DEPARTMENT OF FOOD AND RESOURCE ECONOMICS
UNIVERSITY OF COPENHAGEN

IFRO Working Paper

Panel Data Nonparametric Estimation of Production Risk and Risk Preferences: An Application to Polish Dairy Farms

Tomasz Gerard Czekaj
Arne Henningsen

2013 / 6

IFRO Working Paper 2013 / 6

Panel Data Nonparametric Estimation of Production Risk and Risk Preferences:
An Application to Polish Dairy Farms

Authors: Tomasz Gerard Czekaj, Arne Henningsen

Department of Food and Resource Economics (IFRO)

University of Copenhagen

Rolighedsvej 25

DK 1958 Frederiksberg DENMARK

www.ifro.ku.dk

Panel Data Nonparametric Estimation of Production Risk and Risk Preferences: An Application to Polish Dairy Farms

TOMASZ GERARD CZEKAJ and ARNE HENNINGSEN

Department of Food and Resource Economics, University of Copenhagen,
Rolighedsvej 25, 1958 Frederiksberg C, Denmark
e-mail: tcz@ifro.ku.dk, arne@ifro.ku.dk

October 2013

Abstract

We apply nonparametric panel data kernel regression to investigate production risk, output price uncertainty, and risk attitudes of Polish dairy farms based on a firm-level unbalanced panel data set that covers the period 2004–2010. We compare different model specifications and different approaches for obtaining firm-specific measures of risk attitudes. We found that Polish dairy farmers are risk averse regarding production risk and price uncertainty. According to our results, Polish dairy farmers perceive the production risk as being more significant than the risk related to output price uncertainty.

Keywords: production risk, price uncertainty, nonparametric econometrics, panel data, Polish dairy farms

JEL codes: C14, C23, D24, Q12

1. Introduction

Uncertainty and risk are inherent features of agricultural production (Moschini and Hennessy, 2001).[1] Farmers' revenue is uncertain due to both output price uncertainty and production uncertainty (production risk). The latter is related to the uncertainty of the outcome of biological processes in agricultural production due to weather conditions, plant and animal diseases, natural disasters (droughts, floods etc.), and even climatic changes in the long run. Therefore, a proper analysis of agricultural production technologies and farmers' production decisions should account for these uncertainties, factors that influence the uncertainties, and the producers' attitudes towards risk (risk preferences).

The predominant approach to investigating production risk is based on the parametric stochastic production function which was introduced by Just and Pope (1978).[2] Most studies that use this approach (e.g. Griffiths and Anderson, 1982; Antle, 1983, 1987; Kumbhakar, 1993, 2002b; Tveterås, 1999; Tveterås, Flaten and Lien, 2011; Asche and Tveterås, 1999, to mention only a few.) focus on the risk in production in agriculture or aquaculture.

Other studies (e.g. Sandmo, 1971; Chambers, 1983; Appelbaum and Ullah, 1997; Kumbhakar, 2002b,a; Kumbhakar and Tsionas, 2009) only analyse price risk and ignore production risk or analyse price risk and production risk separately. However, since farmers face both price risk and production risk at the same time, both sources of risk should be analysed simultaneously, because neglecting one of them may result in a misspecified model (Kumbhakar, 2001). Therefore, we use the method proposed by Kumbhakar (2001) that extends the approach of Just and Pope (1978) to simultaneously account for production risk and price uncertainty.

Most of the studies that use the Just and Pope (1978) framework rely on rather strong assumptions, e.g. that the mean production function (and output variance function) are of a distinct (parametric) functional form. It is well known that the specification of the functional form of a production function (or generally a regression function) plays a fundamental role in econometric analyses. If the functional form of the regression function is different from the "true" functional form of the relationship between the dependent

[1] Although the term "risk" often has negative connotations, we use the terms "risk," "uncertainty," and "variability" interchangeably, with all terms meaning that a future state (e.g. a price or a production output) is not known in advance (in the decision making process) and might be smaller or larger than the expected value (see e.g. comments included in the paper of Gardebroek, Chavez and Lansink, 2010).

[2] Alternatively, the "state contingent approach" proposed by Chambers and Quiggin (2000) can also be used to analyse production risk. However, this approach is rarely used in empirical studies mainly because it requires very detailed and large data sets in order to reflect a reasonable number of states of nature (Just, 2003). Rasmussen (2004) suggest estimating "state contingent" production functions for only a few states of nature. A recent discussion on the "state contingent approach" can be found in Shankar (2012).

variable and its covariates, the obtained results will be biased and their relevance questionable.

To avoid a possible misspecification of the functional form of the regression functions in the Just and Pope (1978) framework, Kumbhakar and Tsionas (2009, 2010) proposed to use nonparametric regression methods. Our analysis is based on the approach of Kumbhakar and Tsionas (2009, 2010) but it includes three important improvements. First, although Kumbhakar and Tsionas (2009, 2010) use panel data in both of their studies, they estimate a "pooled model" that does not account for individual and time heterogeneity. In contrast, we use a nonparametric regression method that was proposed by Henderson and Simar (2005) and Racine (2008) which accounts for the panel structure of our data set.

Second, while Kumbhakar and Tsionas (2009, 2010) do not use any formal tests to assess the statistical significances of their estimates, we use the bootstrapping method proposed by Racine (1997) and Racine, Hart and Li (2006) to test the significance of the explanatory variables in the nonparametric regressions.

Third, while Kumbhakar and Tsionas (2009, 2010) use a local-constant kernel estimator, we use a local-linear kernel estimator, as it usually outperforms the local-constant kernel estimator (Li and Racine, 2004).

Finally, we compare different specifications in the nonparametric estimation of the mean production function and compare different approaches to obtain the values of the risk preference functions.

We apply this method to Polish dairy farms. Dairy farmers are affected by production risk, e.g. caused by weather conditions that influence the quantity and quality of the feed produced at the farm or by animal diseases that directly influence the milk production. Dairy farmers in the European Union (EU) have been less affected by milk price volatility and therefore price risk in the past, but this has changed in recent years. The reason is an ongoing process of liberalisation of the dairy sector in the EU. Furthermore, it is anticipated that price uncertainty may even increase in the coming years due to the expected abolition of the milk quota system in the EU in 2015. Therefore, both theoretical and empirical studies on farms' production and price risk are important and relevant.

The rest of the paper is organised as follows. Section 2 presents a theoretical framework for the risk analysis. Section 3 provides the specification of the applied nonparametric econometric model, section 4 describes the data set on Polish dairy farms that is used in our empirical application, and section 5 presents the results of the analysis. Section 6 concludes.

2. Analytical framework

Just and Pope (1978) proposed an analytical framework to investigate the production risk using the model of a stochastic production function. The general form of the Just and Pope (1978) production function is given by:

$$y = f(x, z) + u = f(x, z) + h(x, z)\,\varepsilon, \tag{1}$$

where y is the observed output quantity, x is a vector of variable input quantities $(x_1, \ldots, x_J)$, z is a vector of quasi-fixed input quantities $(z_1, \ldots, z_K)$, $f(.)$ is the mean production function, $h(.) \geq 0$ is the output variability function, u is an additive and possibly heteroskedastic error term with zero mean and variance $(h(.))^2$, and ε is a homoskedastic error term with zero mean and a variance of one. The output variance function $h(x, z)$ is used to analyse the marginal influences of inputs on the variance of production. The marginal effect of the j-th input is risk increasing if the first order partial derivative of the $h(x, z)$ function with respect to j-th input (denoted by $h_j(x, z)$) is positive, or risk decreasing if $h_j(x, z)$ is negative. Consequently, inputs which have an insignificant effect on $h(x, z)$ are risk neutral.

The model of Just and Pope (1978) has been extended to panel data by Griffiths and Anderson (1982). Kumbhakar (1993) implemented the estimation of technical efficiency. Kumbhakar (2001) included price uncertainty (both regarding the output price and the input prices) and recently Kumbhakar and Tsionas (2009) derived models that include production risk, output price uncertainty, and technical efficiency in a nonparametric framework.

An extensive description of the models that generalise the Just and Pope (1978) model to account for both production risk and price uncertainty is given in Kumbhakar (2001) and Kumbhakar and Tsionas (2009). However, for the convenience of the reader, we briefly present the fundamental derivations of this modelling approach here.

In the analyses of the risk in production it is often assumed that producers maximise the expected utility based on anticipated normalised restricted profit:[3]

$$\max_x E\left[U\left(\frac{\pi^e}{p}\right)\right], \tag{2}$$

where π^e is the anticipated restricted profit[4] and p is the output price.

[3] Since the anticipated restricted profit is homogeneous of degree one in input and output prices, it is common to impose the homogeneity condition by normalizing anticipated profit with either the output price or the price of one of the inputs (Kumbhakar and Tsionas, 2009).

[4] "Restricted profit" can also be called "short-run profit" or "gross margin." For simplicity, we use the term "profit" in the remainder of the paper.

2.1. Production risk only

When only the production risk is taken into consideration, the anticipated profit is defined as:

$$\pi^e = p\,y - w'\,x = p\,f(x,z) - w'\,x + p\,h(x,z)\,\varepsilon, \tag{3}$$

where w is a vector of the prices of the variable inputs $(w_1, \ldots, w_J)$.

In this case, the normalised anticipated profit is given by:

$$\frac{\pi^e}{p} = y - \frac{w'x}{p} = f(x,z) - \frac{w'x}{p} + h(x,z)\,\varepsilon = f(x,z) - \widetilde{w}'x + h(x,z)\,\varepsilon, \tag{4}$$

where $\widetilde{w}$ is a vector of normalised input prices with elements $\widetilde{w}_j = w_j/p \;\forall\; j = 1, \ldots, J$.

Under the assumption that the producers maximise their expected utility of normalised anticipated profit $(E\left[U\left(\pi^e/p\right)\right])$, we get the following first-order conditions (FOC):

$$E\left[U'\left(\frac{\pi^e}{p}\right)\left(f_j(x,z) - \widetilde{w}_j + h_j(x,z)\,\varepsilon\right)\right] = 0 \;\forall\; j = 1, \ldots, J, \tag{5}$$

where $U'(\pi^e/p)$ is the marginal utility of anticipated normalised profit, and f_j and h_j denote the first derivatives of the mean production function and the output variability function, respectively, with respect to the j-th variable input.

To derive the risk preference function, one can rewrite (5) in the following way:

$$f_j(x,z) = \widetilde{w}_j - h_j(x,z)\frac{E[U'(\pi^e/p)\,\varepsilon]}{E[U'(\pi^e/p)]} = \widetilde{w}_j - h_j(x,z)\,\theta_1 \;\forall\; j = 1, \ldots, J, \tag{6}$$

where

$$\theta_1 \equiv \frac{E[U'(\pi^e/p)\,\varepsilon]}{E[U'(\pi^e/p)]} \tag{7}$$

is the risk preference function. A positive (negative) value of θ_1 indicates risk averse (risk seeking) producers and $\theta_1 = 0$ indicates risk neutral producers (Chambers, 1983).

As shown by Kumbhakar and Tsionas (2010), the risk preference function (θ_1) can be related to the Arrow-Pratt measure of risk aversion in the following way:[5]

$$\begin{aligned}\theta_1 &\equiv \frac{E\left[U'\left(\pi^e/p\right)\varepsilon\right]}{E\left[U'\left(\pi^e/p\right)\right]} = \frac{E\left[\left(U'\left(\pi^e/p\right) + U''\left(\pi^e/p\right)\cdot h(x,z)\,\varepsilon\right)\varepsilon\right]}{E\left[U'\left(\pi^e/p\right) + U''\left(\pi^e/p\right)\cdot h(x,z)\,\varepsilon\right]} \\ &= \frac{U''(\mu_\pi)h(x,z)}{U'(\mu_\pi)} = -AR(\mu_\pi)\;h(x,z)\end{aligned} \tag{8}$$

where $AR(\mu_\pi) = -U''(\mu_\pi)/U'(\mu_\pi)$ is the Arrow-Pratt measure of risk aversion and $\mu_\pi = E[\pi^e/p]$ is the expected profit.

[5] The proof is similar to the proof of Proposition 1 in the Appendix of Kumbhakar and Tveterås (2003).

The risk premium (RP) is the amount that makes the producer indifferent between the uncertain normalised profit (π^e/p) and the certain profit ($E[\pi^e/p] - RP$). Based on a first-order Taylor series approximation of $U(E[\pi^e/p] - RP)$ at $RP = 0$ and a second-order approximation of $U(E[\pi^e/p])$ at $\pi^e/p = \mu_\pi$, Antle (1987) and Chavas and Holt (1996) show that the risk premium can be obtained by:

$$RP = 0.5 \; AR(\mu_\pi) \; \mathrm{Var}\,(\pi^e/p) \,. \tag{9}$$

Since the variance of the anticipated profit $\mathrm{Var}(\pi^e/p)$ is given by $h(x, z)^2$ in the model that accounts for production risk only, we can write:

$$RP = -0.5 \, \frac{\theta_1}{h(x,z)} \; h(x,z)^2 = -0.5 \; \theta_1 \; h(x,z). \tag{10}$$

2.2. Price risk only

In the case of no production risk (i.e. $h(x, z) = 0$) but uncertain output price, Kumbhakar and Tsionas (2009) follow Zellner, Kmenta and Dreze (1966) and model the anticipated output price as $p^e = p\, e^\eta$, where p is the observed (realised) output price and η is an error term that reflects the uncertainty of the anticipated output price. It is assumed that $E[e^\eta] = 1$ so that the expected value of p^e is the same as the observed price p. Under these assumptions, the anticipated profit is defined as:

$$\pi^e = p^e y - wx = pf(x,z) - w'x + p\, f(x,z)(e^\eta - 1). \tag{11}$$

Hence, the anticipated normalised profit is defined as:

$$\frac{\pi^e}{p} = f(x,z) - \widetilde{w}'x + f(x,z)(e^\eta - 1) = \mu_\pi + f(x,z)\omega_1, \tag{12}$$

where $\omega_1 = (e^\eta - 1)$.

The FOC for the maximisation of the expected utility are:

$$E\left[U'\left(\frac{\pi^e}{p}\right)(f_j(x,z) - \widetilde{w}_j + f_j(x,z)\,\omega_1))\right] = 0 \; \forall \; j = 1, \ldots, J. \tag{13}$$

These conditions may be rewritten as:

$$f_j(x,z)(1 + \theta_2) = \widetilde{w}_j \tag{14}$$

where

$$\theta_2 \equiv \frac{E[U'(\pi^e/p)\omega_1]}{E[U'(\pi^e/p)]} \tag{15}$$

is the risk preference function related to output price uncertainty. A positive (negative) value of θ_2 indicates risk averse (risk seeking) producers and $\theta_2 = 0$ indicates risk neutral producers.

We can use the same motivation as used to relate the risk preference function to the Arrow-Pratt measure of risk aversion and write:[6]

$$\theta_2 \equiv \frac{E[U'(\pi^e/p)\omega_1]}{E[U'(\pi^e/p)]} = \frac{E[(U'(\pi^e/p) + U''(\pi^e/p) \cdot f(x,z)\omega_1)\omega_1]}{E[U'(\pi^e/p) + U''(\pi^e/p) \cdot f(x,z)\omega_1]}$$
$$= \frac{U''(\mu_\pi) f(x,z)}{U'(\mu_\pi)} = -AR(\mu_\pi) f(x,z) \tag{16}$$

Since the variance of normalised profit $\mathrm{Var}(\pi^e/p)$ is given by $(f(x,z))^2 \,\mathrm{Var}(\omega_1)$ in the model with price risk only, the risk premium can be obtained as follows:

$$RP = -0.5 \, \frac{\theta_2}{f(x,z)} \, (f(x,z))^2 \,\mathrm{Var}(\omega_1) = -0.5 \, \theta_2 \, f(x,z) \,\mathrm{Var}(\omega_1). \tag{17}$$

2.3. Both production risk and price risk

Finally, when the model accounts for both production risk and output price uncertainty, the normalised anticipated profit is specified as:

$$\pi^e/p = e^\eta \, y - \tilde{w}'x = f(x,z) - \tilde{w}'x + f(x,z)(e^\eta - 1) + h(x,z)(e^\eta \varepsilon)$$
$$\equiv \mu_\pi + f(x,z)\omega_1 + h(x,z)\omega_2 \tag{18}$$

where $\omega_2 = e^\eta \varepsilon$.[7]

The first-order conditions for the maximisation of the expected utility are:

$$E\left[U'\left(\frac{\pi^e}{p}\right)(f_j(x,z) - \tilde{w}_j + f_j(x,z)\omega_1 + h_j(x,z)\omega_2)\right] = 0 \;\forall\; j = 1, \ldots, J. \tag{19}$$

This can be expressed as:

$$f_j(x,z)(1 + \tilde{\theta}_2) = \tilde{w}_j - h_j(x,z)\,\tilde{\theta}_1 \;\forall\; j = 1, \ldots, J, \tag{20}$$

where

$$\tilde{\theta}_2 \equiv \frac{E[U'(\pi^e/p)\omega_2]}{E[U'(\pi^e/p)]} = -AR(\mu_\pi) f(x,z), \tag{21}$$

and

$$\tilde{\theta}_1 \equiv \frac{E[U'(\pi^e/p)\omega_1]}{E[U'(\pi^e/p)]} = -AR(\mu_\pi) h(x,z), \tag{22}$$

[6] The proof is analogous to the proof of Proposition 1 in the Appendix of Kumbhakar and Tveterås (2003).

[7] Following Kumbhakar and Tsionas (2009) we assume independence of η and ε.

If production risk and output price uncertainty are considered jointly, the variance of normalised anticipated profit $\mathrm{Var}(\pi^e/p)$ is given by $(f(x,z))^2\,\mathrm{Var}(\omega_1)+(h(x,z))^2\,\mathrm{Var}(\omega_2)$. Once estimates of either $\tilde{\theta}_1$ or $\tilde{\theta}_2$ are obtained, the risk premium can be calculated either from:

$$\begin{aligned} RP &= -0.5\,\frac{\tilde{\theta}_1}{h(x,z)}\,\big((f(x,z))^2\,\mathrm{Var}(\omega_1)+h(x,z)^2\,\mathrm{Var}(\omega_2)\big) \\ &= -0.5\,\tilde{\theta}_1\left(\frac{(f(x,z))^2\,\mathrm{Var}(\omega_1)}{h(x,z)}+h(x,z)\,\mathrm{Var}(\omega_2)\right) \end{aligned} \tag{23}$$

or from:

$$\begin{aligned} RP &= -0.5\,\frac{\tilde{\theta}_2}{f(x,z)}\cdot\big((f(x,z))^2\,\mathrm{Var}(\omega_1)+(h(x,z))^2\,\mathrm{Var}(\omega_2)\big) \\ &= -0.5\,\tilde{\theta}_2\left(f(x,z)\,\mathrm{Var}(\omega_1)+\frac{(h(x,z))^2\,\mathrm{Var}(\omega_2)}{f(x,z)}\right). \end{aligned} \tag{24}$$

3. Nonparametric regression model

Although the analysis of risk and uncertainty based on the Just-Pope production function has been studied extensively in the past three decades, most studies use parametric specifications of the production, risk variance and risk preference functions. The only applications of the Just and Pope (1978) approach that use nonparametric regression methods are the studies of Kumbhakar and Tsionas (2009, 2010).

Following Kumbhakar and Tsionas (2009, 2010) we assume that the production technology is of a Just and Pope (1978) form:

$$y_{it} = f(x_{it}, z_{it}) + u_{it} = f(x_{it}, z_{it}) + h(x_{it}, z_{it})\epsilon_{it} \tag{25}$$

where all variables are defined as before but the subscripts i and t denote the individual firm and the time period, respectively.

In order to obtain nonparametric estimates of the mean production function $f(.)$, the output variability function $h(.)$, and the risk preference functions θ_1, θ_2, $\tilde{\theta}_1$, and $\tilde{\theta}_2$, we follow the multi-step nonparametric estimation procedure proposed by Kumbhakar and Tsionas (2009). In the first step, the estimates of the mean production function of the Just and Pope (1978) technology are obtained using nonparametric kernel regression. In the second step, the residuals of the model which was estimated in the first step are regressed on the same set of explanatory variables as in the estimation of the mean production function.[8] Once nonparametric estimates of the mean production function

[8] Just and Pope (1978) showed that the estimation of the risk variance function can be obtained after logarithmic transformation of the (absolute values) of residuals of the mean production function and remain explanatory variables. Kumbhakar and Tsionas (2009) regressed the squares of the residuals (or their absolute values) on the explanatory variables. The approaches of Just and Pope (1978) and

and the output variance function have been obtained, the risk preference functions can be nonparametrically estimated based on the theoretical model. We follow Kumbhakar and Tsionas (2009) who propose the following approach to estimate the risk preference functions. The risk preference function of the model with production risk only (7) is nonparametrically estimated based on the averaged FOC of this model (6):

$$D_1 \equiv \frac{1}{J}\sum_{j=1}^{J}\left[\frac{f_j(x,z)-\widetilde{w}_j}{-h_j(x,z)}\right] = \phi_1(\widetilde{w},x,z)+\xi_1, \tag{26}$$

where ξ_1 is an error term that captures optimisation errors, although these errors are not explicitly stated in the model (Kumbhakar and Tsionas, 2009). Setting the error term ξ_1 to zero, we can obtain the estimates of the risk preference function θ_1 as the predicted values of the regression function $\phi_1(.)$, i.e. $\theta_1 = \hat{\phi}_1(\widetilde{w},x,z)$.

Analogously, the risk preference function of the model with price uncertainty only (15) is nonparametrically estimated based on the averaged FOC of this model (14):

$$D_2 \equiv \frac{1}{J}\sum_{j=1}^{J}\left[\frac{\widetilde{w}_j}{f_j(x,z)}-1\right] = \phi_2(\widetilde{w},x,z)+\xi_2, \tag{27}$$

where ξ_2 is an error term. Setting the error term ξ_2 to zero, we can obtain the estimates of the risk preference function θ_2 as the predicted values of the regression function $\phi_2(.)$, i.e. $\theta_2 = \hat{\phi}_2(\widetilde{w},x,z)$.

A similar approach is used to obtain the risk preference function from the model that incorporates both production risk and output price uncertainty (21, 22). Kumbhakar and Tsionas (2009) propose to first estimate a regression function for the averaged FOC of this model:

$$D_3 \equiv \frac{1}{J}\sum_{j=1}^{J}\left[\frac{\widetilde{w}_j-h_j(x,z)\tilde{\theta}_1}{\widetilde{w}_1-h_1(x,z)\tilde{\theta}_1}\right] = \frac{1}{J}\sum_{j=1}^{J}\frac{f_j}{f_1} = \phi_3(\widetilde{w},x,z)+\xi_3, \tag{28}$$

where ξ_3 is an error term. Using the predicted values from the regression function $\phi_3(\widetilde{w},x,z)$, denoted $\hat{\phi}_3(\widetilde{w},x,z)$, and solving (28) for $\tilde{\theta}_1$ (setting ξ_3 to 0), the nonparametric estimates of the risk preference function related to the production risk $\tilde{\theta}_1$ can be obtained from:

$$\tilde{\theta}_1 = \frac{\sum_{j=1}^{J}\left[\widetilde{w}_j-\hat{\phi}_3(\widetilde{w},x,z)\widetilde{w}_1\right]}{\sum_{j=1}^{J}\left[h_j(x,z)-\hat{\phi}_3(\widetilde{w},x,z)h_1(x,z)\right]} \tag{29}$$

Kumbhakar and Tsionas (2009) are basically equivalent. Most of the studies that investigate the risk using the Just and Pope (1978) framework use the Cobb-Douglas or Translog functional form as an approximation of the "true" output variability function.

Solving the FOC (20) for $\tilde{\theta}_2$, the estimates of $\tilde{\theta}_2$ can be obtained based on the estimates of $\tilde{\theta}_1$:[9]

$$\tilde{\theta}_2 = \frac{\sum_{j=1}^{J} \left[\widetilde{w}_j - h_j(x,z)\tilde{\theta}_1 \right]}{\sum_{j=1}^{J} f_j(x,z)} - 1 \tag{30}$$

The nonparametric estimation of functions $\phi_1(.)$, $\phi_2(.)$ and $\phi_3(.)$ is easily affected by the "curse of dimensionality," especially if the number of variable inputs (J) is large, because the prices of the variable inputs are used along with the quantities of all variable and fixed inputs as explanatory variables so that the number of explanatory variables increases to $2 \cdot J + K$. In order to avoid the (nonparametric) estimation of the risk preference functions, observation-specific values of these functions can be directly obtained by ignoring the optimisation errors that are captured by ξ_1, ξ_2 and ξ_3. In this case, the values of the risk preference function of the model with only production risk, and the values of the risk preference function of the model with only price uncertainty can be obtained by $\theta_1 = D_1$ and $\theta_2 = D_2$ as defined in equations (26) and (27), respectively. The values of the risk preference functions of the model that accounts both for production risk and price uncertainty, $\tilde{\theta}_1$ and $\tilde{\theta}_2$, can be directly derived from equations (29) and (30) by substituting the calculated values of D_3 for the estimated values of $\hat{\phi}_3(\widetilde{w}, x, z)$ in equation (29).

For the nonparametric estimations of the mean production function, the output variance function, and the risk preference functions, we use the local-linear kernel estimator instead of the local constant estimator that was used by Kumbhakar and Tsionas (2009, 2010), because the local-linear estimator usually outperforms the local-constant estimator (Li and Racine, 2004). Furthermore, we extend the approach of Kumbhakar and Tsionas (2009, 2010) to the panel data context.[10]

In order to account for the panel data structure, we follow Henderson and Simar (2005) and Racine (2008) who estimate a fully nonparametric two-ways panel data model by applying the nonparametric regression method proposed by Li and Racine (2004) and Racine and Li (2004) that can handle both continuous and categorical explanatory variables. In this approach to nonparametric panel data estimation, categorical explanatory variables of time and firm ID are used to reflect the panel structure of the data. This estimation method has several advantages in applied production analysis.[11] However, in the context of this paper there are two important advantages. First, this method allows one to estimate the regression function without making strong assumptions regarding the functional form of the relationship between the dependent variable and the explanatory

[9] There is a typo in the formula in Kumbhakar and Tsionas (2009), where the following formula is given: $\tilde{\theta}_2 = \left(\sum_{j=1}^{J} \left[\widetilde{w}_j - h_j(x,z)\tilde{\theta}_1 \right] \right) / \left(\sum_{j=1}^{J} f_j(x,z) - 1 \right)$.

[10] Although Kumbhakar and Tsionas (2009, 2010) use panel data sets (on salmon farming in Norway and rice farming in the Philippines), they estimate "pooled" models that do not account for the panel structure of their data. Kumbhakar and Tsionas (2009) included a time trend but they neglected individual heterogeneity.

[11] For an extensive discussion of this estimation method and its usefulness in applied production analysis see Czekaj and Henningsen (2012, 2013).

variables. Secondly, the method allows us to estimate observation-specific measures of the production technology and the risk preferences (e.g. marginal products, partial production elasticities, risk premiums) without any parametric assumptions. This is especially desirable in the analysis of risk behaviour, because a parametric misspecification affects the error term u so that the analysis of the risk is based on incorrect measures of output variability. Furthermore, we test the statistical significance of the explanatory variables in the nonparametrically estimated mean production function, output variance function, and the risk preference functions using the bootstrapping method proposed by Racine (1997) and Racine, Hart and Li (2006).

4. Data

In this study, we use an unbalanced panel data set from the Polish Farm Accountancy Data Network (FADN) which consists of farms specialising in dairy production[12] in the period 2004 to 2010. Our data set includes 4,650 observations of 736 farms in total.

The dependent variable of the mean production function (Y) is the farms' output which is measured as a quantity index of total agricultural production.[13] Four inputs are used in the regression analyses: labour (L), land (A), intermediate inputs (V), and capital (K). Labour is measured by Annual Work Units (AWU), where 1 AWU equals 2200 hours of work. The total utilised agricultural area in hectares is used as a measure of land input. Intermediate inputs are measured as the sum of the total (deflated) farming overheads (e.g. maintenance, energy, services, other direct inputs) and specific costs (e.g. fodder, medicine, fertilisers, etc.).[14] The capital input is measured as the (deflated) value of total fixed assets excluding the value of land.[15] Descriptive statistics of the regression variables are presented in Table 1.

The average herd size was around 18 milking cows while the average annual milk yield per milking cow was about 4,700 kilograms of milk.[16]

[12] We selected all farms that are classified as specialised in dairy production according to the FADN methodology for at least 5 out of the 7 years covered in the data set.

[13] The quantity index of the aggregated output was calculated by dividing the value of the aggregated production by farm-specific output price indices. The farm-specific price indices of the aggregated output were calculated from regional prices of the main agricultural products that are published by the Central Statistical Office of Poland (GUS, 2012a). using Fisher's "ideal" index formula (defined as the geometric mean of the Paasche and Laspeyres price indices). In order to avoid arbitrarily choosing one year and one region to obtain the "base prices" (p^0), and arbitrarily choosing one observation to obtain the "base quantities" (q^0), we choose the sample means as "base prices" and "base quantities."

[14] Since the prices of the intermediate inputs were unavailable at the individual (farm) and regional level, we used country-level price indices of products, goods and services purchased by private farms in Polish agriculture published by the Central Statistical Office of Poland (GUS, 2012b) to deflate the costs of intermediate inputs.

[15] Similarly to the intermediate inputs, no reliable prices of capital were available at the farm level or the regional level. Therefore, we use country level price indices of investment goods and services purchased by private farms in agriculture in Poland published by the Central Statistical Office of Poland (GUS, 2012b) to obtain the quantity indices of the capital input.

[16] We assume that 1 litre of milk equals 1.031 kg.

Table 1: Descriptive statistics of the sample 2004–2010

Variable	Min	Median	Mean	Max	Std. dev.
Output (Y) [in 1,000 PLN*]	1.606	93.430	116.900	1097.000	93.763
Labour (L) [in AWU]	0.430	1.890	1.886	5.850	0.484
Land (A) [in ha]	2.880	20.080	23.510	295.700	15.646
Intermediate inputs (V) [in 1,000 PLN**]	3.716	45.570	57.100	544.300	45.132
Capital stock (K) [in 1,000 PLN**]	15.080	230.300	283.000	2378.000	210.562
Cattle (average number of cows)	1.620	15.850	17.790	120.600	10.353
Average milk yield (kg per cow per year)	1290.000	4549.000	4734.000	10880.000	1297.832

* = deflated, base = sample average (see footnotes 13),
** = deflated, base = 2004.

We treat land and intermediate inputs as variable inputs and the remaining inputs (labour and capital) as quasi-fixed inputs. While it is standard to treat intermediate inputs as variable inputs, we also treat land as a variable input, although this is less common in the literature. We do this because the farms can adjust their land input by renting land which is the case for 63.2% of the observations. In contrast, we treat labour as a quasi-fixed input because Polish family farms which specialise in dairy production mainly use family labour. For example, in our specific sample, hired labour is only used in 16.6% of the observations while the average share of hired labour in total labour is only 12.7% (6.1% at the median) within the farms that use hired labour. We also treat capital as a quasi-fixed input. This is a common assumption in the analysis of firm level risk when price uncertainty is involved. The reason is that the adjustments in the use of this input may not be observable in a short unbalanced panel data such as the one we have used.

The price of land is obtained as the regional average rent paid by farmers in the sample. Furthermore we use price indices of the intermediate inputs and the aggregated output as described in footnotes 13 and 14.

5. Results

All estimations and calculations were conducted within the statistical software environment "R" (R Development Core Team, 2012) using the add-on package "plm" (Croissant and Millo, 2008) for panel data estimations and the add-on package "np" (Hayfield and Racine, 2008) for nonparametric regression and specification tests.[17]

Since the Just and Pope (1978) production function relies on an additive error structure, we follow Kumbhakar and Tsionas (2009) and estimate the mean production function $f(x, z)$ in levels of the regression variables (output quantity and input quantities).

[17] The R commands used for this analysis are available in Appendix E.

We apply the local-linear kernel regression method proposed by Li and Racine (2004) and Racine and Li (2004) that can handle both continuous and categorical explanatory variables. In addition to the four (continuous) input quantities, we use two additional categorical explanatory variables to account for the panel structure of the data set: the firm ID and the time (Henderson and Simar, 2005). This panel data specification allows for possible individual and time heterogeneity in the data without imposing any assumptions regarding the specification of individual and time effects (Czekaj and Henningsen, 2013).

The estimation results are presented in Table A1 in Appendix A. According to the bootstrap significance test proposed by Racine (1997) and Racine, Hart and Li (2006), not a single input quantity had a statistically significant influence on the output quantity. Furthermore, many estimated values of the risk preference functions and many estimated risk premiums had implausible or even infinite values, because the marginal effects of the input quantities on $f()$ and $h()$ were often close to zero or equal to zero so that fractions with these marginal effects in the denominator became very large or even infinite.

One reason for this might be that the "true" relationship between the input quantities and the output quantity is not similar to a (local-)linear production function, which would imply perfect (local) substitutability between the inputs. Hence, the bandwidths of the input quantities must be very small in order to allow for considerably nonlinear relationships between the input quantities and the output quantity but these bandwidths might be too small to find statistically significant effects. Another problem with the nonparametric estimation with fixed bandwidths and regression variables in levels arises because the values of the regression variables in our data set have very right-skewed distributions so that there are much less observations within the bandwidths for large values of these variables (corresponding to large farms) than there are for small values of these variables (corresponding to small farms). Hence, the regression function could be under-smoothed for large farms and/or over-smoothed for small farms (Czekaj and Henningsen, 2012).

Therefore, we used a logarithmic transformation of all regression variables to estimate the mean production function and the risk variance function.[18] However, the use of the logarithmic transformation of regression variables to estimate the Just and Pope (1978) production function might not be suitable (Tveterås, Flaten and Lien, 2011). Just and Pope (1978) already showed that the logarithmic transformation of equation (25) with an additive error term becomes

$$\ln y_{it} = \ln f(x_{it}, z_{it}) + \ln \left[1 + \frac{h(x_{it}, z_{it})\epsilon}{f(x_{it}, z_{it})} \right]. \tag{31}$$

[18]The logarithmic transformation results in rather evenly distributed regression variables, which is desirable when kernel estimators with fixed bandwidths are used.

The estimation of this model might result in biased estimates because the error term:[19]

$$u^* \equiv \ln \left[1 + \frac{h(x,z)\epsilon}{f(x,z)} \right] \tag{32}$$

is likely to be correlated with the explanatory variables. Similar conclusions can be found in Antle (2010) who analyses the bias in least-squares estimations of multiplicative error models. However, the bias of estimation of the model (25) using the logarithmic form (31) is small if instability in production is relatively small (Just and Pope, 1978). Re-arranging equation (32), we get:

$$\exp(u^*) - 1 = \frac{h(x,z)\epsilon}{f(x,z)}. \tag{33}$$

A first-order Taylor series approximation of the left-hand side of the above equation at $u^* = 0$ results in:

$$u^* \approx \frac{h(x,z)\epsilon}{f(x,z)}. \tag{34}$$

The right-hand side of the above equation is uncorrelated with x and z if the assumptions that are required for an unbiased estimation of the Just-Pope production model with an additive error term (25) are met.[20] Hence, we can conclude that the biases in the estimation of model (31) are negligible if u^* is sufficiently close to zero. Therefore, we check whether u^* is close to zero and whether $u^* \approx \exp(u^*) - 1$ is a suitable approximation. These questions are graphically illustrated in Figures A1 and A2 in Appendix D. Figure A1 shows the scatter plot of u^* vs. $\exp(u^*) - 1$. If $u^* = \exp(u^*) - 1$, all points would be at the 45 degree line. It can be seen that for our dataset most observations (except some outliers) are located very close to or on this line. To further explore the bias, we present histograms of $\exp(u^*) - 1$, u^* and their difference $u^* - (\exp(u^*) - 1)$ in Figure A2. The difference between u^* and $\exp(u^*) - 1$ is less than 0.05 for 0.95% of the observations. This indicates that the bias induced by estimating model (25) using logarithmic transformation (31) is relatively small in the case of our data and model.

Hence, we decided to continue our analysis with the nonparametric estimation of the mean production function with logarithmic input and output quantities. The results of this estimation are presented in the Table 2. The cross-validated bandwidth for the (categorical) time variable is equal to 1, which indicates that this variable is smoothed out (irrelevant). In contrast, the cross-validated bandwidth for the (categorical) farm ID is relatively small, which allows for considerable heterogeneity between the individual farms.[21] The cross-validated bandwidth for the logarithmic intermediate inputs $\ln(V)$ is

[19]For simplicity, we omit subscripts indicating individuals, time and inputs.

[20] The Just-Pope production model with additive error term (25) can be estimated without bias if ε is not correlated with x or z, because in this case the error term $u = h(x,z)\,\varepsilon$ is also uncorrelated with x or z. This also implies that $u^* \approx h(x,z)\,\varepsilon / f(x,z) = u/f(x,z)$ is uncorrelated with x or z.

[21]Parmeter et al. (2012) and Czekaj and Henningsen (2012) obtained similar bandwidths for the ID variable using this estimator to investigate production functions based on panel data.

less than two times the variable's standard deviation[22] which indicates that the estimated nonparametric regression function is nonlinear in this variable. The cross-validated bandwidths for all other (continuous) inputs quantities are very large. This indicates that the mean production function $f(.)$ is linear in $\ln(L)$ for given values of all other explanatory variables ($\ln(A)$, $\ln(V)$, $\ln(K)$, ID, time), linear in $\ln(A)$ for given values of all other explanatory variables ($\ln(L)$, $\ln(V)$, $\ln(K)$, ID, time), and linear in $\ln(K)$ for given values of all other explanatory variables ($\ln(L)$, $\ln(A)$, $\ln(V)$, ID, time). However, as the other variables vary between observations, the gradients $(\partial \ln f(.)/ \ln \partial x_j)$ may also vary between observations, although the bandwidths of some explanatory variables are very large. The estimated gradients and their variation are graphically presented in Figure A3 in Appendix D and are summarised in Table 2.

Table 2: Results of the nonparametric estimation of the mean production function

Dependent variable: $\ln(Y)$				
Regressor	Bandwidth	Gradients		P-Value
		Mean	Std. Dev	
$\ln(L)$	695532.8	0.085	0.188	<0.001
$\ln(A)$	1766718.0	0.086	0.121	0.035
$\ln(V)$	0.699247	0.691	0.163	<0.001
$\ln(K)$	735637.2	0.255	0.109	<0.001
year (ordered)	1.000	0.000	0.000	0.772
ID (unordered)	0.001	0.017	0.153	0.073
$R^2 = 0.947$				

According to the results of the nonparametric regression of the mean production function of the Polish dairy farms, we can conclude that intermediate inputs and capital are the production inputs with the largest partial output elasticities, i.e. the largest relative marginal effects on the output quantity. Labour and land have considerably lower partial output elasticities (around 9% for each of these two inputs). The mean elasticity of scale (the average of the sums of the partial output elasticities) is equal to 1.04 which indicates that the analysed Polish dairy farms on average operate under slightly increasing returns to scale. The bootstrap significance test proposed by Racine (1997) and Racine, Hart and Li (2006) indicates that all (logarithmic) input quantities have a statistically significant effect on the (logarithmic) output quantity. While the effect of the farm ID was statistically significant at the 10% level, the effect of time was insignificant

The estimated *average* partial production elasticities from the model in logarithms (Table 2) are not dissimilar to the *average* partial production elasticities from the model

[22] Parmeter et al. (2012) proposed this as the "rule of thumb for linearity".

in levels (Table A1 in Appendix A). However, the individual partial output elasticities substantially differ between the two models and they are not even correlated.

In order to obtain nonparametric estimates of the output variability function in the second step of the analysis, we use the same regression methods that we used to estimate the mean production function. Following Just and Pope (1978), we regress the absolute values of the residuals of model (25)[23] on the logarithms of all input variables.

The results of the nonparametric local-linear kernel regression of the output variability function are reported in Table 3. Most cross-validated bandwidths considerably differ from the bandwidths selected for the mean production function. The bandwidth for the time variable is still large, but this variable is not smoothed out in this regression, which indicates that the output variability differs between years. The bandwidth for the ID variable is relatively small but considerably larger than in the mean production function model. This means that the individual heterogeneity is larger for the mean output than for the output variability. The bandwidths for $\ln(L)$, $\ln(A)$ and $\ln(V)$ are very large, which indicates that the output variability function is linear in $\ln(L)$, $\ln(A)$ and $\ln(V)$ for given values of the other explanatory variables. In contrast, the output variability function is nonlinear in $\ln(K)$, as its bandwidth is relatively small.

Table 3: Results of the nonparametric estimation of the output variability function

Dependent variable: $\ln\lvert\hat{u}\rvert$				
Regressor	Bandwidth	Gradients Mean	Gradients Std. Dev	P-Value
$\ln(L)$	38019.4	0.182	0.115	0.005
$\ln(A)$	852617.4	0.207	0.094	<0.001
$\ln(V)$	640764.6	0.527	0.104	<0.001
$\ln(K)$	0.533	0.237	0.120	<0.001
year (ordered)	0.759	0.009	0.029	0.010
ID (unordered)	0.077	-0.001	0.103	<0.001
$R^2 = 0.283$				

Since the marginal effects of all four (logarithmic) input quantities on the output variability are positive for almost all farms, we can conclude that all four inputs increase output variability.[24]

[23] We calculate the residuals of model (25) based on the estimation of model (31), where $\hat{u} = \hat{f}(x,z)(1 + \exp(\hat{u}^*))$ are the estimated residuals of model (25), $\hat{f}(.) = \exp(\widehat{\ln f(.)})$ are the predicted output quantities (in levels) based on the estimated model (31), and $\hat{u}^*$ are the estimated residuals from model (31).

[24] As shown in Table 3 and in figure A4, the estimated gradients of the output variability vary across individuals. Therefore, the commonly used Cobb-Douglas functional form of the output variability function would be inappropriate.

We use the two different approaches for obtaining estimates of the risk preferences functions that are described in section 3. In order to investigate the robustness of the risk preference measures, we calculate not only the *averaged* FOC (26 – 28), but also the FOC for individual (variable) inputs, i.e. setting J equal to one and j indicating either land or intermediate inputs in equations (26 – 28). Furthermore, we changed the order of the inputs in equation (28), i.e. $j = 1 =$ intermediate inputs and $j = 2 =$ land instead of $j = 1 =$ land and $j = 2 =$ intermediate inputs.

The results of the nonparametric estimations of the risk preference functions are presented in tables A3 – A12 in Appendix B. Almost all regressors (e.g. quantities of fixed and variable inputs and relative prices of variable inputs) in all these nonparametric regression models are statistically insignificant. The values of the risk preference functions that are directly calculated from equations (26), (27), and (28 – 30) are compared with the values that are derived from the nonparametrically estimated risk preference functions in Tables A13 to A15 in Appendix C

The relative risk premiums obtained from the nonparametrically estimated risk preference functions are rather implausible. For instance, the relative risk premiums obtained from the model that only accounts for production risk are far above the reasonable range, while the risk premiums obtained from the model that accounts both for production risk and output price uncertainty are negative.

Given the lack of statistical significance and the implausibility of the results obtained from the nonparametrically estimated risk preference functions, our further analysis is based on the values of the risk preference functions that are directly calculated from equations (26), (27), and (28 – 30).

We found that the values of the risk preference functions, as well as the relative risk premiums, that were obtained from the averaged FOC do not differ much from those that were obtained from the FOCs that were separately derived for each variable input. Therefore, we only focus on the risk preference functions that were obtained from the averaged FOC. The median values of the corresponding relative risk premiums (RRP) that were calculated from equations (26), (27), and (28 – 30) are presented in Table 4.

Table 4: Median values of estimated Relative Risk Premiums in the years 2004–2010

	Year							
	2004	2005	2006	2007	2008	2009	2010	Total
RRP1 (production risk)	0.333	0.343	0.302	0.233	0.243	0.200	0.192	0.258
RRP2 (output price uncertainty)	0.006	0.006	0.006	0.005	0.005	0.006	0.005	0.005
RRP3 (production risk and output price uncertainty)	0.030	0.091	0.176	0.160	0.191	0.232	0.112	0.150

Based on the values of the risk preferences that are derived from the model that only incorporates production risk, we found that most dairy farms (84.6% of the farms in the sample) are risk averse (the median value of the θ_1 is -3.17). The risk premium (RP) which is the amount that a risk averse individual is willing to pay to insure against profit uncertainty gives a more intuitive interpretation than the values of the risk preference function. Because RP depends on the unit of measurement, it is more appropriate to use the relative risk premium (RRP) which is usually defined as the ratio of the risk premium and a value of average profit. We found that a typical[25] Polish farmer who specialises in dairy production is willing to give up around 25.7% of his or her profit[26] or around 11.2% of the total revenue in order to eliminate production risk. The median values of the estimated RRP based on equation (10) are presented in the first row of Table 4.

Next we investigated farmers' attitudes towards risk when they only face price uncertainty. We use equation (27) to calculate the risk preference function associated with output price uncertainty. The values of θ_2 are negative for 91.4% of the observations, which indicates that most farmers are risk averse. In the model with only price uncertainty, the risk for farmers results from the difference between the anticipated price and the actually realised price. Since the researcher does not usually know the anticipated price, we assume that farmers anticipate that the price in the following period will remain approximately the same as in the current period, i.e. $p_t^e = p_{t-1}$. Since we postulated $p_t^e = p_t\, e^\eta$, we can calculate $\mathrm{Var}(\omega_1) = \mathrm{Var}(e^\eta - 1) = \mathrm{Var}(p_{t-1}/p_t - 1)$, which is 0.0107 in our data set (corresponding to a coefficient of variation of 10.4%).[27] We use equation (17) in order to calculate the risk premiums (RP) and the corresponding relative risk premiums (RRP). The average RRP based on our data was very close to zero (0.005 at the median) which indicates that most farmers are not willing to pay to insure against the risk related to uncertain output prices. The median values of the RRP for the period 2004-2010 are presented in the second row of Table 4.

Finally, we apply the model that facilitates a joint analysis of production risk and output price uncertainty. The risk preference function related to the production risk ($\tilde{\theta}_1$) is negative for 59.3% of the observations, while the risk preference function related to price uncertainty ($\tilde{\theta}_2$) is negative for 62.9% of the observations. Hence, the results from this model confirm our previous findings that most of the analysed Polish dairy farmers are risk averse regarding production risk and the risk related to output price uncertainty. Assuming that e^η and ε are independent and given that the variance of ε is normalised to

[25] We will refer to the median values rather than to the mean values because the calculated RRP are not evenly distributed and the calculation of mean values is affected by outliers.

[26] It is important to emphasise that farmers in the EU and Poland often face negative profit if one excludes the value of agricultural policy support (i.e. direct payments from the Common Agricultural Policy (CAP).

[27] The calculation of the output price indices is described in footnote 13. As different types of output have different price changes, and the farms in the data set have different proportions of the different types of output, the calculated output price indices differ between observations. We calculated $\mathrm{Var}(\omega_1)$ as the variance over all individuals and time periods in the sample, for which lagged prices were available, i.e. $\mathrm{Var}(\omega_1) = \mathrm{Var}(p_{i,t-1}/p_{it} - 1)$, and used the same variance for all observations in the data set. We leave the development of a more sophisticated approach for obtaining $\mathrm{Var}(\omega_1)$ for future research.

one, we obtain $\text{Var}(\omega_2) = \text{Var}(e^{\eta}\epsilon) = \text{Var}(e^{\eta})\,\text{Var}(\epsilon) = \text{Var}(e^{\eta}) = \text{Var}(e^{\eta} - 1) = \text{Var}(\omega_1)$ and use equations (23) and (24) to calculate the risk premiums. The average RRP in the period 2005-2010 was around 0.150, which indicates that a typical farmer in our sample is willing to pay 15.0% of the profit as insurance against both production risk and output price uncertainty. Finally, we found that the average RRP during the years 2007-2010 was around 0.17 which means that the farmers' willingness to pay for insurance against risk increased from 2007. It is worth mentioning that since 2007, milk prices have tended to be more volatile and therefore less predictable for producers. The median values of the RRP for years 2005-2010 are shown in the third row of Table 4. The values of RRP from the model that accounts for both production risk and output price uncertainty are on average considerably lower than in the model that only consider production risk. However, in our opinion, the results obtained from the model that jointly addressed production risk and output price uncertainty are more reliable, because the models that ignore output price uncertainty or ignore production risk might be misspecified.

6. Conclusion

We use nonparametric econometric methods for panel data to analyse production risk and price uncertainty in the Just and Pope (1978) framework. Our analysis is based on Kumbhakar and Tsionas (2009) and we generalise this approach to panel data. Furthermore, we used a more recent kernel estimator and applied significance tests in the nonparametric estimations. Finally, we compare different specifications of the mean production function and compare different approaches to obtaining the values of the risk preference functions.

The advantage of the nonparametric approach in the estimation of the Just and Pope (1978) production model is that it does not impose any parametric specifications (functional forms) on the mean production function and the output variability function. Furthermore, no parametric assumptions are made regarding the risk preference functions and the underlying utility function.

We use three variants of the model: a model that only accounts for production risk, a model that only accounts for output price uncertainty, and a model that accounts for both production risk and output price uncertainty. We apply these models to investigate the risk and risk attitudes in Polish dairy farming based on an unbalanced panel data set of Polish family farms that specialised in dairy production in the years 2004-2010.

We found for our data set that the nonparametric estimation of the Just and Pope (1978) production model with an additive error structure resulted in a model that was insignificant in all explanatory variables. Moreover, although the median values of the estimated marginal products and partial output elasticities from this model were reasonable, a large share of the derived measures of interest (e.g. risk premiums) had implausible values or were infinite.

Therefore, we estimated the mean production function in logarithms. We investigate the (approximation) bias in the model with logarithmic regression variables (and multiplicative error) and concluded that the bias was small. Ultimately, the nonparametric estimation of the mean production function with logarithmic regression variables within the Just and Pope (1978) production model was statistically significant. Furthermore, this model delivered plausible economic results.

We calculated the relative risk premiums (RRP) based on the farmers' risk preference functions. Our nonparametric estimations of the risk preference functions were insignificant in virtually all explanatory variables and most of the derived relative risk premiums were implausible. In contrast, the values of the risk preference functions, which were directly calculated from the FOC of the theoretical model, and the nonparametric estimation results of the mean production function, and the output variability function, were plausible.

The results of the model that only accounts for output variability (production risk) indicate that the majority of the analysed Polish dairy farmers is risk averse. The estimated farm-specific relative risk premiums indicate that on average they are willing to pay around 25.7% of their profit (or 11.2% of the value of total revenues) to insure against production risk.

According to the results obtained from the model that only accounts for output price uncertainty, we observed that although the analysed farmers were risk averse, they were on average unwilling to pay to insure against output price uncertainty in the years 2005-2010. Our analysis revealed that, in spite of volatile milk prices, farmers' willingness to pay to insure against lower profit due to uncertain output prices was very low.

Finally, we estimated the model that incorporates both production risk and output price uncertainty. The results regarding the farmers' risk attitudes were mostly in between the results from the models that only account for production risk or only account for output price uncertainty. All three models consistently indicate that the analysed Polish farmers who specialised in dairy production were risk averse. As the models that ignore output price uncertainty or ignore production risk might be misspecified, we consider the results of the model that accounts for both production risk and output price uncertainty as most reliable.

Acknowledgements

We would like to thank Jeff Racine for his many very helpful comments and suggestions regarding nonparametric regression, Subal Kumbhakar and Mike Tsionas for sending us additional information on their studies, and the Agricultural Accountancy Department of IAFE-NRI (IERiGŻ-PIB) in Warsaw for providing access to the data. We are also grateful to the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG) for financially supporting this research. Of course, all remaining errors are the sole responsibility of the authors.

Appendix

A. Table with detailed results

Table A1: Results of the nonparametric local-linear kernel regression of the mean production function estimated in levels

Dependent variable: Y				
Regressor	Bandwidth	Gradients		P-Value
		Median	Std. Dev	
L	0.967	5.556	1187.834	1
A	4.114	0.359	86.820	1
V	11.009	1.541	78.670	1
K	56.496	0.099	2.452	1
year (ordered)	0.500	0.116	325.802	1
ID (unordered)	0.499	-0.003	158.483	1
		Elasticities		
	Mean	Median	Std. Dev	
L	-0.014	0.124	13.298	
A	-0.379	0.078	33.155	
V	0.565	0.727	10.392	
K	0.215	0.232	3.341	
$R^2 = 0.482$				

Table A2: Results of the nonparametric local-linear kernel regression of the output variability function related to the model estimated in levels

Dependent variable: $\ln\|\hat{u}\|$					
Regressor	Bandwidth	Gradients			P-Value
		Median	Mean	Std. Dev	
$\ln(L)$	0.158	0.190	-0.286	9.510	1
$\ln(A)$	0.193	-0.033	-0.053	24.539	1
$\ln(V)$	0.081	0.256	0.062	28.050	1
$\ln(K)$	0.447	0.284	0.497	8.499	1
year (ordered)	0.430	0.016	0.040	1.287	0.005
ID (unordered)	0.514	0.010	0.003	0.230	0.010
$R^2 = 0.797$					

B. Results of nonparametric estimation of risk preference functions

Table A3: Results of the nonparametric estimation of the risk preference function from the model with only production risk using averaged solutions to FOC

Dependent variable: D_1				
Regressor	Bandwidth	Gradients		P-Value
		Mean	Std. Dev	
$\ln(L)$	271450.9	50.690	7.728	0.248
$\ln(A)$	370916.2	-7.561	0.709	0.800
$\ln(V)$	710654.4	14.500	2.093	0.569
$\ln(K)$	1.473022	-18.47	7.579	0.381
$\ln(\widetilde{w}_A)$	74007.8	-27.28	3.897	0.466
$\ln(\widetilde{w}_V)$	199809.4	24.95	13.867	0.797
year (ordered)	1	0.000	0.000	0.712
ID (unordered)	1	0.000	0.000	0.431
$R^2 = 0.001$				

Table A4: Results of the nonparametric estimation of the risk preference function from the model with only production risk using solutions to FOC for land

Dependent variable: D_{1A}				
Regressor	Bandwidth	Gradients		P-Value
		Mean	Std. Dev	
$\ln(L)$	0.001	0.000	0.000	< 0.001
$\ln(A)$	90491.84	-576.8	32168.63	1
$\ln(V)$	1601302	-275.6	42500.56	1
$\ln(K)$	0.2126243	105.4	14801.46	1
$\ln(\widetilde{w}_A)$	2299712	13.8	31230.11	1
$\ln(\widetilde{w}_V)$	0.04193292	-37.49	2858.028	1
year (ordered)	0.436	3078	180034.3	1
ID (unordered)	1	0.000	0.001	1
$R^2 = 0.868$				

Table A5: Results of the nonparametric estimation of the risk preference function from the model with only production risk using solutions to FOC for intermediate inputs

Dependent variable: D_{1V}				
Regressor	Bandwidth	Gradients		P-Value
		Mean	Std. Dev	
ln(L)	0.387	86.98	41.499	0.113
ln(A)	1523518	20.85	12.641	0.419
ln(V)	1.133	1.770	16.542	0.549
ln(K)	520509.6	-35.44	8.117	0.211
ln($\widetilde{w}_A$)	0.491	-7.214	21.457	0.619
ln($\widetilde{w}_V$)	19367.01	192.1	20.549	0.190
year (ordered)	1	1	0.000	0.323
ID (unordered)	1	0.000	0.000	0.083
$R^2 = 0.009$				

Table A6: Results of the nonparametric estimation of the risk preference function from the model with only price uncertainty using averaged solutions to FOC

Dependent variable: D_2				
Regressor	Bandwidth	Gradients		P-Value
		Mean	Std. Dev	
ln(L)	1.576	0.649	56.765	1
ln(A)	0.510	-3.326	132.023	1
ln(V)	0.619	0.776	101.242	1
ln(K)	0.124	3.704	157.269	1
ln($\widetilde{w}_A$)	0.008	-580.1	25141.66	1
ln($\widetilde{w}_V$)	0.350	591.7	22660.31	1
year (ordered)	0.887	-0.056	3.0279	1
ID (unordered)	0.807	0.005	1.5172	1
$R^2 = 0.792$				

Table A7: Results of the nonparametric estimation of the risk preference function from the model with only price uncertainty using solutions to FOC for land

Dependent variable: D_{2A}				
Regressor	Bandwidth	Gradients		P-Value
		Mean	Std. Dev	
$\ln(L)$	8.956	-	-	0.905
$\ln(A)$	0.349	-	-	1
$\ln(V)$	0.061	-	-	1
$\ln(K)$	0.206	-	-	1
$\ln(\widetilde{w}_A)$	0.184	-	-	0.987
$\ln(\widetilde{w}_V)$	0.047	-	-	0.977
year (ordered)	1		-	0.772
ID (unordered)	1	-	-	0.940
$R^2 = 0.000$				

Table A8: Results of the nonparametric estimation of the risk preference function from the model with only price uncertainty using solutions to FOC for intermediate inputs

Dependent variable: D_{2V}				
Regressor	Bandwidth	Gradients		P-Value
		Mean	Std. Dev	
$\ln(L)$	4.28003	-0.074	1.073	1
$\ln(A)$	1.777198	-0.022	0.453	1
$\ln(V)$	10.59916	0.144	0.586	1
$\ln(K)$	0.4417764	-0.121	0.905	0.561
$\ln(\widetilde{w}_A)$	0.279175	-0.255	9.609	1
$\ln(\widetilde{w}_V)$	0.006	-2.053	62.755	1
year (ordered)	0.3916	0.000	0.127	0.042
ID (unordered)	0.102	0.004	0.391	0.010
$R^2 = 0.849$				

Table A9: Results of the nonparametric estimation of the risk preference function from the model with both production risk and price uncertainty using averaged solutions to FOC and land as a numeraire input ($j = 1 =$ land, $j = 2 =$ intermediate inputs)

Dependent variable: D_{3VA}				
Regressor	Bandwidth	Gradients		P-Value
		Mean	Std. Dev	
$\ln(L)$	814626.1	-0.015	0.273	0.035
$\ln(A)$	0.398	-0.086	0.386	0.363
$\ln(V)$	627433.6	0.066	0.394	0.333
$\ln(K)$	4820067	0.065	0.289	0.095
$\ln(\widetilde{w}_A)$	0.0996	0.047	1.7646	0.241
$\ln(\widetilde{w}_V)$	535934.4	-0.122	4.235	0.110
year (ordered)	1	0.000	0.000	0.105
ID (unordered)	0.001	-0.005	0.405	0.015
$R^2 = 0.930$				

Table A10: Results of the nonparametric estimation of the risk preference function from the model with both production risk and price uncertainty using averaged solutions to FOC and intermediate inputs as a numeraire input ($j = 1 =$ intermediate inputs, $j = 2 =$ land)

Dependent variable: D_{3AV}				
Regressor	Bandwidth	Gradients		P-Value
		Mean	Std. Dev	
$\ln(L)$	1.757	10.67	660.601	0.113
$\ln(A)$	0.526	-11.96	1013.514	0.419
$\ln(V)$	0.731	16.10	629.061	0.549
$\ln(K)$	0.125	30.05	1497.161	0.211
$\ln(\widetilde{w}_A)$	0.008	-9040	324932.7	0.619
$\ln(\widetilde{w}_V)$	1.595	9197	308123.2	0.190
year (ordered)	0.73917	0.318	18.231	0.323
ID (unordered)	1	0.499	17.800	0.083
$R^2 = 0.865$				

Table A11: Results of the nonparametric estimation of the risk preference function from the model with both production risk and price uncertainty using solutions to FOC for land

Dependent variable: D_{3A}				
Regressor	Bandwidth	Gradients		P-Value
		Mean	Std. Dev	
$\ln(L)$	2.695	35.000	1049.810	1
$\ln(A)$	0.484	-15.025	393.237	1
$\ln(V)$	0.334	43.053	437.317	1
$\ln(K)$	0.090	-58.150	2047.369	1
$\ln(\widetilde{w}_A)$	0.105	-251.470	6574.699	0.709
$\ln(\widetilde{w}_V)$	2.635	-477.400	13853.640	0.118
year (ordered)	0.999	0.777	38.387	0.997
ID (unordered)	0.685	0.3682	42.888	0.922
$R^2 = 0.773$				

Table A12: Results of the nonparametric estimation of the risk preference function from the model with both production risk and price uncertainty using solutions to FOC for intermediate inputs

Dependent variable: D_{3V}				
Regressor	Bandwidth	Gradients		P-Value
		Mean	Std. Dev	
$\ln(L)$	1615179	-0.040	0.520	0.085
$\ln(A)$	1008954	-0.190	0.466	0.015
$\ln(V)$	0.567	0.150	0.963	0.003
$\ln(K)$	2842631	0.065	0.549	0.155
$\ln(\widetilde{w}_A)$	0.110	0.110	3.767	0.083
$\ln(\widetilde{w}_V)$	917447.2	-0.246	7.942	0.078
year (ordered)	1	0.000	0.000	0.000
ID (unordered)	0.007	-0.016	0.723	0.033
$R^2 = 0.923$				

C. Comparison of calculated and estimated values of risk preferences and relative risk premiums

Table A13: Median values of risk preference functions and relative risk premiums for Model 1 using calculated and estimated RPF

		T	RRP1	$RRP1_{rev}$
Calculated	A	-2.87	0.243	0.104
	V	-3.24	0.256	0.114
	D1	-3.17	0.258	0.112
Estimated	A	-836	74.2	31.3
	V	-22.20	1.747	0.793
	D1	-26.870	2.086	0.980

Note:
Calculated - refers to values obtained from risk preference functions derived from economic model,
Estimated - refers to values obtained from nonparametric estimation of risk preference functions.
A,V indicates the input used to solve FOC (land or intermediate inputs) RRP1 - relative risk premium for model with production risk only = risk premium/restricted profit
$RRP1_rev$ - relative risk premium (II) for model with production risk only = risk premium/revenue.

Table A14: Median values of risk preference functions and relative risk premiums for Model 2 using calculated and estimated RPF

		T	RRP2	$RRP2_{rev}$
Calculated	A	-0.845	0.009	0.004
	V	-0.239	0.003	0.001
	D2	-0.508	0.005	0.002
Estimated	A	-1.000	0.000	0.002
	V	-0.199	0.002	0.001
	D2	-0.498	0.005	0.002

Note:
RRP2 - relative risk premium for model with price uncertanity only = risk premium/restricted profit
$RRP2_rev$ -relative risk premium (II) for model with price uncertanity only = risk premium/revenue.

Table A15: Median values of risk preference functions and relative risk premiums for Model 3 using calculated and estimated risk preferences functions

		T	T2	RRP3	$RRP3_{rev}$
Calculated	VA	-1	-13	0.150	0.064
	AV	-1	-13	0.150	0.064
	V	-1.390	-16.50	0.190	0.081
	A	-1.10	-13.00	0.150	0.064
Estimated	VA	-3.200	-40	0.480	0.200
	AV	0.600	7.00	-0.070	-0.040
	V	-3.27	-40.50	0.500	0.207
	A	0.420	5	-0.06	-0.026

Note: AV, (VA) - indicates that land (intermediate inputs) variable is used as a numereire input
RRP3 - relative risk premium for model with both production risk and output price uncertanity = risk premium/restricted profit
$RRP3_rev$ - relative risk premium (II) for model with both production risk and output price uncertanity = risk premium/revenue.

D. Figures with detailed results

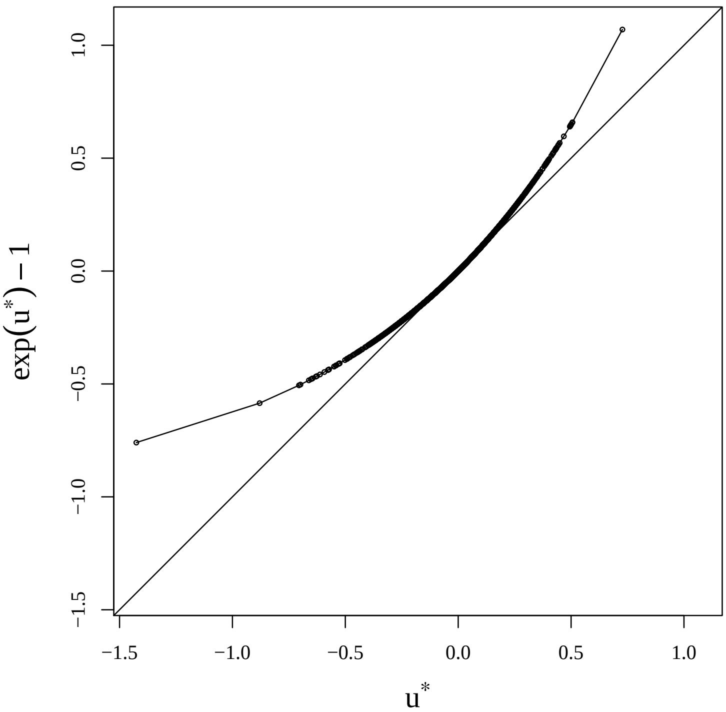
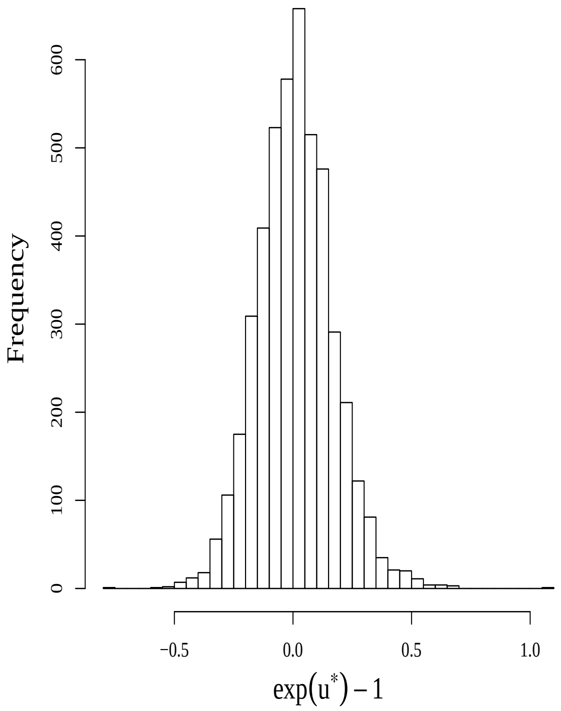
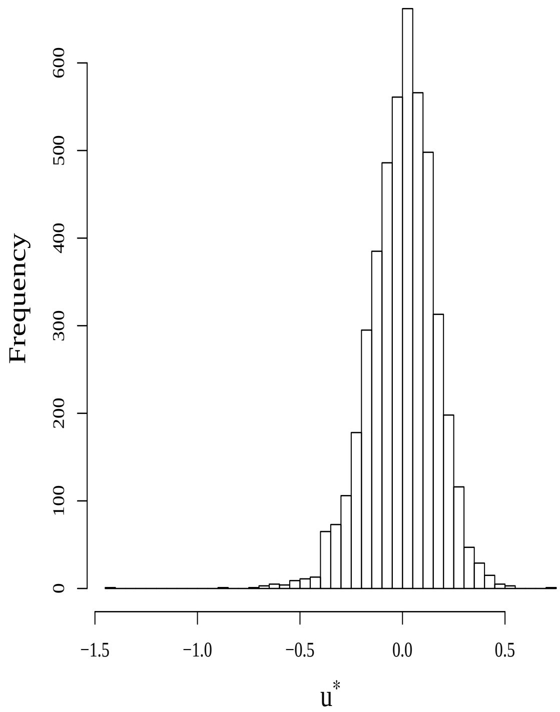


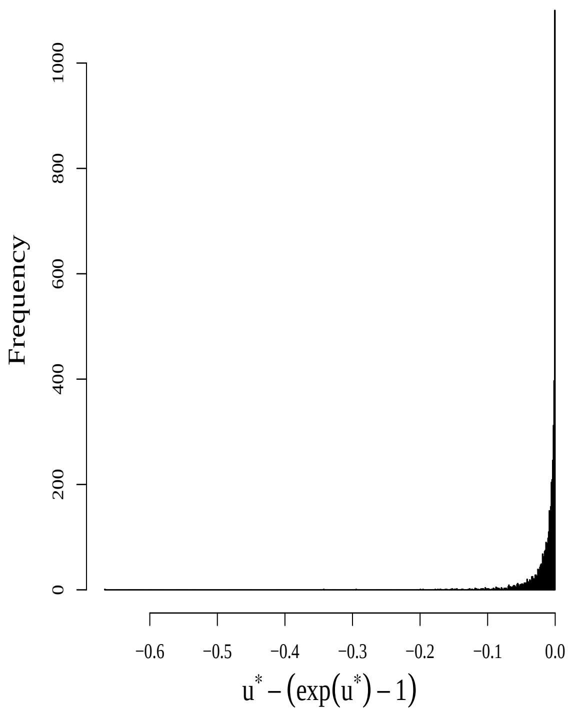
Figure A1: Scatter plot of u^* and $\exp(u^*) - 1$



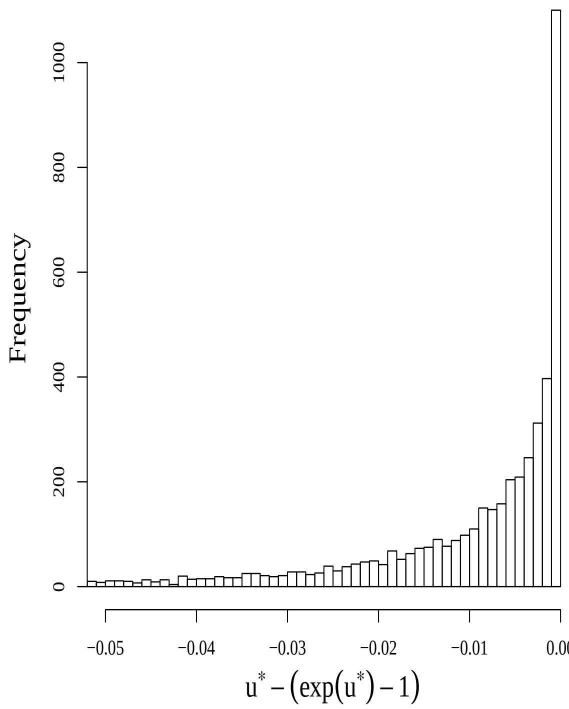
(a)



(b)

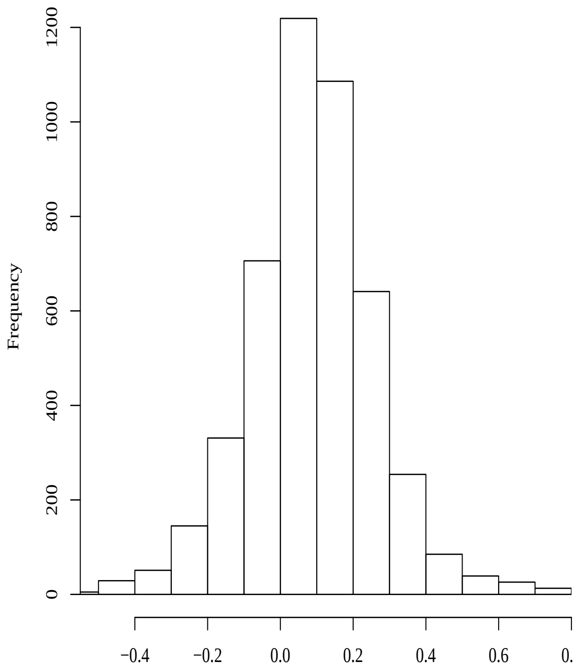


(c)

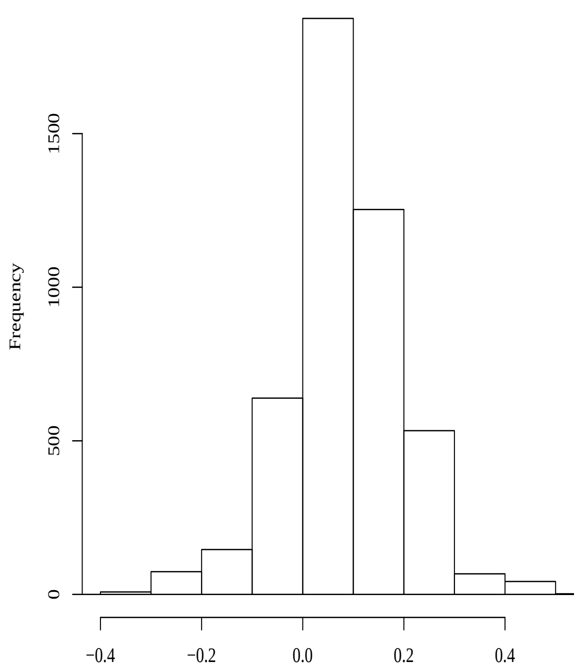


(d)

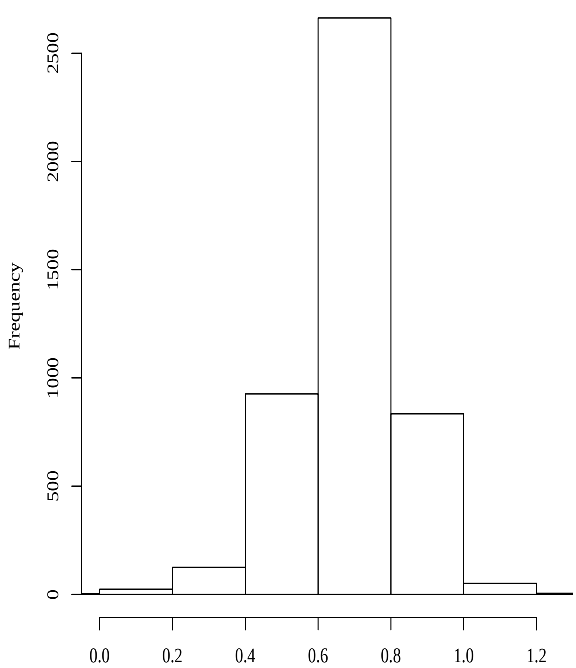
Figure A2: Histograms of $\exp(u^*) - 1$ (A2a), u^* (A2a), the difference $u^* - (\exp(u^*) - 1)$ (A2c), and the difference $u^* - (\exp(u^*) - 1)$ for the interval $[-0.05, 0]$ (A2d).



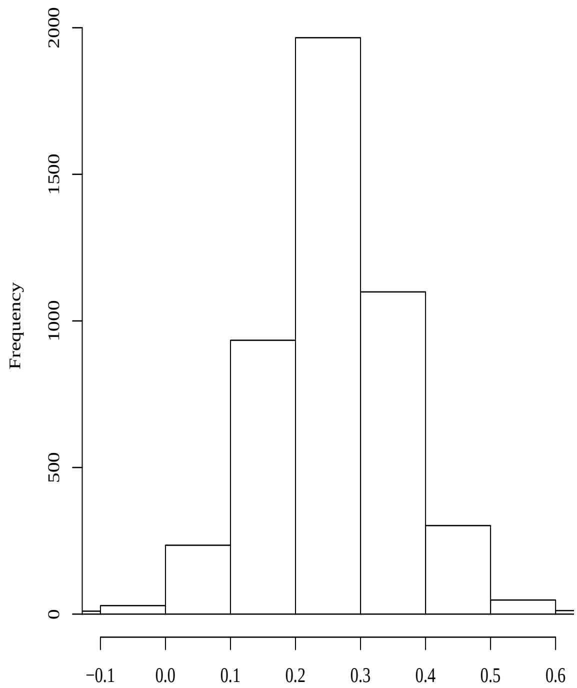
(a) Labour



(b) Land

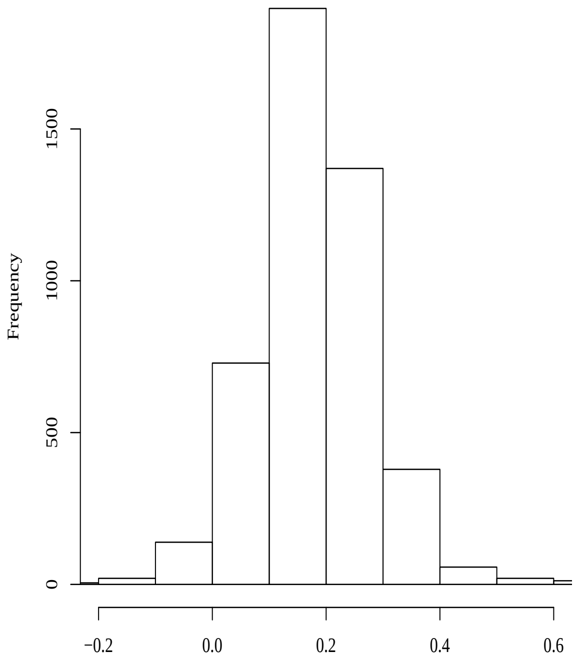


(c) Intermediate inputs

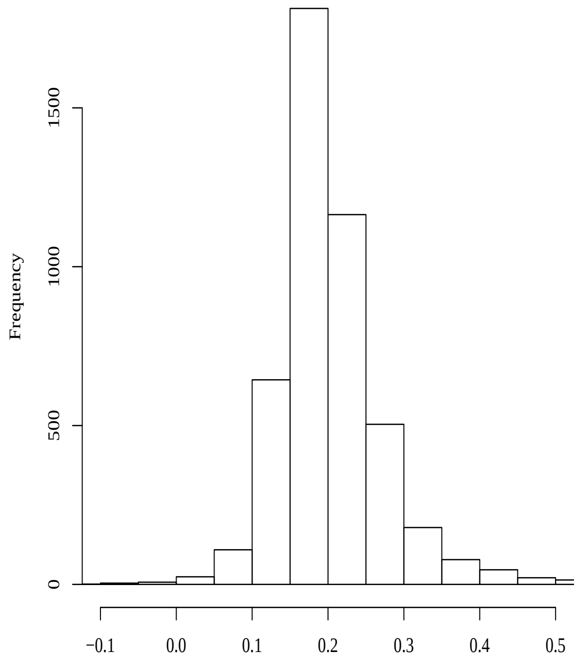


(d) Capital

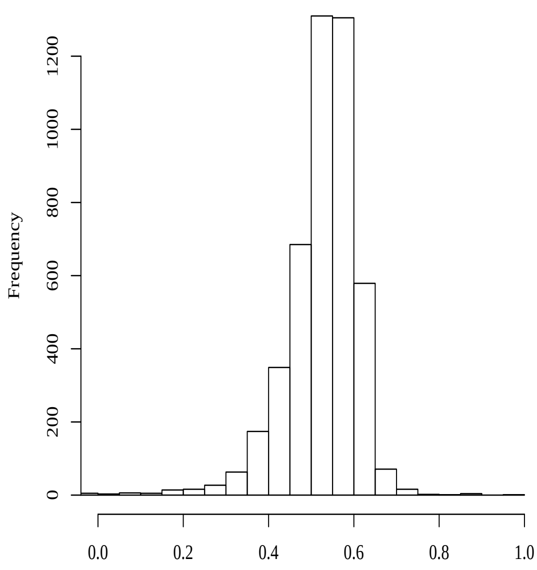
Figure A3: Partial output elasticities of the mean production function ($f(x, z)$) for the model in logarithms.



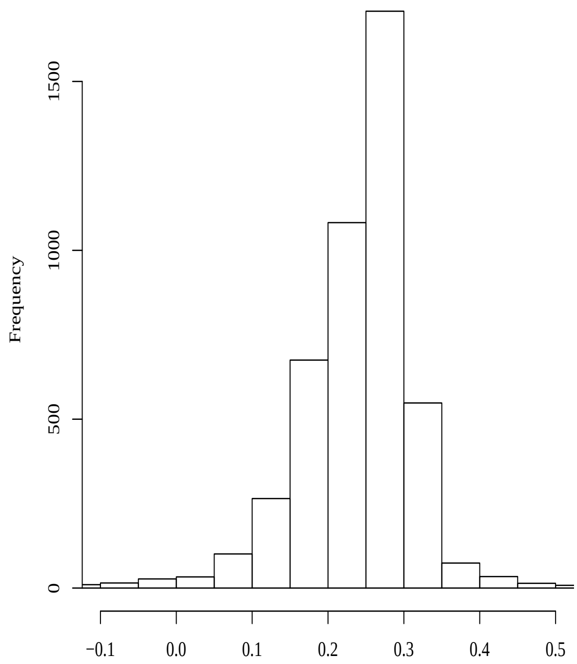
(a) Labour



(b) Land

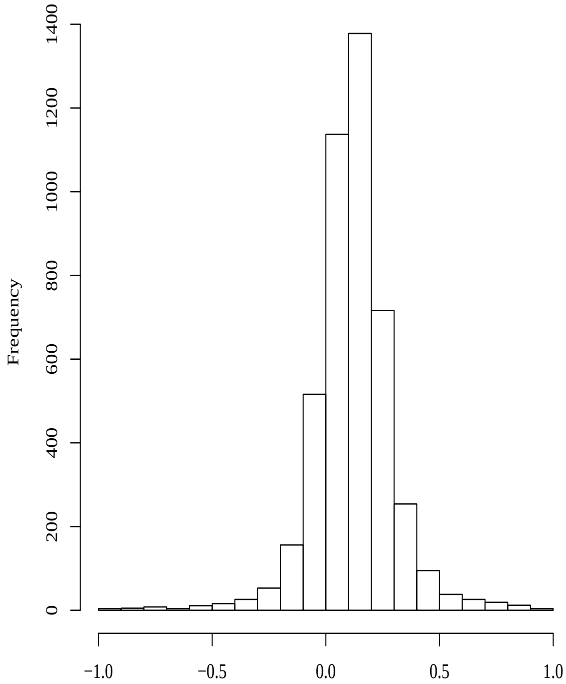


(c) Intermediate inputs

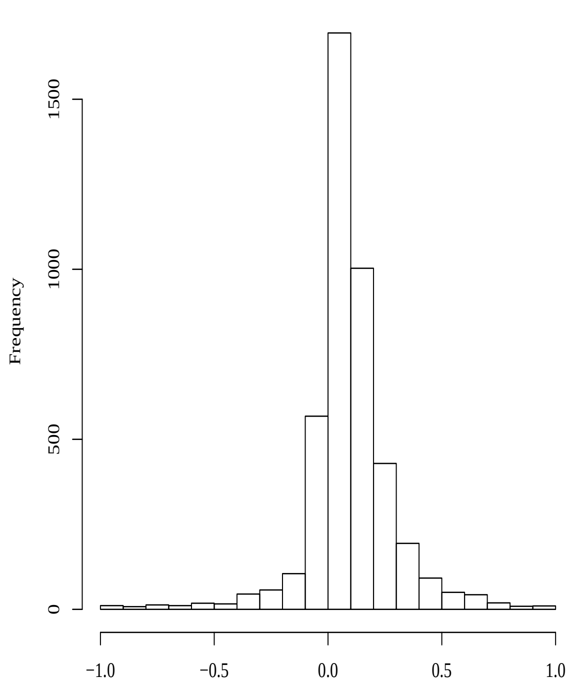


(d) Capital

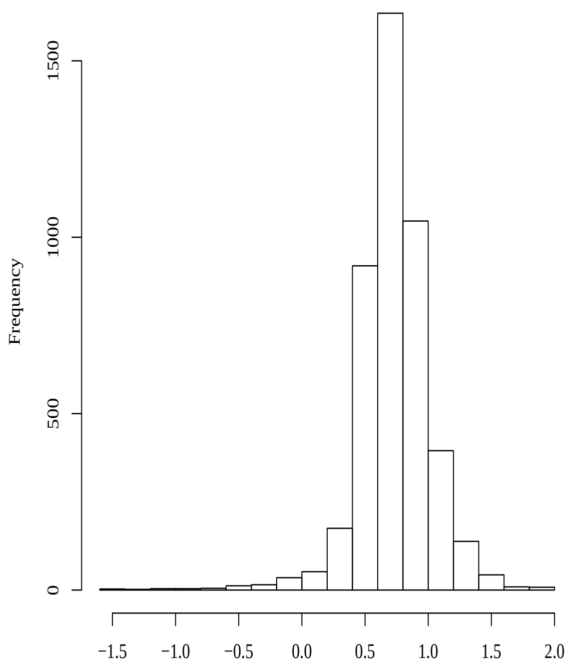
Figure A4: Partial output elasticities of the output variability function ($h(x, z)$) for the model in logarithms.



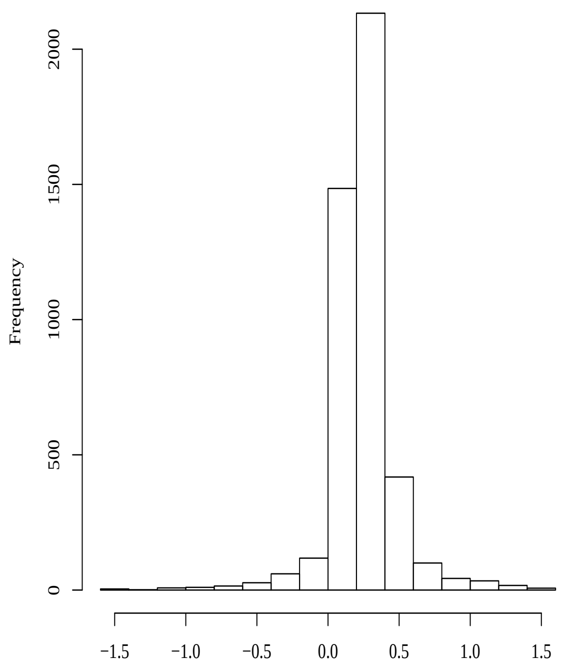
(a) Labour



(b) Land

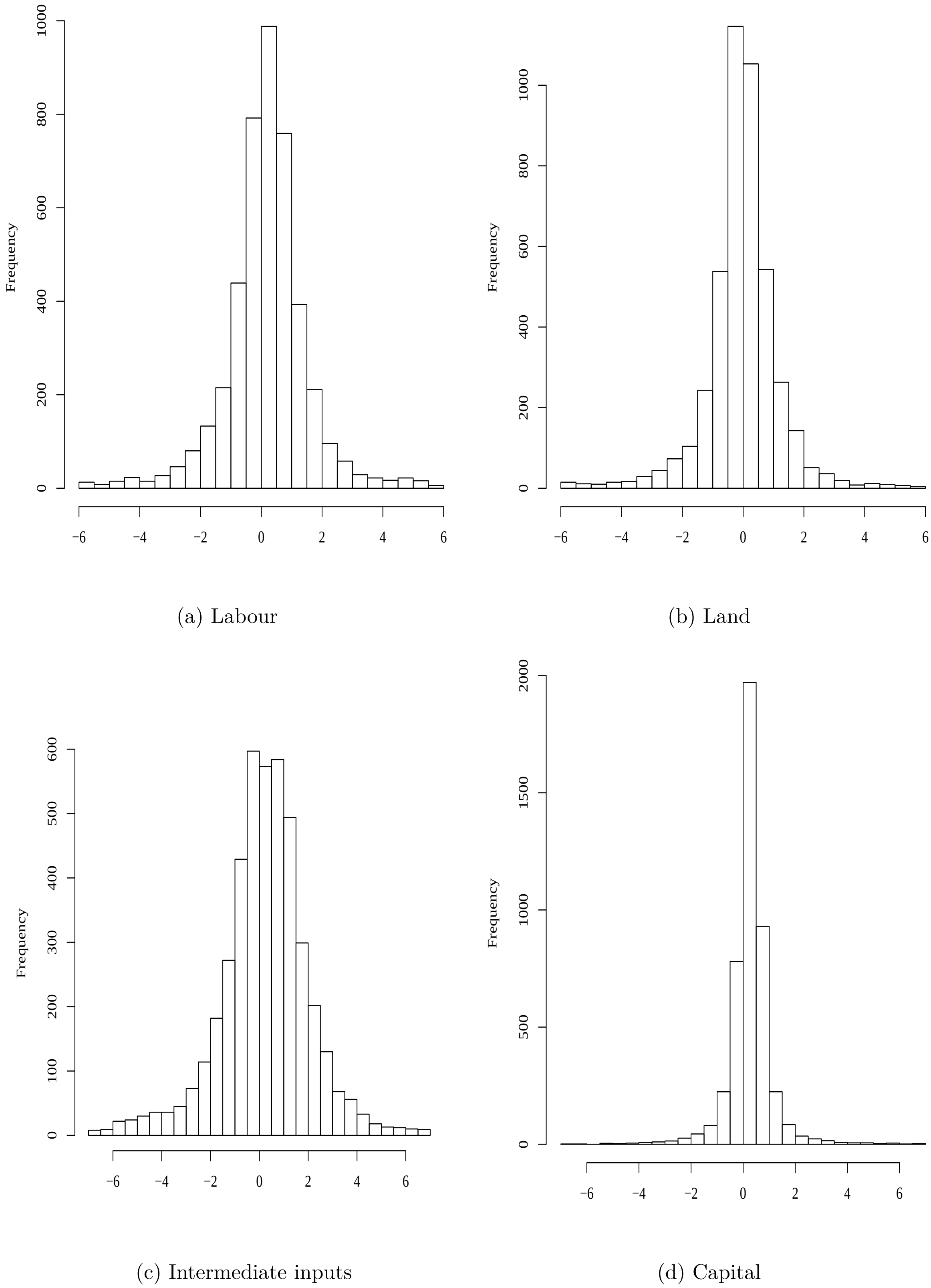


(c) Intermediate inputs



(d) Capital

Figure A5: Partial output elasticities of the mean production function ($f(x, z)$) for the model in levels.



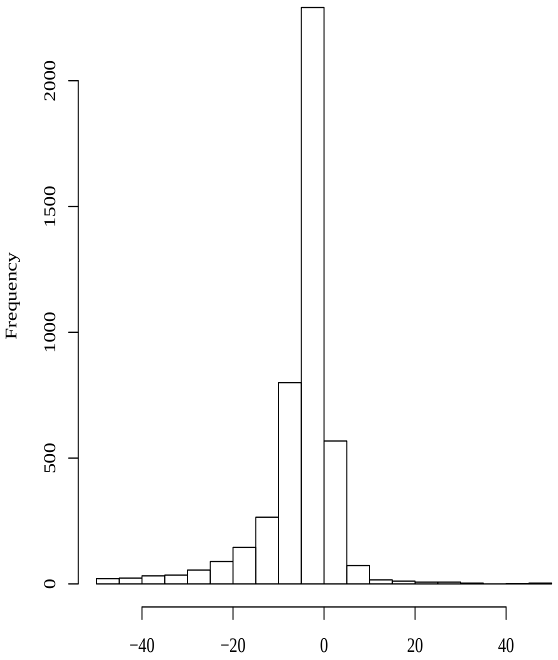
(a) Labour

(b) Land

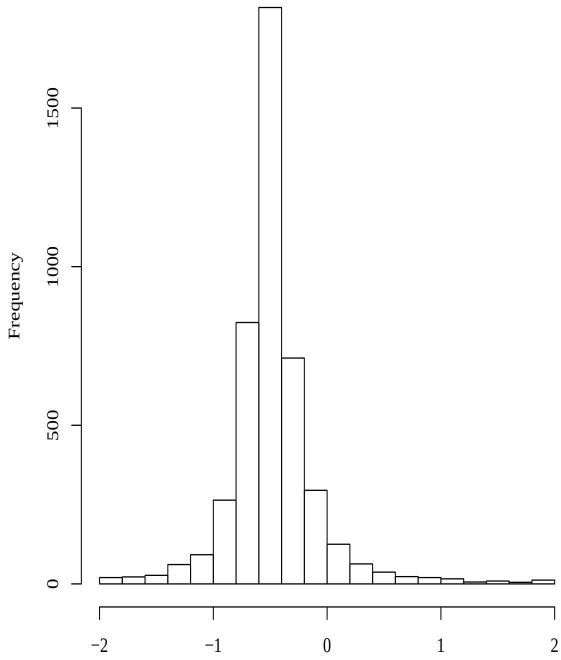
(c) Intermediate inputs

(d) Capital

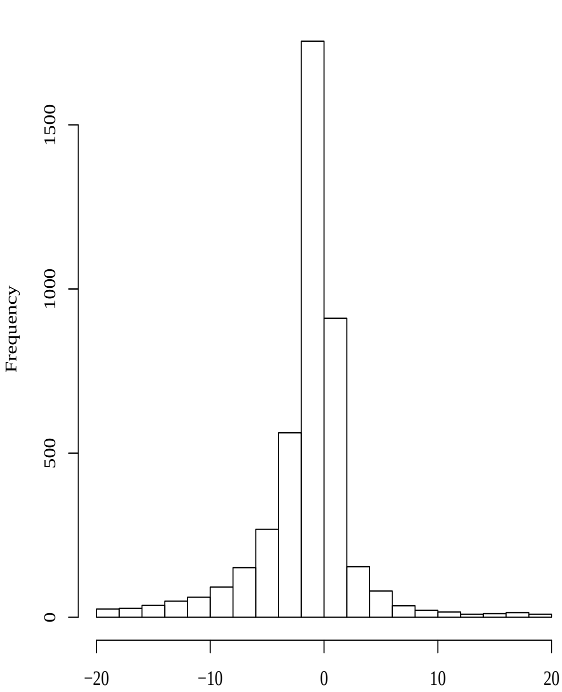
Figure A6: Partial output elasticities of the output variability function ($h(x, z)$) for the model in levels.



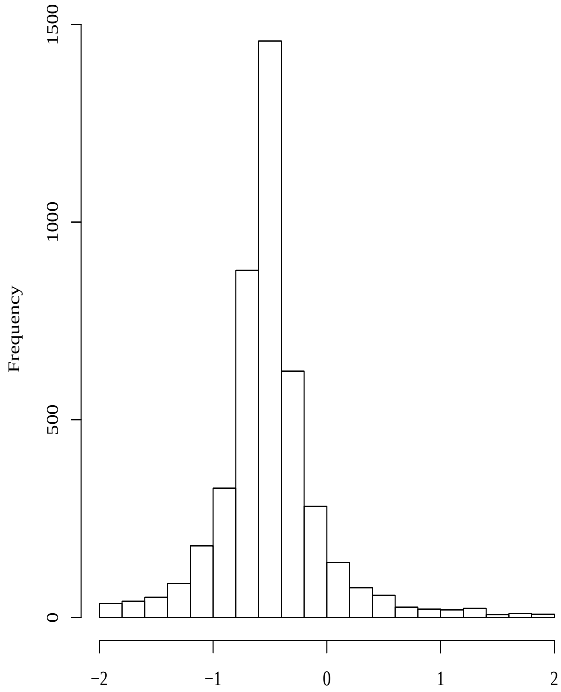
(a) θ_1



(b) θ_2

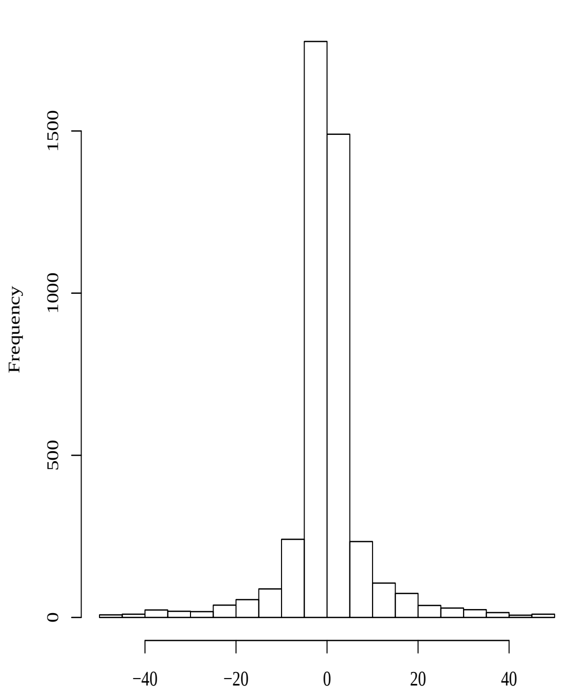


(c) $\tilde{\theta}_1$

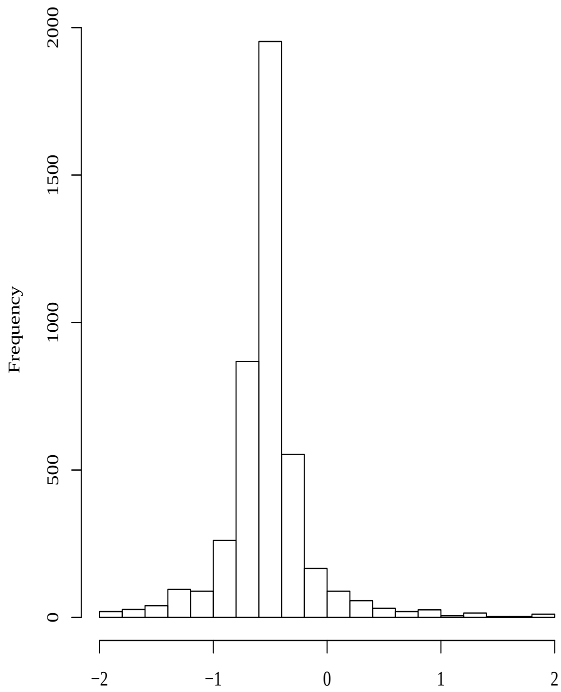


(d) $\tilde{\theta}_2$

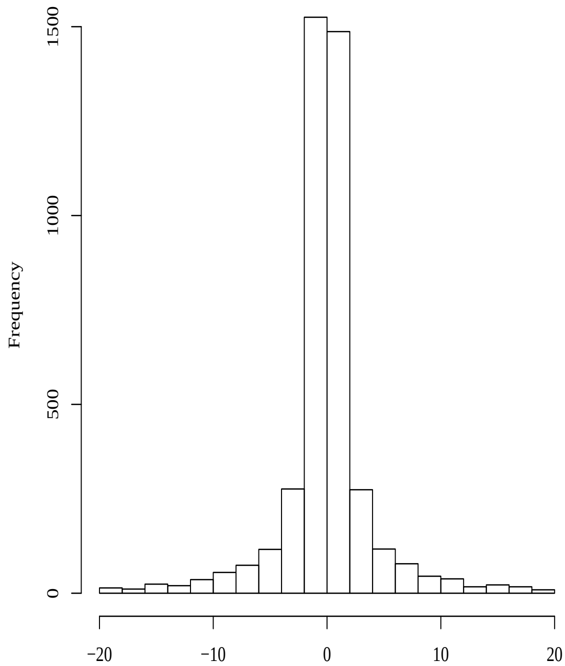
Figure A7: Histograms of Risk Preference functions from the model in logarithms.



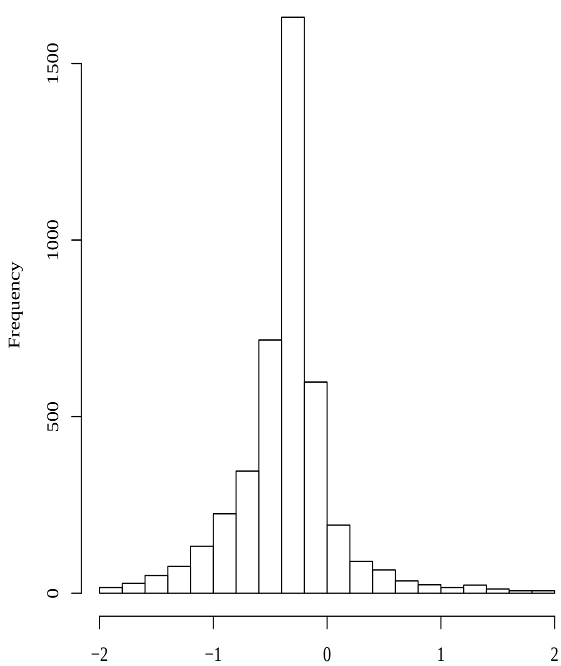
(a) θ_1



(b) θ_2

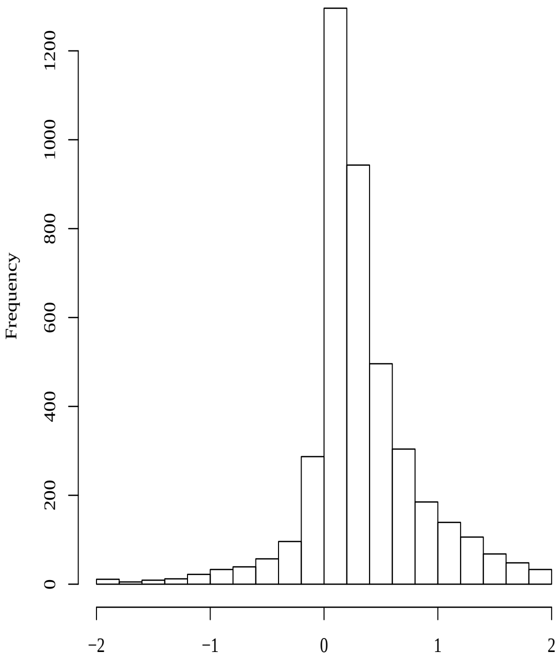


(c) $\tilde{\theta}_1$

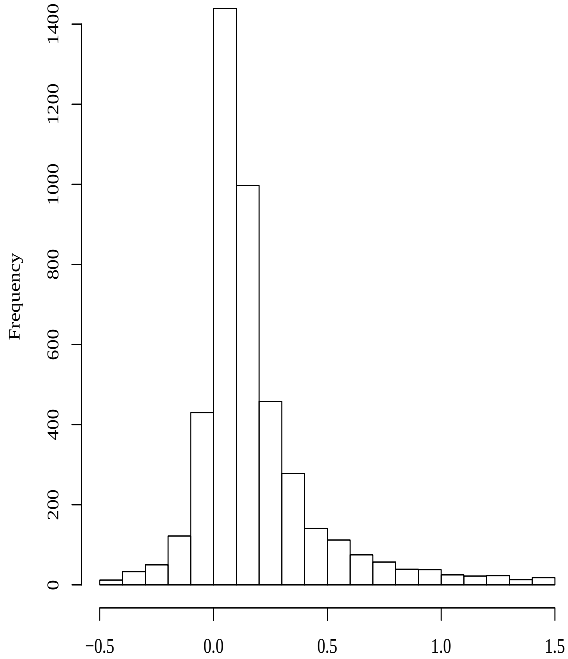


(d) $\tilde{\theta}_2$

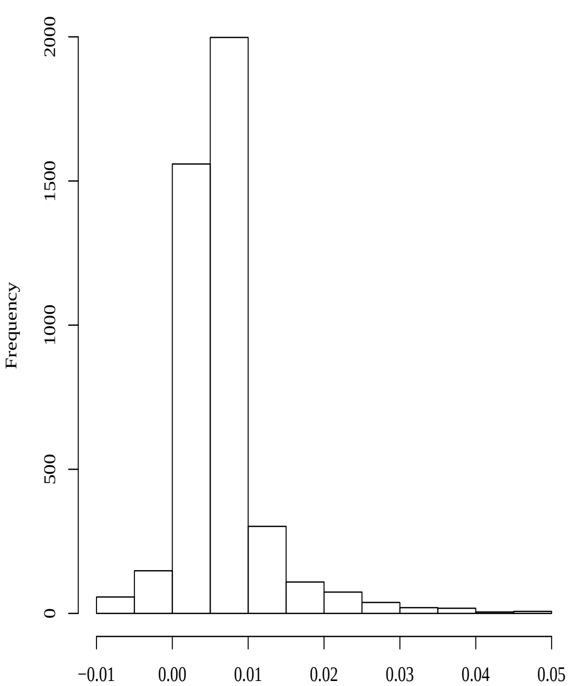
Figure A8: Histograms of Risk Preference functions from the model in levels.



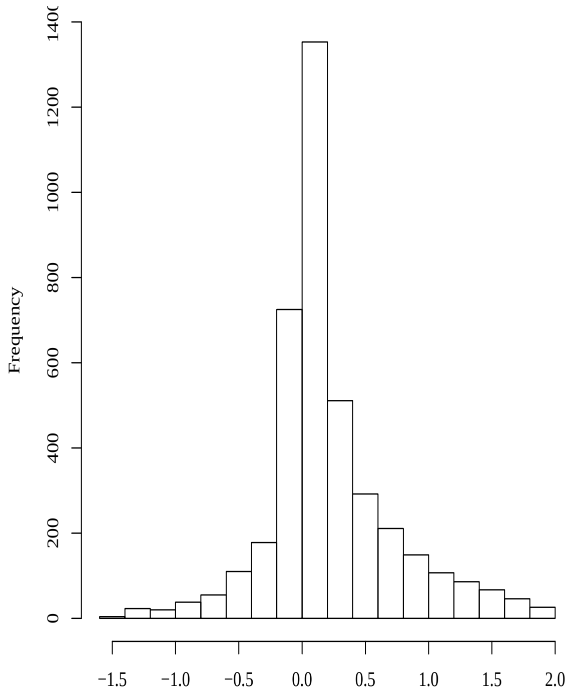
(a) RRP11



(b) RRP11

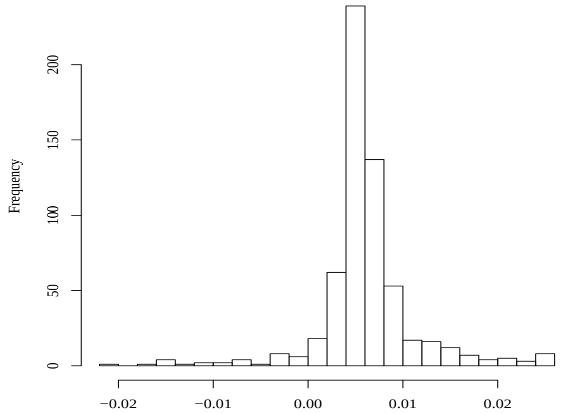


(c) RRP2

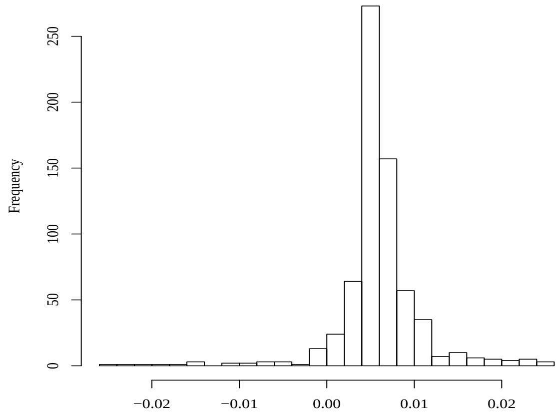


(d) RRP3

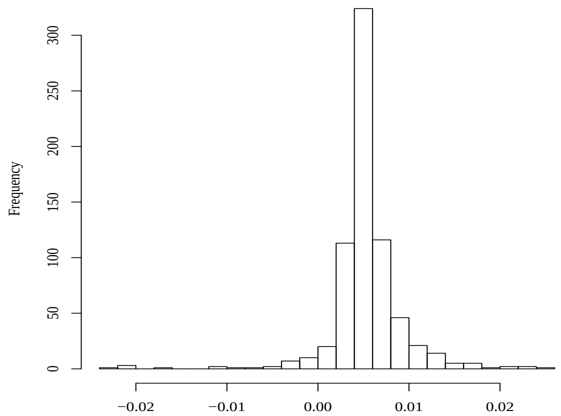
Figure A9: Histograms of Risk Preference functions from the model in logarithms.



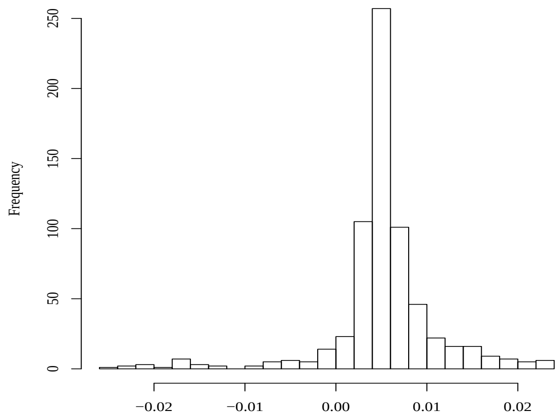
(a) RRP2 2005



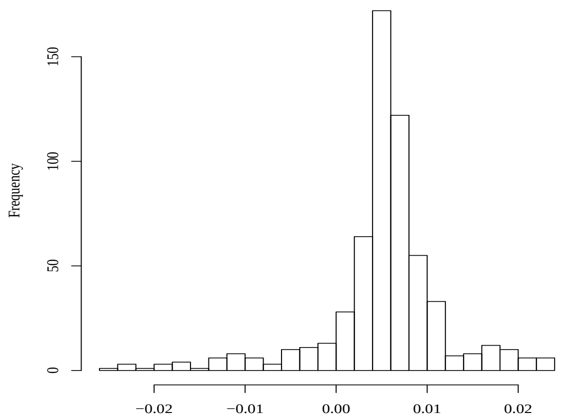
(b) RRP2 2006



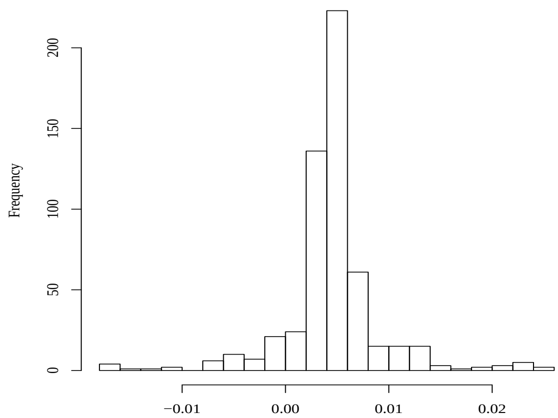
(c) RRP2 2007



(d) RRP2 2008

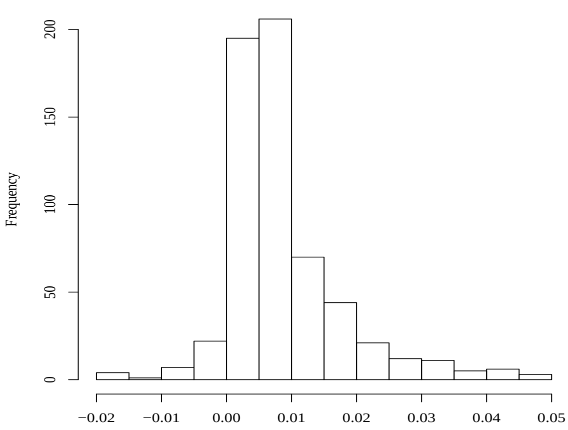


(e) RRP2 2009

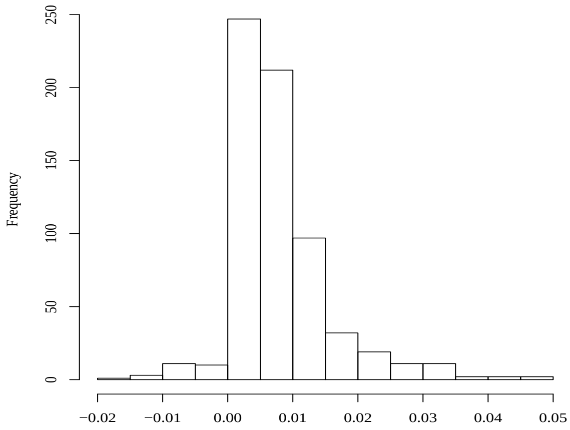


(f) RRP2 2010

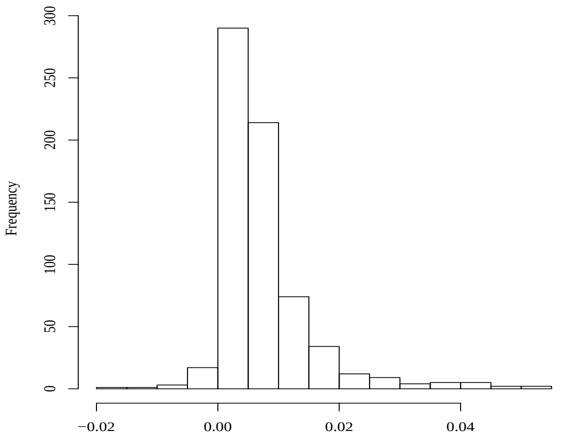
Figure A10: Histograms of Relative Risk Premiums from the model with price uncertenity in the years 2005-2010 from the model in logarithms.



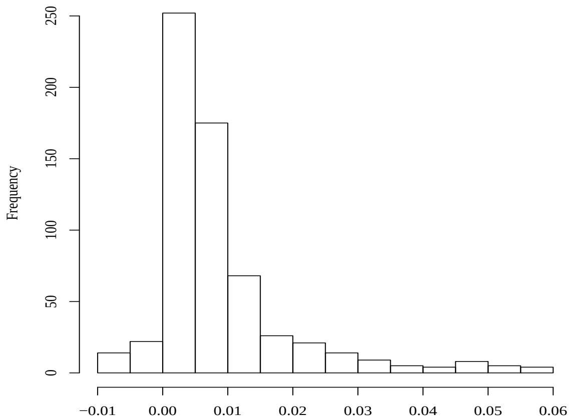
(a) RRP2 2005



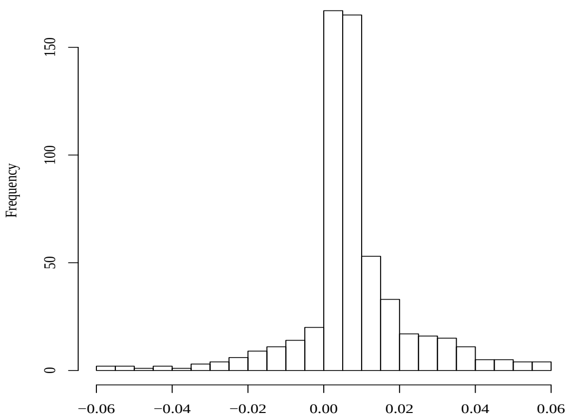
(b) RRP2 2006



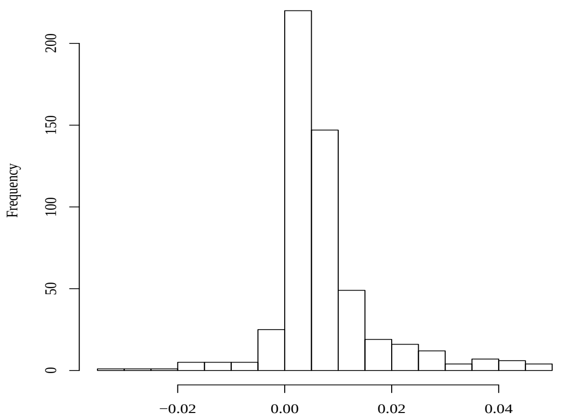
(c) RRP2 2007



(d) RRP2 2008

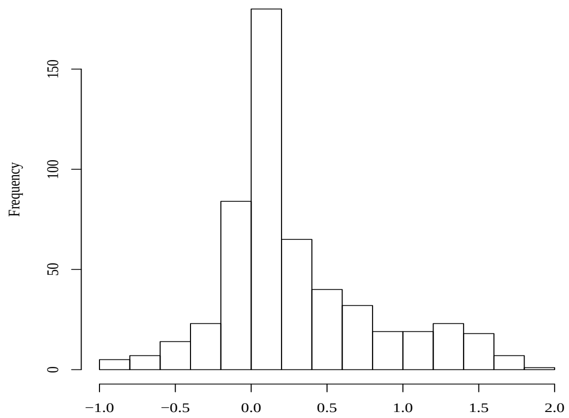


(e) RRP2 2009

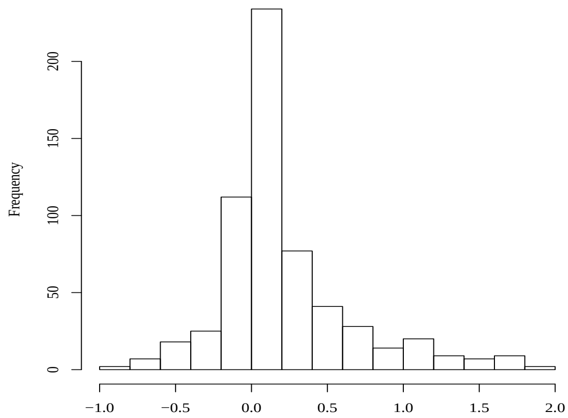


(f) RRP2 2010

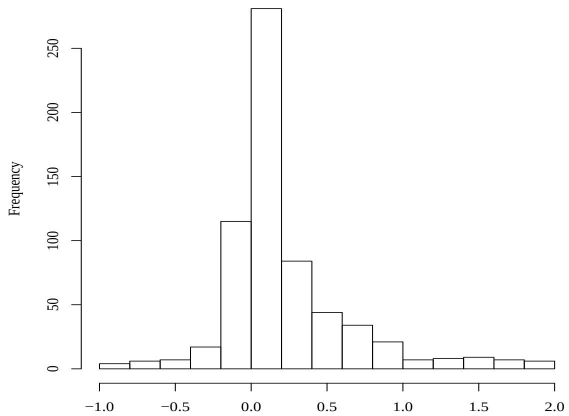
Figure A11: Histograms of Relative Risk Premiums from the model with price uncertenity in the years 2005-2010 from the model in levels.



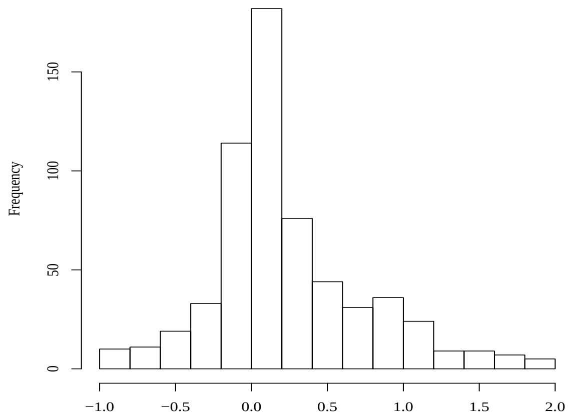
(a) RRP3 2005



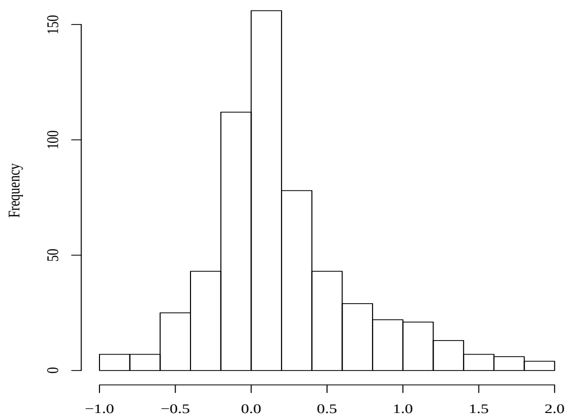
(b) RRP3 2006



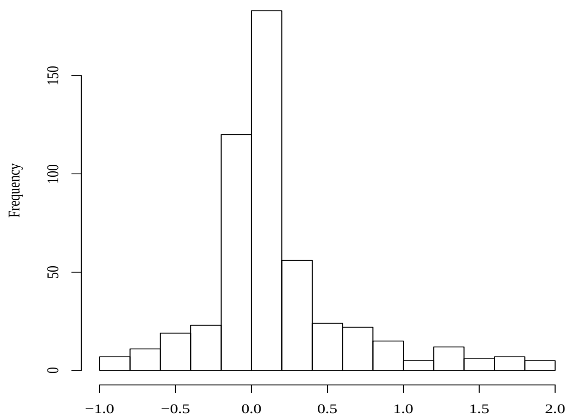
(c) RRP3 2007



(d) RRP3 2008

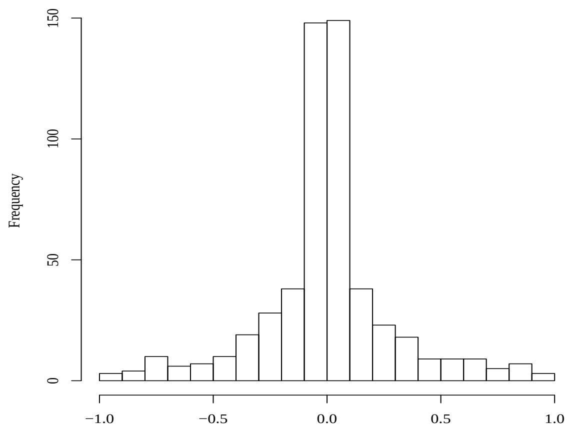


(e) RRP3 2009

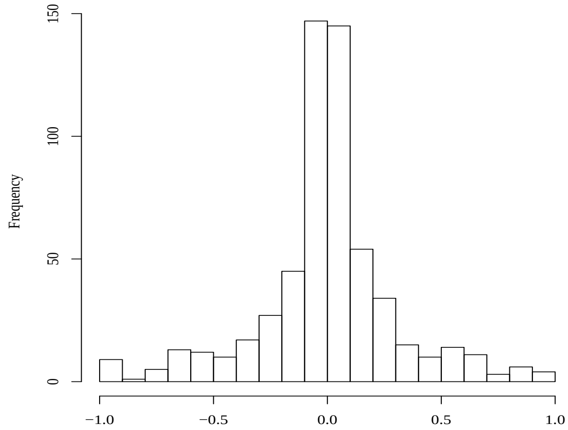


(f) RRP3 2010

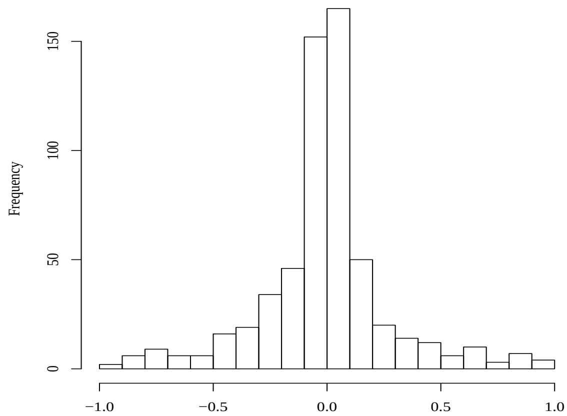
Figure A12: Histograms of Relative Risk Premiums from the model with production variability and price uncertainty in the years 2005-2010 from the model in logarithms.



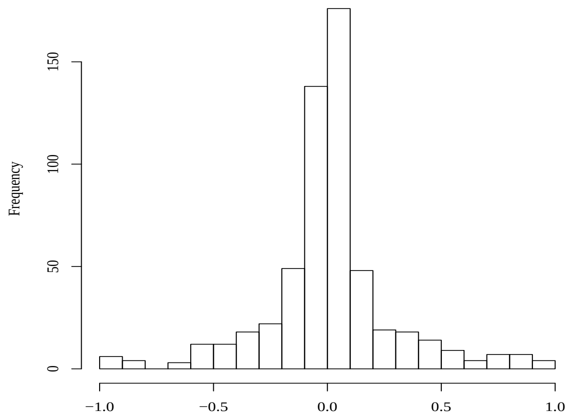
(a) RRP3 2005



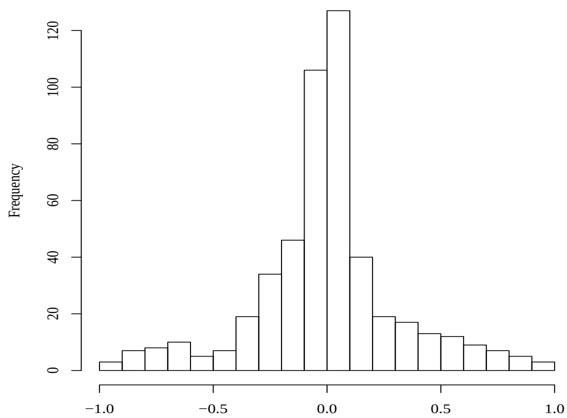
(b) RRP3 2006



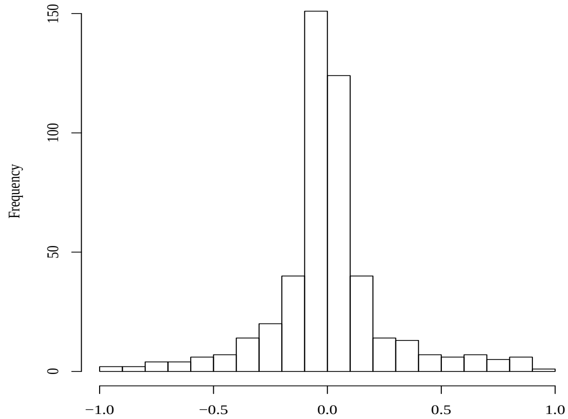
(c) RRP3 2007



(d) RRP3 2008



(e) RRP3 2009



(f) RRP3 2010

Figure A13: Histograms of Relative Risk Premiums from the model with production risk and price uncertainty in the years 2005-2010 from the model in levels.

E. R code

E.1. Estimation in levels

```
##Nonparametric estimation of the mean production function of the Just and Pope
##technology in levels of regression variables
##
##Cross-validation bandwidth selection
bw.f <- npregbw( Y ~ L + A + V + K + T_of + ID,
                     regtype = "ll",
                     bwmethod = "cv.aic",
                     ckertype = "epanechnikov",
                     ukertype = "liracine",
                     okertype = "liracine",
                     data = PLF_TF56 )
##Estimation of the nonparametric model
model.np.f <- npreg( bws = bw.f,
                         data = PLF_TF56,
                         gradients = TRUE,
                         residuals = TRUE )
summary( model.np.f )
##Bootstrap significance test
sigtest.model.np.f <- npsigtest( model.np.f, boot.num = 399 )
summary( sigtest.model.np.f )

##Nonparametric estimation of the risk variance function of the Just and Pope
##technology (mean production function estimated in levels, risk variance
##function estimated in logarithms)

##Calculate dependent variable
PLF_TF56$Absu <- abs(model.np.f$resid)
PLF_TF56$lAbsu <- log(PLF_TF56$Absu)

##Cross-validation bandwidth selection
bw.g <- npregbw( lAbsu ~ lL + lA + lV + lK + T_of + ID,
                                         regtype = "ll",
                                         bwmethod = "cv.aic",
                                         ckertype = "epanechnikov",
                                         ukertype = "liracine",
                                         okertype = "liracine",
                                         data = PLF_TF56 )
##Estimation of the nonparametric model
model.np.g <- npreg( bws = bw.g,
                         data = PLF_TF56,
                         gradients = TRUE,
                         residuals = TRUE )
summary( model.np.g )
```

```
##Bootstrap significance test
sigtest.model.np.g <- npsigtest( model.np.g, boot.num = 399 )
summary( sigtest.model.np.g )

##Nonparametric estimation of the risk preference function (phi) for Model 3
##Calculate D3
PLF_TF56$D3 <- (0.5* ((model.np.f$grad[,3] / model.np.f$grad[,3]) +
                      (model.np.f$grad[,2] / model.np.f$grad[,3]))
##Cross-validation bandwidth selection
bw.D3 <- npregbw( D3 ~ lL + lA + lV + lK + lAwp + lVwp + T_of + ID,
                      regtype = "ll",
                      bwmethod = "cv.aic",
                      ckertype = "epanechnikov",
                      ukertype = "liracine",
                      okertype = "liracine",
                      data = PLF_TF56),
##Estimation of the nonparametric model
model.np.D3 <- npreg( bws = bw.D3,
                          data = PLF_TF56,
                          gradients = TRUE,
                          residuals = TRUE )
summary( model.np.D3 )
sigtest.model.np.D3 <- npsigtest( model.np.D3, boot.num = 399 )
##Bootstrap significance test
summary( sigtest.model.np.D3 )

##Estimation of risk preference functions and relative risk premiums for
##Models 1-3
#calculate marginal products of the mean production function of the
#Just and Pope technology
PLF_TF56$fL <- model.np.f$grad[,1]
PLF_TF56$fA <- model.np.f$grad[,2]
PLF_TF56$fV <- model.np.f$grad[,3]
PLF_TF56$fK <- model.np.f$grad[,4]
PLF_TF56$fHat <- fitted(model.np.f)
#calculate marginal effects of the risk function
#of the  Just and Pope technology
PLF_TF56$gL <- model.np.g$grad[,1] * ((exp(PLF_TF56$lAbsu)) / PLF_TF56$L)
PLF_TF56$gA <- model.np.g$grad[,2] * ((exp(PLF_TF56$lAbsu)) / PLF_TF56$A)
PLF_TF56$gV <- model.np.g$grad[,3] * ((exp(PLF_TF56$lAbsu)) / PLF_TF56$V)
PLF_TF56$gK <- model.np.g$grad[,4] * ((exp(PLF_TF56$lAbsu)) / PLF_TF56$K)
PLF_TF56$gHat <- exp(fitted(model.np.g))
summary(PLF_TF56$gA)
summary(PLF_TF56$gK)
PLF_TF56$resid.model.np.f <- resid(model.np.f)
#calculate elasticities of the mean production function
#of the Just and Pope technology
```

```
PLF_TF56$fL_ela <- model.np.f$grad[,1] / (PLF_TF56$YQ / PLF_TF56$L)
PLF_TF56$fA_ela <- model.np.f$grad[,2] / (PLF_TF56$YQ / PLF_TF56$A)
PLF_TF56$fV_ela <- model.np.f$grad[,3] / (PLF_TF56$YQ / PLF_TF56$V)
PLF_TF56$fK_ela <- model.np.f$grad[,4] / (PLF_TF56$YQ / PLF_TF56$K)
#calculate elasticities of the risk function
#of the Just and Pope technology
PLF_TF56$gL_ela <- model.np.g$grad[,1]
PLF_TF56$gA_ela <- model.np.g$grad[,2]
PLF_TF56$gV_ela <- model.np.g$grad[,3]
PLF_TF56$gK_ela <- model.np.g$grad[,4]
#Drop observations with marginal effects of mean production function and risk
#variance function equal to zero
PLF_TF56 <- subset(PLF_TF56,
                   PLF_TF56$gA != 0 &
                   PLF_TF56$gV != 0 &
                   PLF_TF56$fA != 0 &
                   PLF_TF56$fV != 0 )
#normalise prices of variable inputs (A, L)
PLF_TF56$Awp <- PLF_TF56$Aw / PLF_TF56$Y_P
PLF_TF56$Vwp <- PLF_TF56$Vw / PLF_TF56$Y_P
####################  MODEL1 1 (only production risk) #########################
#calculate risk preference function Theta1 (D1) for Model 1
PLF_TF56$d1.2 <- (PLF_TF56$fA - PLF_TF56$Awp) / ( - PLF_TF56$gA)
PLF_TF56$d1.3 <- (PLF_TF56$fV - PLF_TF56$Vwp) / ( - PLF_TF56$gV)
PLF_TF56$D1 <- 0.5 * (PLF_TF56$d1.2 + PLF_TF56$d1.3)
#calculate RP for Model 1
PLF_TF56$RP1 <- -0.5 * PLF_TF56$D1 * fitted(model.np.g)
summary(PLF_TF56$RP1)
#calculate RP for Model 1
PLF_TF56$RRP1 <- PLF_TF56$RP1 / (PLF_TF56$Y_P * PLF_TF56$YQ-
                   (PLF_TF56$Aw * PLF_TF56$A + PLF_TF56$Vw * PLF_TF56$V))
summary(PLF_TF56$RRP1)
#calculate "modified" RRP (RP1 / total revenue) for Model 1
PLF_TF56$RRP1_revenue  <- PLF_TF56$RP1 / (PLF_TF56$Y_P * PLF_TF56$YQ)
summary(PLF_TF56$RRP11)

##################  MODEL 2 (only output price uncertainty) ###################
#calculate eta
library(plm)
PLF_TF56pd <- pdata.frame(PLF_TF56, c("ID","T"), drop = FALSE)
PLF_TF56pd$y_p <- PLF_TF56pd$Y_P
PLF_TF56pd$y_p_lag <- lag(PLF_TF56pd$y_p, 1)
PLF_TF56pd$eta <- log(PLF_TF56pd$y_p_lag / PLF_TF56pd$y_p)
PLF_TF56$eta <- as.vector(PLF_TF56pd$eta)
#calculate variances of omega1 and omega2
PLF_TF56$omega1 <- var((exp(PLF_TF56$eta) - 1), na.rm = TRUE)
PLF_TF56$omega2 <- PLF_TF56$omega1
```

```
#calculate risk preference function Theta2 (D2) for Model 2
PLF_TF56$d2.2 <- (PLF_TF56$Awp / PLF_TF56$fA) - 1
PLF_TF56$d2.3 <- (PLF_TF56$Vwp / PLF_TF56$fV) - 1
PLF_TF56$D2 <- 0.5 * (PLF_TF56$d2.2 + PLF_TF56$d2.3)
summary(PLF_TF56$D2)
#calculate RP for Model 2
PLF_TF56$RP2 <- -0.5 * PLF_TF56$D2 * fitted(model.np.f) * PLF_TF56$omega1
#calculate RP for Model 2
PLF_TF56$RRP2 <- PLF_TF56$RP2 / (PLF_TF56$Y_P * PLF_TF56$YQ-
                   (PLF_TF56$Aw * PLF_TF56$A + PLF_TF56$Vw * PLF_TF56$V))
summary(PLF_TF56$RRP2)
#calculate "modified" RRP (RP2 / total revenue) for Model 2
PLF_TF56$RRP22 <- PLF_TF56$RP2 / (PLF_TF56$Y_P * PLF_TF56$YQ)
summary(PLF_TF56$RRP22)

########  MODEL 3 (both production risk and output price uncertainty)  #########
#Calculate risk preference functions Theta1 (D3_T1.1) and Theta2 (D3_T1.1)
#for Model 3
PLF_TF56$D3_T1.1 <- (PLF_TF56$Awp - fitted(model.np.D3) * PLF_TF56$Vwp) /
                      (PLF_TF56$gA - fitted(model.np.D3) * PLF_TF56$gV)
PLF_TF56$D3_T2.1 <- ((PLF_TF56$Awp - PLF_TF56$gA * PLF_TF56$D3_T1.1) /
   (PLF_TF56$fA)) - 1
#calculate RP for Model 3
PLF_TF56$RP3 <- -0.5 * (PLF_TF56$D3_T1.1 / fitted(model.np.g)) *
                      ((fitted(model.np.f)^2) * PLF_TF56$omega1 +
                       (fitted(model.np.g)^2) * PLF_TF56$omega2)
#calculate RP for Model 3
PLF_TF56$RRP3 <- PLF_TF56$RP3 / (PLF_TF56$Y_P * PLF_TF56$YQ -
                  (PLF_TF56$Aw * PLF_TF56$A + PLF_TF56$Vw * PLF_TF56$V))
summary(PLF_TF56$RRP3)
#calculate "modified" RRP (RP3 / total revenue) for Model 3
PLF_TF56$RRP31 <- PLF_TF56$RP3 / (PLF_TF56$Y_P * PLF_TF56$YQ)
summary(PLF_TF56$RRP31)
```

E.2. Estimation in logs

```
##Nonparametric estimation of the mean production function of the Just and Pope
##technology in logarithms of regression variables
##Cross-validation bandwidth selection
bw.f.log <- npregbw(lY ~ lL + lA + lV + lK + T_of +ID,
                    regtype = "ll",
                    bwmethod = "cv.aic",
                    ckertype = "epanechnikov",
                    ukertype = "liracine",
                    okertype = "liracine",
                    data = PLF_TF56)
##Estimation of the nonparametric model
model.np.f.log <- npreg(bws = bw.f.log,
                        data = PLF_TF56,
                        gradients = TRUE,
                        residuals = TRUE)
summary(model.np.f.log)
##Bootstrap significance test
sigtest.model.np.f.log <- npsigtest(model.np.f.log, boot.num = 399)
summary(sigtest.model.np.f.log)

##Nonparametric estimation of the risk variance function of the Just and Pope
##technology in logarithms of regression variables
##Calculate dependent variable
PLF_TF56$lAbsu <- log(abs((exp(model.np.f.log$resid) - 1)*
                          exp(fitted(model.np.f.log))))
##Cross-validation bandwidth selection
bw.g.log <- npregbw(lAbsu ~ lL + lA + lV + lK + T_of +ID,
                                              regtype = "ll",
                                              bwmethod = "cv.aic",
                                              ckertype = "epanechnikov",
                                              ukertype = "liracine",
                                              okertype = "liracine",
                                              data = PLF_TF56)
##Estimation of the nonparametric model
model.np.g.log <- npreg(bws = bw.g.log,
                        data = PLF_TF56,
                        gradients = TRUE,
                        residuals = TRUE)
summary(model.np.g.log)
##Bootstrap significance test
sigtest.model.np.g.log <- npsigtest(model.np.g.log, boot.num = 399)
summary(sigtest.model.np.g.log)
##Nonparametric estimation of the risk preference functions
##Model 1
################################################################################
```

```
##
##Nonparametric estimation of the risk preference functions
##using avaraged sollutions to F.O.C.
PLF_TF56$d1.2 <-(PLF_TF56$fA-PLF_TF56$Awp)/(-PLF_TF56$gA)
PLF_TF56$d1.3 <-(PLF_TF56$fV-PLF_TF56$Vwp)/(-PLF_TF56$gV)
calculate risk preference function Theta1 (D1) for Model 1
PLF_TF56$D1 <- 0.5*(PLF_TF56$d1.2+PLF_TF56$d1.3)

bw.D1.log <- npregbw(D1 ~ lL + lA + lV + lK + lAwp + lVwp  +T_of +ID,
     regtype="ll",
     bwmethod="cv.aic",
     ckertype = "epanechnikov",
     ukertype = "liracine",
     okertype = "liracine",
     data=PLF_TF56)
model.np.D1.log <- npreg(bws=bw.D1.log,
     data=PLF_TF56,
     gradients = TRUE,
     residuals = TRUE)
sigtest.model.np.D1.log <- npsigtest(model.np.D1.log,
  boot.num=399)
sigtest.model.np.D1.log
###
##Nonparametric estimation of the risk preference functions
using sollution to F.O.C. with respect to land variable (A)
bw.d1.2.log <- npregbw(d1.2 ~ lL + lA + lV + lK + lAwp + lVwp  +T_of +ID,
  regtype="ll",
  bwmethod="cv.aic",
  ckertype = "epanechnikov",
  ukertype = "liracine",
  okertype = "liracine",
  data=PLF_TF56)

model.np.d1.2.log <- npreg(bws=bw.d1.2.log,
  data=PLF_TF56,
  gradients = TRUE,
  residuals = TRUE)
sigtest.model.np.d1.2.log <- npsigtest(model.np.d1.2.log,
  boot.num=399)

bw.d1.3.log <- npregbw(d1.3 ~ lL + lA + lV + lK + lAwp + lVwp  +T_of +ID,
  regtype="ll",
  bwmethod="cv.aic",
  ckertype = "epanechnikov",
  ukertype = "liracine",
  okertype = "liracine",
  data=PLF_TF56)
```

```
model.np.d1.3.log <- npreg(bws=bw.d1.3.log,
  data=PLF_TF56,
  gradients = TRUE,
  residuals = TRUE)
sigtest.model.np.d1.3.log <- npsigtest(model.np.d1.3.log,
  boot.num=399)
sigtest.model.np.d1.3.log
##Model 2
##Nonparametric estimation of the risk preference functions
##using avaraged sollutions to F.O.C.
PLF_TF56$d2.2 <- (PLF_TF56$Awp/PLF_TF56$fA)-1
PLF_TF56$d2.3 <- (PLF_TF56$Vwp/PLF_TF56$fV)-1
#calculate risk preference function Theta2 (D2) for Model 2
PLF_TF56$D2 <- 0.5*(PLF_TF56$d2.2+PLF_TF56$d2.3)

bw.D2.log <- npregbw(D2 ~ lL + lA + lV + lK + lAwp + lVwp  +T_of +ID,
  regtype="ll",
  bwmethod="cv.aic",
  ckertype = "epanechnikov",
  ukertype = "liracine",
  okertype = "liracine",
  data=PLF_TF56)
bw.D2.log
model.np.D2.log <- npreg(bws=bw.D2.log,
  data=PLF_TF56,
  gradients = TRUE,
  residuals = TRUE)
model.np.D2.log
sigtest.model.np.D2.log <- npsigtest(model.np.D2.log,
  boot.num=399)
sigtest.model.np.D2.log
##
##Nonparametric estimation of the risk preference functions
##using sollution to F.O.C. with respect to land variable (A)
bw.d2.2.log <- npregbw(d2.2 ~ lL + lA + lV + lK + lAwp + lVwp  +T_of +ID,
  regtype="ll",
  bwmethod="cv.aic",
  ckertype = "epanechnikov",
  ukertype = "liracine",
  okertype = "liracine",
  data=PLF_TF56)

model.np.d2.2.log <- npreg(bws=bw.d2.2.log,
  data=PLF_TF56,
  gradients = TRUE,
  residuals = TRUE)
```

```
sigtest.model.np.d2.2.log <- npsigtest(model.np.d2.2.log,
  boot.num=399)
##Nonparametric estimation of the risk preference functions
##using sollution to F.O.C. with respect to variable inputs variable (V)
bw.d2.3.log <- npregbw(d2.3 ~ lL + lA + lV + lK + lAwp + lVwp  +T_of +ID,
  regtype="ll",
  bwmethod="cv.aic",
  ckertype = "epanechnikov",
  ukertype = "liracine",
  okertype = "liracine",
  data=PLF_TF56)

model.np.d2.3.log <- npreg(bws=bw.d2.3.log,
  data=PLF_TF56,
  gradients = TRUE,
  residuals = TRUE)
sigtest.model.np.d2.3.log <- npsigtest(model.np.d2.3.log,
  boot.num=399)
sigtest.model.np.d1.3.log

##Nonparametric estimation of the risk preference functions
##using avaraged sollutions to F.O.C. with V as a numeraire (D3_VA)
PLF_TF56$D3_VA <- 0.5* ((model.np.f.log$grad[,3]*(PLF_TF56$YQ/PLF_TF56$V))/
    (model.np.f.log$grad[,3]*(PLF_TF56$YQ/PLF_TF56$V))+
    (model.np.f.log$grad[,2]*(PLF_TF56$YQ/PLF_TF56$A))/
    (model.np.f.log$grad[,3]*(PLF_TF56$YQ/PLF_TF56$V))

bw.D3_VA.log <- npregbw(D3_VA ~ lL + lA + lV + lK + lAwp + lVwp  +T_of +ID,
  regtype="ll",
  bwmethod="cv.aic",
  ckertype = "epanechnikov",
  ukertype = "liracine",
  okertype = "liracine",
  data=PLF_TF56)
bw.D3_VA.log
model.np.D3_VA.log <- npreg(bws=bw.D3_VA.log,
  data=PLF_TF56,
  gradients = TRUE,
  residuals = TRUE)
model.np.D3_VA.log
sigtest.model.np.D3_VA.log <- npsigtest(model.np.D3_VA.log,
  boot.num=399)
sigtest.model.np.D3_VA.log
##using avaraged sollutions to F.O.C. with A as a numeraire (D3_AV)
PLF_TF56$D3_AV <- 0.5* ((model.np.f.log$grad[,2]*(PLF_TF56$YQ/PLF_TF56$A))/
    (model.np.f.log$grad[,2]*(PLF_TF56$YQ/PLF_TF56$A))+
    (model.np.f.log$grad[,3]*(PLF_TF56$YQ/PLF_TF56$V))/
```

```
    (model.np.f.log$grad[,2]*(PLF_TF56$YQ/PLF_TF56$A)))

bw.D3_AV.log <- npregbw(D3_AV ~ lL + lA + lV + lK + lAwp + lVwp  +T_of +ID,
  regtype="ll",
  bwmethod="cv.aic",
  ckertype = "epanechnikov",
  ukertype = "liracine",
  okertype = "liracine",
  data=PLF_TF56)
bw.D3_AV.log
model.np.D3_AV.log <- npreg(bws=bw.D3_AV.log,
  data=PLF_TF56,
  gradients = TRUE,
  residuals = TRUE)
model.np.D3_AV.log
sigtest.model.np.D3_AV.log <- npsigtest(model.np.D3_AV.log,
  boot.num=399)
sigtest.model.np.D3_AV.log
##
##Nonparametric estimation of the risk preference functions
##using sollution to F.O.C. with respect to variable inputs variable (V)
PLF_TF56$D3_V <-(model.np.f.log$grad[,2]*(PLF_TF56$YQ/PLF_TF56$A))/
  (model.np.f.log$grad[,3]*(PLF_TF56$YQ/PLF_TF56$V))

bw.D3_V.log <- npregbw(D3_V ~ lL + lA + lV + lK + lAwp + lVwp  +T_of +ID,
  regtype="ll",
  bwmethod="cv.aic",
  ckertype = "epanechnikov",
  ukertype = "liracine",
  okertype = "liracine",
  data=PLF_TF56)
bw.D3_V.log
model.np.D3_V.log <- npreg(bws=bw.D3_V.log,
  data=PLF_TF56,
  gradients = TRUE,
  residuals = TRUE)
model.np.D3_V.log
sigtest.model.np.D3_V.log <- npsigtest(model.np.D3_V.log,
  boot.num=399)
sigtest.model.np.D3_V.log
##
##Nonparametric estimation of the risk preference functions
##using sollution to F.O.C. with respect to land variable (A)
PLF_TF56$D3_A <- (model.np.f.log$grad[,3]*(PLF_TF56$YQ/ PLF_TF56$V))/
  (model.np.f.log$grad[,2]*(PLF_TF56$YQ/ PLF_TF56$A))

bw.D3_A.log <- npregbw(D3_A ~ lL + lA + lV + lK + lAwp + lVwp  +T_of +ID,
```

```
  regtype="ll",
  bwmethod="cv.aic",
  ckertype = "epanechnikov",
  ukertype = "liracine",
  okertype = "liracine",
  data=PLF_TF56)

model.np.D3_A.log <- npreg(bws=bw.D3_A.log,
  data=PLF_TF56,
  gradients = TRUE,
  residuals = TRUE)
sigtest.model.np.D3_A.log <- npsigtest(model.np.D3_A.log,
  boot.num=399)

sigtest.model.np.D3_VA.log <- npsigtest(model.np.D3_VA.log,
  boot.num=399)
##
## "Calculated" risk preferences and relative risk premiums
## for  MODEL 1 (only production risk)
#risk preference function Theta1 (D1) for Model 1
PLF_TF56$d1.2 <-(PLF_TF56$fA-PLF_TF56$Awp)/(-PLF_TF56$gA)
summary(PLF_TF56$d1.2)
PLF_TF56$d1.3 <-(PLF_TF56$fV-PLF_TF56$Vwp)/(-PLF_TF56$gV)
summary(PLF_TF56$d1.3)
PLF_TF56$D1 <- 0.5*(PLF_TF56$d1.2+PLF_TF56$d1.3)
summary(PLF_TF56$D1)
length(PLF_TF56$D1[PLF_TF56$D1<0])/length(PLF_TF56$D1)
#RP for the model with only production risk
PLF_TF56$RP1 <- -0.5*PLF_TF56$D1*(exp(fitted(model.np.g.log)))
summary(PLF_TF56$RP1)
#RRP for the model with only production risk
PLF_TF56$RRP1 <- PLF_TF56$RP1/(PLF_TF56$Y_P*PLF_TF56$YQ-
    (PLF_TF56$Aw*PLF_TF56$A+PLF_TF56$Vw*PLF_TF56$V))
summary(PLF_TF56$RRP1)
#RRP for the model with only production risk (RP1/total revenue)
PLF_TF56$RRP1_rev <- PLF_TF56$RP1/(PLF_TF56$Y_P*PLF_TF56$YQ)
summary(PLF_TF56$RRP11)

#RP for the model with only production risk
#based on the F.O.C solved for A
PLF_TF56$RP1_d1.2 <- -0.5*PLF_TF56$d1.2*(exp(fitted(model.np.g.log)))
summary(PLF_TF56$RP1_d1.2)
#cRRP for the model with only production risk
#based on the F.O.C solved for A
PLF_TF56$RRP1_d1.2 <- PLF_TF56$RP1_d1.2/(PLF_TF56$Y_P*PLF_TF56$YQ-
    (PLF_TF56$Aw*PLF_TF56$A+PLF_TF56$Vw*PLF_TF56$V))
summary(PLF_TF56$RRP1_d1.2)
```

```
#RRP for the model with only production risk (RP1/total revenue)
#based on the F.O.C solved for A
PLF_TF56$RRP1_rev_d1.2  <- PLF_TF56$RP1_d1.2/(PLF_TF56$Y_P*PLF_TF56$YQ)
summary(PLF_TF56$RRP1_rev_d1.2)

#RP for the model with only production risk
#based on the F.O.C solved for V
PLF_TF56$RP1_d1.3 <- -0.5*PLF_TF56$d1.3*(exp(fitted(model.np.g.log)))
summary(PLF_TF56$RP1_d1.3)
#RRP for the model with only production risk
#based on the F.O.C solved for V
PLF_TF56$RRP1_d1.3 <- PLF_TF56$RP1_d1.3/(PLF_TF56$Y_P*PLF_TF56$YQ-
    (PLF_TF56$Aw*PLF_TF56$A+PLF_TF56$Vw*PLF_TF56$V))
summary(PLF_TF56$RRP1_d1.3)
#RRP for the model with only production risk (RP1/total revenue)
#based on the F.O.C solved for V
PLF_TF56$RRP1_rev_d1.3  <- PLF_TF56$RP1_d1.3/(PLF_TF56$Y_P*PLF_TF56$YQ)
summary(PLF_TF56$RRP1_rev_d1.3)
##
## "Estimated" risk preferences and relative risk premiums
## for  MODEL 1 (only production risk)

#RP for the model with only production risk
#using the estimated nonparametric risk preferences
#(avaraged for 2 variable inputs: A,V)
# all.equal(PLF_TF56$D1 , fitted(model.np.D1.log)+resid(model.np.D1.log),
# check.attributes = FALSE)
summary(fitted(model.np.D1.log))
PLF_TF56$RP1_est <- -0.5*fitted(model.np.D1.log)*(exp(fitted(model.np.g.log)))
summary(PLF_TF56$RP1_est)
#RRP for the model with only production risk
PLF_TF56$RRP1_est <- PLF_TF56$RP1_est/(PLF_TF56$Y_P*PLF_TF56$YQ-
    (PLF_TF56$Aw*PLF_TF56$A+PLF_TF56$Vw*PLF_TF56$V))
summary(PLF_TF56$RRP1_est)
#RRP for the model with only production risk (RP1/total revenue)
PLF_TF56$RRP1_rev_est  <- PLF_TF56$RP1_est/(PLF_TF56$Y_P*PLF_TF56$YQ)
summary(PLF_TF56$RRP1_rev_est)
#RP for the model with only production risk
#using the estimated nonparametric risk preferences (only A)
all.equal(PLF_TF56$d1.2 , fitted(model.np.d1.2.log)+resid(model.np.d1.2.log),
check.attributes = FALSE)
summary(PLF_TF56$d1.2)
summary(fitted(model.np.d1.2.log))
#RP for the model with only production risk
PLF_TF56$RP1_d1.2_est <- -0.5*fitted(model.np.d1.2.log)*(exp(fitted(model.np.g.log)))
summary(PLF_TF56$RP1_d1.2_est)
#RRP for the model with only production risk
```

```
PLF_TF56$RRP1_d1.2_est <- PLF_TF56$RP1_d1.2_est/(PLF_TF56$Y_P*PLF_TF56$YQ-
    (PLF_TF56$Aw*PLF_TF56$A+PLF_TF56$Vw*PLF_TF56$V))
summary(PLF_TF56$RRP1_d1.2_est)
#RRP for the model with only production risk (RP1/total revenue)
PLF_TF56$RRP1_rev_d1.2_est  <- PLF_TF56$RP1_d1.2_est/(PLF_TF56$Y_P*PLF_TF56$YQ)
summary(PLF_TF56$RRP1_rev_d1.2_est)
#RP for the model with only production risk
#using the estimated nonparametric risk preferences (only V)
all.equal(PLF_TF56$d1.3 , fitted(model.np.d1.3.log)+resid(model.np.d1.3.log),
check.attributes = FALSE)
#RP for the model with only production risk
PLF_TF56$RP1_d1.3_est <- -0.5*fitted(model.np.d1.3.log)*(exp(fitted(model.np.g.log)))
summary(fitted(model.np.d1.3.log))
summary(PLF_TF56$RP1_d1.3_est)
#calculate RRP for the model with only production risk
PLF_TF56$RRP1_d1.3_est <- PLF_TF56$RP1_d1.3_est/(PLF_TF56$Y_P*PLF_TF56$YQ-
    (PLF_TF56$Aw*PLF_TF56$A+PLF_TF56$Vw*PLF_TF56$V))
summary(PLF_TF56$RRP1_d1.3_est)
#calculate RRP for the model with only production risk (RP1/total revenue)
PLF_TF56$RRP1_rev_d1.3_est  <- PLF_TF56$RP1_d1.3_est/(PLF_TF56$Y_P*PLF_TF56$YQ)
summary(PLF_TF56$RRP1_rev_d1.3_est)
##
## "Calculated" risk preferences and relative risk premiums
##  for MODEL 2 (only output price uncertainty)
#calculate eta
library(plm)
PLF_TF56pd <- pdata.frame(PLF_TF56, c("ID","T"), drop = FALSE)
PLF_TF56pd$y_p <- PLF_TF56pd$Y_P
PLF_TF56pd$y_p_lag <- lag(PLF_TF56pd$y_p, 1)
PLF_TF56pd$eta <- log(PLF_TF56pd$y_p_lag/PLF_TF56pd$y_p)
PLF_TF56$eta <- as.vector(PLF_TF56pd$eta)
#calculate variances of omega1 and omega2
PLF_TF56$omega1 <- var((exp(PLF_TF56$eta)-1), na.rm=TRUE)
PLF_TF56$omega2 <-PLF_TF56$omega1
# #calculate risk preference function Theta2 (D2) for Model 2
# PLF_TF56$d2.2 <- (PLF_TF56$Awp/PLF_TF56$fA)-1
# PLF_TF56$d2.3 <- (PLF_TF56$Vwp/PLF_TF56$fV)-1
# PLF_TF56$D2 <- 0.5*(PLF_TF56$d2.2+PLF_TF56$d2.3)
summary(PLF_TF56$D2)
#RP for the Model 2
PLF_TF56$RP2 <- -0.5*PLF_TF56$D2*exp(fitted(model.np.f.log))*PLF_TF56$omega1
#RRP2 for the Model 2
PLF_TF56$RRP2 <- PLF_TF56$RP2/(PLF_TF56$Y_P*PLF_TF56$YQ-
    (PLF_TF56$Aw*PLF_TF56$A+PLF_TF56$Vw*PLF_TF56$V))
summary(PLF_TF56$RRP2)
#"modified" RRP2 (RP2/total revenue) for Model 2
PLF_TF56$RRP2_rev <- PLF_TF56$RP2/(PLF_TF56$Y_P*PLF_TF56$YQ)
```

```
summary(PLF_TF56$RRP2_rev)


all.equal(PLF_TF56$d2.2 , fitted(model.np.d2.2.log)+resid(model.np.d2.2.log),
check.attributes = FALSE, tol=0.001)
#RP for for Model 2
#using the estimated nonparametric risk preferences (only A)
summary(fitted(model.np.d2.2.log))
PLF_TF56$RP2_d2.2 <- -0.5*PLF_TF56$d2.2*exp(fitted(model.np.f.log))*PLF_TF56$omega1
#RRP2 for the Model 2
PLF_TF56$RRP2_d2.2 <- PLF_TF56$RP2_d2.2/(PLF_TF56$Y_P*PLF_TF56$YQ-
    (PLF_TF56$Aw*PLF_TF56$A+PLF_TF56$Vw*PLF_TF56$V))
summary(PLF_TF56$RRP2_d2.2)
#calculate "modified" RRP2 (RP2/total revenue) for Model 3
PLF_TF56$RRRP2_d2.2_rev <- PLF_TF56$RP2_d2.2/(PLF_TF56$Y_P*PLF_TF56$YQ)
summary(PLF_TF56$RRP22)

#RP for for Model 2
#using the estimated nonparametric risk preferences (only V)
PLF_TF56$RP2_d2.3 <- -0.5*PLF_TF56$d2.3*exp(fitted(model.np.f.log))*PLF_TF56$omega1
#RRP2 for the Model 2
PLF_TF56$RRP2_d2.3 <- PLF_TF56$RP2_d2.3/(PLF_TF56$Y_P*PLF_TF56$YQ-
    (PLF_TF56$Aw*PLF_TF56$A+PLF_TF56$Vw*PLF_TF56$V))
summary(PLF_TF56$RRP2_d2.3)
#calculate "modified" RRP2 (RP2/total revenue) for Model 3
PLF_TF56$RRP2_d2.3_rev <- PLF_TF56$RP2_d2.3/(PLF_TF56$Y_P*PLF_TF56$YQ)
summary(PLF_TF56$RRP2_d2.3_rev)

## "Estimated" risk preferences and relative risk premiums
##  MODEL 2 (only output price uncertainty)
#RP for the Model 2
PLF_TF56$RP2_est <- -0.5*fitted(model.np.D2.log)*exp(fitted(model.np.f.log))*
                          PLF_TF56$omega1
summary(fitted(model.np.D2.log))
#RRP2 for the Model 2
PLF_TF56$RRP2_est <- PLF_TF56$RP2_est/(PLF_TF56$Y_P*PLF_TF56$YQ-
    (PLF_TF56$Aw*PLF_TF56$A+PLF_TF56$Vw*PLF_TF56$V))
summary(PLF_TF56$RRP2_est)
#calculate "modified" RRP2 (RP2/total revenue) for Model 3
PLF_TF56$RRP2_rev_est <- PLF_TF56$RP2_est/(PLF_TF56$Y_P*PLF_TF56$YQ)
summary(PLF_TF56$RRP2_rev_est)

#RP for for Model 2
#using the estimated nonparametric risk preferences (only A)
all.equal(PLF_TF56$d2.2, fitted(model.np.d2.2.log)+resid(model.np.d2.2.log),
check.attributes = FALSE)
summary(fitted(model.np.d2.2.log))
```

```
#RP for the Model 2
PLF_TF56$RP2_d2.2 <- -0.5*fitted(model.np.d2.2.log)*exp(fitted(model.np.f.log))*
                          PLF_TF56$omega1
summary(PLF_TF56$RP2_d2.2)
#RRP2 for the Model 2
PLF_TF56$RRP2_d2.2 <- PLF_TF56$RP2_d2.2/(PLF_TF56$Y_P*PLF_TF56$YQ-
                      (PLF_TF56$Aw*PLF_TF56$A+PLF_TF56$Vw*PLF_TF56$V))
summary(PLF_TF56$RRP2_d2.2)
#calculate "modified" RRP2 (RP2/total revenue) for Model 2
PLF_TF56$RRRP2_d2.2_rev <- PLF_TF56$RP2_d2.2/(PLF_TF56$Y_P*PLF_TF56$YQ)
summary(PLF_TF56$RRP22)

#RP for for Model 2
#using the estimated nonparametric risk preferences (only V)
all.equal(PLF_TF56$d2.3, fitted(model.np.d2.3.log)+resid(model.np.d2.3.log),
check.attributes = FALSE)
#RP for the Model 2
summary(fitted(model.np.d2.3.log))
PLF_TF56$RP2_d2.3_est <- -0.5*fitted(model.np.d2.3.log)*
                          exp(fitted(model.np.f.log))*PLF_TF56$omega1
#RRP2 for the Model 2
PLF_TF56$RRP2_d2.3_est <- PLF_TF56$RP2_d2.3_est/(PLF_TF56$Y_P*PLF_TF56$YQ-
                           (PLF_TF56$Aw*PLF_TF56$A+PLF_TF56$Vw*PLF_TF56$V))
summary(PLF_TF56$RRP2_d2.3_est)
#calculate "modified" RRP2 (RP2/total revenue) for Model 3
PLF_TF56$RRP2_d2.3_rev_est <- PLF_TF56$RP2_d2.3_est/(PLF_TF56$Y_P*PLF_TF56$YQ)
summary(PLF_TF56$RRP2_d2.3_rev_est)

## "Calculated" risk preferences and relative risk premiums
## for MODEL 3 (both production risk and output price uncertainty)
##
## using avaraged sollutions to F.O.C. with V as a numeraire (D3_VA)
#Calculate risk preference functions Theta1 (D3_T1.1) and Theta2 (D3_T1.1)
#for Model 3

PLF_TF56$D3_T1_VA <-((PLF_TF56$Awp - PLF_TF56$D3_VA*PLF_TF56$Vwp)
  + (PLF_TF56$Vwp - PLF_TF56$D3_VA*PLF_TF56$Vwp))/
  ((PLF_TF56$gA - PLF_TF56$D3_VA*PLF_TF56$gV)
    + (PLF_TF56$gV - PLF_TF56$D3_VA*PLF_TF56$gV))
summary(PLF_TF56$D3_T1_VA)

length(PLF_TF56$D3_T1_VA[PLF_TF56$D3_T1_VA<0])/length(PLF_TF56$D3_T1_VA)


PLF_TF56$D3_T2_VA <- ((PLF_TF56$Awp-PLF_TF56$gA*PLF_TF56$D3_T1_VA+PLF_TF56$Vwp-
                       PLF_TF56$gV*PLF_TF56$D3_T1_VA)/
                       (PLF_TF56$fA+PLF_TF56$fV))-1
```

```
summary(PLF_TF56$D3_T2_VA)
length(PLF_TF56$D3_T2_VA[PLF_TF56$D3_T2_VA<0])/length(PLF_TF56$D3_T2_VA)

PLF_TF56$RP3_VA <- -0.5*(PLF_TF56$D3_T1_VA/exp(fitted(model.np.g.log)))*
  (((exp(fitted(model.np.f.log))^2)*PLF_TF56$omega1)+
      ((exp(fitted(model.np.g.log))^2)*PLF_TF56$omega2))
summary(PLF_TF56$RP3_VA)

PLF_TF56$RRP3_VA <- PLF_TF56$RP3_VA/(PLF_TF56$Y_P*PLF_TF56$YQ-
    (PLF_TF56$Aw*PLF_TF56$A+PLF_TF56$Vw*PLF_TF56$V))
summary(PLF_TF56$RRP3_VA)

summary(PLF_TF56$RRP3_VA[PLF_TF56$T==6])

#calculate "modified" RRP (RP3/total revenue) for Model 3
PLF_TF56$RRP3_rev_VA <- PLF_TF56$RP3_VA/(PLF_TF56$Y_P*PLF_TF56$YQ)
summary(PLF_TF56$RRP3_rev_VA)

##
##using avaraged sollutions to F.O.C. with A as a numeraire (D3_AV)

PLF_TF56$D3_T1_AV <-((PLF_TF56$Vwp - PLF_TF56$D3_AV*PLF_TF56$Awp)
  + (PLF_TF56$Awp - PLF_TF56$D3_AV*PLF_TF56$Awp))/
  ((PLF_TF56$gV - PLF_TF56$D3_AV*PLF_TF56$gA)
    + (PLF_TF56$gA - PLF_TF56$D3_AV*PLF_TF56$gA))
summary(PLF_TF56$D3_T1_AV)


PLF_TF56$D3_T2_AV <-PLF_TF56$D3_T1_AV*(exp(fitted(model.np.f.log))/
    exp(fitted(model.np.g.log)))

PLF_TF56$RP3_AV <- -0.5*(PLF_TF56$D3_T1_AV/exp(fitted(model.np.g.log)))*
  (((exp(fitted(model.np.f.log))^2)*PLF_TF56$omega1)+
      ((exp(fitted(model.np.g.log))^2)*PLF_TF56$omega2))
summary(PLF_TF56$RP3_AV)

PLF_TF56$RRP3_1_AV <- PLF_TF56$RP3_AV/(PLF_TF56$Y_P*PLF_TF56$YQ-
    (PLF_TF56$Aw*PLF_TF56$A+PLF_TF56$Vw*PLF_TF56$V))
summary(PLF_TF56$RRP3_1_AV)

#calculate "modified" RRP (RP3/total revenue) for Model 3
PLF_TF56$RRP3_rev_1_AV <- PLF_TF56$RP3_AV/(PLF_TF56$Y_P*PLF_TF56$YQ)
summary(PLF_TF56$RRP3_rev_1_AV)

##using sollution to F.O.C. with respect to variable inputs (V)
PLF_TF56$D3_T1_V_calc <- (PLF_TF56$Awp - PLF_TF56$D3_V*PLF_TF56$Vwp)/
  (PLF_TF56$gA - PLF_TF56$D3_V*PLF_TF56$gA)
```

```
summary(PLF_TF56$D3_T1_V_calc)

PLF_TF56$D3_T2_V_calc <-PLF_TF56$D3_T1_V_calc*(exp(fitted(model.np.f.log))/
    exp(fitted(model.np.g.log)))
summary(PLF_TF56$D3_T2_V_calc)

PLF_TF56$RP3_V_calc <- -0.5*(PLF_TF56$D3_T1_V_calc/exp(fitted(model.np.g.log)))*
  (((exp(fitted(model.np.f.log))^2)*PLF_TF56$omega1) +
      ((exp(fitted(model.np.g.log))^2)*PLF_TF56$omega2))
summary(PLF_TF56$RP3_V_calc)

# #calculate "modified" RRP (RP3/total revenue) for Model 3
PLF_TF56$RRP3_V_calc <- PLF_TF56$RP3_V_calc/(PLF_TF56$Y_P*PLF_TF56$YQ-
    (PLF_TF56$Aw*PLF_TF56$A+PLF_TF56$Vw*PLF_TF56$V))
summary(PLF_TF56$RRP3_V_calc)

PLF_TF56$RRP3_V_rev_calc <- PLF_TF56$RP3_V_calc/(PLF_TF56$Y_P*PLF_TF56$YQ)
summary(PLF_TF56$RRP3_V_rev_calc)

##using sollution to F.O.C. with respect to land (A)
PLF_TF56$D3_T1_A_calc <-((PLF_TF56$Vwp - PLF_TF56$D3_A*PLF_TF56$Awp)/
    (PLF_TF56$gV - PLF_TF56$D3_A*PLF_TF56$gV))
summary(PLF_TF56$D3_T1_A_calc)

PLF_TF56$D3_T2_A_calc <-PLF_TF56$D3_T1_A_calc*(exp(fitted(model.np.f.log))/
    exp(fitted(model.np.g.log)))
summary(PLF_TF56$D3_T2_A_calc)

PLF_TF56$RP3_A_calc <- -0.5*(PLF_TF56$D3_T1_A_calc/exp(fitted(model.np.g.log)))*
  (((exp(fitted(model.np.f.log))^2)*PLF_TF56$omega1) +
      ((exp(fitted(model.np.g.log))^2)*PLF_TF56$omega2))
summary(PLF_TF56$RP3_A_calc)


PLF_TF56$RRP3_A_calc <- PLF_TF56$RP3_A_calc/ (PLF_TF56$Y_P*PLF_TF56$YQ -
    (PLF_TF56$Aw*PLF_TF56$A + PLF_TF56$Vw*PLF_TF56$V))
summary(PLF_TF56$RRP3_A_calc)
# #calculate "modified" RRP (RP3/total revenue) for Model 3
PLF_TF56$RRP3_A_rev_calc <- PLF_TF56$RP3_A_calc/ (PLF_TF56$Y_P*PLF_TF56$YQ)
summary(PLF_TF56$RRP3_A_rev_calc)

### "Estimated" risk preferences and relative risk premiums
### for MODEL 3 (both production risk and output price uncertainty)
## using avaraged sollutions to F.O.C. with V as a numeraire (D3_VA)
all.equal(PLF_TF56$D3_VA , fitted(model.np.D3_VA.log)+resid(model.np.D3_VA.log),
check.attributes = FALSE)
```

```
PLF_TF56$D3_T1_VA_est <-((PLF_TF56$Awp - fitted(model.np.D3_VA.log)*PLF_TF56$Vwp)
  + (PLF_TF56$Vwp - fitted(model.np.D3_VA.log)*PLF_TF56$Vwp))/
  ((PLF_TF56$gA - fitted(model.np.D3_VA.log)*PLF_TF56$gV)
    + (PLF_TF56$gV - fitted(model.np.D3_VA.log)*PLF_TF56$gV))
summary(PLF_TF56$D3_T1_VA_est)

PLF_TF56$D3_T2_VA_est <-PLF_TF56$D3_T1_VA_est*(exp(fitted(model.np.f.log))/
    exp(fitted(model.np.g.log)))
summary(PLF_TF56$D3_T2_VA_est)

PLF_TF56$RP3_VA_est <- -0.5*(PLF_TF56$D3_T1_VA_est/exp(fitted(model.np.g.log)))*
  (((exp(fitted(model.np.f.log))^2)*PLF_TF56$omega1)+
      ((exp(fitted(model.np.g.log))^2)*PLF_TF56$omega2))
summary(PLF_TF56$RP3_VA_est)


PLF_TF56$RRP3_VA_est <- PLF_TF56$RP3_VA_est/(PLF_TF56$Y_P*PLF_TF56$YQ-
    (PLF_TF56$Aw*PLF_TF56$A+PLF_TF56$Vw*PLF_TF56$V))
summary(PLF_TF56$RRP3_VA_est)
#calculate "modified" RRP (RP3/total revenue) for Model 3
PLF_TF56$RRP3_rev_VA_est <- PLF_TF56$RP3_VA_est/(PLF_TF56$Y_P*PLF_TF56$YQ)
summary(PLF_TF56$RRP3_rev_VA_est)

summary(PLF_TF56$RRP3_rev_VA_est)

##using avaraged sollutions to F.O.C. with A as a numeraire (D3_AV)
all.equal(PLF_TF56$D3_AV , fitted(model.np.D3_AV.log)+resid(model.np.D3_AV.log),
check.attributes = FALSE)

PLF_TF56$D3_T1_AV_est <-((PLF_TF56$Vwp - fitted(model.np.D3_AV.log)*PLF_TF56$Awp)
  + (PLF_TF56$Awp - fitted(model.np.D3_AV.log)*PLF_TF56$Awp))/
  ((PLF_TF56$gV - fitted(model.np.D3_AV.log)*PLF_TF56$gA)
    + (PLF_TF56$gA - fitted(model.np.D3_AV.log)*PLF_TF56$gA))
summary(PLF_TF56$D3_T1_AV_est)


PLF_TF56$D3_T2_AV_est <-PLF_TF56$D3_T1_AV_est*(exp(fitted(model.np.f.log))/
    exp(fitted(model.np.g.log)))
summary(PLF_TF56$D3_T2_AV_est)
PLF_TF56$RP3_AV_est <- -0.5*(PLF_TF56$D3_T1_AV_est/exp(fitted(model.np.g.log)))*
  (((exp(fitted(model.np.f.log))^2)*PLF_TF56$omega1)+
      ((exp(fitted(model.np.g.log))^2)*PLF_TF56$omega2))
summary(PLF_TF56$RP3_AV_est)

PLF_TF56$RRP3_AV_est <- PLF_TF56$RP3_AV_est/(PLF_TF56$Y_P*PLF_TF56$YQ-
    (PLF_TF56$Aw*PLF_TF56$A+PLF_TF56$Vw*PLF_TF56$V))
summary(PLF_TF56$RRP3_AV_est)
```

```
#calculate "modified" RRP (RP3/total revenue) for Model 3
PLF_TF56$RRP3_rev_AV_est <- PLF_TF56$RP3_AV_est/(PLF_TF56$Y_P*PLF_TF56$YQ)
summary(PLF_TF56$RRP3_rev_AV_est)


##using sollution to F.O.C. with respect to variable inputs (V)
all.equal(PLF_TF56$D3_V , fitted(model.np.D3_V.log)+resid(model.np.D3_V.log),
check.attributes = FALSE)
PLF_TF56$D3_T1_V_est <- (PLF_TF56$Awp -fitted(model.np.D3_V.log)*PLF_TF56$Vwp)/
  (PLF_TF56$gA -fitted(model.np.D3_V.log)*PLF_TF56$gA)
summary(PLF_TF56$D3_T1_V_est)
PLF_TF56$D3_T2_V_est <-PLF_TF56$D3_T1_V_est*(exp(fitted(model.np.f.log))/
    exp(fitted(model.np.g.log)))

summary(PLF_TF56$D3_T2_V_est)
PLF_TF56$RP3_V_est <- -0.5*(PLF_TF56$D3_T1_V_est/exp(fitted(model.np.g.log)))*
  (((exp(fitted(model.np.f.log))^2)*PLF_TF56$omega1) +
      ((exp(fitted(model.np.g.log))^2)*PLF_TF56$omega2))
summary(PLF_TF56$RP3_V_est)

#calculate "modified" RRP (RP3/total revenue) for Model 3
PLF_TF56$RRP3_V_est <- PLF_TF56$RP3_V_est/ (PLF_TF56$Y_P*PLF_TF56$YQ-
    (PLF_TF56$Aw*PLF_TF56$A + PLF_TF56$Vw*PLF_TF56$V))
summary(PLF_TF56$RRP3_V_est)

PLF_TF56$RRP3_V_rev_est <- PLF_TF56$RP3_V_est/(PLF_TF56$Y_P*PLF_TF56$YQ)
summary(PLF_TF56$RRP3_V_rev_est)

##using sollution to F.O.C. with respect to land (A)
all.equal(PLF_TF56$D3_A , fitted(model.np.D3_A.log)+resid(model.np.D3_A.log),
check.attributes = FALSE)
PLF_TF56$D3_T1_A_est <-((PLF_TF56$Vwp - fitted(model.np.D3_A.log)*PLF_TF56$Awp)/
    (PLF_TF56$gV - fitted(model.np.D3_A.log)*PLF_TF56$gV))
summary(PLF_TF56$D3_T1_A_est)
PLF_TF56$D3_T2_A_est <-PLF_TF56$D3_T1_A_est*(exp(fitted(model.np.f.log))/
    exp(fitted(model.np.g.log)))

summary(PLF_TF56$D3_T2_A_est)
PLF_TF56$RP3_A_est <- -0.5*(PLF_TF56$D3_T1_A_est/ exp(fitted(model.np.g.log)))*
  (((exp(fitted(model.np.f.log))^2)*PLF_TF56$omega1) +
      ((exp(fitted(model.np.g.log))^2)*PLF_TF56$omega2))
summary(PLF_TF56$RP3_A_est)

PLF_TF56$RRP3_A_est <- PLF_TF56$RP3_A_est/ (PLF_TF56$Y_P*PLF_TF56$YQ -
    (PLF_TF56$Aw*PLF_TF56$A+PLF_TF56$Vw*PLF_TF56$V))
summary(PLF_TF56$RRP3_A_est)
```

```
#calculate "modified" RRP (RP3/total revenue) for Model 3
PLF_TF56$RRP3_A_rev_est <- PLF_TF56$RP3_A_est/(PLF_TF56$Y_P*PLF_TF56$YQ)
summary(PLF_TF56$RRP3_A_rev_est)
```

E.3. Graphs illustrating the bias of the model in logs

```
library( "miscTools" )
compPlot(model.np.f.log$resid, exp(model.np.f.log$resid) - 1,
         xlab = (expression(u^paste(symbol("\052")))),
         ylab = (expression(exp(u^paste(symbol("\052"))) - 1)),
         pch = 1, cex = 0.5, cex.axis = 1, cex.lab = 1.5 )

hist(exp(model.np.f.log$resid) - 1, 50, freq= TRUE, main = "",
     xlab = (expression(exp(u^paste(symbol("\052"))) - 1)), ylab = "Frequency",
     cex.axis = 1, cex.lab = 1.5)

hist(model.np.f.log$resid, 50, freq= TRUE, main = "",
     xlab = (expression(u^paste(symbol("\052")))),
     ylab =("Frequency") ,
     cex.axis = 1, cex.lab = 1.5)

diff <- (model.np.f.log$resid - (exp(model.np.f.log$resid) - 1))
hist(diff, 500, freq= TRUE, main = "", xlab = (expression(nu - upsilon)))

hist(model.np.f.log$resid - (exp(model.np.f.log$resid) - 1), 500,
     freq= TRUE, main = "",
     xlab = expression(u^paste(symbol("\052")) - (exp(u^paste(symbol("\052"))) - 1)),
     cex.axis = 1, cex.lab = 1.5)

hist(model.np.f.log$resid - (exp(model.np.f.log$resid) - 1), 500,
     freq= TRUE, main = "",
     xlab = expression(u^paste(symbol("\052")) - (exp(u^paste(symbol("\052"))) - 1)),
     xlim = c( - 0.05, 0),
     cex.axis = 1, cex.lab = 1.5)
```

F. Derivation of $\tilde{\theta}_1$ and $\tilde{\theta}_2$ for Model 3

F.1. Derivation of $\tilde{\theta}_1$ and $\tilde{\theta}_2$ for Model 3 for J variable inputs

Following Kumbhakar and Tsionas (2009), we rewrite the FOC as

$$\delta_j \equiv \frac{f_j}{f_1} = \frac{\widetilde{w}_j - h_j\tilde{\theta}_1}{\widetilde{w}_1 - h_1\tilde{\theta}_1} \tag{35}$$

$$D_3 \equiv \frac{1}{J}(1 + \sum_{j=2}^{J}\delta_j) = \frac{1}{J}\sum_{j=1}^{J}\frac{f_j}{f_1} = \frac{1}{J}\sum_{j=1}^{J}\frac{\widetilde{w}_j - h_j\tilde{\theta}_1}{\widetilde{w}_1 - h_1\tilde{\theta}_1} = \varphi^{28} \tag{36}$$

$$J\varphi = \sum_{j=1}^{J}\frac{\widetilde{w}_j - h_j\tilde{\theta}_1}{\widetilde{w}_1 - h_1\tilde{\theta}_1} \tag{37}$$

$$J\varphi = \frac{\sum_{j=1}^{J}\widetilde{w}_j - h_j\tilde{\theta}_1}{\widetilde{w}_1 - h_1\tilde{\theta}_1} \tag{38}$$

$$J\varphi(\widetilde{w}_1 - h_1\tilde{\theta}_1) = \sum_{j=1}^{J}\widetilde{w}_j - h_j\tilde{\theta}_1 \tag{39}$$

$$J\varphi\widetilde{w}_1 - J\varphi h_1\tilde{\theta}_1 = \sum_{j=1}^{J}\widetilde{w}_j - h_j\tilde{\theta}_1 \tag{40}$$

$$\sum_{j=1}^{J}h_j\tilde{\theta}_1 - J\varphi h_1\tilde{\theta}_1 = \sum_{j=1}^{J}\widetilde{w}_j - J\varphi\widetilde{w}_1 \tag{41}$$

$$\tilde{\theta}_1(\sum_{j=1}^{J}h_j - J\varphi h_1) = \sum_{j=1}^{J}\widetilde{w}_j - J\varphi\widetilde{w}_1 \tag{42}$$

$$\tilde{\theta}_1 = \frac{\sum_{j=1}^{J}\widetilde{w}_j - J\varphi\widetilde{w}_1}{\sum_{j=1}^{J}h_j - J\varphi h_1} = \frac{\sum_{j=1}^{J}[\widetilde{w}_j - \varphi\widetilde{w}_1]}{\sum_{j=1}^{J}[h_j - \varphi h_1]} \tag{43}$$

For J=2:

$$\tilde{\theta}_1 = \frac{\sum_{j=2}^{J}\widetilde{w}_j - 2\varphi\widetilde{w}_1}{\sum_{j=2}^{J}h_j - J\varphi h_1} = \frac{\widetilde{w}_1 + \widetilde{w}_2 - 2\varphi\widetilde{w}_1}{h_1 + h_2 - 2\varphi h_1} = \frac{\widetilde{w}_1(1-2\varphi) + \widetilde{w}_2}{h_1(1-2\varphi) + h_2} \tag{44}$$

$$\tilde{\theta}_1 = \frac{\sum_{j=1}^{J}[\widetilde{w}_j - \varphi\widetilde{w}_1]}{\sum_{j=1}^{J}[h_j - \varphi h_1]} = \frac{\widetilde{w}_1 - \varphi\widetilde{w}_1 + \widetilde{w}_2 - \varphi\widetilde{w}_1}{h_1 - \varphi h_1 + h_2 - \varphi h_1} = \frac{\widetilde{w}_2 + \widetilde{w}_1(1-2\varphi)}{h_2 + h_1(1-2\varphi)} = \frac{\widetilde{w}_2 - \delta_2\widetilde{w}_1}{h_2 - \delta_2 h_1} \tag{45}$$

$$\varphi = \frac{1}{J}(1 + \sum_{j=2}^{J}\delta_j) = \frac{1}{2}(1 + \delta_2) \tag{46}$$

[28] There is probably a typo in Kumbhakar and Tsionas (2009), where the first part of equation is given by: $D_3 \equiv 1 + \sum_{j=2}^{J}\delta_j = \frac{1}{J}\sum_{j=1}^{J}\frac{f_j}{f_1}$. However, this equality does not hold.

where,

$$\delta_2 = \frac{f_2}{f_1} \tag{47}$$

.

Then:

$$\frac{\widetilde{w}_2 + \widetilde{w}_1(1 - 2 \cdot \frac{1}{2}(1+\delta_2))}{h_2 + h_1(1 - 2 \cdot \frac{1}{2}(1+\delta_2))} = \frac{\widetilde{w}_2 + \widetilde{w}_1(1 - 1 - \delta_2)}{h_2 + h_1(1 - 1 - \delta_2)} = \frac{\widetilde{w}_2 - \delta_2 \widetilde{w}_1}{h_2 + \delta_2 h_1} \tag{48}$$

F.2. Derivation of $\tilde{\theta}_1$ and $\tilde{\theta}_2$ for Model 3 for 2 variable inputs

FOC of 2 variable inputs case can be written as:

$$f_1(x,z)(1+\tilde{\theta}_2) = \widetilde{w}_1 - h_1\tilde{\theta}_1 \tag{49}$$

$$f_2(x,z)(1+\tilde{\theta}_2) = \widetilde{w}_2 - h_2\tilde{\theta}_1 \tag{50}$$

We can solve the system of FOCs in the following way:

$$\frac{f_2}{f_1} = \frac{\widetilde{w}_2 - h_2\tilde{\theta}_1}{\widetilde{w}_1 - h_1\tilde{\theta}_1} \tag{51}$$

Substituting $\frac{f_2}{f_1}$ with δ_2

$$\delta_2 = \frac{\widetilde{w}_2 - h_2\tilde{\theta}_1}{\widetilde{w}_1 - h_1\tilde{\theta}_1} \tag{52}$$

And solving (35) for $\tilde{\theta}_1$:

$$\delta_2(\widetilde{w}_1 - h_1\tilde{\theta}_1) = \widetilde{w}_2 - h_2\tilde{\theta}_1 \tag{53}$$

$$\delta_2(\widetilde{w}_1 - h_1\tilde{\theta}_1) + h_2\tilde{\theta}_1 = \widetilde{w}_2 \tag{54}$$

$$\delta_2\widetilde{w}_1 - \delta_2 h\tilde{\theta}_1 + h_2\tilde{\theta}_1 = \widetilde{w}_2 \tag{55}$$

$$(h_2 - \delta_2 h_1)\tilde{\theta}_1 = \widetilde{w}_2 - \delta_2\widetilde{w}_1 \tag{56}$$

$$\tilde{\theta}_1 = \frac{\widetilde{w}_2 - \delta_2\widetilde{w}}{h_2 - \delta_2 h_1} \tag{57}$$

References

Antle, J. M. (1983). Incorporating risk in production analysis. *American Journal of Agricultural Economics* 65: 1099–1106.

Antle, J. M. (1987). Econometric estimation of producers' risk attitudes. *American Journal of Agricultural Economics* 69: 509–522.

Antle, J. M. (2010). Do economic variables follow scale or location-scale distributions? *Ame* 92: 196–204.

Appelbaum, E. and Ullah, A. (1997). Estimation of moments and production decisions under uncertainty. *Review of Economics and Statistics* 79: 631–637.

Asche, F. and Tveterås, R. (1999). Modeling production risk with a two-step procedure. *JournalofAgricultural and Resource Economics* 24: 424–439.

Chambers, R. G. (1983). Scale and productivity measurement under risk. *The American Economic Review* 73: 802–805.

Chambers, R. G. and Quiggin, J. (2000). *Uncertainty, Production, Choice, and Agency: the State-Contingent Approach*. Cambridge University Press, Cambridge.

Chavas, J.-P. and Holt, M. T. (1996). Economic behavior under uncertainty: A joint analysis of risk preferences and technology. *The Review of Economics and Statistics* 78: 329–335.

Croissant, Y. and Millo, G. (2008). Panel data econometrics in R: The plm package. *Journal of Statistical Software* 27: 1–43.

Czekaj, T. and Henningsen, A. (2012). Comparing Parametric and Nonparametric Regression Methods for Panel Data: the Optimal Size of Polish Crop Farms. FOI Working Paper 2012/12, Institute of Food and Resource Economics, University of Copenhagen.

Czekaj, T. and Henningsen, A. (2013). Panel Data Specifications in Nonparametric Kernel Regression: An Application to Production Functions. IFRO Working Paper 2013/5, Department of Food and Resource Economics, University of Copenhagen.

Gardebroek, C., Chavez, M. D. and Lansink, A. O. (2010). Analysing production technology and risk in organic and conventional dutch arable farming using panel data. *Journal of Agricultural Economics* 61: 60–75.

Griffiths, W. E. and Anderson, J. R. (1982). Using time-series and cross-section data to estimate a production function with positive and negative marginal risks. *Journal of the American Statistical Association* 77: 529–536.

GUS (2012a). Bank Danych Lokalnych / Local Data Bank. Główny Urząd Statystyczny / Central Statistical Office of Poland, http://www.stat.gov.pl/bdlen/ [accessed 09-October-2012].

GUS (2012b). Prices in the national economy - january 2012. Główny Urząd Statystyczny / Central Statistical Office of Poland, http://www.stat.gov.pl/cps/rde/xbcr/gus/PUBL_pt_prices_national_economy_01m_2012.zip [accessed 05-March-2012].

Hayfield, T. and Racine, J. S. (2008). Nonparametric econometrics: The np package. *Journal of Statistical Software* 27: 1–32.

Henderson, D. J. and Simar, L. (2005). A Fully Nonparametric Stochastic Frontier Model for Panel Data. Working Paper 0519, Departmentment of Economics, State University of New York at Binghamton.

Just, R. E. (2003). Risk research in agricultural economics: Opportunities and challenges for the next twenty-five years. *Agricultural Systems* 75: 123–159.

Just, R. E. and Pope, R. D. (1978). Stochastic specification of production functions and economic implications. *Journal of Econometrics* 7: 67–86.

Kumbhakar, S. C. (1993). Production risk, technical efficiency, and panel data. *Economics Letters* 41: 11–16.

Kumbhakar, S. C. (2001). Risk preferences under price uncertainties and production risk. *Communications in Statistics-Theory and Methods* 30: 1715–1735.

Kumbhakar, S. C. (2002a). Risk preference and productivity measurement under output price uncertainty. *Empirical Economics* 27: 461–472.

Kumbhakar, S. C. (2002b). Risk preferences and technology: A joint analysis. *Marine Resource Economics* 17: 77–89.

Kumbhakar, S. C. and Tsionas, E. G. (2009). Nonparametric estimation of production risk and risk preference functions. In Li, Q. and Racine, J. S. (eds), *Nonparametric Econometric Methods*. Emerald Group Publishing Limited, Advances in Econometrics *25*, 335–375.

Kumbhakar, S. C. and Tsionas, E. G. (2010). Estimation of production risk and risk preference function: a nonparametric approach. *Annals of Operations Research* 176: 369–378.

Kumbhakar, S. C. and Tveterås, R. (2003). Risk preferences, production risk and firm heterogeneity. *Scandinavian Journal of Economics* 105: 275–293.

Li, Q. and Racine, J. S. (2004). Cross-validated local linear nonparametric regression. *Statistica Sinica* 14: 485–512.

Moschini, G. and Hennessy, D. A. (2001). Chapter 2 uncertainty, risk aversion, and risk management for agricultural producers. In Gardner, B. L. and Rausser, G. C. (eds), *Agricultural Production*. Elsevier, Handbook of Agricultural Economics *1, Part A*, 88–153.

Parmeter, C. F., Sun, K., Henderson, D. J. and Kumbhakar, S. C. (2012). Regression and inference under smoothness restrictions.

R Development Core Team (2012). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0.

Racine, J. S. (1997). Consistent significance testing for nonparametric regression. *Journal of Business and Economic Statistics* 15: 369–379.

Racine, J. S. (2008). Nonparametric econometrics: A primer. *Foundations and Trends in Econometrics* 3: 1–88.

Racine, J. S., Hart, J. and Li, Q. (2006). Testing the significance of categorical predictor variables in nonparametric regression models. *Econometric Reviews* 25: 523–544.

Racine, J. S. and Li, Q. (2004). Nonparametric estimation of regression functions with both categorical and continuous data. *Journal of Econometrics* 119: 99–130.

Rasmussen, S. (2004). Optimizing Production under Uncertainty: Generalisation of the State-Contingent Approach and Comparison of Methods for Empirical Application. Tech. rep., The Royal Veterinary and Agricultural University, Food and Resource Economic Institute.

Sandmo, A. (1971). On the theory of the competitive firm under price uncertainty. *American Economic Review* 61: 65–73.

Shankar, S. (2012). Production economics in the presence of risk. *The Australian Journal of Agricultural and Resource Economics* 56: 597–620.

Tveterås, R. (1999). Production risk and productivity growth: Some findings for norwegian salmon aquaculture. *Journal of Productivity Analysis* 12: 161–179.

Tveterås, R., Flaten, O. and Lien, G. (2011). Production risk in multi-output industries: Estimates from norwegian dairy farms. *Applied Economics* 43: 4403–4414.

Zellner, A., Kmenta, J. and Dreze, J. (1966). Specification and estimation of cobb-douglas production function models. *Econometrica* 34: 784–795.