Enhancing predictive performance of vulnerability to poverty estimates

Maike Hohberg*[1], Katja Landau[1], Thomas Kneib[1], Stephan Klasen[1], and Walter Zucchini[1]

[1] *University of Goettingen*

November 1, 2016

Abstract

This paper analyzes two modifications to a simple measure of vulnerability as expected poverty. Firstly, in order to model income we use distributional regression which relates each parameter of the conditional income distribution to the covariates. Secondly, instead of defining a household as vulnerable if its probability of being poor in the next period is larger than 0.5, we construct a vulnerability line as vulnerability threshold. For this purpose, we employ the receiver operating characteristic considering prerequisites in terms of the true positive rate. Using long-term panel data from Germany, we build both mean and distributional regression models with the established 0.5 probability cutoff and our defined vulnerability line. Predictive performance is assessed by proper scoring rules. We find that our new cutoff method considerably increases predictive performance while placing the income regression model into the distributional regression framework improves predictions for some years.

Keywords: vulnerability to poverty, distributional regression, gamlss, ROC, scoring rules

JEL classification: C13, C18, C52, I32

*Correspondence to: Maike Hohberg, Chair of Statistics and Econometrics, University of Goettingen, Humboldtallee 3, 37073 Goettingen, Germany, mhohber@uni-goettingen.de

1 Introduction

Knowing which households are vulnerable to poverty and which are not can guide social policy makers on how to efficiently allocate resources in order to prevent households from falling into poverty at some future date. Although closely related, poverty and vulnerability to poverty are two different concepts. Poverty refers to a state at a (static) point in time usually measured *ex post* using household income or expenditure surveys whereas vulnerability to poverty refers to a potential state in the future, i.e. an occurrence that may or may not happen in future (Moser, 1998). Therefore, unlike poverty, vulnerability has the nature of a probability forecast, an *ex ante* assessment of poverty risk. The need for such a forecast results directly from the fact that many households - though not living in poverty at the moment - are extremely vulnerable to events such as job loss, unexpected expenditures, economic downturns, and weather phenomena, and can therefore easily fall below the poverty line in the near future. Policy-makers who want to reduce poverty are therefore more interested in knowing which people are at risk of poverty in future than knowing who happened to be poor in the past.

Though in the recent years a few empirical applications of vulnerability to poverty measures evaluated predictive performance of their estimates (Bergolo, Cruces, & Ham, 2012; Celidoni, 2013; Feeny & McDonald, 2016; Jha & Dang, 2010; Ligon & Schechter, 2004; Zhang & Wan, 2009), very little attention has been paid to their improvement. One reason is that assessing and improving the accuracy of probability forecasts requires knowledge of the outcome, that is whether or not the household did become poor. Hence, such an analysis relies on panel data which is not always available. Therefore, most authors have been concerned with developing vulnerability measures in the context of cross-sectional data (e.g. Chaudhuri, Jalan, & Suryahadi, 2002; Christiaensen & Subbarao, 2005; Günther & Harttgen, 2009; Jha & Dang, 2010; Suryahadi & Sumarto, 2003). However, the increased availability of good quality panel data motivates to take a closer look on how to improve predictive performance of vulnerability estimates.

Ideally, vulnerability to poverty correctly identifies households which will be poor at some point in the future while minimizing the number of households that are classified as vulnerable but will not be poor. In addition to correct identification, for practical relevance it is desirable to keep a vulnerability measure comprehensible and relatively easy to implement with data that is widely available or can easily be collected. One popular approach considers vulnerability as expected poverty (VEP), i.e. the probability of a household falling into poverty in a future period (Chaudhuri, 2003; Chaudhuri et al., 2002; Pritchett, Suryahadi, & Sumarto, 2000). In order to improve this existing measure, we present two modifications which meet the desirable demands: The first proposal is related to the form of the regression model empirical researchers use to model income, consumption or any other simplistic measure of welfare. The second aspect concerns the probability threshold which is used to classify an individual or a household as vulnerable or not vulnerable.

Although the underlying conceptional framework of vulnerability to poverty estimates might be different, most of the empirical applications rely on some kind of regression model with a simple measure of welfare, such as per capita income, as the dependent variable. Our first modification aims at improving the predictive performance of these regression models by embedding them into the framework of distributional regression. This flexible framework allows to link each parameter of a conditional distribution to a structured additive predictor (Klein, Kneib, Lang, & Sohn, 2015; Stasinopoulos & Rigby, 2007). Hence, instead of relying only on the (log)normal distribution and simply modeling location (e.g. the mean), other distributions are possible and parameters such as scale and shape can vary (potentially in a non-linear way) with covariates.

Once a regression model with income or consumption as the dependent variable is built, VEP classifies households as vulnerable or not vulnerable depending on their predicted income. Traditionally, households are (arbitrarily) classified as vulnerable if their probability of being poor in the future is equal or greater than 0.5 or, alternatively, above the observed poverty rate (e.g. Chaudhuri et al., 2002; Pritchett et al., 2000). Though this classification has the advantage of simplicity, it does not always perform well in terms prediction (Bergolo et al., 2012; Celidoni, 2013). As a second modification, we construct a vulnerability to poverty line (VPL), which is used as a new vulnerability cutoff, employing the Receiver Operating Characteristic (ROC). The VPL marks the threshold under which households are declared vulnerable and hence likely to be poor in the next period. Although the ROC is generally a diagnostic instrument for *ex post* evaluation of binary prediction, it allows us to take account of specific *ex ante* requirements in terms of a minimum true positive rate (TPR). We thus directly address recent criticism of the traditional vulnerability threshold raised by e.g. Bergolo et al. (2012) and McCarthy, Brubaker, and La Fuente (2016) and propose a endogenous cutoff which directly addresses predictive precision.

In the context of vulnerability to poverty, the ROC has been already used to compare predictions of a range of vulnerability measures (Celidoni, 2013). Our contribution, however, substantially differs from this kind of analysis. Instead of using the ROC in order to *assess* performance, we construct a new vulnerability cutoff to *improve* performance. On the other hand, we propose evaluating predictions of vulnerability estimates by means of proper scoring rules as these in contrast to the ROC curve allow us to make direct comparisons between different models.

Even though vulnerability measures are mostly applied to data from developing countries, we rely on a high-quality long-term panel to be able to analyze predictive performance of our modification for several years. Hence, we use data of 15 years of the German Socio-Economic Panel (SOEP)[1], which offers comprehensive coverage of household characteristics and income enabling us to retrospectively observe whether a household did become poor or not. We find that our new cutoff method significantly improves predictive

[1]Socio-Economic Panel (SOEP), data of the years 1993-2008, version 26, SOEP, 2010, doi: 10.5684/soep.v26.

performance while placing the income regression model into the distributional regression framework does yield additional benefit for some years.

The structure of this paper is as follows: Section 2 briefly reviews the empirical approaches to measure vulnerability as expected poverty, and highlights the issues that the propositions made in this paper aim to enhance. Section 3 presents the estimation strategy for predicting vulnerability. Predictive precision and estimation results are discussed in Section 4 while Section 5 concludes.

2 Measuring vulnerability as expected poverty

The literature on the empirical assessment of vulnerability is traditionally divided into three strands: vulnerability as expected poverty (VEP), vulnerability as expected utility (VEU), and vulnerability as exposure to risk (VER).[2] The latter one, VER, retrospectively measures if an observed shock reduced welfare (for an application see e.g. Skoufias & Quisumbing, 2005). The second approach, VEU, accounts for risk preferences and defines vulnerability as the difference between a utility derived from a certainty equivalent at which the household would not be vulnerable and the expected utility derived from possible states in the future (e.g. Ligon & Schechter, 2003). Besides being difficult to interpret, this approach has been criticized for being dependent on the choice of a utility function and risk aversion parameter (Celidoni, 2013; Christiaensen & Subbarao, 2005; Gaiha & Imai, 2008). Finally, VEP considers vulnerability as the probability of a household falling into poverty in a future period. Most applications of this approach draw on Chaudhuri et al. (2002) who estimate a regression model with consumption as the dependent variable and a covariate dependent, heteroskedastic variance of the error term. As we are primarily concerned with measuring and predicting, we will base our analysis on the expected poverty concept as it is easily comprehensible, interpretable, forward looking (in contrast to VER), and has been widely applied (in contrast to VEU). Easy implementation is also the reason why we stick to regression analyses and propose modifications that can be combined with other existing concepts rather than presenting yet another vulnerability measure.

Vulnerability as expected poverty defines vulnerability of an individual or a household h at time t as the probability that some measure of welfare (usually income, expenditures, or consumption) y falls below the poverty line z at time $t+1$. That is

$$V_{ht} = \Pr(y_{h,t+1} < z) \tag{1}$$

Pritchett et al. (2000) were probably among the first ones who designed a measure to anticipate poverty

[2] See Klasen and Waibel (2013) for a comprehensive review.

status. In their paper, for an n-year horizon, a household h is defined to be vulnerable if the probability that the household income will fall below the poverty line z in at least one of the subsequent n-years exceeds a pre-specified cutoff p.

To empirically estimate this probability, most applications follow Chaudhuri et al. (2002). Using cross-sectional data, it is assumed that consumption is generated by

$$\ln y_h = X_h\beta + e_h \tag{2}$$

where y_h is consumption expenditure, X_h are household characteristics, and e_h is the error term capturing idiosyncratic shocks under the assumption of being identically and independently distributed over time. Variance is allowed to vary with covariates across households implying a relationship between higher volatility in consumption and poverty risk. The variance of e_h is given by

$$\sigma^2_{e,h} = X_h\theta \tag{3}$$

Estimating β and θ via a three-step feasible generalized least squares (FGLS) procedure (Amemiya, 1977), yields expected consumption and variance

$$\hat{\mathrm{E}}[\ln y_h|X_h] = X_h\hat{\beta} \tag{4}$$

$$\widehat{\mathrm{Var}}[\ln y_h|X_h] = \hat{\sigma}^2_{e,h} = X_h\hat{\theta} \tag{5}$$

Under the assumption that (log) consumption is normally distributed, the probability of being poor, i.e. the vulnerability level, will be

$$\widehat{\Pr}(\ln y_h < \ln z|X_h) = \Phi\left(\frac{\ln z - X_h\hat{\beta}}{\sqrt{X_h\hat{\theta}}}\right) \tag{6}$$

The household is then classified as vulnerable if this probability is equal or greater than 0.5. The brief sketch of the VEP approach shows that its interpretation is quite easy as it is expressed in monetary terms.[3] The need for an enhanced approach results from three major drawbacks: Firstly, as Celidoni (2013) points out, the welfare measure is always assumed to be (log)normally distributed. While this simplifies estimation and inference, scholars such as McDonald and Ransom (2008) and Sohn, Klein, and Kneib (2014) have provided evidence that in many application income or expenditures do not behave

[3]Some modifications to this model are available which will not be considered further since in order to demonstrate our suggestions, it is desirable to keep the model simple. These modifications include for example differentiating between covariate and idiosyncratic shocks on household and community level (Günther & Harttgen, 2009), accounting for depth of poverty (Hoddinott & Quisumbing, 2003), allowing for different risk sensitivity (Calvo & Dercon, 2013), and using an individual reference line which depends on the current living standard instead of a general poverty line (Dutta, Foster, & Mishra, 2011).

(log)normally. Current software packages and computational techniques easily allow to use distributions other than the (log) normal distribution.

Secondly, closely related to the first assumption, consumption or income regression models used for vulnerability measures usually rely on regression models that model a linear relationship between covariates and the expected mean and the variance of the income distribution. However, if we depart from the normality assumption, parameters beyond that should be modeled to capture the full effect on the whole conditional income or consumption distribution. To our knowledge, no previous paper in the vulnerability literature departs from the normality assumption and models non-linear effects on parameters beyond variance.

The third drawback concerns the cutoff p which determines whether a household is classified as vulnerable or not. In most application cases it is set at 0.5 (e.g. Chaudhuri et al., 2002; Günther & Harttgen, 2009; Nguyen, Jolly, Bui, Chuong T. P. N., & Le, 2015; Novignon, Nonvignon, Mussa, & Chiwaula, 2012; Zereyesus, Embaye, Tsiboe, & Amanor-Boadu, 2016) that is a household is vulnerable if the probability of earning an income (or having a consumption level) below the poverty line in $t+1$ is equal or above 50 percent. However, this neglects the variability a household faces since if the expected income (or consumption) equals the poverty line, the normal density function in equation 6 yields 0.5 independent of the standard deviation. (McCarthy et al., 2016). Alternatively, Chaudhuri et al. (2002) also proposes the observed poverty rate as cutoff. Both methods are arbitrary and can largely influence the vulnerability predictions. Since in practice it is beneficial to classify a large amount of households correctly, this cutoff should be optimized.

We tackle all of these drawbacks by two modifications: Distributional regression is aiming at the first and second point. Within this flexible framework, different distributions can be chosen to model income, and all parameters of this distribution are related to a structured additive predictor which can incorporate non-linear effects. Our approach to use the ROC to determine the vulnerability cutoff improves the third drawback of the traditional method and directly addresses the need expressed in Bergolo et al. (2012) of an endogenously chosen vulnerability threshold taking account of prediction errors in order to improve targeting of social policy programs.

3 Estimation strategy

3.1 Data and variables

To demonstrate our modifications we use the Socio-Economic Panel (SOEP) which is a panel study of German households starting in 1984 and carried out by the German Institute for Economic Research

(DIW), Berlin. For our application case, we used the version "Soepv26" and included observations of private households covering years from 1993 to 2008. The covariates in the regression models included variables that are aggregated at the household level as well as characteristics on the household head. Households with incomplete information of the household head were excluded. We used household-level cross-sectional weights and staying probabilities provided in the dataset. Equivalence income is computed using the modified OECD equivalence scales[4] and adjusted for inflation using 2005 as basis year.

Special care is needed to account for the different timing structure of retrospective income and prospective household characteristics (see e.g. Frick, Jenkins, Lillard, Lipps, & Wooden, 2008). Since the income reported by household members in the SOEP survey year in fact refers to the income in the previous year, incomes for year t has to be extracted from the records of year $t+1$. However, household composition can change from year to year and and thus affects the equivalence income. We therefore use the household composition in the year in which the income accrued, and not from the survey year.[5] Hence, we base our analysis on four-year panels: two waves are used to build a model, one wave to verify predictions and an additional wave is needed due to the one-year delay in availability of income information. Although "even millionaires are vulnerable to poverty" (Pritchett et al., 2000), we assume that households whose real annual income per adult equivalent exceeds 100,000 EUR are not at risk of becoming poor in the immediate future and excluded them from our analysis. We also removed households which were not represented in all four years. Equivalence income is used as dependent variable and several household characteristics as explaining factors (Table 4 in the Appendix provides an overview). The information extracted at the household level includes age structure, i.e. the number of children and elderly in the household, and variables related to the employment situation of the household members, namely the number of full-time employees and its quadratic term.[6] The selected covariates related to the household head include sex, marital status, education, employment sector (or unemployed) and status of ownership of residence.

3.2 A distributional regression model for income

Measuring vulnerability as expected poverty basically consists of two major steps: The first one estimates expected income based on a regression approach and the second step translates the results of this model

[4]These scales take the number of household members and their age into account. Weights of 1, 0.5, and 0.3 are assigned to the household head, other household members above the age of 14, and children below the age of 14, respectively (see e.g. Atkinson, 2002; Krause & Ritz, 2006; Stauder & Hüning, 2004). In the analysis carried out by Celidoni (2013) the choice of equivalent scales does not significantly change the evaluation results of the predictive performance of different vulnerability measures.

[5]This introduces some bias if individuals, who accrued income in the previous year, have joined or left the household but if we chose to consider the current year household composition, a similar and arguably more problematic bias would occur as household composition and income-earning do not refer to the same year.

[6]We expect that a higher number of full-working household members is generally associated with a higher household equivalence income. However, if the number is unusually large, this might be due to a low household income forcing members to work that would perhaps in richer households study or stay at home.

into a measure of vulnerability. While this section considers the first step, the next one takes care of the second one.

Chaudhuri et al. (2002) formulates a model in which both the expected mean and variance are covariate dependent. To facilitate inference, log income is often assumed to follow a normal distribution. However, sometimes other distributions such as the Generalized Beta, the Singh-Maddala, or the dagum distribution can provide a better fit (e.g. McDonald & Ransom, 2008; Sohn et al., 2014). These distributions have other parameters than location and scale and in the distributional regression framework, all of them can vary with covariates. Hence, we would not only model the expected mean but the whole conditional income distribution.

This is exactly our first suggestion. We argue that one could improve predictive accuracy of existing vulnerability measures by placing the income regression in the framework of structured additive distributional regression (Klein et al., 2015; Stasinopoulos & Rigby, 2007). That is, the conditional income distribution is given by a density conditioned on parameters θ_k of which each of the K parameters is itself dependent on the explanatory variables. We thus write

$$g_k(\theta^{(k)}) = \eta^{(k)} = X^{(k)}\beta^{(k)} + \sum_{j=1}^{J^{(k)}} s_j(z_j^{(k)}) \tag{7}$$

where g_k is a link function, $\eta^{(k)}$ the predictor for the kth parameter, the matrix X_k contains covariates which are assumed to have a linear effect, and $s_j(z_{j,k})$ are smooth functions of J continuous covariates z which have non-linear effects. More precisely, for the covariate *past income* we relax the restrictive assumption of a linear effect and use P(enalised)-splines (Eilers & Marx, 1996) to flexibly model its relationship to the dependent variable. As conditional distributions, we use in addition to the normal, the Singh-Maddala distribution which has been shown to provide a good fit to the SOEP income data (Biewen & Jenkins, 2002; Selezneva & van Kerm, 2016). For the two parameters of the normal distribution, we thus have

$$\hat{\mu} = \eta^{(\mu)} = X^{(\mu)}\beta^{(\mu)} + \sum_{j=1}^{J^{(\mu)}} s_j(z_j^{(\mu)}) \tag{8}$$

$$\log(\hat{\sigma}) = \eta^{(\sigma)} = X^{(\sigma)}\beta^{(\sigma)} + \sum_{j=1}^{J^{(\sigma)}} s_j(z_j^{(\sigma)}) \tag{9}$$

$$\tag{10}$$

with a log link employed for σ to ensure positivity and a multiplicative connection. For the three parameter Singh-Maddala distribution we have

$$\log(\hat{a}) = \eta^{(a)} = X^{(a)}\beta^{(a)} + \sum_{j=1}^{J^{(a)}} s_j(z_j^{(a)}) \tag{11}$$

$$\log(\hat{b}) = \eta^{(b)} = X^{(b)}\beta^{(b)} + \sum_{j=1}^{J^{(b)}} s_j(z_j^{(b)}) \tag{12}$$

$$\log(\hat{q}) = \eta^{(q)} = X^{(q)}\beta^{(q)} + \sum_{j=1}^{J^{(q)}} s_j(z_j^{(q)}) \tag{13}$$

(14)

This model formulation differs to the traditional approach in that it makes a different distributional assumption and hence models three instead of two parameters. Additionally, we add non-linear effects. While the common approach uses a three step Feasible Generalized Least Squares (FGLS) estimator to estimate the model (8), distributional regression models are estimated via a back-fitting algorithm that maximizes the penalized likelihood. In this way, parameters are estimated simultaneously in contrast to the step wise FGLS approach. The methodology is implemented in the `gamlss` package in R, and described in Stasinopoulos and Rigby (2007).

In addition to these two models, we estimate simpler versions of them, in order to analyze the benefits of incorporating past income as an additional covariate, modeling the variance, using the Singh-Maddala instead of the normal distribution, and incorporating non-linear effects. That is, in total we use six different models for the income regression (Table 1).

[Place table 1 here]

Model M3 serves as our comparison model, similar to the one formulated in Chaudhuri et al. (2002). To facilitate comparability we also apply the `gamlss` algorithm to this model. M1 and M2 serve as a check if one can reach similar predictive performance if using a much simpler model. Model M1 is a simple linear model with income as dependent variable but does not include past income. This might be relevant if only cross sectional data is available. M2 is also a simple linear model but includes past income as a covariate. Models M4-M6 check if additional complexity improves performance. Models M4 and M6 incorporate past income in a non-linear way. Model M3 and M4 assume a normal distribution with all parameters depending on covariates, while M5 and M6 assume a Singh-Maddala distribution for incomes.

In each model, the covariate vector X contains the variables described in Section 3.1 except that for parameters σ, b, and q the square of *full-time work* is not included.

3.3 Constructing a vulnerability cutoff using the ROC

The ROC is a well-established instrument to quantify and compare the accuracy of binary diagnostic techniques in many fields (e.g. Egan, 1975; Fawcett, 2006; Thompson & Zucchini, 1989)[7]. In the case of vulnerability, we assess how accurately different methods predict the households which will become poor and those that will not. The diagnosis proceeds by first fitting a model to estimate the future income of each household. Then a cutoff (the vulnerability line) is specified such that households whose predicted income fall below the cutoff are classified as vulnerable; the others are classified as non-vulnerable. Thus, four different diagnosis-outcome combinations are possible: true positives, false positives, true negatives, and false negatives.

The ROC is then simply a graph in which the false positive rate (FPR) is plotted on the X-axis against the true positive rate (TPR) on the ordinate. The result is a non-decreasing function with start point (0,0) and end point at the point (1,1). Roughly speaking, the faster the curve approaches the level tpr=1 the better the diagnostic method; the method is perfect if its ROC reaches TPR=1 straight away, i.e. it passes through the point (0,1).

By varying the cutoff point, we can balance the rates at which the two types of error (false negatives and false positives) occur. We use this characteristic of the ROC to contruct a vulnerability line which satisfies some prespecified criteria and is hence less arbitrary then a vulnerability line set at 0.5 probability of becoming poor in the next period.

The idea is to compare the diagnostic performance of a model over the entire range of possible cutoffs which classifies a household as vulnerable or not. We then declare the cutoff as our vulnerability line which has led to the best predictive performance of our model.[8] The criteria used to choose the optimal cutoff is straightforward: we aim at achieving a prescribed TPR. Suppose, for example, that we wish to identify at least 80 percent of all households that will be poor in the coming year. The value TPR=0.8 determines retrospectively the vulnerability to poverty line which in turn determines the FPR. If the resulting FPR is high, the absolute number of "false alarms" might well be too large for the model being of practical use.

The reason why we require a prescribed TPR while minimizing the FPR as opposed to the other way round, i.e. pre-specifying a FPR, is a practical one: We argue that a (benevolent) policy maker is more interested in targeting almost all vulnerable households (at the cost of false positives) than in missing poor households.

[7] For example, in medical research it is used to quantify the trade-off between the *sensitivity* or true positive rate (TPR, i.e. the probability to correctly diagnose a diseased patient) and the "specificity" (the probability to correctly diagnose a non-diseased patient).

[8] See Landau (2012) for an extensive introduction to using the ROC to measure vulnerability to poverty including profiles of vulnerable households, extensions to an n-year period and interval income data, and analyses of macroeconomic variables.

To put it more concretely, we first use information for two years to estimate a model (current income as a function of past covariates and past income) which is then used to forecast the income of each household in the third year. To determine which households are actually poor in the third year, we used the observed incomes and a predefined poverty line. We then order the predicted incomes and use each one as a hypothetical vulnerability line. Households with incomes below the hypothetical vulnerability line will be declared vulnerable the others as not vulnerable. For each possible cutoff, we then construct the ROC by comparing the actual poor with vulnerable and non-vulnerable households. As the vulnerability line increases, both the TPR and FPR will increase. For the purpose of illustration, we focus on vulnerability lines that leads to a TPR of 80 percent[9], i.e. in order to correctly identify 80 per cent of the poor households in year three, all households with a predicted income (for year 3) below the selected vulnerability line must be classified as vulnerable. Or, to put it differently, the vulnerability line is selected such that 80 percent of the households that are poor in the third year would have been classified as vulnerable at the end of the second year. For a specified TPR, the model with the smallest FPR is the best as it identifies the specified proportion of households that are poor in year 3 but leads to the smallest number of false alarms.[10]

Regarding the choice of a poverty line, it is common practice in Germany to use a relative poverty line, set at 60 percent of the median per capita income (see e.g. Celidoni, 2013; Stauder & Hüning, 2004). As vulnerability refers to poverty in the next period, it is necessary to specify the value of the poverty line in the next period. To avoid the additional source of uncertainty that arises when forecasting the poverty line, we use the relative poverty line of the current year, which ranges from about 10,000 to 11,700 EUR annual income and yields poverty estimates between around 10 and 14 percent in each year.

With the method described so far, we determine a cutoff *ex post* such that we reach a TPR of 80 percent when comparing predicted and observed incomes for a specific year. However, our aim is to estimate vulnerability as a forward looking perspective on poverty. Hence, we must fix a vulnerability line *ex ante* in a way that we still reach a high TPR. We achieve this simply by considering past vulnerability lines. That is, we predict incomes for a specific year based on our models, then we use the VPL we constructed for the previous year to classify households into vulnerable or not vulnerable. As the classification is now based on previous VPLs, not only the FPR but also the TPR will vary between our different models. Alternatively to using the VPL of the previous year as it is done in this analysis, one might also consider using the average of VPLs of all previous years.

[9]Other prerequisites can be chosen here depending on the preference of the policy maker.

[10]The above refers to the estimation of one-year-ahead vulnerability but can be easily extended to n-year-ahead vulnerability. For example, in the two-year case a household is (retrospectively) defined as poor if it is poor in *at least one* of the two years subsequent to the vulnerability forecast. Clearly, accuracy decreases as n increases.

3.4 Proper scoring rules to assess predictive performance

Relatively little attention has been paid in the literature to the predictive performance of vulnerability estimates with some notable exception who evaluate predictions of vulnerability estimates with real data. They mostly rely on the TPR or proportions of TPR and FPR (Bergolo et al., 2012; Celidoni, 2013; Feeny & McDonald, 2016; Jha & Dang, 2010; Zhang & Wan, 2009). We follow the literature by calculating TPR and FPR for all six models under the two different cutoffs and extend it by the use of proper scoring rules to assess the models' performance.[11]

A scoring rule is a function $S(\hat{\pi}, r)$ assigning a value to the event that r is observed under a predictive distribution $\hat{\pi}$. In our case, $r_h \in \{0, 1\}$ is the (true) classification whether a household h is vulnerable (i.e. poor in the next period) $r_h = 1$ or not $r_h = 0$. The first rule we use to evaluate our model is the logarithmic scoring rule which is based on the log-likelihood contributions and applying it to our binary prediction yields:

$$\text{log score} = \frac{1}{n}\sum_{h=1}^{n} r_h \cdot \log(\hat{\pi}_h) + (1 - r_h) \cdot \log(1 - \hat{\pi}_h) \tag{15}$$

where we average over all households $h = 1, \ldots, n$, $\hat{\pi}_h$ is the probability of household h of having a predicted income below the vulnerability line . One disadvantage of the logarithmic score is that it uses only one of the probabilities of the predictive distribution and is hence outlier sensitive (Kneib et al., 2007). This problem is addressed by two other scores:

$$\text{Brier score} = \frac{1}{n}\sum_{h=1}^{n}(\hat{\pi}_h - r_h)^2 \tag{16}$$

is the Brier Score. Here, the optimal predictive distribution in the true category is compared to $\hat{\pi}$. For binary outcomes the Brier score equals the mean squared error of the forecast. The spherical score for binary predictions is given by

$$\text{spherical score} = \frac{1}{n}\sum_{1}^{n} \frac{\hat{\pi}_h^{r_h} \cdot (1 - \hat{\pi}_h)^{(1-r_h)}}{\sqrt{\hat{\pi}_h^2 + (1 - \hat{\pi}_h)^2}} \tag{17}$$

which draws a relation to the norm of the predictive distribution.

[11] Proper and strictly proper scoring rules are discussed in Gneiting and Raftery (2007) and for example applied in Kneib, Baumgartner, and Steiner (2007) to assess models of consumer choice behavior.

4 Results

4.1 Regression results for one year

To assess vulnerability as described in the previous section and to be able to compare forecasts of different strategies, we firstly estimate the six presented models. This estimation is done for all of the 15 years in our dataset. Table 2 presents the results for M1 and M2 for year 1996 which will be used to construct the VPL in 1997 (see Appendix for M3, M4, M5, M6). Unfortunately, for some years M4 and the algorithm implemented in the `gamlss` package do not converge due to using non-linear effects. For the selected year, estimating M4 is possible. Note that coefficients only from the linear models can be interpreted directly. Most effects in the linear model are as expected: past income is highly significant and explains most of the variations in income, children in the household tend to decrease income, income increases with education, mostly with higher education, households headed by women have lower income on average, more full time working members increase income only until a certain number of members as very large households tend to be poorer. Perhaps most striking is the result that elderly in the household increase income, which might be explained by relatively high pensions in Germany in the 1990s, and that widowed households tend to have a higher income, which might be explained by specific pensions for widowed.

[Place table 2 here]

After running the regressions for each year, we predict expected incomes $\hat{y}$ for the following year. These predicted incomes are sorted in an ascending order and each income is then used as a potential vulnerability line for which TPR and FPR based on the predicted and actually observed incomes are calculated. This procedure is repeated for each of the 15 years in our data set. Figure 1 presents the ROC curve for year 1997 and the six different models. Only the model which is not using past income as a covariate (M1) is performing significantly worse compared to the others. The other five models perform very similarly. Remarkably, the simple linear model with constant variances (M2) does not perform much worse in terms of ROC than the more complex ones. This suggests that current income is a very good predictor of income for the coming year.

[Place figure 1 about here]

4.2 Constructing a vulnerability line using the ROC

After calculating TPR and FPR for each year, we pick that predicted income which leads to a TPR of 80 percent and declare it the VPL of that specific year. Figure 2 shows the VPLs for every year. The simple linear model M1 without past income and M5 stands apart as the associated VPLs are higher compared to the other models.

Though each model consequently leads to a TPR of 80 percent, they differ in terms of FPR. Figure 3 presents the FPRs associated with each model and the selected VPL. It seems that M2, M3, M4, and M6 perform very similarly with FPRs higher than 0.1 but smaller than 0.25. Only M1 has an unacceptably high FPR. Remember that the calculated VPL and FPR in this first step only concern the current year and are not yet predictions for the future. This will be considered in the next section.

[Place figure 2 about here]

[Place figure 3 about here]

4.3 Evaluating predictions

As vulnerability is often considered an *ex ante* measure of future poverty, we use the calculated VPLs (of the previous year) to make predictions about poverty status in the next year.[12]. More precisely, households with a predicted income in $t+1$ below the most recent VPL are considered as vulnerable. Related TPR and FPR are presented in Figure 4(b),(d), and (f) for M1, M2, and M6 respectively and in Figure 5 in the Appendix. Note that the TPR is now not exactly at 80 percent as we use the previous VPL and not the current actual one (which we do not know if we want to forecast poverty status). Each prediction for each year is represented by a tuple (FPR, TPR).[13] The better a model performs the more tuple should lie in the 4th quadrant which corresponds to a FPR below 20 percent and a TPR of 80 percent. For VPLs associated with a TPR of about 80 percent most models lead good results in terms of a low FPR. Only M1 is performing much worse than the others and M5 does have some outlier years with low TPRs or high FPRs. Distinguishing the performance of models M2-M6 seems hardly possible when using this form of presentation, we thus have a closer look on these models in the next section where we calculate predictive scores.

Regarding our second modification, i.e. using a different cutoff method, Figure 4(a),(b), and (e) and Figure 6 in the Appendix show that the "traditional" cutoff method does not perform acceptably. The method classifies only very few households as vulnerable. That is, while yielding a low FPR, it is not able to identify those households that will be poor in the next period. This is in contrast to Zhang and Wan (2009) who found that the 0.5 probability performs well in predicting poverty status in rural China but it is in line with Bergolo et al. (2012) who found weak predictive performance for data from Argentina and Chile and Celidoni (2013) who also does not classify the method as a "high performer". .

Nonetheless, the low TPRs for some years are conspicuous and we put further effort in investigating why we obtain these striking differences in predictive performance. Looking at the effects of past income on

[12] Alternative approaches are possible here: One could use moving averages of previous VPLs and/or extend poverty predictions to a n-year horizon.

[13] The first year we can make a prediction for is 1995 as we require the VPL of the previous year as cutoff and the first available cutoff is the VPL in 1994.

current income[14], we find that for the most part the relationship is roughly linear whereas non-linearities occur especially at the bottom end of the income distribution where the effects are much lower and can even be negative. These effects are not accounted for in the models assuming a linear relationship (M2, M3, M5). The past income regression coefficient for these models is higher than the coefficient would be when only looking at the poor. On the other hand, models that do incorporate non-linear effects (M4, M6) also have difficulties in predicting very low incomes as our dataset only comprises about 10-14 percent poor households. Hence, the poor observations only have a small impact on the overall models' prediction ability and thus the models predict overly optimistic incomes resulting in too few households being classified as vulnerable.

Our new cutoff method is able to mitigate these weaknesses of the models since the cutoff is determined endogenously and directly aiming at improving predictive performance. In contrast, the traditional method relies heavily on the model specification and its prediction abilities which can be weak especially at the lower extremes of the income distribution.[15] Since the 0.5 probability was advocated for developing countries, where the share of the poor is generally higher than in Germany, it is possible that the traditional cutoff performs better in other settings. This means that the old cutoff method cannot be readily applied to every country context. One might argue that the 0.5 cutoff simply needs to be set lower in these cases.[16] But then again, what is the optimal cutoff probability? It is basically an arbitrary decision and therefore we argue that our suggestion to relate the cutoff to VPLs based on the ROC does a much better job in predicting poverty status and in mitigating weak prediction abilities of the underlying income model while at the same time it is easy to grasp and implement.

[Place figure 4 about here]

[14]The Appendix includes a plot of the M4 model for 1996 as an example

[15]Note that excluding outliers at the lower end of the distribution to improve model predictions was not an option for us since it is specifically the poor who we are interested in.

[16]We tested several other probability cutoff and to reach a similar TPR of around 80 percent for all model the cutoff lies between 0.2 and 0.1 which is far away from the original 0.5 and leads mostly to FPRs well above 20 percent.

4.4 Scores

We now use proper scoring rules to evaluate the performance of our models further. That is, we evaluate the probabilities of a household to have an income below the VPL. Hence, we will only be able to compare our six models under the new cutoff method. It is not possible to make comparisons of the same model with the two different cutoffs since under the old cutoff the VPL is set as a probability itself. Hence, proper scoring rules under the old cutoff would require to calculate the probability of being vulnerable as a probability of being poor. This clearly is a disadvantage of the cutoff method: Though we can evaluate binary predictions as in Section 4.3, we are not able to calculate predictive scores.

Table 3 shows the calculated scores using our new cutoff method for each year and each of the six models separately. A model performs better if the log score and the spherical score are large and the Brier score is small. In each year, M3, M4, or M6 perform best. If computation of M4 was possible, it outperforms the other models. For around half of the years, M6 yields better predictions than M3. The fact that M5 performs relatively bad compared to the other models while M6 is doing well indicates that including non-linear effects of past income plays a more important role than the choice of the distribution. Our results also confirm the importance of modeling the variance as covariate dependent because the heteroskedastic model M2 performs lower compared to the distributional models. Modeling parameters beyond variance as in M5 and M6 seems to yield only minor differences which might be due to using data from an industrialized country with a social safety net which most likely reduces shocks on households. Since modeling scale and shape parameters should account for idiosyncratic shocks, it is possible that for Germany those idiosyncratic shocks are either less prevalent than in developing countries or households are able to cope better with them. Nonetheless, distributional regression is an attractive alternative to estimate vulnerability to poverty as it estimates mean and variance simultaneously and is able to incorporate non-linear effects which seem to be key when looking at incomes at the lower end of the distribution.

[Place table 3 here]

5 Conclusion

Vulnerability is an important concept for researchers and policymakers if their interest not only lies in a static snap-shot of current poverty but they are worried about future development and aim at preventing households to fall into poverty in the future. The majority of the empirical literature follows the concept of vulnerability as expected poverty which defines a household as vulnerable if its probability of earning an income less than the poverty line in the next period is higher or equal than 0.5.

This value was arbitrarily set and only little attention has been paid to the predictive performance of this measure. This paper therefore proposes two modifications: Firstly, we use distributional regression to model all parameters of an income distribution instead of estimating mean and variance separately with as introduced by Chaudhuri et al. (2002). Secondly, to address recent criticism of the traditional vulnerability threshold, we propose a different cutoff method to differentiate between vulnerable and non-vulnerable households. Both suggestions are implemented using household panel data from Germany to be able to evaluate predictions for many years and to ensure data quality.

For the first suggestion we place our regression analysis into the framework of structured additive distributional regression models where each parameter is dependent on the covariates and not only the mean. We find that for some years using distributional regression does improve predictive performance mainly due to incorporating non-linear effects of past income. We suggest that effects of modeling shape parameters are relatively small due to the use of data from an industrialized countries where social safety nets allow households to cope with idiosyncratic shocks. Future research can certainly contribute by investigating under which circumstances modeling parameters beyond mean and variance becomes necessary.

For the second suggestion, we develop a new method to determine a vulnerability cutoff. For each year we calculate VPLs using the ROC such that 80 percent of the vulnerable households are correctly identified. To make a prediction *ex ante* we have to use a previous VPL. We find that in terms of predictive performance our new cutoff method outperforms the traditional approach where the (arbitrary) cutoff is set at 50 percent probability of being poor in the next period. Since our new cutoff method is constructed to improve predictive performance, can mitigate weaknesses in the income generating model specification, and is easy to implement, it is a useful tool if researchers or policymakers have a panel data set at hand and are particularly interested in correctly identifying vulnerable households rather than in measuring overall vulnerability.

References

Amemiya, T. (1977). The Maximum Likelihood and the Nonlinear Three-Stage Least Squares Estimator in the General Nonlinear Simultaneous Equation Model. *Econometrica*, *45*(4), 955.

Atkinson, A. B. (2002). *Social indicators: The EU and social inclusion.* Oxford: Oxford University Press.

Bergolo, M., Cruces, G., & Ham, A. (2012). Assessing the predictive power of vulnerability measures:evidence from panel data for argentina and chile. *Journal of Income Distribution*, *21*(1), 28-64. Retrieved from `http://EconPapers.repec.org/RePEc:jid:journl:y:2012:v:21:i:1:p:28-64`

Biewen, M., & Jenkins, S. P. (2002). *Accounting for poverty differences between the United States, Great Britain, and Germany* (No. 311). Berlin: Berlin, Germany : DIW Berlin, German Institute for Economic Research.

Calvo, C., & Dercon, S. (2013). Vulnerability to individual and aggregate poverty. *Social Choice and Welfare*, *41*(4), 721–740.

Celidoni, M. (2013). Vulnerability to Poverty: An Empirical Comparison of Alternative Measures. *Applied Economics*, *45*, 1493–1506.

Chaudhuri, S. (2003). *Assessing vulnerability to poverty: concepts, empirical methods and illustrative examples* (Mimeo). Columbia University.

Chaudhuri, S., Jalan, J., & Suryahadi, A. (2002). *Assessing Household Vulnerability to Poverty from Cross-sectional Data: A Methodology and Estimates from Indonesia* (Tech. Rep. No. 0102-52). New York: Department of Economics, Columbia University.

Christiaensen, L. J., & Subbarao, K. (2005). Towards an Understanding of Household Vulnerability in Rural Kenya. *Journal of African Economies*, *14*(4), 520–558.

Dutta, I., Foster, J., & Mishra, A. (2011). On measuring vulnerability to poverty. *Social Choice and Welfare*, *37*(4), 743–761.

Egan, J. P. (1975). *Signal detection theory and ROC analysis.* New York, NY: Acad. Pr.

Eilers, P. H. C., & Marx, B. D. (1996). Flexible smoothing with B -splines and penalties. *Statistical Science*, *11*(2), 89–121. doi: 10.1214/ss/1038425655

Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters*, *27*(8), 861–874.

Feeny, S., & McDonald, L. (2016). Vulnerability to multidimensional poverty: Findings from households in melanesia. *The Journal of Development Studies*, *52*(3), 447-464. doi: 10.1080/00220388.2015.1075974

Frick, J. R., Jenkins, S. P., Lillard, D. R., Lipps, O., & Wooden, M. (2008). Die internationale Einbettung des Sozio-oekonomischen Panels (SOEP) im Rahmen des Cross-National Equivalent File (CNEF). *Vierteljahrshefte zur Wirtschaftsforschung*, *77*(3), 110–129.

Gaiha, R., & Imai, K. (2008). *Measuring Vulnerability and Poverty* (No. 40/2008).

Gneiting, T., & Raftery, A. E. (2007). Strictly Proper Scoring Rules, Prediction, and Estimation. *Journal of the American Statistical Association*, *102*(477), 359–378.

Günther, I., & Harttgen, K. (2009). Estimating Households Vulnerability to Idiosyncratic and Covariate Shocks: A Novel Method Applied in Madagascar. *World Development*, *37*(7), 1222–1234.

Hoddinott, J., & Quisumbing, A. (2003). *Methods for Microeconometric Risk and Vulnerability Assessments* (No. 0324). Washington, DC.

Jha, R., & Dang, T. (2010). Vulnerability to Poverty in Papua New Guinea in 1996. *Asian Economic Journal*, *24*(3), 235–251.

Klasen, S., & Waibel, H. (2013). *Vulnerability to Poverty.* London: Palgrave Macmillan UK. doi: 10.1057/9780230306622

Klein, N., Kneib, T., Lang, S., & Sohn, A. (2015). Bayesian Structured Additive Distributional Regression with an Application to Regional Income Inequality in Germany. *Annals of Applied Statistics*, *9*(2), 1024–1052. doi: 10.1214/15-AOAS823

Kneib, T., Baumgartner, B., & Steiner, W. J. (2007). Semiparametric multinomial logit models for analysing consumer choice behaviour. *AStA Advances in Statistical Analysis*, *91*(3), 225–244.

Krause, P., & Ritz, D. (2006). EU-Indikatoren zur sozialen Inklusion in Deutschland. *Vierteljahrshefte zur Wirtschaftsforschung*, *75*(1), 152–173. doi: 10.3790/vjh.75.1.152

Landau, K. (2012). *Messung der Vulnerabilität der Armut: Eine statistische Analyse mit deutschen Paneldaten* (Dissertation). Universität Göttingen, Göttingen.

Ligon, E., & Schechter, L. (2003). Measuring Vulnerability. *The Economic Journal*, *113*(March).

Ligon, E., & Schechter, L. (2004). *Evaluating different approaches to estimating vulnerability* (No. 30159).

McCarthy, N., Brubaker, J., & La Fuente, A. d. (2016). *Vulnerability to Poverty in rural Malawi* (No. WPS7769). Washington, DC. Retrieved from `http://www-wds.worldbank.org/external/default/WDSContentServer/WDSP//`

McDonald, J. B., & Ransom, M. (2008). The Generalized Beta Distribution as a Model for the Distribution of Income: Estimation of Related Measures of Inequality. In D. Chotikapanich (Ed.), *Modeling Income Distributions and Lorenz Curves* (Vol. 5, pp. 147–166). New York, NY: Springer-Verlag New York.

Moser, C. O. (1998). The asset vulnerability framework: Reassessing urban poverty reduction strategies. *World Development*, *26*(1), 1–19. doi: 10.1016/S0305-750X(97)10015-8

Nguyen, K. A. T., Jolly, C. M., Bui, Chuong T. P. N., & Le, T. H. (2015). Climate change, rural household food consumption and vulnerability: The case of Ben Tre province in Vietnam. *Agricultural Economics Review*, *16*(2), 95–109.

Novignon, J., Nonvignon, J., Mussa, R., & Chiwaula, L. S. (2012). Health and vulnerability to poverty in Ghana: evidence from the Ghana Living Standards Survey Round 5. *Health economics review*, *2*, 11. doi: 10.1186/2191-1991-2-11

Pritchett, L., Suryahadi, A., & Sumarto, S. (2000). *Quantifying Vulnerability to Poverty: A Proposed Measure, Applied to Indonesia.* The World Bank.

Selezneva, E., & van Kerm, P. (2016). A distribution-sensitive examination of the gender wage gap in Germany. *The Journal of Economic Inequality*, *14*(1), 21–40. doi: 10.1007/s10888-016-9320-z

Skoufias, E., & Quisumbing, A. R. (2005). Consumption insurance and vulnerability to poverty: A synthesis of the evidence from bangladesh, ethiopia, mali, mexico and russia. *The European Journal of Development Research*, *17*(1), 24-58. doi: 10.1080/09578810500066498

Sohn, A., Klein, N., & Kneib, T. (2014). A New Semiparametric Approach to Analysing Conditional Income Distributions. *SOEPpapers on Multidisciplinary Panel Data Research*, *No. 676*.

Stasinopoulos, D. M., & Rigby, R. A. (2007). Generalized Additive Models for Location Scale and Shape (GAMLSS) in R. *Journal of Statistical Software*, *23*(7). doi: 10.18637/jss.v023.i07

Stauder, J., & Hüning, W. (2004). *Die Messung von Äquivalenzeinkommen und Armutsquoten auf der Basis des Mikrozensus* (No. 13).

Suryahadi, A., & Sumarto, S. (2003). Poverty and Vulnerability in Indonesia Before and After the Economic Crisis. *Asian Economic Journal*, *17*(1), 45–64.

Thompson, M. L., & Zucchini, W. (1989). On the statistical analysis of ROC curves. *Statistics in Medicine*, *8*(10), 1277–1290.

Zereyesus, Y. A., Embaye, W. T., Tsiboe, F., & Amanor-Boadu, V. (2016). *Participation in non-farm work and vulnerability to food poverty of households in northern Ghana.* Boston, MA.

Zhang, Y., & Wan, G. (2009). How Precisely Can We Estimate Vulnerability to Poverty? *Oxford Development Studies*, *37*(3), 277–287.

A Appendix

[Place table 4 here]

[Place table 5 here]

[Place table 6 here]

[Place figure 5 here]

[Place figure 6 here]

[Place figure 7 here]

Tables

Table 1: Overview of six income regression models

Model	Discription	
M1:	$\hat{y}_{h,t} = X_{h,t-1}\hat{\beta}$	linear model, normal distribution of log income, no past income
M2:	$\hat{y}_{h,t} = X_{h,t-1}\hat{\beta} + y_{h,t-1}\hat{\gamma}$	linear model, normal distribution of log income, past income
M3:	$\log(\hat{\mu}) = X^{(\mu)}_{h,t-1}\hat{\beta}^{(\mu)} + y_{h,t-1}\hat{\gamma}^{(\mu)}$ $\log(\hat{\sigma}) = X^{(\sigma)}_{h,t-1}\hat{\beta}^{(\sigma)}$	distributional regression, normal distribution of log income, past income
M4:	$\log(\hat{\mu}) = X^{(\mu)}_{h,t-1}\hat{\beta}^{(\mu)} + s(y_{h,t-1})$ $\log(\hat{\sigma}) = X^{(\sigma)}_{h,t-1}\hat{\beta}^{(\sigma)}$	distributional regression, normal distribution of log income, non-linear past income
M5:	$\log(\hat{a}) = X^{(a)}_{h,t-1}\hat{\beta}^{(a)} + y_{h,t-1}\hat{\gamma}^{(a)}$ $\log(\hat{b}) = X^{(b)}_{h,t-1}\hat{\beta}^{(b)}$ $\log(\hat{q}) = X^{(q)}_{h,t-1}\hat{\beta}^{(q)}$	distributional regression, Singh-Maddala distribution, past income
M6:	$\log(\hat{a}) = X^{(a)}_{h,t-1}\hat{\beta}^{(a)} + s(y_{h,t-1})$ $\log(\hat{b}) = X^{(b)}_{h,t-1}\hat{\beta}^{(b)}$ $\log(\hat{q}) = X^{(q)}_{h,t-1}\hat{\beta}^{(q)}$	distributional regression, Singh-Maddala distribution, non-linear past income

Table 2: Regression results for M1 and M2 and year 1996

	M1			M2		
log(equincome)	β	$se(\beta)$	PVal	β	$se(\beta)$	PVal
(Intercept)	9.642	0.049	0.000	4.450	0.114	0.000
log(past equincome)				0.537	0.011	0.000
b18	-0.121	0.008	0.000	-0.048	0.007	0.000
o60	0.061	0.012	0.000	0.017	0.010	0.089
famstat: divorced	-0.052	0.022	0.016	-0.026	0.018	0.143
famstat: separated	0.083	0.043	0.054	-0.042	0.038	0.276
famstat: single	0.025	0.018	0.168	0.015	0.015	0.339
famstat: widowed	0.118	0.021	0.000	0.052	0.017	0.003
fullworking	8.778	0.686	0.000	3.137	0.583	0.000
fullworking^2	-3.590	0.562	0.000	-1.637	0.496	0.001
industry: agriculture	0.019	0.063	0.760	-0.030	0.051	0.556
industry: bank,insurance	0.371	0.040	0.000	0.129	0.034	0.000
industry: construction and mining	0.141	0.027	0.000	0.003	0.023	0.901
industry: energy	0.258	0.063	0.000	0.098	0.052	0.062
industry: manufacturing	0.187	0.024	0.000	0.067	0.021	0.001
industry: services	0.220	0.023	0.000	0.085	0.019	0.000
industry: trade	0.146	0.027	0.000	0.057	0.023	0.012
industry: transport	0.170	0.037	0.000	0.035	0.031	0.259
owner: main tenant	-0.203	0.012	0.000	-0.073	0.011	0.000
owner: sub tenant	-0.209	0.035	0.000	-0.015	0.027	0.577
educ: 9/10th	0.125	0.047	0.009	0.107	0.040	0.007
educ: higher educ	0.452	0.048	0.000	0.246	0.040	0.000
educ: higher voc	0.287	0.050	0.000	0.177	0.042	0.000
educ: voc + abi	0.216	0.053	0.000	0.109	0.045	0.015
educ: voc or abi	0.147	0.046	0.001	0.103	0.039	0.008
sex: female	-0.062	0.014	0.000	-0.049	0.012	0.000
no. obs	5421			5421		

Source: own elaboration based on SOEP data

Table 3: Scores to evaluate the six models' predictions under new cutoff method

	log score						Brier score						spherical score					
	M1	M2	M3	M4	M5	M6	M1	M2	M3	M4	M5	M6	M1	M2	M3	M4	M5	M6
1995	-0.75	-0.66	-0.57	-0.51	-0.77	-0.66	0.27	0.22	0.18	0.17	0.26	0.21	0.69	0.75	0.79	0.81	0.70	0.77
1996	-0.74	-0.49	-0.42		-0.85	-0.63	0.26	0.16	0.13		0.30	0.21	0.70	0.82	0.85		0.66	0.76
1997	-0.72	-0.43	-0.38		-0.72	-0.55	0.26	0.13	0.11		0.25	0.18	0.70	0.85	0.88		0.71	0.79
1998	-0.78	-0.56	-0.51		-0.58	-0.56	0.28	0.19	0.16		0.20	0.18	0.67	0.79	0.81		0.77	0.79
1999	-0.63	-0.35	-0.29		-0.73	-0.45	0.22	0.10	0.08		0.26	0.15	0.75	0.89	0.91		0.70	0.83
2000	-0.73	-0.54	-0.49		-0.43	-0.33	0.26	0.18	0.16		0.14	0.10	0.70	0.80	0.82		0.84	0.89
2001	-0.71	-0.41	-0.38		-0.43	-0.36	0.25	0.13	0.12		0.14	0.11	0.71	0.86	0.87		0.84	0.88
2002	-0.69	-0.43	-0.37		-0.54	-0.40	0.25	0.14	0.11		0.19	0.13	0.72	0.85	0.87		0.79	0.86
2003	-0.63	-0.36	-0.32		-0.45	-0.30	0.22	0.11	0.10		0.15	0.09	0.75	0.88	0.90		0.83	0.90
2004	-0.69	-0.38	-0.34		-0.59	-0.32	0.24	0.12	0.10		0.20	0.10	0.72	0.87	0.89		0.77	0.89
2005	-0.65	-0.32	-0.29	-0.28	-0.61	-0.33	0.23	0.10	0.08	0.08	0.21	0.10	0.74	0.90	0.91	0.91	0.76	0.89
2006	-0.69	-0.42	-0.35	-0.33	-0.54	-0.32	0.24	0.13	0.11	0.10	0.18	0.10	0.72	0.86	0.88	0.89	0.79	0.89
2007	-0.62	-0.30	-0.28	-0.27	-0.41	-0.27	0.21	0.09	0.08	0.08	0.14	0.08	0.76	0.90	0.91	0.91	0.85	0.91
2008	-0.57	-0.34	-0.31	-0.29	-0.42	-0.28	0.19	0.10	0.09	0.09	0.14	0.09	0.78	0.88	0.90	0.90	0.84	0.90

Note: A good forecast maximizes the log score and the spherical score and minimizes the Brier score.
Source: own elaboration based on SOEP data

Table 4: Description of the variables used in the regression. The baseline levels for the categorical variables are indicated in boldface.

variables	description
past equincome	equivalence income of the household in the previous year
sex	sex of household head (**male**, female)
b18 o60	number of individuals aged below 18 number of individuals above 59
famstat	family status of the household head (5 levels: **married**; single; widowed; divorced; separated)
industry	employment sector of the household head (10 levels: **not employed**; agriculture; energy; mining; manufacturing; construction; trade; transport; bank, insurance; services)
owner	owner status of residence (3 levels: **owner**; main-tenant; sub-tenant)
educ	highest level of education (ISCED classification) of the household head (6 levels: **no degree**; 9/10th grade; "Abitur" (12th/13th grade) or vocational training; "Abitur" and vocational training; higher vocational training; higher education)
fullworking fullworking^2	number of household members in full-time employment quadratic term

Table 5: Regression results for M3 and M4 and year 1996

	M3		*M4*	
log(equincome)	β_μ	β_σ	β_μ	β_σ
(Intercept)	3.992	-1.498	4.128	-1.356
log(past equincome)	0.588			
b18	-0.040	-0.168	-0.025	-0.199
o60	0.011	-0.278	0.015	-0.291
famstat: divorced	-0.018	-0.103	-0.024	-0.074
famstat: separated	-0.040	0.171	-0.041	0.287
famstat: single	0.023	0.109	0.024	0.197
famstat: widowed	0.041	-0.185	0.027	-0.514
fullworking	2.330	-0.379	1.802	-0.351
fullworking^2	-0.822		-0.664	
industry: agriculture	-0.032	0.094	-0.018	0.108
industry: bank,insurance	0.085	-0.144	0.052	-0.155
industry: construction and mining	-0.003	-0.294	-0.005	-0.292
industry: energy	0.071	-0.602	0.055	-0.609
industry: manufacturing	0.057	-0.437	0.049	-0.439
industry: services	0.077	-0.229	0.070	-0.228
industry: trade	0.059	0.076	0.048	0.034
industry: transport	0.041	-0.662	0.035	-0.699
owner: main tenant	-0.070	0.076	-0.054	-0.040
owner: sub tenant	-0.008	0.722	-0.024	0.485
educ: 9/10th	0.061	-0.234	0.068	-0.482
educ: higher educ	0.191	-0.167	0.156	-0.345
educ: higher voc	0.121	-0.175	0.119	-0.246
educ: voc + abi	0.082	-0.691	0.079	-0.843
educ: voc or abi	0.055	-0.203	0.061	-0.317
sex: female	-0.041	0.150	-0.036	0.180
no. obs	5421		5421	

Note: The `gamlss` package gives a warning that it used the option qr to calculate standard errors. However, as these are not reliable (Stasinopoulos & Rigby, 2007) and we are only interested in the coefficients to make the predictions, we decided to not report them. M4 includes non-linear effects for past income.
Source: own elaboration based on SOEP data

Table 6: Regression results for M5 and M6 and year 1996

	M5			M6		
equincome	β_a	β_b	β_q	β_a	β_b	β_q
(Intercept)	2.623	1.142	1.465	2.488	1.258	1.063
past equincome	0.033			0.034		
b18	-0.056	0.097	-0.082	-0.038	0.107	-0.065
o60	-0.054	0.203	-0.244	-0.048	0.214	-0.220
famstat: divorced	-0.104	-0.107	-0.182	-0.094	-0.063	-0.227
famstat: separated	0.079	-0.368	0.323	0.060	-0.423	0.306
famstat: single	-0.012	-0.146	-0.164	-0.021	-0.104	-0.191
famstat: widowed	-0.077	0.140	-0.320	0.012	0.115	0.034
fullworking	4.271	0.094	0.161	3.376	0.079	0.165
fullworking^2	-1.405			-1.098		
industry: agriculture	-0.181	0.127	-0.636	-0.107	0.051	-0.375
industry: bank,insurance	-0.128	0.279	-0.731	-0.096	0.192	-0.504
industry: construction and mining	-0.139	0.272	-0.494	-0.118	0.243	-0.372
industry: energy	-0.087	0.434	-0.654	-0.115	0.531	-0.708
industry: manufacturing	-0.085	0.354	-0.477	-0.072	0.345	-0.402
industry: other						
industry: services	-0.077	0.302	-0.512	-0.063	0.290	-0.453
industry: trade	-0.070	0.129	-0.446	-0.063	0.126	-0.391
industry: transport	0.003	0.380	0.027	0.007	0.334	0.124
owner: main tenant	-0.104	0.090	-0.184	-0.068	0.102	-0.101
owner: sub tenant	0.063	-0.195	0.261	0.083	-0.218	0.271
educ: 9/10th	-0.183	0.301	-0.952	-0.143	0.289	-0.787
educ: higher educ	-0.081	0.267	-0.870	-0.065	0.278	-0.757
educ: higher voc	-0.163	0.327	-1.128	-0.157	0.307	-1.030
educ: voc + abi	-0.078	0.338	-0.478	-0.051	0.263	-0.349
educ: voc or abi	-0.118	0.267	-0.699	-0.102	0.246	-0.599
sex: female	0.023	-0.032	0.201	0.016	-0.036	0.189
no. obs		5421		5421		

Note: The `gamlss` package gives a warning that it used the option qr to calculate standard errors. However, as these are not reliable (Stasinopoulos & Rigby, 2007) and we are only interested in the coefficients to make the predictions, we decided to not report them. M6 includes non-linear effects for past income.
Source: own elaboration based on SOEP data

Figures

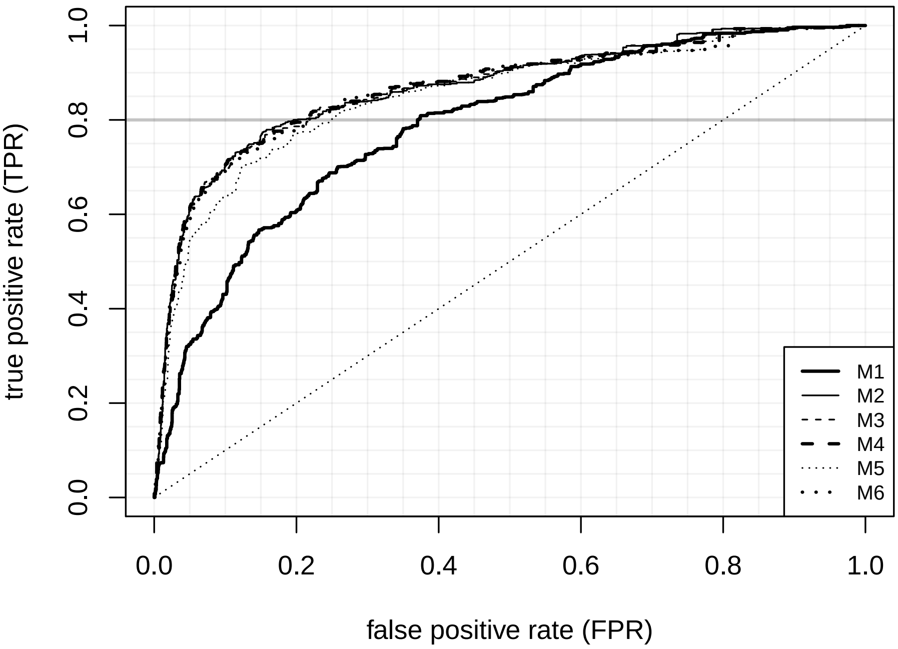


Figure 1: ROC for 1997 using several models. Each predicted income is used as potential VPL for which TPR and FPR are calculated, i.e. each predicted income is associated with a specific FPR and TPR and represents one point in the graph.
Source: own elaboration based on SOEP data

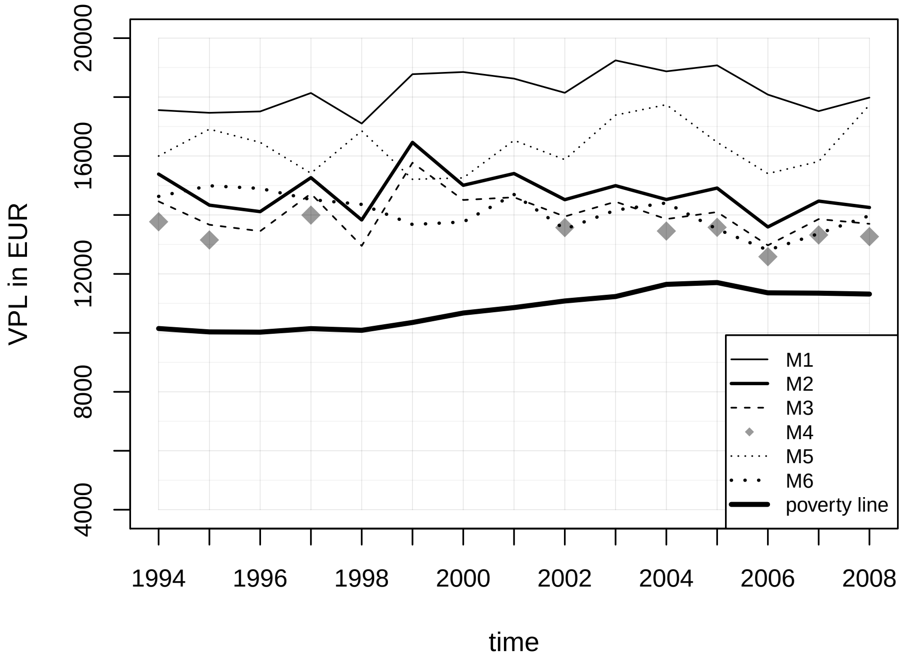


Figure 2: VPL at 80% TPR over time for different models
Source: own elaboration based on SOEP data

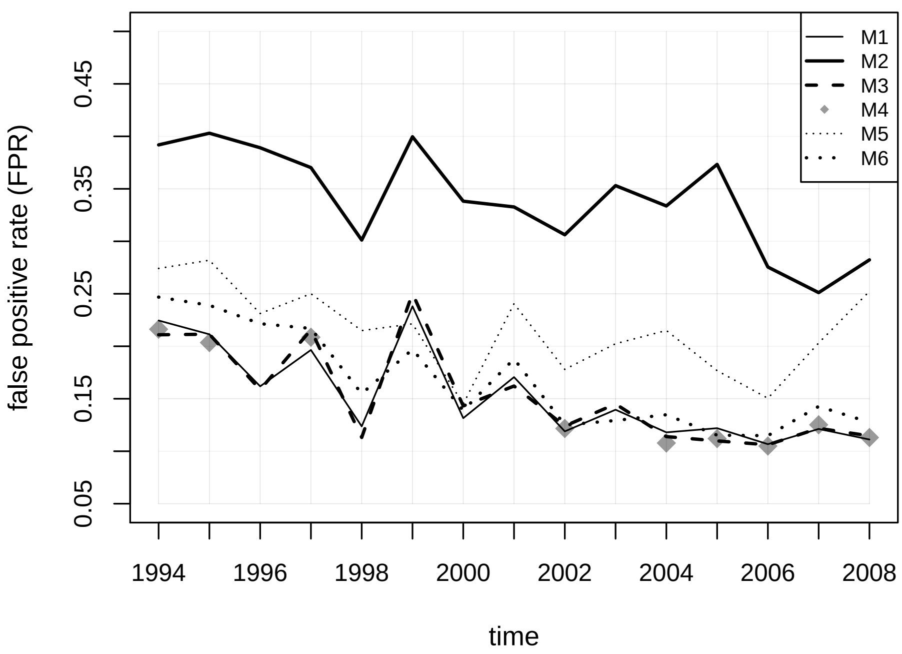
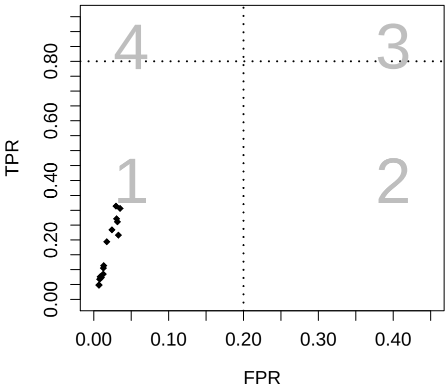
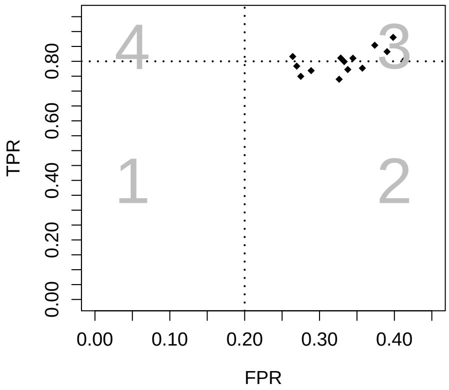


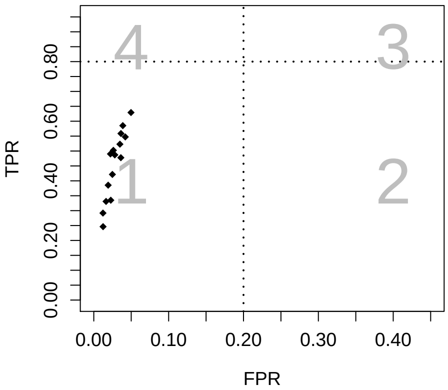
Figure 3: FPR at 80% TPR over time for different models
Source: own elaboration based on SOEP data



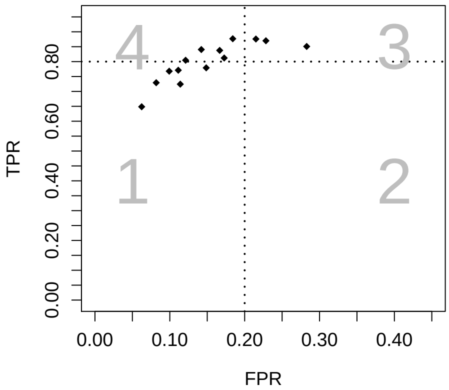
(a) M1, cutoff at 0.5



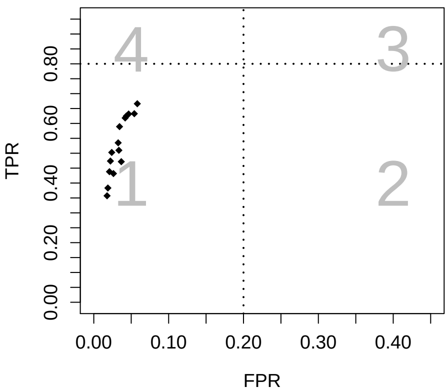
(b) M1, new cutoff



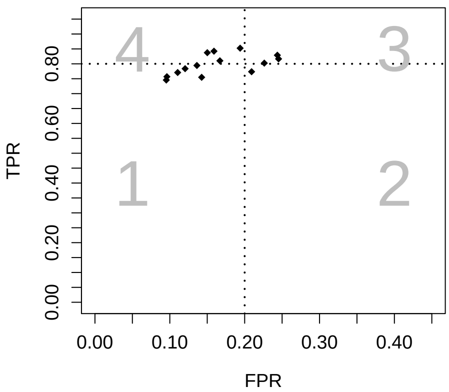
(c) M2, cutoff at 0.5



(d) M2, new cutoff

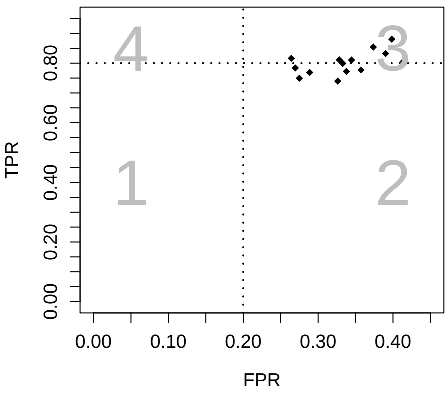


(e) M6, cutoff at 0.5

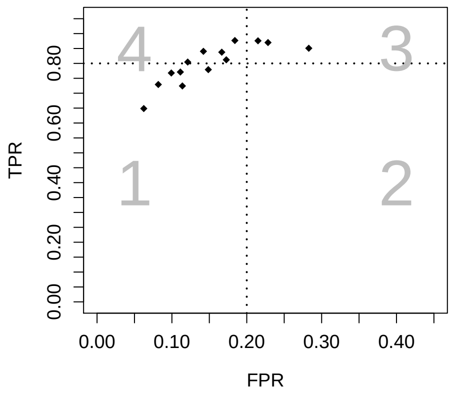


(f) M6, new cutoff

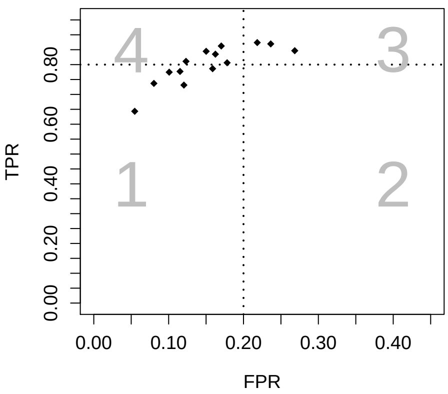
Figure 4: Plots of accuracy for different models. Left panel: cutoff at 0.5 probability of being poor. Right panel: new cutoff using the ROC. Each prediction for each year is represented by a tuple. Best predictions lie in the 4th quadrant, worst predictions in the 2nd quadrant.
Source: own elaboration based on SOEP data



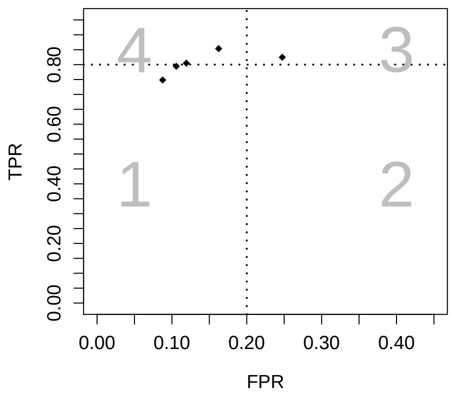
(a) M1



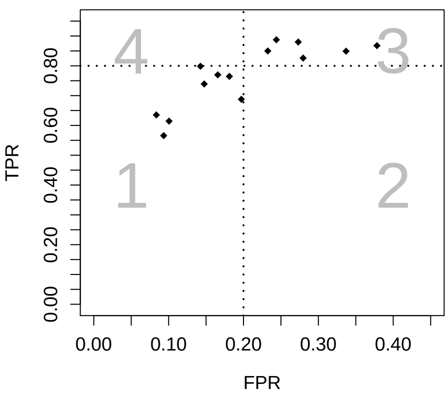
(b) M2



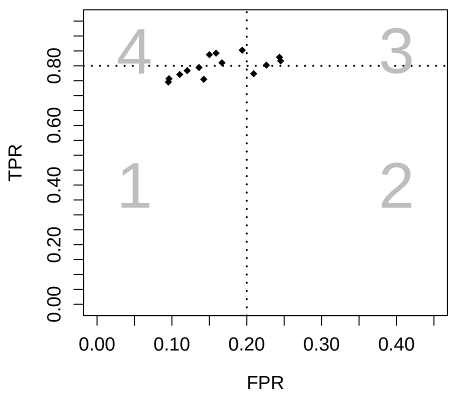
(c) M3



(d) M4

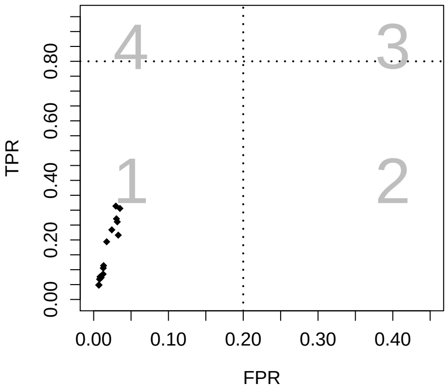


(e) M5

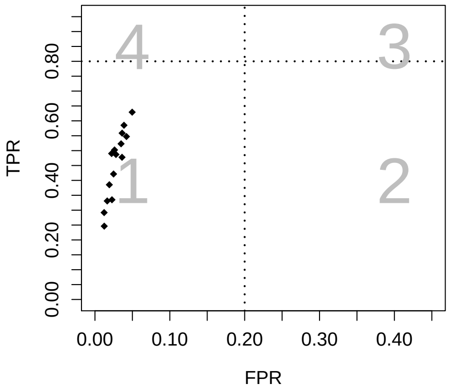


(f) M6

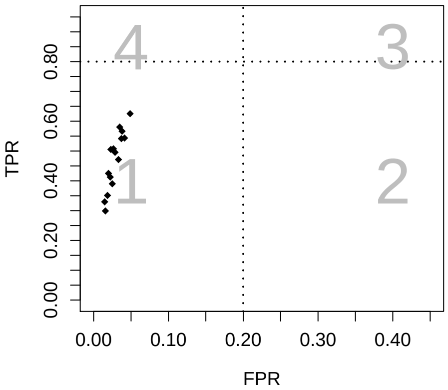
Figure 5: Plots of accuracy for different models and VPL at about 80%. Each prediction for each year is represented by a point. Best predictions lie in the 4th quadrant, worst predictions in the 2nd quadrant. *Source:* own elaboration based on SOEP data



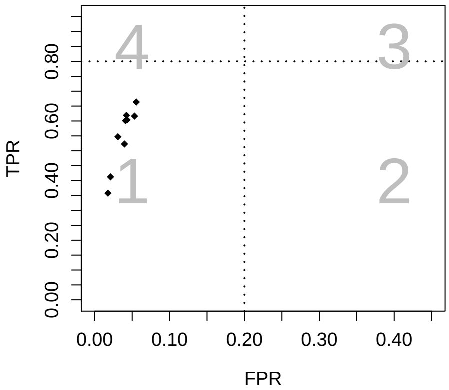
(a) M1



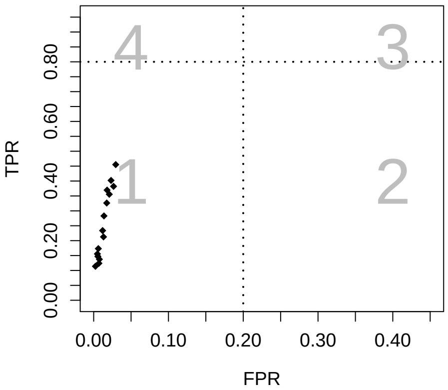
(b) M2



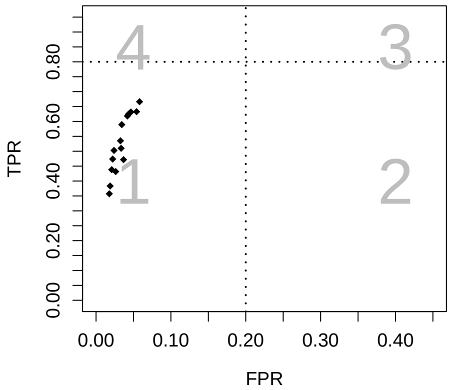
(c) M3



(d) M4



(e) M5



(f) M6

Figure 6: Plots of accuracy for different models and cutoff at 0.5 probability of being poor. Each prediction for each year is represented by a point. Best predictions lie in the 4th quadrant, worst predictions in the 2nd quadrant.
Source: own elaboration based on SOEP data

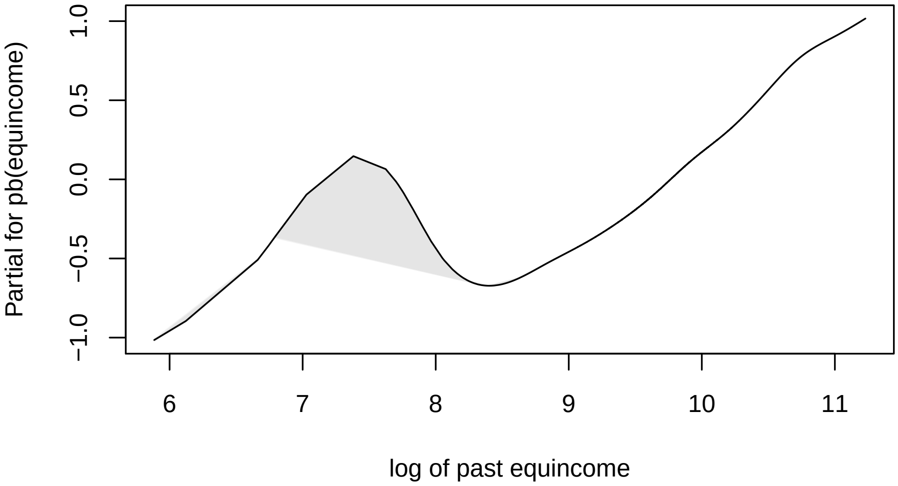


Figure 7: Non-linear effects of past income (1995) on current income (1996) for model M4.
Source: own elaboration based on SOEP data