
Figures 6 and 7 present the time series of temperature and precipitation aligned with the loan classification based on altitude groups. We observe some stable cyclical patterns with respect to temperature. In contrast, precipitation does not show a stable pattern; specifically, there was a sharp increase in precipitation during 2010 and 2011 in most zones of production. To facilitate the interpretation of the marginal climate effects on the statistical model, we transform the temperature and precipitation data into standard deviations from the monthly historical median measure. This transformation has two advantages: first, it preserves any seasonal patterns in the time series, and second, the use of the median reduces the contamination effect of outlying observations, specifically the rainfall data. In the estimation of the model, we also consider lagged measures of rainfall and temperature in order to account for the crop cycle. Therefore, we consider 6 month, 12 month and 48 month lags and choose the best performance of the lagged versions of the climate data.

Figure 6: Average Rainfall According to Altitude Groups *(monthly data for 2002–12)*



Source: Authors' elaboration.

Figure 7: Temperature According to Altitude Groups *(monthly data for 2002–12)*



Source: Authors' elaboration.

3. The Structural Credit Risk Model

The default risk model has its roots in Merton (1974). In the last 10 years, this model has been at the center of the literature on portfolio credit risk modeling. The most general version of the Merton model considers the asset value of a firm $i = 1, .., N$ V_i as a latent stochastic variable. Let the pair $(V_i, \epsilon_i) \sim N(0, I_2)$ $\forall t = 1, \ldots, T$. These firms belong to the portfolio of an investor (say, a rural bank) that wishes to model the default dependence across the portfolio. With this in mind, the investor considers a multi-factor dynamic model (Castro, 2012; Koopman et al., 2011) as the underlying structure behind the dynamics and dependence across the asset values for the firms that belong in the portfolio:

$$V_{i,t} = \omega_{i,0} f_t^{macro,ppi} + \sum_{p=1}^{P} \omega_{i,j}{}' F_{t,j}^{pvol} + \sum_{l=1}^{L} \omega_{i,l}{}' F_{t,l}^{clim} + \sqrt{1 - \sum_{p=1}^{P} \omega_{i,j}^2 - \sum_{l=1}^{L} \omega_{i,l}^2 \epsilon_{i,t}} \qquad t = 1, \ldots, T. \tag{1}$$

Let $\tilde{f}_t = (f_t^{macro,ppi}, F_{t,1}^{pvol}, \ldots, F_{t,P}^{pvol}, F_{t,1}^{clim}, \ldots, F_{t,L}^{clim})$ collect the set of common macro factors, the product-specific price volatility factor $F_{t,p}^{pvol}$, and the region-specific climate factor $F_{t,l}^{clim}$. Collecting the weight vector in ω_i such that $\omega_i'\omega_i < 1$,[2] we can then re-write expression 1 in the following simple form:

$$V_{i,t} = \omega_i'\tilde{f}_t + \sqrt{1 - \omega_i'\omega_i}\epsilon_{i,t} \tag{2}$$

If $V_{i,t}$ falls below a predetermined threshold μ_i (related to the level of debt) then a particular event is triggered. For capital adequacy purposes, the most important event is default. Then, the time-varying conditional default probability is

$$\pi_{i,t} = P(\epsilon_{i,t} < \frac{\mu_i - \omega_i'\tilde{f}_t}{\sqrt{1 - \omega_i'\omega_i}} | \{\tilde{f}_1, \dots, \tilde{f}_t\}) \tag{3}$$

Therefore, the probability of default of the elements of the rural portfolio is a function of the macro factor and, more importantly, commodity price and climate effects.

Although the model is indexed for a particular firm (loan), in practice estimation of the parameters is performed on a more aggregate scale. One way to reduce dimensionality is to define a set of homogenous risk classes, i.e., to group firms (or loans in the portfolio) by some identifiable characteristic, such as by crop and/or the geographic region where production takes place.

4. The Default Risk Model

The process of adding up firms or loans to obtain the homogenous risk classes creates a set of default counts for such groups (in other words, the cross sections). Let $k_{p,z,t}$ denote the number of loans for agricultural produce p in region z that are still active at the start of period t, and let $y_{p,z,t}$ denote the number of defaulting loans in period t.[3] Then $y_{p,z,t}$ can be considered a realization of a binomial distribution conditional on the realization of a set of factors $\tilde{f}_t$

[2]This particular weighting scheme guarantees that the asset value process, V_i, follows a standard normal distribution, which is a standard assumption in the portfolio credit risk literature. See Koopman et al. (2011).
[3]Note that (p,z) denotes the unit of analysis of the cross-section.

$$y_{p,z,t}|\tilde{f}_t \sim Binomial(k_{p,z,t}, \pi_{p,z,t}),$$
$$p = 1, \dots, P, z = 1, \dots, Z, t = 1, \dots, T. \tag{4}$$

where $\pi_{p,z,t}$ is the conditional probability of default such that $\pi_{p,z,t} = (1 + e^{-\theta_{p,z,t}})^{-1}$. This function guarantees that $\pi_{p,z,t}$ is indeed a probability (link function) and maps the probability to the factors, as in expression 3. Note that a set of restrictions guarantees the identification of all of the parameters of the statistical model. Furthermore, there is mapping between the statistical model parameters and those of the structural model (see Castro, 2012; Koopman et al., 2011).

$$\theta_{p,z,t} = \lambda_0 + \lambda_{1,p} + \lambda_{2,z} + \beta_0 f_t^{macro} + \beta_1 f_t^{ppi} + \sum_{p=1}^{P} \gamma_p{}' F_{t,p}^{pvol} + \sum_{z=1}^{Z} \delta_z{}' F_{t,z}^{rain} + \sum_{z=1}^{Z} \theta_z{}' F_{t,z}^{temp} \tag{5}$$

This is not the only feasible specification of the model. One could think of a setup where climate factors and price volatility have the same effect on all of the units of analysis.

$$\theta_{p,z,t} = \lambda_0 + \lambda_{1,p} + \lambda_{2,z} + \beta_0 f_t^{macro} + \beta_1 f_t^{ppi} + \gamma' F_{t,p}^{pvol} + \delta' F_{t,p}^{rain} + \delta' F_{t,p}^{temp} \tag{6}$$

The previous expressions (4,5) represent a generalized linear model (GLM). This particular GLM is designed for proportion data, that is, the number of defaulted vs. non-defaulted loans. There are two main assumptions in the model. First, the error follows a binomial distribution, and second, the link function is logistic. The link function is the transformation that guarantees that the marginal effects of the risk factors on the response variable can be interpreted as the change in the probabilities of the event of interest (default). Estimation of the parameters of the model can be performed by writing the likelihood function and using an iterative method for maximizing the likelihood (McCullagh and Nelder, 1989).[4]

4.1. Estimation Results

We have collected a dataset on counts of exposures and defaults from loans of the biggest

[4]We performed the estimation of the model using the glm library of R.

rural bank in Colombia (monthly data from December 2005 to November 2012). The proportion data on defaulted and non-defaulted $(y_{p,z,t}, k_{p,z,t} - y_{p,z,t})$ loans over the period of analysis is organized in two levels: the produce and the geographical zone of production (altitude zones). Therefore, this paired variable of interest has three indices: produce(p), zone(z) and month(t). These levels or indices represent homogenous risk groups, with two important implications: first, any loan within such a group has a similar behavior (in terms of creditworthiness), and second, the random variables that they represent are conditionally independent given the realization of a set of common and group-specific risk factors (see expression 4).

Table 2 presents the parameter estimates for the specifications of the linear factor model (expressions 5 and 6). When comparing the Akaike information criteria (based on the evaluation of the likelihood or deviance) we find a significant difference in the log-likelihood of the two specifications 5 and 6. However, a χ^2-test based on the analysis of explained variance indicates that the specification that includes group-specific slopes with respect to volatility and climate risk factors (model in expression 5) is preferred to a model with constant slope coefficients with respect to such factors (the result is significant at $percent1$ level).[5]

Although the observable risk factors are statistically significant at $percent5$ level, their economic significance is small. However, in order to interpret the marginal effects of the factors on the outcome variable (proportion of defaults), we use the inverse of the logit function. This transformation of the estimated level and slope coefficients provide the conditional probability of default given a particular category or risk factor.

[5]Because of the result of the test, we only report the estimation based on expression 5. Further results are available upon request to the corresponding author.

Table 2: Parameter Estimates, Zone of Production Based on Altitude Groups

	Estimate	z value		Estimate	z value
λ 0	-0,0604155	-3,966	ϒ Ric	0,011068	17,966
λ Ric	-1,4990673	-91,113	ϒ Chi	0,0010607	0,586
λ Chi	-0,6629319	-39,092	ϒ Ban	-0,0052589	-3,653
λ Ban	-1,3320171	-68,354	ϒ For	0,002724	5,506
λ For	-2,1940309	-65,069	ϒ Coc	-0,0022925	-3,742
λ Coc	-2,6082562	-166,342	ϒ Cof	0,0118726	27,857
λ Cof	-2,9649269	-192,181	ϒ Sug	-0,0025602	-1,067
λ Sug	-2,2462785	-50,665	ϒ Rub	0,010165	1,98
λ Rub	-4,432348	-76,271	ϒ Flo	0,0034989	4,588
λ Flo	-1,7511114	-60,119	ϒ Bee	-0,0133951	-31,252
λ Bee	-2,1837187	-144,066	ϒ Cor	0,0133955	13,047
λ Cor	-1,1007276	-64,712	ϒ Pal	0,0064426	3,181
λ Pal	-3,6713582	-111,272	ϒ Pot	0,0041913	10,516
λ Pot	-1,6358277	-90,756	ϒ Pla	0,0020533	4,879
λ Pla	-1,7126201	-108,658	ϒ Por	0,0072484	17,659
λ Por	-1,8388967	-119,582	ϒ Tab	0,0494964	8,071
λ Tab	-3,7688036	-85,134	δ 51-100	-0,0145933	-4,554
λ 51-100	0,3027413	73,665	δ 101-500	0,0058512	4,193
λ 101-500	0,1204871	45,298	δ 501-1200	-0,0068642	-4,344
λ 501-1200	-0,1977453	-70,81	δ 1201-1800	0,0220469	11,542
λ 1201-1800	-0,5284594	-192,749	δ 1801-2200	-0,0046369	-2,442
λ 1801-2200	-0,7617442	-213,159	δ 2201-3000	-0,0078769	-4,683
λ 2201-3000	-0,9284286	-284,884	δ >3000	0,0534892	19,991
λ >3000	-0,7453008	-164,948	ϴ 51-100	-0,0068955	-2,374
β pca	0,1897542	361,324	ϴ 101-500	-0,0741987	-36,874
β ppi	0,3367142	29,161	ϴ 501-1200	-0,0755523	-35,631
ϒ	-0,0033713	-8,664	ϴ 1201-1800	-0,022657	-10,933
δ	-0,0067497	-4,838	ϴ 1801-2200	0,0536568	11,049
ϴ	0,1202476	68,019	ϴ 2201-3000	-0,0167048	-6,324
			ϴ >3000	-0,008375	-3,32

Source: Authors' elaboration.

5. Portfolio Credit Risk Model for a Rural Bank

The elements of a credit risk portfolio model are: exposures, probabilities of default, and the loss given the default or recovery rate (see de Servigny and Renault, 2004). With these elements we are able to obtain the loss distribution. Most credit risk model follow a bottom-up approach, where the idea is to aggregate the credit risk of all individual instruments in the portfolio and then use some risk measure to quantify unexpected losses and relate these losses to a capital surcharge. Among these elements perhaps the most important is the probability of default or the default risk model that allow us to obtain a conditional probability of default $PD_{p,z}^{cond}$. The default risk model derived and estimated in Section 4 permits two crucial elements in the credit risk portfolio model. First, it provides a functional relationship between the probability of default and its determinants (such as macroeconomic conditions and commodity prices, among other factors). Second, it provides one method to incorporate dependence across the exposures in the portfolio. The reason for this is that if two exposures have similar determinants of their probabilities of default, then these probabilities could be expected to jointly deteriorate if this determinant is stressed.

5.1. The Conditional Default Probabilities

The portfolio of the rural bank is composed of homogenous risk groups defined along two dimensions (in addition to the time dimension, t): produce, p, and zone of production, z (Table 1). From the bank record we obtain the average exposures within the portfolio, $E_{p,z}$. This represents the average size of the loan granted to the agricultural project for produce p and zone of production z. This average is obtained from the historical data or the most recent (1 year) data on the portfolio. The rural bank may provide some estimate of the recovery rate on the loans in case there is a default, $RR_{p,z}$. The recovery rate denotes the percentage of the value of loan that one can expect to obtain after a default; therefore, $RR_{p,z} \in (0,1)$ and the loss-given default is defined as $1 - RR_{p,z}$. The recovery value of a defaulted loan comes from government guarantees, land, or any other collateral that can support the loan. Unfortunately, the rural bank does not have specific information on the expected recovery rate for the elements of the portfolio. In such cases, one can assume a flat, predetermined recovery rate for all exposures in the portfolio $R_{p,z} = rr \quad \forall p,z$. Estimates of the unconditional (historical) default probability, $PD_{p,z}^{hist}$, can also be obtianed from the bank record.

$$\bar{PD}_{p,z} = \frac{1}{T}\sum_{t=1}^{T} PD_{p,z,t} = \frac{1}{T}\sum_{t=1}^{T} \frac{No.\ defaulted\ obligations_{p,z,t}}{No.\ obligations_{p,z,t}}$$

Although it is possible to simulate losses using unconditional default probabilities, it is more informative to do so using conditional default probabilities. The main reason, as mentioned in the introduction, is that these probabilities are conditional on the realization of the risk factor. Therefore, a distribution resulting from a given scenario for the risk factors can be generated. In order to functionally relate the conditional default probabilities and the risk factor, we need to use the estimated default risk model, 4. Recalling expressions 4 and 5, the conditional default probability is determined by

$$PD_{p,z}^{cond}(\boldsymbol{F}) = \frac{1}{1 + exp^{-\theta_{p,z}(\mathbf{F})}}$$

where $\theta_{p,z}(\boldsymbol{F})$,

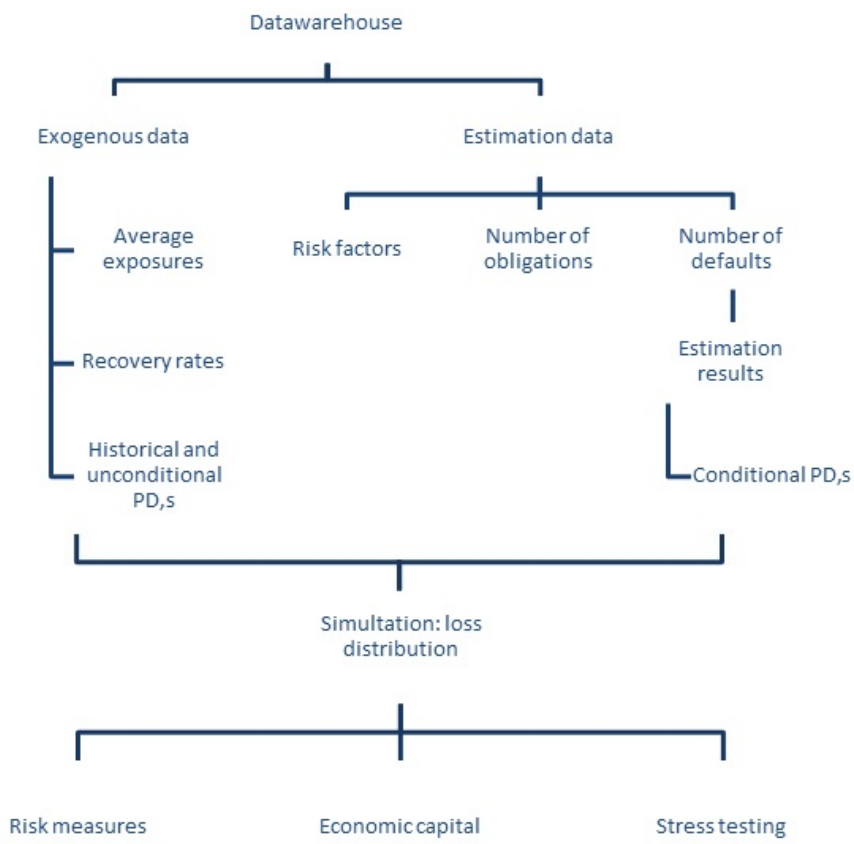
$$\theta_{p,z}(\boldsymbol{F}) = \lambda_0 + \lambda_p + \lambda_z + \beta_0 f^{macro} + \beta_0 f^{ppi} + \gamma_0 F^{pvol} + \gamma_p F^{pvol}$$
$$+\delta_0^{rain} F^{rain} + \delta_z^{rain} F^{rain} + \delta_0^{temp} F^{temp} + \delta_z^{temp} F^{temp}$$

where $\boldsymbol{F}(.) := (f^{macro}, f^{ppi}, F^{pvol}, F^{rain}, F^{temp})$ denotes the vector of observable factors. These risk factors account for: macroeconomic uncertainty, producer price variations, commodity price volatility, and climate factors (rainfall and temperature), respectively. As illustrated in Figure 4, the first two factors affect all of the positions in the portfolio, whereas the last three are related to a specific crop and zone of production. All of the observable risk factor can be evaluated at the mean ($F(\text{mean})$), under average conditions, giving rise to what we call the sequel to the baseline results. By the same token, the risk factors can be evaluated at another level that denotes stressed or a particular scenario for the factors ($F(\text{stress})$). These stressed factors in turn provide the opportunity to build stress probabilities of default for all of the positions in the portfolio.

5.2. Simulating the Loss Distribution

Most financial institutions have a large portfolio and are therefore required to maintain a well-organized information system regarding the loans that make up their portfolio. The most efficient portfolio credit risk models are those that are directly linked to the bank's information system, which enables them rapidly and thoroughly to provide reports on the creditworthiness of the portfolio. A schematic example of such a risk engine is presented in Figure 8.

Figure 8: Estimation and Simulation Engine to Derive Loss Distribution



Source: Authors' elaboration.

This example depicts the steps taken to go from the loan data to the loss distribution and the economic capital.

1. The starting point is the data warehouse, where the bank centralizes all of the information on the loans in the portfolio. For this exercise, the following information is needed: the peso value of average exposures, the recovery rates, the number of obligations, and the number of defaults. The bank regularly has an area in charge of monitoring and maintaining an updated system on observable risk factors.
2. The default probability model uses the information on the number of obligations, the number of defaults, and the historical information on risk factors to estimate

the parameters that relate the risk factors to the default process. The main output at this stage is an estimated equation that can provide, given an scenario for the observable factors, the conditional default probabilities.

3. At the simulation stage, first there is a probability model that receives the number of obligations and the default probabilities (conditional or unconditional) and generates draws on the number of defaults expected out of a given number of loans granted in that segment of the portfolio. Let s denote the index for the simulation, $s = 1, \dots, S$.

$$y_{p,z}^{s} \sim Binomial(k_{p,z}, PD_{p,z}) \tag{7}$$

where $k_{p,z}$ is the number of obligations and $PD_{p,z}$ can be defined as the conditional or unconditional default probabilities. These two parameters determine a draw from the binomial distribution. This draw represents the number of defaulted obligations within the segment of the portfolio associated with produce p and zone of production z. Let $Y_{p,z} := \{y_{p,z}^{s}\}_{s=1}^{S}$ denote the succession of these draws, that is, the total number of simulated defaults for that particular element in the portfolio.

4. The default counts generated by the probability model, the average exposures, and the recovery rates provide the value of the net losses in this segment of the portfolio

$$\mathrm{l}_{\mathrm{p,z}}^{\mathrm{s}} = \mathrm{E}_{\mathrm{p,z}}(1 - \mathrm{RR}_{\mathrm{p,z}})\mathrm{y}_{\mathrm{p,z}}^{\mathrm{s}} \tag{8}$$

Let $L_{p,z} := \{l_{p,z}^{s}\}_{s=1}^{S}$ denote the succession of losses. $L_{p,z} \sim \hat{f}_{p,z}$ where $\hat{f}_{p,z}$ denotes the empirical loss distribution.

5. The simulated individual losses on every position in the portfolio are aggregated to obtain the overall loss distribution of the portfolio

$$l^{s} = \sum_{p}\sum_{z} l_{p,z}^{s}$$

Let $L := \{l^s\}_{s=1}^{S}$ denote the succession of portfolio losses. $L \sim \hat{f}$ where $\hat{f}$ denotes the empirical loss distribution. Note that the previous expression implies independence across the losses of the homogenous risk groups. In practical terms it also implies no diversification gain from holding and providing loans across different crops or regions. This is therefore quite a conservative estimate of total losses. This is, however, a strong assumption for draws $y_{p,z}^{s}$ based on the historical unconditional PD^{hist} default probability. But this is not the case for draws based on the conditional PD^{cond} default probability, since $y_{p,z}^{s}$ are conditionally independent given the realization of the factors. This is an important element to keep in mind when obtaining loss distributions based on conditional or unconditional default probabilities.

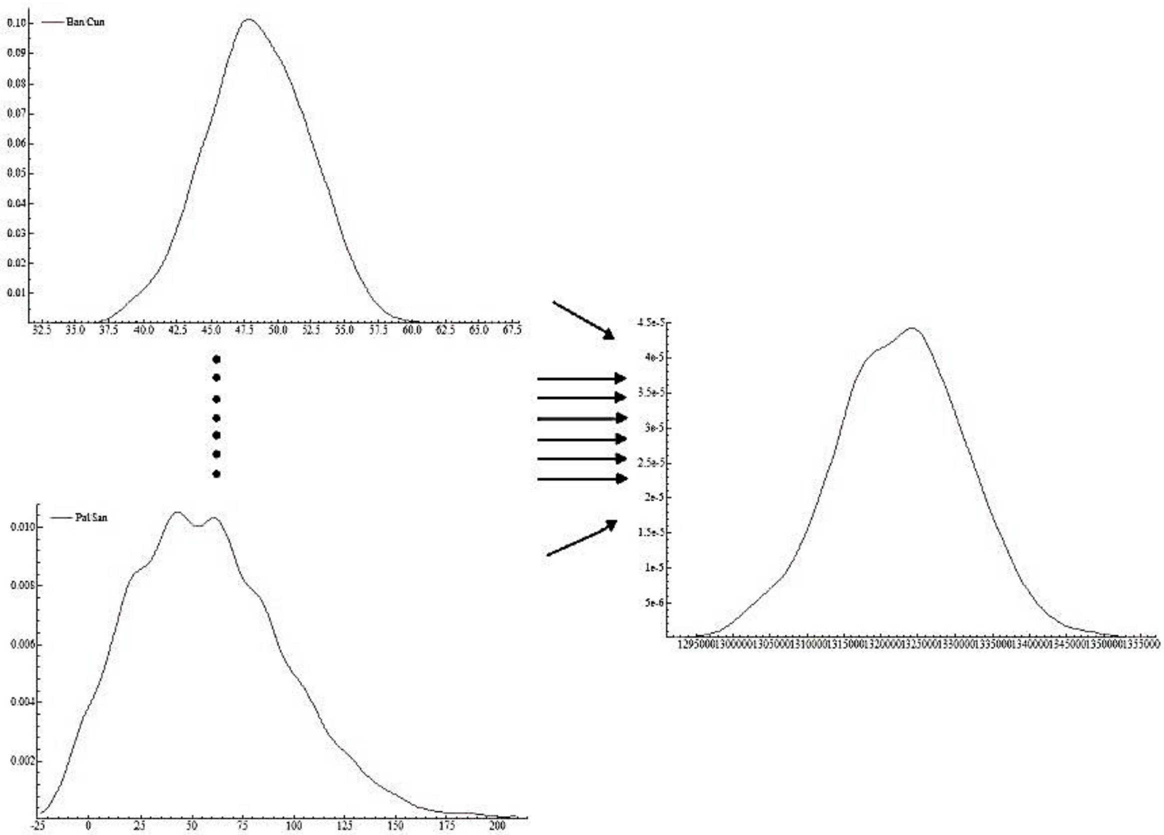
6. After we obtained the empirical loss distributions (at the portfolio level $\hat{f}$ and/or specific exposure levels $\hat{f}_{p,z}$, we can apply standard risk measures such as value-at-risk ($VaR_L(\tau)$) or expected shortfall ($ES_L(\tau)$) to quantify tail risk. Economic capital is interpreted as the amount of capital buffer that a bank needs to set aside to avoid being insolvent with τ confidence level. Most of the time it is calculated at the one-year horizon. A common market practice is to use the following estimates depending on the risk measure chosen.

$$EC_{VaR} = VaR_L(\tau) - \bar{L}$$
$$EC_{ES} = ES_L(\tau) - \bar{L}$$

where $\bar{L}$ is the average loss (center of the loss distribution).

At the end of the simulation process, we arrive at a distribution for the overall losses of the financial institution. This distribution can be based on average conditions or stress conditions based on particular scenarios that management considers relevant. An example of the type of picture to be expected when plotting the loss distribution is presented in Figure 9.

Figure 9: Distribution of Losses for the Positions in the Portfolio and the Overall Loss Distribution of the Portfolio.



Source: Authors' elaboration.
Notes: The support of all of the distribution is currency expressed in billions of pesos (COP).

6. Baseline Results: Empirical Application

This section contains an example based on the data from the rural bank's portfolio and the observed factors to illustrate the procedure for simulating the loss distribution (for both the individual position and the overall portfolio) and estimating the risk measures and the economic capital.

As explained in the previous section, the first step is to obtain the information required for the simulation. The altitude group portfolio is composed of 115 positions or homogenous risk groups determined by the pairing of produce and zones of production. Columns (3) to (5) in Table 3 present a sample (17 homogenous risk groups) of the information from the portfolio used in simulation: average exposure (value in portfolio in billion COP), the average number of obligations that are in the homogenous risk group, the historical unconditional probability of default. All of the averages are taken with respect to the time series information. The table represents the average behavior of the homogenous risk groups in the portfolio from December 2005 to November 2012.

Column (6), conditional probability of default, is also considered in the simulation. However, we do not estimate this probability of default using the historical data but rather using the default probability model. The conditional default probability is obtained using the observed factors (macroeconomic, price volatility, and climate) evaluated at their historical average levels. In particular, climate information say temperature is considered to be at its historical (seasonal) median level. In general, columns (5) and (6) of Table 3 show that historical probabilities and conditional probabilities of default look very similar. This implies that the estimated model, in Section 4, is able to provide an accurate picture of the average default risk conditions within the rural portfolio.

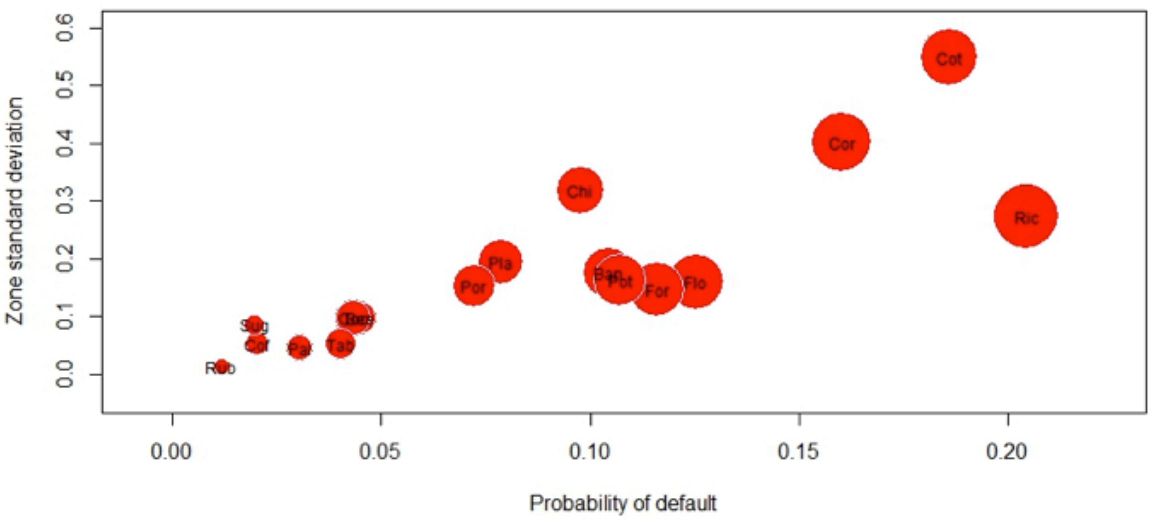
Table 3: Sample of Data for Simulation: Zone of Production Based on Altitude Groups

	2	3	4	5	6	7	8	9
Produce	**Zone**	**Value in Portfolio**	**No. Of Obligations**	**PD Historical**	**PD Conditional**	**ES(0.99)**	**Economic Capital (ES)**	**EC(ES)/Exposure**
Cot	101-500	14,3	186	65%	58%	889,4	123,2	4,6%
Ric	501-1200	14,5	322	25%	20%	606,5	130,5	2,8%
Chi	1201-1800	3,8	717	30%	28%	432,7	55,8	2,1%
Ban	501-1200	4,6	411	17%	21%	253,2	52,7	2,8%
For	2201-3000	0,8	26	13%	5%	1,7	1,1	5,3%
Coc	0-50	7,9	1313	11%	9%	563,5	118,7	1,1%
Cof	1201-1800	216,1	54071	5%	4%	257832,2	14189,2	0,1%
Sug	1201-1800	5,3	81	7%	7%	32,9	17,9	4,2%
Rub	101-500	15,5	176	2%	2%	67,2	43,0	1,6%
Flo	>3000	1,7	86	25%	11%	14,0	6,2	4,2%
Bee	501-1200	198,2	29419	10%	10%	306397,3	14112,4	0,2%
Cor	101-500	3,0	403	45%	34%	239,1	37,2	3,1%
Pal	51-100	33,5	152	3%	4%	230,4	117,3	2,3%
Pot	1801-2200	2,3	375	22%	12%	68,1	17,6	2,1%
Pla	501-1200	15,0	4053	21%	16%	5215,7	414,0	0,7%
Por	501-1200	11,4	3689	18%	15%	3512,6	331,5	0,8%
Tab	2201-3000	0,3	66	6%	2%	0,7	0,5	2,4%

Source: Authors' elaboration.
Notes: Value in portfolio in billions of COP (exposure); number of obligations; historical and conditional default probability; expected shortfall and economic capital in billions of COP; economic capital as percentage of full exposure.

As observed, there is important variation in the historical and conditional default probabilities across the portfolio, ranging from 65 percent and 58 percent in cotton at zone 101-500, to 2 percent in rubber at zone 101-500. To get a complete picture of the variation, Figure 10 reports the average historical default probability of produce in the x-axis against the standard deviation across the zone of production. In addition, the size of the bubbles also reflects the distance between the largest and the smallest probability of default across the zones of production but within the same crop. The results are interesting in terms of the risk of the positions in the portfolio. We find that high probability of default produce, such as cotton, rice, and corn, also have considerable variation across the zones of production of these products, whereas low probability of default produce such as rubber and palm oil show little variation across zones of production. In terms of the number of obligations and the average size of the exposures, the most important positions in this sample are coffee in zone 1201-1800 and livestock in zone 501-1200.

Figure 10: Average Historical Default Probabilities for each Crop and Variation across Zone of Production Altitude



Source: Authors' elaboration.
Note: The abbreviations (inside the bubbles) correspond to the crops mentioned in Table 1.

In addition to the information on the portfolio or the estimated default risk model, we make an assumption about the recovery rate of $50\,percent$ in all of the positions in the portfolio. In other words, when there is a default, the rural bank is able to recover $50\,percent$ of the value of the exposure. This is an arbitrary assumption that we must make since there is no historical information on the recovery rate on defaulted loans in this rural portfolio.

The following are used to obtain the baseline losses of the portfolio: the average exposure, the number of obligations, the assumption on the recovery rate, and—in particular—the conditional default probability (estimated from the factors at their historical average level). With the average number of obligations and the estimated default probability, we use expression 7 to take draws on the number of defaulted obligations $y_{p,z}^{s}$. We then use these draws, the value of the average exposure, and the assumed recovery rate to get a realization of the losses $l_{p,z}^{s}$. To obtain the loss distribution of each position in the portfolio, we use ($S = 1000$) simulations. Columns (7) to (9) in Table 3 contain some of the results of the simulation by position in the portfolio. Column (7) indicates the maximum one-year loss given by the expected shortfall evaluated at the 99 percent confidence level. Column (8) reports the estimated economic capital based on the previous tail risk measure. Finally, column (9) reports a blunt estimate of the ratio of unexpected losses over the full exposure. This full exposure is determined by multiplying the average exposure by the number of obligations. In general, in the subsample, this unexpected loss ratio is below 5 percent except for forest crops in zone 2201-3000. Even though we only present the subsample of the 17 positions in the portfolio, we derive the loss distribution and risk measures for all 115 positions. Although economic capital at the position level is indicative of the unexpected losses on this position, we should refrain from using this estimation to determine the capital contribution of each position to the buffer for the full portfolio. The reason is that economic capital allocation should depend on the composition of the whole portfolio and not on the individual losses that each position may face.

Figure 9 contains an example of the simulated loss distribution for bananas and palm oil for zones 501-1200 and 51-100, respectively. The shapes of these distributions are quite different from one other, reinforcing the specificity of the risk inherent in each of these rural projects. The figure also depicts the process of adding together the individual losses for each position in order to obtain the overall losses of the entire portfolio (picture on the right-hand side). Once we have the loss distribution (at both levels: individual and overall) we can use the risk measures such as VaR ($\tau = 99\ percent$) and expected shortfall ($\tau = 99\ percent$) to determine the maximum loss that we should observe most of the time.

7. Scenarios and Stress Testing: Empirical Application

Simulation of the loss distribution can also be used for stress testing. The crucial element for informed stress testing is the development of a series of scenarios based on the observable risk factors. For the current exercise, we build five scenarios to compare against the baseline results. The baseline results were constructed using the following assumptions:

- The recovery rate is set at 50 percent.
- Macroeconomic conditions are determined to be 1.6 percent annual growth in agricultural production and 5 percent growth in the agricultural producer price index.
- Price volatility is set to its historical average.
- Since the climate variables are measured as standard deviations from their monthly median estimates, at the baseline they have less than half of a standard deviation from their historical levels. Whether this deviation is positive or negative depends on the historical value for each zone of production.

By modifying the previous assumptions, we generate the following five scenarios:

1. Sce1: Macroeconomic conditions are worsened by setting agricultural production at 0.2 percent annual growth and 19 percent growth in the agricultural producer price index. Price volatility is set to its 95 percentile. Climate variables are set at one or two standard deviations above their historical levels. We choose these deviations because they reflect, in most zones of production, the situation observed in the second semester of 2010 and the first semester of 2011, when climate conditions (especially rainfall) were severe across Colombia.
2. Sce2: All variables are set at baseline level but the recovery rate is increased to 70 percent rather than 50 percent. The purpose of this scenario is primarily to stimulate the use of an agricultural insurance scheme. There are many different ways to mitigate counterparty risk in a portfolio and hence increase the recovery rate on a particular loan once a default event is triggered, such as increasing the collateral requirements or increasing the effectiveness of legal procedures. However, an increasingly popular scheme is to provide crop yield insurance, or any type of insurance designed for agricultural projects. The increase in the recovery rate assumed in the portfolio credit risk model could possibly come from the availability of such insurance schemes.
3. Sce3: only considers the price volatility effects of scenario 1.

4. Sce4: only considers the increase in rainfall from scenario 1.
5. Sce5: only considers the increase or decrease in temperature from scenario 1.

From the baseline and scenarios we derive conditional default probability $PD_{p,z}^{cond}$ and then use these as a primary input to simulate losses across the rural portfolio.

The results of the stress test implied by the scenarios in the rural portfolio are presented in Table 4 and Figure 11. The results convey three main messages. First, as expected, a substantial increase in the recovery rate (Scenario 2) reduces economic capital by around 40 percent. Important steps should be taken to mitigate losses once a default occurs. One such scheme is to implement some sort of agriculture-oriented insurance. We saw that default is not an unlikely event in agricultural portfolios (it is even highly likely in some crops, such as cotton). Second, weather-related risk is probably more important that price volatility; commodity price-related shocks increase economic capital by 12.2 percent, while weather-related shocks increase economic capital by 20 percent. Third, there is no conclusive evidence on whether one particular weather shock, i.e., rainfall or temperature, is more important than another in determining the default and hence economic capital. From the estimation (of the default risk model) and the stressed scenarios, they seem to be equally important.

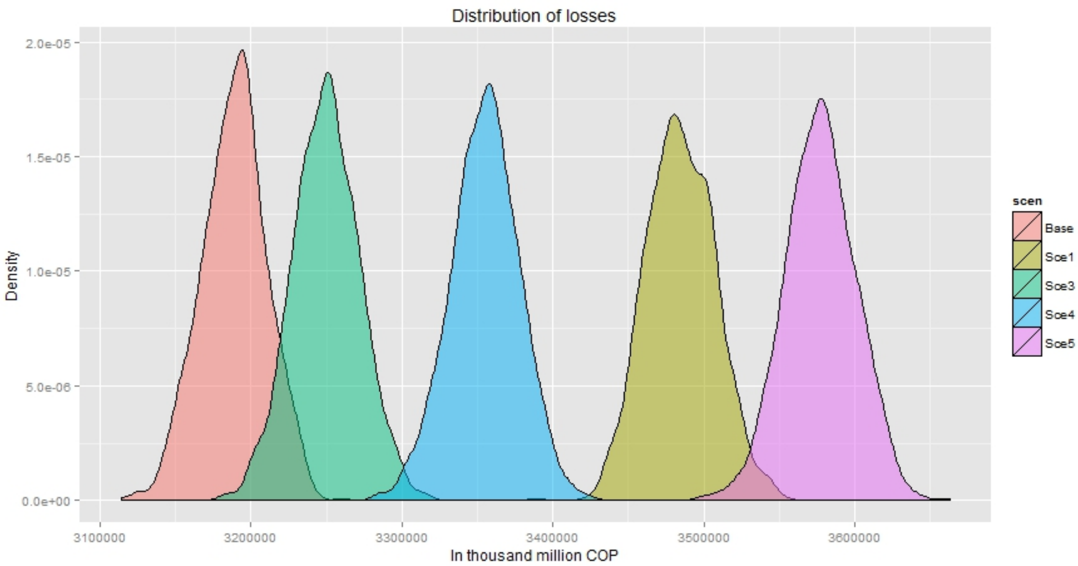
As observed in Figure 11, the different scenarios are not having a particularly strong effect on the tails of the distribution; rather, they shift the location of the distribution. This is consistent with the fact that the shocks that have been assigned in the different scenarios are to specific factors rather than common factors.

Table 4: Percentage Change in Tail Risk Measure (expected shortfall) and Economic Capital of Scenarios with respect to Baseline Measure

	Altitude	
	ES(0.99)	**EconomicCapital(ES)**
Sce1	9,4%	17,6%
Sce2	-39,7%	-23,9%
Sce3	2,1%	12,2%
Sce4	5,4%	16,7%
Sce5	12,3%	19,7%

Source: Authors' elaboration.

Figure 11: Loss Distributions of Baseline and Stressed Scenarios, Zone of Production Altitude



Source: Authors' elaboration.

8. Conclusions

Rural banks face important challenges in terms of risk management. Many of them are inherited from their specific charter (sometimes government-sponsored) to support the development of the agricultural sector. In theory, they are specialized institutions, unlike commercial banks. The reality is that each agricultural project supported by a loan is sensitive to many risk factors (e.g., price of inputs, demand, weather conditions, uncertainty of spot price of produce). We use this fact to create a risk map and design a portfolio credit risk model that can be accommodated to the portfolio of an existing rural bank. In particular, our design is interested mainly in the effects of commodity price volatility (the final price at which the farmer is able to sell his produce in the spot market) and the effect of climate variation (temperature and rainfall).

In general, the literature on agricultural risk management is more normative than practical, and there are very few applications in actual portfolios. We design and estimate a default risk model that is able to adequately replicate (in-sample) the average behavior of historical default probabilities of the rural portfolio. Furthermore, we explain how to use the model to derive estimates of the unexpected losses of the rural portfolio and estimate the economic capital required to cover such losses. This portfolio credit risk model allows us to perform stress testing on the parameters used in the simulation. From the stress test we can compare the relevance of the different risk factors as determinant of the economic capital. The results of the stress test indicate that climate risk factors are more important than commodity price volatility. However, we cannot precisely determine whether temperature or rainfall is more important. Further studies should focus on the relationship between regional variation in rainfall (and to a lesser extent temperature) and crop yields. This could be very important for the success of an agricultural project and hence the financial viability (from the point of view of a rural bank) of granting a loan to such project.

References

Behrens, A. and G. Pederson. 2007. "An Analysis of Credit Risk Migration Patterns of Agricultural Loans. *Agricultural Finance Review* 67: 87–98.

Castro, C. 2012. "Confidence Sets for Asset Correlation in Portfolio Credit Risk." *Revista de Economia del Rosario* 15(1): 19–58.

de Servigny, A. and O. Renault. 2004. *The Standard & Poor's Guide to Measuring and Managing Credit Risk*. First Edition. New York, NY: McGraw-Hill.

Demey, P., J.-F. Jouanin, and C. Roget. 2004. "Maximum Likelihood Estimate of Default Correlation." *Risk* (November): 104–8.

Escalante, C., P. Barry, T. Park, and E. Demir. 2004. "Farm-level and Macroeconomic Determinants of Farm Credit Migration Rates." *Agricultural Finance Review*, 64: 135–49.

Fama, E. and K. French. 1993. "Common Risk Factors in the Returns on Stocks and Bonds." *Journal of Financial Economics* 33(1): 356

Geman, H. 2005. *Commodities and Commodity Derivatives. Modelling and Pricing for Agriculturals, Metals and Energy*. Hoboken, NJ: John Wiley & Sons, Ltd.

Gilbert, C. and C. Morgan. 2010. Review food price volatility. *Philosophical Transactions of the Royal Society*, 365(1554): 3023–34.

Gloy, B. A., E. L. LaDue, and M. A. Gunderson. 2005. "Credit Risk Migration and Downgrades Experienced by Agricultural Lenders." *Agricultural Finance Review* 65(1): 1–16.

Gordy, M. 2000. "A Comparative Anatomy of Credit Risk Models. *Journal of Banking and Finance* 24(1-2): 119–49.

Gupton, G., C. Finger, and M. Bhatia. 1997. *CreditMetrics.* Technical Document. New York, NY: J.P. Morgan & Co., Inc.

Hess, U., K. Richter, and A. Stoppa. 2004. *Weather Risk Management for Agriculture and Agrobusiness in Developing Countries.* Washington, DC and Rome, Italy: IFC, World Bank and ProcomAgr.

Kahn, B. and D. Zaks. 2009. *Investing in Agriculture: Far-reaching Challenge, Significant Opportunity.* Frankfort, Germany: Deutsche Bank Climate Change Advisors.

Katchova, A. and P. Barry. 2005. "Credit Risk Models and Agricultural Lending." *American Journal of Agricultural Economics* 87(1) (February): 194–205.

Koopman, S., A. Lucas, and B. Schwaab. 2011. “Modeling Frailty-Correlated Defaults Using many Macroeconomic Covariates.” *Journal of Econometrics* 162: 312–325.

Koopman, S., A. Lucas, and R. Daniels. 2008. A Non-Gaussian Panel Time Series Model for Estimating and Decomposing Default Risk. *Journal of Business and Economic Statistics* 26(4): 510–525.

Larson, D., J. Anderson, and P. Varangis. 2004. “Policies on Managing Risk in Agricultural Markets.” *The World Bank Research Observer* 19(2): 199–230.

McCullagh, P. and J. Nelder. 1989. *Generalized Linear Models*. Chapman & Hall.

Merton, R. 1974. “On the Pricing of Corporate Debt: The Risk Structure of Interest Rates.” *The Journal of Finance* 29(2): 449–70.

OECD (Organisation for Economic Co-operation and Development). 2011. *Risk Management in Agriculture: What Role for Governments?* Paris, France: OECD.

Pederson, G. and L Zech. 2009. “Assessing Credit Risk in an Agricultural Loan Portfolio.” *Canadian Journal of Agricultural Economics* 57:169–85.

Ross, S. 1976. “The Arbitrage Theory of Capital Asset Pricing.” *Journal of Economic Theory* 13(3): 341–360.

Schaffnit-Chatterjee, C. (2010. “Risk Management in Agriculture: Toward Market Solutions in the EU.” *Deutsche Bank Research. Current Issues*.