Panel Data Estimation of Labour Supply Equations for Dutch Dairy Farmers

D.L. (Daan) Ooms¹ and A.R. (Alastair) Hall²

¹ Agricultural Economics and Rural Policy, Social Sciences Group, Wageningen University, Wageningen, The Netherlands
² Economics Department, North Carolina State University, Raleigh NC, USA


February 2-5, 2005
Parma, Italy

Copyright 2005 by D.L. (Daan) Ooms and A.R. (Alastair) Hall. All rights reserved. Readers may make verbatim copies of this document for non-commercial purposes by any means, provided that this copyright notice appears on all such copies.
Panel Data Estimation of Labour Supply Equations for Dutch Dairy Farmers

D.L. (Daan) Ooms* and A.R. (Alastair) Hall **

Abstract

This research focuses on the estimation of labour supply equations for Dutch dairy farmers. Off-farm labour supply is characterised by the fact that only half of the farm households supply off-farm labour. This results in a sample selection estimation problem that has to be taken into account in estimation. In this research we use the panel data sample selection estimation method of Wooldridge (1995) to estimate the off-farm labour supply equation. Simulation results suggest that the decoupledness assumption of direct income payment in CAP reform might hold. At least there is a negligible effect through labour supply.

Introduction

This research focuses on the estimation of on- and off-farm labour supply equations for Dutch dairy farmers. Off-farm labour supply of Dutch dairy farmers is characterised by the fact that only half of the farm households supply off-farm labour. We assume that the farmers that supply off-farm labour is not a representative sample from the Dutch dairy farmers’ population. This results in a sample selection estimation problem that has to be taken into account in estimation. The Dutch Agricultural Research Institute (LEI) collects data for Dutch farm households. Every year a group of farm households leaves the data set and a new group enters. The farm households are in the data set for a number of subsequent years. This results in a rotating panel data set. A panel data set gives information on changes over time within the households in the panel. Panel data estimation methods use this information in estimation. In this research we use the panel data sample selection estimation method of Wooldridge (1995) to estimate the off-farm labour supply equation. This estimation method is based on the fixed effects panel data estimation method for linear models of Mundlak (1978). We will use this method to estimate the linear on-farm labour supply equation. On June 26 2003 the European ministers of agriculture agreed on a reform of the Common Agricultural Policy (CAP) (European Commission, 2003). The reform entails a reduction in intervention prices for skimmed milk powder of 15% (in three yearly steps of 5% from 2004 to 2006) and 25% for butter (three yearly steps

* Agricultural Economics and Rural Policy, Social Sciences Group, Wageningen University, Wageningen, The Netherlands.
** Economics Department, North Carolina State University, Raleigh NC, USA.
of 7% from 2004 to 2006 and 4% in 2007). Moreover, a prolongation of the milk quota system until 2014/15 is agreed in combination with a milk quota increase (in the Netherlands) of 1.5% in three yearly steps of 0.5% starting in 2006. There will be no further quota increases, but a market report to look at further quota increases will be presented once the reform is fully implemented. As a further precaution against a possible surge in butter intervention post EU enlargement, a new butter intervention ceiling was set at 70,000t in 2004/05, with an annual cut of 10,000t until it reaches 30,000t in 2008/09. Finally, a decoupled direct income payment of €35,50 per tonne is proposed. In the Netherlands there is a strong interest in the possible effects of dairy policy reform given the economic importance of the dairy industry. The effects of this policy change is simulated with the estimated on- and off-farm labour supply equations.

The remainder of this paper contains the following. First, a theoretical derivation of on- and off-farm labour supply equations is given. After that we describe a general panel data sample selection model and explain the estimation method for it introduced by Wooldridge (1995). The LEI data set does not contain all variables needed based on the theoretical derivation. We describe how we deal with these data limitations. After that we give a description of the data we do have at our disposal. Next we describe the empirical analyses and policy simulations based on the estimated on- and off-farm labour supply equations. Finally, we give a brief summary and some conclusions.

**Theoretical model**

The following theoretical model is based on the household utility model of Huffman (1980). Labour supply decisions of dairy farm household \( i \) at time \( t \) are assumed to be the result of maximising utility \( (u_i) \) received from a vector of consumption goods and services \( (c_{i,t}) \) and home time \( (t_{h,i}) \) given a vector of utility shifting household characteristics \( (z_{h,i}) \) and a vector of other variables influencing the households’ decision making environment \( (o_i) \),

\[
u_i = u(c_{i,t}, t_{h,i}, z_{h,i}, o_i)
\]

(1)

where \( u(\cdot) \) is a utility function that is the same for all households. Differences between the utility levels of households come from the different choices made with respect to the elements of the utility function. Total time endowment \( (t^o_i) \) is allocated between farm labour \( (t_{f,i}) \), off-farm labour \( (t_{of,i}) \) and home time. Which results in the time constraint:

\[
t^o_i = t_{f,i} + t_{of,i} + t_{h,i}, \quad t_{of,i} \geq 0
\]

(2)

The time constraint is a strict equality because home time is defined to be the difference between total time and labour time. Home time consists of leisure, household work, etc. A non-negativity constraint is imposed on off-farm labour because it may be zero. The value of household consumption is defined as the product of consumption goods and services with the
price of consumption goods and services \( (p_c) \). We assume prices to be the same for all households and only differ between time periods. Consumption is constrained by total income. Total income consists of income from the farm \( (y_{f,t}) \), off-farm labour income \( (y_{of,t}) \) and other, non-farm non-labour income \( (y_{n,t}) \).

\[
y_{f,t} + y_{of,t} + y_{n,t} = p_c c_t
\]  

(3)

Farm income is defined as production \( (q_a) \) times output price \( (p) \) minus variable input \( (g_a) \) times variable input price \( (v) \) minus paid fixed costs associated with production \( (k) \):

\[
y_{f,t} = p q_a - v g_a - k
\]  

(4)

Off-farm labour income is defined as off-farm labour time times off-farm wage \( (w_{of,t}) \):

\[
y_{of,t} = w_{of,t} t_{of,t}
\]  

(5)

Notice that we assume wages to be farm and time specific. Farm output \( (q_a) \) is produced using household labour, variable inputs and fixed inputs \( (z) \):

\[
q_a = f(q_a, g_a; z)
\]  

(6)

\( f( ) \) is a vector production function. Combining equations (3) to (6) results in:

\[
p_i q_i f(q_a, g_a; z) - v g_a - k_s + w_{of,t} t_{of,t} + y_{n,t} = p_c c_t
\]  

(7)

We assume the household maximises (1) subject to (7) and (2) by choosing the elements of the choice set \( c_t, t_{of,t} \). The Kuhn-Tucker first-order conditions are:

\[
\frac{\partial u}{\partial c_t} = \lambda_1 p_c,
\]  

(8)

\[
\frac{\partial u}{\partial t_{of,t}} = \lambda_2,
\]  

(9)

\[
\lambda_1 p_i \frac{\partial q_i}{\partial t_{of,t}} - \lambda_2 = 0,
\]  

(10)

\[
\lambda_1 w_{of,t} - \lambda_2 \leq 0,
\]  

(11)
\[ \lambda \left( \frac{\partial q_i}{\partial g_i} p_i - v_i \right) = 0 \]  

(12)

plus equations (7) and (2) where \( \lambda_i \) is the marginal utility of income and \( \lambda_z \) is the marginal utility of time. If an interior solution exists (i.e. off-farm labour supply is non-zero) the first part of equation (11) holds as equality. This together with equation (10) can be solved to yield:

\[ \frac{\partial u_i}{\partial c_{i,a}} = -\frac{P_{i,a}}{w_{i,a}} \]  

(13)

and

\[ \frac{\partial q_i}{\partial g_i} p_i' = w_{i,a} \]  

(14)

Equation (13) implies that the marginal rate of substitution between leisure and consumption goods is equal to the ratio of the consumption good price to the wage rate. Equation (14) implies that the marginal product of farm labour is equal to off-farm wage. A corner solution in this research occurs if off-farm labour supply is zero. Because for this research there is a particular interest in labour decisions, we apply the solution approach for the system of Kuhn-Tucker conditions of Kimhi and Lee (1996). They carried out two steps. First, they expressed the two endogenous labour time variables \( (t_{i,a} \) and \( t_{f,a} \)) as a function of the other endogenous variables (in our case: \( c_{i,a}, t_{i,a}, g_i \) and the Lagrange multipliers). Second, they inserted the reduced-form equations of the other endogenous variables into the structural endogenous labour supply equations. We assume that Dutch dairy farmers make their labour decisions recursively by first choosing the amount of labour needed on the farm and dividing remaining time between off-farm labour and leisure. Therefore we add a third step to the two steps of Kimhi and Lee (1996) in which the reduced form equation for off-farm labour supply is inserted in the on-farm labour supply equation. This results in the following recursive system of supply equations for off-farm labour \( s_{i,a} \) and on-farm labour \( s_{f,a} \):

\[ t_{i,a} = s_{i,a} \left( t_{i,a}, w_{i,a}, P_{i,a}, p_i, v_i, k_i, y_{i,a}, z_{i,a}, a_i, t_{i,a}^0 \right) \]  

(15)

\[ t_{f,a} = s_{f,a} \left( w_{i,a}, P_{i,a}, p_i, v_i, k_i, y_{i,a}, z_{i,a}, a_i, t_{f,a}^0 \right) \]  

(16)

\( s_{i,a} ( \cdot ) \) and \( s_{f,a} ( \cdot ) \) are the off-farm and farm labour supply functions, respectively. Not all farmers in western agriculture supply off-farm labour. It is assumed that the group of farmers that do supply off-farm labour is not a representative sample of all farmers. This, together with the availability of panel data, calls for panel data sample selection estimation methods.
Panel data sample selection

This section gives, first, a description of a general panel data sample selection model. After that the panel data sample selection estimation methods of Wooldridge (1995) is described. This method is a fixed effects extension of the sample selection estimation method for cross-section data introduced by Heckman (1979). A simple extension in which a fixed effect estimator is used for both steps in the Heckman approach does not produce consistent estimates (Wooldridge (2002) pp. 582-583).

A panel data sample selection model

The following gives a description of a general panel data sample selection model that is useful for the explanation of panel data sample selection methods and for the empirical application in a later section of this paper. Consider the model:

\[ y_{i} = d_{i} \cdot y_{i}^{*} = d_{i} \cdot (x_{i}' \beta + \alpha_{i} + \varepsilon_{i}) = x_{i}' \beta + \alpha_{i} + \varepsilon_{i} \quad (17) \]

\[ d_{i} = 1 \{ w_{i} \gamma + \eta_{i} - u_{i} \geq 0 \} \quad (18) \]

Here, \( 1 \{ \} \) is an indicator function that has the value 1 if the comparison between brackets is true and 0 otherwise. \( \beta \) and \( \gamma \) are parameter vectors to be estimated. The latent \( y_{i}^{*} \) is the dependent variable of interest that is only observed if \( d_{i} = 1 \). The explanatory variables in equation (17), \( x_{i}' \), might be observed for both values of \( d_{i} \). However, from the two estimation methods, described below, it becomes clear that it is not restrictive to equate \( x_{i} \) to 0 if \( d_{i} = 0 \). The explanatory variables in equation (18), \( w_{i}' \), are assumed to be observed for both values of \( d_{i} \). \( \alpha_{i} \) and \( \eta_{i} \) are unobservable farm specific effects and \( \varepsilon_{i} \) and \( u_{i} \) unobservable error terms. From equation (17) follows that \( \alpha_{i} \) and \( \varepsilon_{i} \) are 0 if \( d_{i} = 0 \). Notice that the unobserved effect \( \alpha_{i} \) in (17) depends on \( d_{i} \) and that it has a subscript \( i \) for this reason. For notational convenience, a vector \( \xi_{i} = (w_{i}, x_{i}', \alpha_{i}, \eta_{i}) \) containing all observed and unobserved explanatory variables of period \( t \) and \( s \) for individual \( i \). From equation (17) and (18) two problems for estimation arise. First, the presence of the unobservable effect \( \alpha_{i} = d_{i} \cdot \alpha_{i}' \). And, second, the potential sample selection bias caused by the possibility that \( d_{i} \) does not select a representative sample from the underlying population. Panel data sample selection estimation methods have to deal with these two problems.
Wooldridge (1995)

The panel data sample selection estimation method considered in this research is based on Wooldridge (1995). This paragraph provides a short description of this estimator. We mentioned before that a simple extension of the Heckman approach to a fixed effect panel data situation does not work. Wooldridge (1995) shows that an extension based the panel data approach of Chamberlain (1980) does work. He also suggests the use of simplifications of Chamberlain’s approach. One of these simplifications is the approach of Mundlak (1978). Based on this approach and the fact that in a Heckman sample selection approach only positive observations are used in estimation of the structural equation, equation (17) and (18) change to become:

\[ y_{it} = x_{it} \beta + \gamma T + \delta \lambda_{it} + \nu_{it}, \quad \text{for all } i \text{ and } t \text{ for which } d_{it} = 1 \]  
\[ d_{it} = 1 \left\{ \frac{w_{it} \gamma + \mu - u_{it}}{w_{it} \gamma + \mu} \geq 0 \right\} \]  

Here, \( \gamma T \) and \( \mu \) are parameter vectors to be estimated. \( \delta \) is a parameter to be estimated associated with the estimated Mills ratio:

\[ \delta = \frac{\phi \left( w_{it} \gamma + \mu \right)}{\Phi \left( w_{it} \gamma + \mu \right)} \]  

Here, \( \phi(\cdot) \) and \( \Phi(\cdot) \) are the density of the standard normal distribution and its cumulative density, respectively. The panel data sample selection estimation method based on Wooldridge (1995) consists of four steps:

1. Estimate equation (24) with pooled probit\(^2\).
2. Create the estimated Mills ratio using equation (25) for all \( i \) and \( t \) for which \( d_{it} = 1 \).
3. Estimate equation (23) with pooled OLS\(^3\) for all \( i \) and \( t \) for which \( d_{it} = 1 \).
4. Correct the estimated variance of the estimated parameters in equation (23) for individual heteroscedasticity.

---

1 The interested reader is referred to Wooldridge (1995) and Wooldridge (2002, pp. 581-583) for a more elaborate description of this estimator.
2 Pooled probit is probit in which all observations are treated as coming from one cross section.
3 Pooled OLS is OLS in which all observations are treated as coming from one cross section.
The last step is necessary because pooled OLS does not explicitly take the panel structure of the data into account. This does not influence the parameter estimates in the Mundlak (1978) approach, but it does influence the estimated variance of the estimated parameters. Appendix A gives the formulas for the variance correction.

**Data limitations**

Equation (17) and (18) show that labour supply is expressed in time units and is, amongst others, explained by individual wages. The data set of Dutch dairy farmers at our disposal contains on-farm hours. However, it does not contain off-farm hours and individual wages. This section explains how we deal with these limitations.

**Off-farm labour supply**

In stead of off-farm hours, our data set contains off-farm income. This paragraph explains how it is possible to estimate the parameters in the off-farm labour supply equation, using off-farm income in stead of hours. To this end we write equation (17) as:

\[ t_{of, \alpha} = \exp\left( \beta_1 + \beta_\alpha \ln w_{of, \alpha} + x_{-of, \alpha} \beta_{-x} + \epsilon_\alpha \right) \]  \hspace{1cm} (26)

Here, \( w_{of, \alpha} \) represents the wage and \( x_{-of, \alpha} \) is the vector of all explanatory variables except off-farm wage. \( \epsilon_\alpha \) is an error term with expectation zero. \( \beta_1 \) is a farm specific effect, \( \beta_{-x} \) is the parameter to associated with the log of wage and the vector \( \beta_{-x} \) contains parameters associated with the other explanatory variables. For notational convenience, the content of \( x_{-of, \alpha} \) is not specified in this section.

Off-farm income \( y_{of, \alpha} \) is by definition:

\[ y_{of, \alpha} = t_{of, \alpha} w_{of, \alpha} \]  \hspace{1cm} (27)

Taking the natural logarithm of (26) and (27) gives:

\[ \ln t_{of, \alpha} = \beta_1 + \beta_\alpha \ln w_{of, \alpha} + x_{-of, \alpha} \beta_{-x} + \epsilon_\alpha \]  \hspace{1cm} (28)
\[ \ln y_{of, \alpha} = \ln t_{of, \alpha} + \ln w_{of, \alpha} \]  \hspace{1cm} (29)

Replacing \( \ln t_{of, \alpha} \) by its equation obtained from rewriting (29) results in:

\[ \ln y_{of, \alpha} - \ln w_{of, \alpha} = \beta_1 + \beta_\alpha \ln w_{of, \alpha} + x_{-of, \alpha} \beta_{-x} + \epsilon_\alpha \]  \hspace{1cm} (30)

and
\[ \ln y_{off} = \beta_i + \beta_r (\alpha + 1) \ln w_{off} + x_{off} \beta_c + \epsilon_o \]  

This derivation shows the possibility to estimate the parameters of equation (26), in which labour time is the dependent variable, using the logarithm of labour income as dependent variable in stead. The only thing that has to be taken into account is that the estimated parameter on \( \ln w_{off} \) is not \( \beta_r \) but \( (\beta_r + 1) \).

**Off-farm wage**

In stead of farm specific off-farm wages, our data set contains national wages for labourers in the agricultural sector \( w_{i} \). Data on these wages differ between periods, not between farms. We assume that this variable is a good approximation for the off-farm wage rate for farmers in The Netherlands\(^4\). Using this national wage rate for individual farmers implies a measurement error \( (\xi_e) \). We assume:

\[ w_{off} = w_i \xi_e \quad \text{with } \xi_e > 0 \quad \text{and } \mathbb{E}(\xi_e) = 1 \]  

Inserting (33) in (31) gives:

\[ \ln y_{off} = \beta_i + (\beta_r + 1)(\ln w_i + \ln \xi_e) + x_{off} \beta_c + \epsilon_o \]  

Since \( \mathbb{E}(\xi_e) = 1 \), \( \mathbb{E}(\ln \xi_e) = 0 \) and (33) can be rewritten to become:

\[ \ln y_{off} = \beta_i + (\beta_r + 1)\ln w_i + x_{off} \beta_c + \tau_o \]  

where \( \tau_o = (\beta_r + 1)\ln \xi_e + \epsilon_o \) has expectation zero. This derivation shows that the measurement error caused by using \( \ln w_i \) in stead of \( \ln w_{off} \) in equation (31) disappears in the new error term \( \tau_o \) in equation (34). It has to be noted that the measurement error can influence parameter estimates of variables that explain the difference in individual wages like e.g. education. These variables possibly explain some of the measurement error. This will be taken into account while explaining estimation results later on. Parameter estimates for variables that are not correlated with off-farm wage are not influenced by the measurement error.

**Data**

This section gives a description of the data used in estimation. The farm specific data come from the Dutch Agricultural Research Institute (LEI) unbalanced rotating panel data set of Dutch farms. A farm is classified to be a dairy farm if its returns consist for 50% or more of

\(^4\) Compare with Woldehanna et al. (2000) in which the same assumption is made in a model for off-farm labour supply of Dutch arable farmers.
milk revenues. The data set consists of 6338 observations on 1307 farms. The period investigated is from 1987/88 until 1999/00. National data come from Statistics Netherlands (CBS). Off-farm labour is represented by off-farm income. The total number of family hours worked on the farm represents on-farm labour. Off-farm wage is represented by the national index of wages for agricultural hired labour. 1991 is the base year for this and subsequent indices. Price variables, influencing farm income, are the milk price index; a Thonqvist price index for other output and a Thonqvist index for non-factor input. Other output contains marketable crops, veal, pigs, poultry and other farm revenues. Non-factor input contains, amongst others, feed and veterinary costs. Paid factor costs contain paid rent and other costs of financing. Non-farm non-labour income is a monetary value. It includes, amongst other, income from externally allocated capital and income from social allowances. Land is expressed in the number of hectares used by the farmer. Machinery is the average value of machinery over the year divided by the Thonqvist price index for machinery. Debt is the total value of short and long-term debt. Assets are represented by their value calculated by the LEI. Unemployment is expressed as the national unemployment rate. From the LEI data set it is possible to derive the importance of different activities as a percentage of total activity. This is based on output and the allocation of resources. The percentage for milk production is used as the specialisation rate of dairy farmers. Household variables used are number of household members; a dummy for the presence of a successor and a dummy variable indication the education level of the head of the household. See Appendix B for summary statistics of the data.

**Empirical analysis**

This section gives estimation results for the off-farm labour supply equation and the on-farm labour supply equation. The estimation method for the off-farm labour decision and off-farm labour supply equation is described above in the panel data sample selection section. Since all farm households supply labour to the farm, the on-farm labour supply equation (16) can be estimated by linear panel data estimation methods. Equation (15) and (16) constitute a recursive system of equations. We choose to estimate equation (16) in a way that is based on the same idea as the way equation (15) is estimated. The Wooldridge (1995) approach we use is based on pooled OLS where the farm specific unobserved effect is approximated in the way of Mundlak (1978). We also estimate equation (16) with pooled OLS where the farm specific effect is approximated in the way of Mundlak (1978). The estimated variance of the parameter estimates is adjusted for farm specific heterogeneity in the same way as is done in the Wooldridge (1995) estimation method. An advantage of the munlak approach is that it allows for time invariant explanatory variables. This is not the case in the more well known fixed effects estimation methods.

We perform a Hausman test\(^5\) to test our assumption of exogeneity of on-farm labour supply in the off-farm labour supply equation. The Hausman test requires that we have at least one explanatory variable in the on-farm labour supply equation that does not appear in the off-

\(^5\) See e.g. Wooldridge (2002, p.119) for a description of this test.
farm labour supply equation. We choose to use the on-farm specialisation rate for milk production as the variable that does appear in the on-farm labour supply and the off-farm labour decision equation, but not in the off-farm labour supply equation. The results of Weiss and Briglauer (2000) suggest that off-farm labour is a diversification choice in the reduction of risk. A high on-farm specialisation might increase the propensity to work off-farm to reduce risk. We assume that this effect works mainly through the diversification decision. Based on this assumption we only include on-farm specialisation in the off-farm labour decision equation and not in the off-farm labour supply equation. A negative effect of labour specialisation on on-farm labour use is an often-found result in economic research.

**Results**

Table 1 gives the estimation results for the off- and on-farm labour supply equations.

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Dependent variable</td>
<td>Log off-farm labour</td>
</tr>
<tr>
<td></td>
<td>estimate</td>
</tr>
<tr>
<td>Constant</td>
<td>96.19</td>
</tr>
<tr>
<td>Log on-farm labour</td>
<td>-0.92</td>
</tr>
<tr>
<td>Log off-farm wage</td>
<td>2.05</td>
</tr>
<tr>
<td>Log milk price</td>
<td>-0.01</td>
</tr>
<tr>
<td>Log other output price</td>
<td>-0.44</td>
</tr>
<tr>
<td>Log input price</td>
<td>2.27</td>
</tr>
<tr>
<td>Factor costs</td>
<td>-0.49</td>
</tr>
<tr>
<td>Non-farm non-labour income</td>
<td>-0.41</td>
</tr>
<tr>
<td>Land</td>
<td>1.11</td>
</tr>
<tr>
<td>Machinery</td>
<td>-1.36</td>
</tr>
<tr>
<td>Debt over asset ratio</td>
<td>0.19</td>
</tr>
<tr>
<td>Unemployment rate</td>
<td>-0.02</td>
</tr>
<tr>
<td>Specialisation</td>
<td></td>
</tr>
<tr>
<td>Number of household members</td>
<td>-0.00</td>
</tr>
<tr>
<td>Successor dummy</td>
<td>0.75</td>
</tr>
<tr>
<td>Education</td>
<td>-0.29</td>
</tr>
<tr>
<td>Mills ratio</td>
<td>2.98</td>
</tr>
<tr>
<td>Hausman test for exogeneity</td>
<td></td>
</tr>
</tbody>
</table>

* indicates significance at the 10% or smaller level.
Before describing the estimation results we focus on two assumptions made in the theoretical model section. First, we focus on the assumption that the group of farmers that supply off-farm labour is not a representative sample of all farmers. Based on this assumption we choose to use sample selection estimation methods. The Mills ratio used for sample selection correction is significant at the 10% level. Therefore, we conclude that our sample selection assumption is correct. Second, we focus on the assumption that Dutch dairy farmers make their labour decisions recursive by first choosing the amount of labour needed on the farm and dividing remaining time between off-farm labour and leisure. This assumption implies exogeneity of on-farm labour supply in the on-farm labour supply equation. The Hausman test statistic is not significant. Therefore, the exogeneity assumption of on-farm labour in the off-farm labour equation can not be rejected.

Now we know that two major assumptions we made are justified, we can focus on the parameter estimates. The parameter estimate for off-farm labour supply in the off-farm labour supply equation is negative and significant. Since both the on-and off-farm labour variables are expressed in logarithmic form the parameter estimate is an elasticity. The parameter estimate indicates that, ceteris paribus, 92% of decrease in on-farm labour is allocated to off-farm labour. Off-farm wage is significant and has an expected sign. The variables determining income from farming all have expected signs in the on-farm labour equation. However, only the parameter for factor costs is significant. The fact that changes in milk price do not have a significant impact on on-farm labour supply even though milk is the major output is not unexpected. Since milk is produced under a production quota, the shadow price for labour does not depend on output price changes. It can depend on input price changes through a substitution or complementarity effect. However, apparently this effect either does not exist or the inputs are both substitutes and complements to labour and the effect is cancelled in estimation. In the off-farm labour supply equation, the parameter for factor costs has an unexpected sign. However this parameter is not significant and, therefore, the sign does not matter. The parameter for non-farm non-labour income is negative in both equations. This reflects the effect of an increase in the consumption of leisure if income increases. Non-farm non-labour income has a larger effect on off-farm labour supply than on on-farm labour supply and its parameters are significant. Land has a significantly positive impact on on-farm labour supply. It has a significantly positive effect on off-farm labour supply. This indicates that larger farms supply more off-farm labour and is in correspondence with other estimation results for off-farm labour supply (see e.g. Goodwin and Holt (2002) and Ahituv and Kimhi (2002)). Also machinery has a significantly positive impact on on-farm labour supply. Interesting to see is that machinery has a significantly negative effect on off-farm labour supply. Apparently large farmers (based on the amount of land) that supply off-farm labour have a relatively small amount of machinery. This can be explained by the broad use of agricultural contract workers in The Netherlands that can be hired with machinery. Apparently large farmers that supply off-farm work, make more use of these contract workers. The debt over asset ratio has a significantly positive effect on on-farm labour indicating that farmers work harder on-farm when debt increases. However, this effect can not be found for off-farm labour supply. Apparently Dutch dairy farmers do no use off-farm income to relieve debt. The unemployment rate does not have a significant effect on labour supply. Specialisation in milk production has a negative effect on on-farm labour input,
as expected. However, this effect is not significantly different from zero. Number of household members, the existence of a successor and education level have a significantly positive effect on on-farm labour supply. For off-farm labour supply only the existence of a successor has an effect significantly different from zero. This effect is positive. This can be explained by the existence of Dutch dairy farms on which farm income is not enough for both the owner family and the successor (and his family).

Appendix C gives the parameter estimates for the off-farm labour decision equation for the pooled probit estimator used in the Wooldridge (1995) estimation. These results are not of primary interest in this paper. Therefore, we do not describe them extensively. The overall impression is that there are not many significant parameters. None of the parameters associated with time varying variables are significantly different from zero. On the other hand, all parameters associated with time invariant variables are significant. From these estimations and the estimations for the labour supply equations, follows that these time invariant household variables are important for labour supply.

Policy Simulations

On June 26 2003 the European ministers of agriculture agreed on a reform of the Common Agricultural Policy (CAP) (European Commission, 2003). The reform entails a reduction in intervention prices for skimmed milk powder of 15% (in three yearly steps of 5% from 2004 to 2006) and 25% for butter (three yearly steps of 7% from 2004 to 2006 and 4% in 2007). Moreover, a prolongation of the milk quota system until 2014/15 is agreed in combination with a milk quota increase (in the Netherlands) of 1.5% in three yearly steps of 0.5% starting in 2006. There will be no further quota increases, but a market report to look at further quota increases will be presented once the reform is fully implemented. As a further precaution against a possible surge in butter intervention post EU enlargement, a new butter intervention ceiling was set at 70,000t in 2004/05, with an annual cut of 10,000t until it reaches 30,000t in 2008/09. Finally, a decoupled direct income payment of €35.50 per tonne is proposed. In the Netherlands there is a strong interest in the possible effects of dairy policy reform given the economic importance of the dairy industry.

With the model estimated in this research, it is only possible to simulate a change at ones of the milk price and the direct income payments. This is a limitation of the model. Later on quota will be introduced in the model as an explanatory variable. This is a small deviation from the theoretical model, but gives the opportunity to simulate the effect of the quota increase. Simulating the price change and the direct income payments change leads to an on average decrease of on farm labour of 8.26% and an on average increase of off-farm labour of 11.28%. This last figure is the on average increase of farmers that are already supplying off-farm labour. Output price theoretically does not have an impact on the shadow price of labour. We also do not find an effect of milk price on labour supply. Therefore, it is a bit strange to simulate the effect of a milk price cut on labour supply. Only simulating the effect of a change in direct payments results in an average on-farm labour decrease of 0.09% and an average off-farm labour supply decrease of 0.28%. This indicates that the change in direct income payment have a
very small effect on labour supply. This is interesting since the payments are assumed decoupled. From this simulation follows that this assumption could hold. At least there is a negligible effect through labour supply.

**Summary and conclusions**

This research focuses on the estimation of on- and off-farm labour supply equations for Dutch dairy farmers. Off-farm labour supply of Dutch dairy farmers is characterised by the fact that only half of the farm households supply off-farm labour. We assume that the farmers that supply off-farm labour is not a representative sample from the Dutch dairy farmers’ population. This results in a sample selection estimation problem that has to be taken into account in estimation. The Dutch Agricultural Research Institute (LEI) collects data for Dutch farm households. Every year a group of farm households leaves the data set and a new group enters. The farm households are in the data set for a number of subsequent years. This results in a rotating panel data set. A panel data set gives information on changes over time within the households in the panel. Panel data estimation methods use this information in estimation. In this research we use the panel data sample selection estimation method of Wooldridge (1995) to estimate the off-farm labour supply equation. This estimation method is based on the fixed effect panel data estimation method for linear models of Mundlak (1978). We will use this method to estimate the linear on-farm labour supply equation. The effects of CAP reform policy changes for labour supply of Dutch dairy farmers is simulated with the estimated on- and off-farm labour supply equations. The income compensation part of the policy change is assumed decoupled from production decisions. Simulation results suggest that this assumption could hold. At least there is a negligible effect through labour supply.

**References**


Appendix A

Variance correction formulas

This Appendix gives the formulas for the variance correction for the Wooldridge (1995) estimator in the case that the number of variables in $X_u$ is smaller than the number of variables in $W_u$. This is the case in our research.

\[
\hat{\omega}_u = \left( \hat{\mathbf{x}}_u, \bar{\mathbf{x}}_u, \hat{\mathbf{X}}_u \right) \tag{A.1}
\]

\[
\hat{\kappa} = (\hat{\beta}', \hat{\theta}', \hat{\delta}) \tag{A.2}
\]

\[
\hat{v}_u = y_u - \hat{\omega}_u \hat{\kappa} \tag{A.3}
\]

\[
\hat{p}_i = \sum_{t=1}^{T} d_u \hat{\omega}_u' \hat{v}_u \tag{A.4}
\]

\[
\hat{A} = n^{-1} \sum_{i=1}^{n} \sum_{t=1}^{T} d_u \hat{\omega}_u' \hat{\omega}_u \tag{A.5}
\]

\[
\hat{B} = n^{-1} \sum_{i=1}^{n} \hat{p}_i \hat{p}_i' \tag{A.6}
\]

\[
\text{vár}(\hat{\kappa}) = \hat{A}^{-1} \hat{B} \hat{A}^{-1} / n \tag{A.7}
\]

Equation (A.7) gives the corrected variance estimates of the parameter estimates for equation (20).
Appendix B

Data

<table>
<thead>
<tr>
<th>Variable</th>
<th>Dimension</th>
<th>Mean</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Off-farm income (&gt; 0 for 46.2 of observations)</td>
<td>1000 Euro</td>
<td>5.087</td>
<td>7.765</td>
</tr>
<tr>
<td>On-farm labour</td>
<td>Hours</td>
<td>4068</td>
<td>1484</td>
</tr>
<tr>
<td>Milk price index</td>
<td>1991 = 100</td>
<td>98.44</td>
<td>5.15</td>
</tr>
<tr>
<td>Other output price index</td>
<td>1991 = 100</td>
<td>100.08</td>
<td>11.11</td>
</tr>
<tr>
<td>Input price index</td>
<td>1991 = 100</td>
<td>100.42</td>
<td>4.71</td>
</tr>
<tr>
<td>Factor costs</td>
<td>1000 Euro</td>
<td>18.608</td>
<td>18.776</td>
</tr>
<tr>
<td>Non-farm non-labour income</td>
<td>1000 Euro</td>
<td>10.291</td>
<td>8.821</td>
</tr>
<tr>
<td>Land</td>
<td>Hectares</td>
<td>35.264</td>
<td>19.820</td>
</tr>
<tr>
<td>Machinery</td>
<td>1000 Euro</td>
<td>77.371</td>
<td>51.463</td>
</tr>
<tr>
<td>Dept</td>
<td>Percentage of assets</td>
<td>27.45</td>
<td>18.02</td>
</tr>
<tr>
<td>Unemployment rate</td>
<td>Percentage</td>
<td>6.19</td>
<td>1.28</td>
</tr>
<tr>
<td>Specialisation in milk</td>
<td>Percentage</td>
<td>7.39</td>
<td>9.43</td>
</tr>
<tr>
<td>Household members</td>
<td>Number</td>
<td>4.65</td>
<td>1.91</td>
</tr>
<tr>
<td>Successor</td>
<td>Percentage</td>
<td>4.67</td>
<td></td>
</tr>
<tr>
<td>Education</td>
<td>Dummy</td>
<td>2.48</td>
<td>0.67</td>
</tr>
</tbody>
</table>
Appendix C

Off-farm labour decision estimation results

This Appendix gives the table with estimation results for the off-farm labour decision equation for both the pooled probit estimator used in the Wooldridge (1995) estimation.

Table C1. Estimation results for off-farm labour supply decision

<table>
<thead>
<tr>
<th>Estimation method:</th>
<th>Pooled probit</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>estimate</td>
</tr>
<tr>
<td>Constant</td>
<td>78.78</td>
</tr>
<tr>
<td>Log on-farm labour</td>
<td>-0.28</td>
</tr>
<tr>
<td>Log off-farm wage</td>
<td>0.28</td>
</tr>
<tr>
<td>Log milk price</td>
<td>0.82</td>
</tr>
<tr>
<td>Log other output price</td>
<td>-0.36</td>
</tr>
<tr>
<td>Log input price</td>
<td>1.11</td>
</tr>
<tr>
<td>Factor costs</td>
<td>0.65</td>
</tr>
<tr>
<td>Non-farm non-labour income</td>
<td>-0.11</td>
</tr>
<tr>
<td>Land</td>
<td>0.25</td>
</tr>
<tr>
<td>Machinery</td>
<td>-0.38</td>
</tr>
<tr>
<td>Debt over asset ratio</td>
<td>0.13</td>
</tr>
<tr>
<td>Unemployment rate</td>
<td>-0.00</td>
</tr>
<tr>
<td>Specialisation</td>
<td>0.01</td>
</tr>
<tr>
<td>Number of household members</td>
<td>0.02</td>
</tr>
<tr>
<td>Successor dummy</td>
<td>0.14</td>
</tr>
<tr>
<td>Education</td>
<td>-0.17</td>
</tr>
</tbody>
</table>

* indicates significance at the 10% or smaller level.