SUBJECTIVE ASSESSMENT OF FINANCIAL DISTRESS IN PURCHASING A SUFFICIENT AMOUNT OF FOOD. ECONOMETRIC ANALYSIS OF POLISH MICRODATA

Hanna Dudek

Warsaw University of Life Sciences – SGGW

ABSTRACT

The paper analyses subjective aspects of food poverty in Poland. It deals with households’ assessment of financial difficulties in purchasing a sufficient amount of food in the period 2009–2015. The study is based on Social Diagnosis data. Its purpose is to identify the socio-economic factors affecting financial distress among Polish households. The study also aims to test whether the probability of experiencing financial difficulties is persistent over time. In econometric analysis binary choice models for panel data are applied. The findings state that apart from equivalent incomes and owned savings, loans or debts, factors having a significant impact on the final results are places of residence and biological types of households.

Key words: financial distress, food poverty, panel data, binary output models

INTRODUCTION

There is no one single definition of poverty, but most of them are focused on the inability to meet basic needs at a satisfactory level [Drewnowski 1997, Lemmi and Panek 2016]. All traditional lists of immediate “basic needs” include food, thus, a lot of research devoted to poverty examines the access to a sufficient amount of this good. Recently, much attention in developing as well as developed countries has been paid to the phenomenon of food poverty. By this term is understood “an inability to afford, or to have access to, food to make up a healthy diet” [Maslen et al. 2013] or “the insufficient economic access to an adequate quantity and quality of food to maintain a nutritionally satisfactory, socially acceptable diet” [O’Connor et al. 2016]. There is a shortage of detailed studies on food poverty in Poland. Thus, this study is carried out to get some insights into this field.

In poverty researches, two approaches are applied: the subjective and the objective one. In the analyses on subjective poverty, information on the opinion of the individuals about their situation is used. This approach deals with the subjective view that the households have of their situation as opposed to the objective one that only uses measurable variables. In other words, in the objective approach, the status of individuals can be verified by documentary evidence and is not based on subjective judgment by the respondent [Atkinson et al. 2002], while subjective poverty is defined by examining who people consider to be poor [Nandori 2011].

This study tries to improve the understanding of subjective aspects of food poverty in Poland through estimation of binary choice models for panel data. More precisely, this research deals with econometric analysis of the financial difficulties of households to purchase enough food. Thus, the purpose of this study is an identification of socio-economic factors affecting financial distress among Polish households. It also aims to test whether the probability of experiencing financial difficulties of households to purchase enough food is persistent over time.

**EMPIRICAL DATA**

This study is based on data completed in the framework of the survey *Social Diagnosis* which took place in 2000–2015 [Council for Social Monitoring 2015]. Two questionnaires are used in the survey – for individuals and for households. In this study, data from the second questionnaire is used. The survey questionnaire includes the question: Can your household afford to buy a sufficient amount of the following food items? Provide the answers for each of the following items separately: vegetables and vegetable preserves; fruit and fruit preserves; meat (including poultry); meat and poultry products; fish and fish products; butter and other edible fats; milk; dairy products; sugar; confectionary (sweets, chocolate etc.).

Respondents could choose an answer: yes or no. The aim of this work is to identify the households that could not afford to buy a sufficient amount of at least one of the ten featured group of products. Thus, in econometric models dependent variable is a binary variable taking a value 1 if household reported any financial difficulties in purchasing the food and a value 0 if the household did not indicate any problems in this assessment.

The *Social Diagnosis* research is a panel study. Each subsequent wave involves all available households from the previous wave and households from a new representative sample. So far, eight waves have been conducted: in 2000, 2003, 2005, 2007, 2009, 2011, 2013 and 2015 [Czapiński and Panek 2015]. Approximately 70% of the households surveyed in a given year, participated in the next wave of the research. In the study, the data regarding years 2009–2015 is analysed. Such choice of period is due to the fact that the sample size significantly increased from 3,000–4,000 households in 2000–2005 to around 12,000 households in 2009–2015. Moreover, in 2007, more than a half of the households did not reply to the investigated question, thus, the data with respect to this year has to be omitted in the this study.

The analysis aims to check whether the selected socio-economic factors affected the fact that a household reported financial difficulties in purchasing a sufficient amount of food in at least one of the featured group of products.

**METHODS**

Using a latent variable framework, the binary choice model for a panel of data would be written as [Greene 2012]:

\[
y^*_i = x_i^T \beta + u_i + \epsilon_i, \quad i = 1, \ldots, n, \quad t = 1, \ldots, T_i
\]

where:

\[
\gamma_{it} = \begin{cases} 
1 & \text{if } y^*_i \geq 0 \\
0 & \text{if } y^*_i < 0
\end{cases}
\]

\(^1\) This paper is an extension of study presented in [Dudek 2016] by taking into account a broader set of explanatory variables and considering a broader group of models.
\( \beta \) – a vector of parameters to be estimated, \( \beta^T = [\beta_0, \beta_1, ..., \beta_k] \);

\( y^*_t \) – a latent (unobserved) variable for \( i \)-th household in \( t \)-th year;

\( y_t \) – a value of observed binary variable for \( i \)-th household in \( t \)-th year;

\( u_i \) – an unobserved, individual specific effect for \( i \)-th household;

\( e_{it} \) – an error term for \( i \)-th household in \( t \)-th year, \( e_{it} \sim IID(0, \sigma_e) \);

\( i \) – indexes households;

\( t \) – indexes time period;

\( n \) – number of households;

\( T_i \) – number of observations for \( i \)-th households.

In econometrics for panel data there is a distinction between “random” and “fixed” effects models by the relationship between \( u_i \) and \( x_{it} \). The assumption that \( u_i \) is unrelated to \( x_{it} \) produces the random effects model, otherwise fixed effects model should be applied [Baltagi 2005, Jaba et al. 2016]. Both approaches are fraught with difficulties and unconventional estimation problems. On the one hand, estimation of the random effects model requires strong assumptions about the individual specific effects. On the other hand, the fixed effects model encounters an incidental parameters problem that renders the maximum likelihood estimator inconsistent even when the model is properly specified, moreover, there cannot be any time invariant explanatory variables in a fixed effects binary choice model [Greene and Hensher 2010]. Fixed effects models must be excluded in this study, because several explanatory variables are time invariant. Thus, random effects models are estimated. Such models assume that \( u_i \) and \( e_{it} \) are independent random variables with [Greene and Hensher 2010]:

\[
E[e_{it} | X] = 0; \quad \text{cov}(e_{it}, e_{is}) = \sigma_e^2 \quad \text{if} \quad i = j \quad \text{and} \quad t = s; \quad 0 \quad \text{otherwise} \tag{2}
\]

\[
E[u_i | X] = 0; \quad \text{cov}(u_i, u_j | X) = 0 \quad \text{if} \quad i \neq j, \quad \text{Var}[u_i | X] = \sigma_u^2 \tag{3}
\]

\[
\text{cov}(e_{it}, u_j | X) = 0 \quad \text{for all} \quad i, t, j \tag{4}
\]

where: \( X \) indicates all the exogenous data in the sample, \( x_{it} \) for all \( i \) and \( t \).

Then

\[
\text{cov}(w_{it}, w_{js}) = \sigma_u^2 \rho = \text{corr}(w_{it}, w_{js}) = \frac{\sigma_u^2}{1 + \sigma_u^2} \tag{5}
\]

where: \( w_{it} = u_i + e_{it} \).

Parameter \( \rho \) is the proportion of the total variance contributed by the panel-level variance component. When it equals zero, the binary panel model reduces to the pooled binary model. The conditional probability that \( y \) equals one is given by the formula:

\[
P(y_{it} = 1 | x_{it}, \beta, u) = P(y^*_t \geq 0 | x_{it}, \beta, u) = P(-e_{it} < x_{it}^T \beta + u) = F(x_{it}^T \beta + u) \tag{6}
\]

where: \( F \) denotes a cumulative distribution function (cdf) of \( -e_{it} \).

---

1 Pooled binary model does not contain in formula (1) component \( u_i \) – individual specific effect for \( i \)-th household.
Various functions for $F$ have been suggested in the literature. The most common ones are:

- the logistic cdf, i.e. $F(z) = \Lambda(z) = \frac{\exp(z)}{1 + \exp(z)}$, yielding the logit model;
- the standard normal cdf, i.e. $F(z) = \Phi(z)$, yielding the probit model;
- extreme-value (Gumbel) cdf, i.e. $F(z) = \exp[-\exp(-z)]$, yielding complementary log-log model.

The less-used complementary log-log model is an alternative to logit and probit analysis and is typically applied when one of the outcomes (the positive or negative) are rare [Cameron and Trivedi 2005].

The parameters $\beta_1, \beta_2, ..., \beta_k$ in considered binary choice models are not easy to interpret directly. One can determine the marginal effect of a change in an explanatory variable on the conditional probability that $y_t = 1$.

According to the formula (6) the marginal effect of a given variable, say $X_j$, are given by:

$$
\frac{\partial P(y_{it} = 1|\mathbf{x}_{it}, \beta, u_i)}{\partial x_{jit}} = \frac{\partial F'(x_{it}^T\beta + u_i)}{\partial (x_{it}^T\beta)}
$$

where: $F'(x_{it}^T\beta + u_i) = \frac{\partial F(x_{it}^T\beta + u_i)}{\partial (x_{it}^T\beta)}$

$x_{jit}$ – a value of $j$-th explanatory variable for $i$-th household in $t$-th year.

Hence, the significance and the direction of the marginal effects may be analysed simply by examining the significance and sign of $\beta_j$.

The customary estimation method of random effects models is a maximum likelihood method. Applying this method, it is commonly assumed that individual specific effects $u_i$ are normally distributed, with $u_i \sim N(0, \sigma_u^2)$.

The log-likelihood is given by formula:

$$
\log L = \sum_{i=1}^{n} \log \left[ \prod_{t=1}^{T} G(y_{it}, \alpha + x_{it}^T\beta + \sigma_v v_i) \right] \varphi(v_i) dv_i
$$

where: $\varphi$ – a density function for standard normal distribution, $v_i = \frac{u_i}{\sigma_u}$

$G(y_{it}, \alpha + x_{it}^T\beta + \sigma_v v_i) = F(\{x_{it}^T\beta + \sigma_v v_i\})$, if $y_{it} = 1$

and $G(y_{it}, \alpha + x_{it}^T\beta + \sigma_v v_i) = 1 - F(\{x_{it}^T\beta + \sigma_v v_i\})$ otherwise.

Maximization of the log-likelihood (8) with respect to parameters $\beta$ and $\sigma_u^2$ requires computation of the inner integrals, for which there is no analytical solution, thus, numerical methods have to be used. The most common approach is to use quadrature methods [Cameron and Trivedi 2005].

In poverty analysis sometimes the current state of poverty has been modeled as a function of lagged poverty [Giarda 2013, Alem et al. 2014]. This approach requires the use of a dynamic binary choice model. Such a model for a panel of data that explicitly allows for lagged effects would be written as [Verbeek 2008]:

$$
y^*_n = x_{nT}^T\beta + \gamma_{nT-1} + u_n + \epsilon_n
$$

with $y^*_n = 1$, if $y^*_n > 0$ and 0 otherwise.

---

4 Note that unlike logit and probit the complementary log-log model is asymmetrical, therefore formula (6) yields to $P(y_t = 1|\mathbf{x}_{it}, \beta, u_i) = 1 - \exp[-\exp(-x_{it}^T\beta + u_i)]$. 

---
In the dynamic binary choice model $\gamma > 0$ indicates positive state dependence, i.e. the \textit{ceteris paribus} probability that $y_{it} = 1$ is larger, if $y_{i,t-1} = 1$ [Verbeek 2008]. In order to get consistent parameter estimates of the model (9), Wooldridge approach is applied [2005].

In this study, both types of models: static model given by formula (1) and dynamic model given by the formula (9) are estimated.

**RESULTS AND DISCUSSION**

A decrease of percentage of the households that indicated financial distress in purchasing enough food on time in question is found. In 2009 about 28% of the households had such distress, while in 2015 – only about 19%.

In order to identify households that showed financial difficulties in purchasing enough food, various socio-economic factors are taken into account: demographic structure of the households, class of the place of residence, incomes, savings and debts. Akaike and Bayesian information criteria are used to compare alternative models with various sets of explanatory variables.

Random effects models with various variants of cdf: logit, probit and complementary log-log have been estimated. No meaningful differences between values of maximum likelihood function for these models have been found. Moreover, signs of parameters $\beta$ in all models indicate the same direction of impacts of socio-economic factors. Thus, the paper presents only the results for random effects probit models (Table 1). Table 2 presents description to measured variables. Computations are performed using Stata 14 statistical software package.

It is evident that most of the explanatory variables are statistically significant at 0.01 level. Moreover, the results of estimation of parameters $\sigma_u$ and $\rho$ confirm the presence of the unobserved individual specific effects $u_i$ in formulas (1) and (9).

Based on results of estimation of the dynamic model (9) evidence of state dependence is found, that is, the probability of experiencing financial difficulties in purchasing a sufficient amount of food at time ($t$) positively depends upon the probability of having experienced financial fragility at time ($t - 1$). As expected, higher income and having savings reduced the probability of such difficulties, while having loans or credit – increased it\(^5\). This finding can be explained by the fact that in most cases, savings and debts are the main liquid assets that can be used as a substitute for current income if the income level decreases or the level of spending increases [Kośny and Piotrowska 2013].

According to obtained results, living in the middle-sized towns improved a perception of own financial situation, comparing to other places of residence. Taking into account type of household, with married couples without children as a reference type, it is found that, the probability of experiencing financial difficulties in purchasing a sufficient amount of food was greater among one-person households and single-parent families and was lower among married couples with children. There are no statistical significant differences in this assessment between married couples and non-family multi-person households. These results indicate that psychological components in a subjective assessment of own financial situation are very important.

It is difficult to compare the obtained results with the findings of other research, since the literature lacks studies regarding determinants of financial difficulties in purchasing a sufficient amount of food in Poland. One can only refer to research of various authors on the subjective assessment of the financial conditions of Polish households. It should be mentioned that Kasprzyk [2016] stated that the main factors influencing such assessment are incomes and owned savings, which found confirmation in the present study. The results regarding place of residence mentioned in the literature are not unambiguous; for instance, Kasprzyk [2016] found that the place of residence is of little influence on the subjective assessment of own financial situation, whereas Dudek and Landmesser [2012] stated that the probability of higher levels of income satisfaction of households in the countryside is lower than in the case of households in towns. Taking into account type of household, findings

\(^5\) All presented interpretation were made under \textit{ceteris paribus} assumption.
Table 1. Results of estimation of random effects probit models

<table>
<thead>
<tr>
<th>Variable</th>
<th>Model given by eq. (1)</th>
<th>Model given by eq. (9)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>est.</td>
<td>SE</td>
</tr>
<tr>
<td>Lagged $y$</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Logarithm of income</td>
<td>–1.322***</td>
<td>0.026</td>
</tr>
<tr>
<td>Savings</td>
<td>–0.618***</td>
<td>0.025</td>
</tr>
<tr>
<td>Debts</td>
<td>0.146***</td>
<td>0.023</td>
</tr>
<tr>
<td>Class of place of residence</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Very big town</td>
<td>0.330***</td>
<td>0.065</td>
</tr>
<tr>
<td>Big town</td>
<td>0.190***</td>
<td>0.064</td>
</tr>
<tr>
<td>Middle-sized town</td>
<td>ref.</td>
<td>ref.</td>
</tr>
<tr>
<td>Small town</td>
<td>0.215***</td>
<td>0.056</td>
</tr>
<tr>
<td>Very small town</td>
<td>0.257***</td>
<td>0.059</td>
</tr>
<tr>
<td>Village</td>
<td>0.259***</td>
<td>0.053</td>
</tr>
<tr>
<td>Type of household</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MC without children</td>
<td>ref.</td>
<td>ref.</td>
</tr>
<tr>
<td>MC with 1 child</td>
<td>–0.106***</td>
<td>0.039</td>
</tr>
<tr>
<td>MC with 2 children</td>
<td>–0.293***</td>
<td>0.041</td>
</tr>
<tr>
<td>MC with 3+ children</td>
<td>–0.173***</td>
<td>0.050</td>
</tr>
<tr>
<td>Single-parent</td>
<td>0.216***</td>
<td>0.041</td>
</tr>
<tr>
<td>Multi-family</td>
<td>–0.166***</td>
<td>0.047</td>
</tr>
<tr>
<td>One-person</td>
<td>0.411***</td>
<td>0.036</td>
</tr>
<tr>
<td>Non-family</td>
<td>0.030</td>
<td>0.099</td>
</tr>
<tr>
<td>Year</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2009</td>
<td>0.210***</td>
<td>0.028</td>
</tr>
<tr>
<td>2011</td>
<td>0.123***</td>
<td>0.028</td>
</tr>
<tr>
<td>2013</td>
<td>0.127***</td>
<td>0.027</td>
</tr>
<tr>
<td>2015</td>
<td>ref.</td>
<td>ref.</td>
</tr>
<tr>
<td>Constant</td>
<td>8.114***</td>
<td>0.191</td>
</tr>
<tr>
<td>$\sigma_u$</td>
<td>1.033***</td>
<td>0.021</td>
</tr>
<tr>
<td>$\rho$</td>
<td>0.516***</td>
<td>0.010</td>
</tr>
</tbody>
</table>

* means statistical significance at 0.10; ** statistical significance at 0.05; *** statistical significance at 0.01.

Source: Author’s own computation.

Table 2. List and description of explanatory variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Income</td>
<td>real equivalent income over the period of study (for further explanation see Czapinski and Panek [2015])</td>
</tr>
<tr>
<td>Savings</td>
<td>1 if household has savings, 0 otherwise</td>
</tr>
<tr>
<td>Debts</td>
<td>1 if household has loans or credit, 0 otherwise</td>
</tr>
<tr>
<td>Class of place of residence</td>
<td>the class of place of residence is divided into urban and rural areas, with urban areas further subdivided by resident size units</td>
</tr>
<tr>
<td>very big town</td>
<td>1 if town over 500,000 residents, 0 otherwise</td>
</tr>
<tr>
<td>big town</td>
<td>1 if town with 200,000–500,000, 0 otherwise</td>
</tr>
<tr>
<td>middle-sized town</td>
<td>1 if town with 100,000–200,000 residents, 0 otherwise</td>
</tr>
<tr>
<td>small town</td>
<td>1 if town with 20,000–100,000 residents, 0 otherwise</td>
</tr>
<tr>
<td>very small town</td>
<td>1 if town up to 20,000 residents, 0 otherwise</td>
</tr>
<tr>
<td>village</td>
<td>1 if rural areas, 0 otherwise</td>
</tr>
<tr>
<td>Household type</td>
<td>household type is established on the basis of the number of families and biological family type</td>
</tr>
<tr>
<td>MC without children</td>
<td>1 if married couples (MC) with no children, 0 otherwise</td>
</tr>
<tr>
<td>MC with 1 child</td>
<td>1 if married couples (MC) with one child, 0 otherwise</td>
</tr>
<tr>
<td>MC with 2 children</td>
<td>1 if married couples (MC) with two children, 0 otherwise</td>
</tr>
<tr>
<td>MC with 3+ children</td>
<td>1 if married couples (MC) with three or more children, 0 otherwise</td>
</tr>
<tr>
<td>single-parent</td>
<td>1 if single-parent families, 0 otherwise</td>
</tr>
<tr>
<td>multi-family</td>
<td>1 if multi-family households, 0 otherwise</td>
</tr>
<tr>
<td>one-person</td>
<td>1 if non-family one-person households, 0 otherwise</td>
</tr>
<tr>
<td>non-family</td>
<td>1 if non-family multi-person households, 0 otherwise</td>
</tr>
<tr>
<td>Year</td>
<td>data regarding to 2009–2015 is analysed</td>
</tr>
<tr>
<td>2009</td>
<td>1 if year is 2009, 0 otherwise</td>
</tr>
<tr>
<td>2011</td>
<td>1 if year is 2011, 0 otherwise</td>
</tr>
<tr>
<td>2013</td>
<td>1 if year is 2013, 0 otherwise</td>
</tr>
<tr>
<td>2015</td>
<td>1 if year is 2015, 0 otherwise</td>
</tr>
</tbody>
</table>

Source: Author’s own computation.
of this study are confirmed by other authors. Ulman and Šoltés [2015] found that the greatest risk of subjective monetary poverty affects one-person and single-parent households.

It should be emphasized that this study can be seen as a first step towards a measurement of subjective aspects of food poverty. In future research on determinants of financial distress in purchasing a sufficient amount of food, various characteristics of the members of the household should be taken into account, among others: education, age, gender and labour market status.

CONCLUDING REMARKS

The study undertakes the issue of financial distress in purchasing a sufficient amount of food. It uses the data from Social Diagnosis survey. This data has an important advantage: approximately 70% of the households surveyed in a given year, participated in the next wave of the research, therefore, this type of data can be treated as a panel data. A distinctive feature of panel data modelling is inclusion of unobserved heterogeneity, which is typically interpreted as the individual specific effect of latent factors on the dependent variable.

In econometric analysis binary choice models with random-effect case are estimated. It is found that apart from “financial” reasons, such as achieved incomes, having savings, loans or credit, class of place of residence and biological types of households have an important influence on the perception of financial distress in purchasing a sufficient amount of food. Moreover, the results indicate that such perception among Polish households is persistent over time.

The issue of financial distress in purchasing a sufficient amount of food should be constantly monitored. The obtained findings could be used in creating a social policy supporting vulnerable households.

REFERENCES


OCENA SUBIEKTYWNYCH TRUDNOŚCI FINANSOWYCH W ZAKRESIE NABYWANIA WYSTARCZAJĄCEJ ILOŚCI ŻYWNOŚCI – ANALIZA EKONOMETRYCZNA POLSKICH MIKRODANYCH

STRESZCZENIE


Słowa kluczowe: trudności finansowe, ubóstwo żywnościowe, dane panelowe, modele zmiennej binarnej