Things that make us different:

Analysis of deviance with time-use data

Jorge González Chapela

Department of Economics – University of Girona

Carrer de la Universitat de Girona 10, E-17071, Girona, Spain

Email: Jorge.gonzalez@udg.edu, tel. +34 972 418781, fax. +34 972 418032

This version: March 2013

Abstract

The constrained, non-Normal nature of time-use data poses a challenge to ordinary analysis of variance. This paper investigates a computationally simple variance decomposition technique suitable for those data. As a by-product of the analysis, a measure of fit for systems of time-demand equations is proposed that possesses several useful properties.

Keywords: Analysis of deviance; *R*-squared; time allocation

JEL codes: C52, J22

1 Introduction

It is hardly surprising that the collection and statistical analysis of time-use observations has become so prevalent in the social sciences. As individual and collective behavior is reflected in the allocation of time, time-use observations represent a significant resource for analyzing the determinants and consequences of people's acts.[1] Besides, to the extent that evaluated time use provides a useful means of measuring individual and social welfare (e.g., see [21]), inequality in the use of time will increasingly concern social scientists and policy makers.

As part of the investigation of the main features of the allocation of time, researchers must often carry out multivariate analyses of variance (or covariance) on time-use observations. Yet, since these data take typically the form of vector arrays of activity times in which a given amount of time is divided, the error sums of squares and cross products matrix is singular, what precludes performing an ordinary analysis. Hence, researchers must perform activity-by-activity analyses of variance (e.g., see [11, 13, 39]), but this practice makes it harder to discover the main factors underlying the allocation of time as a whole.

Time-use observations also pose a challenge to two alternative approaches to multivariate analysis of variance. As time-use measurements describe quantitatively the parts of a whole, they constitute an example of compositional data (CODA), and a unified statistical methodology for the analysis of CODA is well developed (see [1, 31]). Essentially, CODA analysts transform the original observations onto the Euclidean space and apply classical methods of analysis on the transformed data. Although the transformation is confined to strictly positive observations, which is rarely the case with time-use data, zero-replacement techniques are available that deal with zero elements in the data matrix.[2]

[1] Good collections of such analyses are [16, 32].

[2] For a complete survey of zero-replacement methods see [27].

Nevertheless, a proper application of zero-replacement techniques requires knowing the reason behind the zeros (do they represent real zeros or are the result of inaccurate measurement?), an information that, at best, is only partially gathered by the standard time-use data collection instrument. Stephens [38] avoids the zero-replacement difficulty by taking the square root of each activity time. He then develops an analysis of variance methodology assuming that the so-transformed observations follow a von Mises distribution, which arises from original observations that are realizations of a Normal random vector of unit length [33]. But although the evidence tends to support the normality assumption in the regression analysis of share equations [41], vectors of time-use observations do not generally have unit length.

In this paper I investigate a variance decomposition technique suitable for independent time-use observations. In contrast with previous methodologies I work with the original time-use measurements, but transform the metric with which the discrepancy between these measurements and those predicted by the model being fitted is evaluated. Thus, no special adjustments are needed to handle zero values. My main assumption, discussed in Section 3, is that the conditional mean and conditional variance function of a sample of time-use observations can be represented by those of the Multinomial distribution.

The paper is organized as follows. I start off in Section 2 by reviewing the specification and estimation of complete systems of time-demand equations. Because of its robustness to distributional failure and computational simplicity, the multinomial logit specification and associated quasi-likelihood estimator proposed in [29] is advocated as an attractive statistical approach for time-use data. The statistical theory needed for carrying out variance decompositions within this framework is set out in Section 3. Following [28], the statistical deviance is used as a generalized measure of discrepancy between a sample of time-use observations and a set of fitted values. In Section 4, and as a by-product of the preceding analysis, an R-squared measure for systems of time-demand equations is proposed that possesses several useful properties. This R-squared is an extension of Hauser's [18] pseudo-R^2 for multinomial regression models that can be computed using quasi-likelihood statistics, instead of maximum likelihood statistics as in Hauser's original formulation. Section 5 illustrates these methods on a sample of time-use observations. Section 6 concludes.

2 Specification and Estimation of Complete Systems of Time-Demand Equations

Let the total time analyzed, T, be classified into M mutually exclusive and exhaustive-of-T activities. t_m is the observed amount of time spent on activity m, whereas $\mathbf{x} = (1, x_2, \ldots, x_K)$ represents a $1 \times K$ random vector of explanatory variables containing qualitative factors and quantitative measurements. The linear functional form for the regression of t_m on $\mathbf{x}$ is given by

$$E(t_m | \mathbf{x}) = \mathbf{x}\boldsymbol{\beta}_m, \quad m = 1, \ldots, M, \tag{1}$$

where $\boldsymbol{\beta}_m$ is a $K \times 1$ vector of unknown parameters. As is well known, the constraint $\sum_{m=1}^{M} t_m = T$ precludes carrying out multivariate analyses of covariance in this linear setup, which must be replaced with equation-by-equation analyses.

The most common alternative to (1) in the time use literature has been the multivariate Tobit model (e.g., see [4, 22, 23]):

$$t_m^* = \mathbf{x}\boldsymbol{\beta}_m + \varepsilon_m, \quad m = 1, \ldots, M, \tag{2}$$

$$\boldsymbol{\varepsilon} \equiv (\varepsilon_1, \ldots, \varepsilon_M)' \sim N(\mathbf{0}, \boldsymbol{\Omega}), \tag{3}$$

where t_m^* is the latent amount of time spent on activity m, ε_m is an error term, and $t_m = \max(0, t_m^*)$. In this context, $E(t_m | \mathbf{x}) \geq 0$, whereby predicted times, $\hat{t}_m$, cannot be negative. Complications arise from the need to accommodate the cross-equation restrictions $\sum_{m=1}^{M} t_m^* = T$ and $\sum_{m=1}^{M} \hat{t}_m = T$. Generally, one of the equations in (2) is deleted before estimation [36], but the multivariate Tobit estimator is not invariant to the equation being deleted (see e.g. [8]).

For systems of equations in which the components of the multivariate dependent variable are non-negative, may take on certain values with positive probability, and add up to a constant, Wales and Woodland [40] developed two alternative econometric specifications estimated by Maximum Likelihood (ML). Although both approaches yield parameter estimates with good econometric properties if the assumed conditional density is correctly specified, the technical and computational complexities involved seem to have limited their use in practice. As to behavioral analysis of the allocation of time, only [35] has applied one of Wales and Woodland's specifications. Those complexities provoked the development of alternative two-step censored demand models, which are reviewed in [8]. These two-step models, however, overlooked the cross-equation restrictions $\sum_{m=1}^{M} t_m^* = T$ and $\sum_{m=1}^{M} \hat{t}_m = T$.

Papke and Wooldridge [30] developed an attractive specification as well as a simple quasi-likelihood estimator for a dependent variable bounded between 0 and 1. More recently, Mullahy and Robert [29] have generalized Papke and Wooldridge's approach to the context of complete systems of time-demand equations where the total time analyzed is normalized to 1. The population regression considered in [29] is of the multinomial logit form,

$$E\left(y_m \middle| \mathbf{x}\right) = \frac{\exp\left(\mathbf{x}\boldsymbol{\beta}_m\right)}{\sum_{k=1}^{M} \exp\left(\mathbf{x}\boldsymbol{\beta}_k\right)}, \quad m = 1,\ldots,M, \tag{4}$$

where $y_m = t_m / T$, $m = 1,\ldots,M$. This nonlinear specification ensures that $\hat{y}_m$ lies between 0 and 1, that $\sum_{m=1}^{M} \hat{y}_m = 1$, and that the partial effect of x_j on $E\left(y_m \middle| \mathbf{x}\right)$ is not constant but dependent on $\mathbf{x}$. Another interesting feature is that equation (4) is well defined even if every

y_m takes on 0 or 1 with positive probability. The normalization $\boldsymbol{\beta}_1 = \mathbf{0}$ is generally imposed for identification purposes.

A particular quasi-likelihood method is advocated in [29] to estimate the parameters of (4). The multinomial logit log-likelihood function

$$l(\mathbf{b}) = \sum_{m=1}^{M} y_m \left(\mathbf{x}\mathbf{b}_m - \ln\left(\sum_{k=1}^{M} \exp(\mathbf{x}\mathbf{b}_k) \right) \right), \tag{5}$$

where $\mathbf{b} = (\mathbf{0}', \mathbf{b}_2', \ldots, \mathbf{b}_M')'$ is a generic element of the parameter space, is an objective function associated with the linear exponential family (LEF) of probability distributions. Given a sample of N independent observations $\{(\mathbf{y}_i, \mathbf{x}_i) : i = 1, \ldots, N\}$, where $\mathbf{y}_i = (y_{i1}, \ldots, y_{iM})'$, the quasi-maximum likelihood estimator (QMLE) of $\boldsymbol{\beta} = (\mathbf{0}', \boldsymbol{\beta}_2', \ldots, \boldsymbol{\beta}_M')'$, $\hat{\boldsymbol{\beta}}$, obtained from the maximization problem

$$\max_{\mathbf{b}} \sum_{i=1}^{N} l_i(\mathbf{b}) \tag{6}$$

is consistent for $\boldsymbol{\beta}$ and asymptotically normal provided that (4) holds.[3] In other words, although the conditional probability distribution of the random vector $\mathbf{y}$ cannot be Multinomial, if its conditional mean is correctly specified the fact that the assumed probability distribution is linear exponential makes the QMLE to have satisfying econometric properties regardless of the true conditional distribution of $\mathbf{y}$. The correct specification of the conditional mean, i.e. $E\left(y_m - \frac{\exp(\mathbf{x}\boldsymbol{\beta}_m)}{\sum_{k=1}^{M} \exp(\mathbf{x}\boldsymbol{\beta}_k)} \middle| \mathbf{x} \right) = 0, \quad m = 1, \ldots, M$, can be tested by an m-

[3] A good exposition of the properties of QML estimators is provided in [14].

test of $E\left(\left(y_m - \frac{\exp(\mathbf{x}\boldsymbol{\beta}_m)}{\sum_{k=1}^{M}\exp(\mathbf{x}\boldsymbol{\beta}_k)}\right)\mathbf{z}\right) = \mathbf{0}, \quad m = 1,\ldots,M$, where $\mathbf{z}$ may be a function of $\mathbf{x}$ but cannot equal $\mathbf{x}$ (see [42] for further details).

Although the QMLE is technically more complex than the equation-by-equation Ordinary Least Squares (OLS) estimator, it benefits from a more sensible functional form for the conditional mean (the OLS estimator, for instance, does not guarantee that $\hat{t}_m$ will lie in the unit interval), and is not much more difficult to compute than OLS, as it can be implemented using minor modifications of ordinary multinomial logit estimation algorithms (already incorporated in, for example, the estimation command *fmlogit* of the statistics package STATA®). More importantly for our purposes, the QMLE offers the possibility of analyzing the determinants of the allocation of time as a whole. In comparison with the MLEs proposed by Wales and Woodland, the multinomial logit QMLE is technically much easier and more robust. Although the stochastic component in any of Wales and Woodland's formulations is assumed to follow a Multivariate Normal distribution, the derived distribution of observed shares does not generally follow a linear exponential probability model. As a result, misspecification of any aspect of the density would lead to inconsistency of the MLE.

The asymptotic covariance matrix of the multinomial logit QMLE shares the general shape of the QMLE variance matrix given in [14]. But if observations are independent and

$$V\left(\mathbf{y}_i \middle| \mathbf{x}_i\right) = \sigma^2 \mathbf{V}_i, \tag{7}$$

where σ^2 denotes a dispersion parameter, $\mathbf{V}_i$ represents a variance function with *mk*th element $p_{im}\left(\delta_{imk} - p_{ik}\right)$,

$$p_{im} = \frac{\exp(\mathbf{x}_i\boldsymbol{\beta}_m)}{\sum_{k=1}^{M}\exp(\mathbf{x}_i\boldsymbol{\beta}_k)}, \tag{8}$$

and δ_{imk} is an indicator variable equal to 1 if $m = k$ and equal to 0 if $m \neq k$, the asymptotic covariance matrix of $\hat{\boldsymbol{\beta}}$ could be simply estimated as

$$-\hat{\sigma}^2 \hat{\mathbf{A}}^{-1}, \tag{9}$$

where, following [28],

$$\hat{\sigma}^2 = \left((M-1)(N-K)\right)^{-1} \sum_{im} \left(y_{im} - \hat{p}_{im}\right)^2 \Big/ \hat{p}_{im}\left(1-\hat{p}_{im}\right) \tag{10}$$

$$\hat{\mathbf{A}} = -\sum_{i=1}^{N} \left(\hat{\mathbf{V}}_i \otimes \mathbf{x}_i' \mathbf{x}_i\right), \tag{11}$$

the symbol $\otimes$ denoting the Kronecker product. In other words, if (7) holds the multinomial logit QMLE is efficient in the class of all QMLEs in the LEF.

3 Variance Decomposition in a System of Time-Demand Equations

The literature on generalized linear models has extended the analysis of variance to certain non-linear contexts based on the concept of deviance. Let $f_{\mathbf{y}_i}$ and $f_{\mathbf{p}}$ denote two absolutely continuous probability distributions associated to the random vector $\mathbf{y}$. $f_{\mathbf{y}_i}$ is centered at a realization of $\mathbf{y}$, whereas $f_{\mathbf{p}}$ is centered at $E\left[\mathbf{y}|\mathbf{x}\right] = \mathbf{p}$. The Kullback-Leibler (KL) divergence between $f_{\mathbf{y}_i}$ and $f_{\mathbf{p}}$ is simply given by

$$K\left(\mathbf{y}_i, \mathbf{p}\right) = 2\ln\left(f_{\mathbf{y}_i}\left(\mathbf{y}_i\right) \big/ f_{\mathbf{p}}\left(\mathbf{y}_i\right)\right) \tag{12}$$

when f belongs to the LEF of probability distributions [9]. Expression (12) measures the deviance between the probability distributions $f_{\mathbf{y}_i}$ and $f_{\mathbf{p}}$ at the point $\mathbf{y}_i$. Given N independent observations and a particular data-generating process for the data ($f_{\mathbf{p}}$), the estimated deviance between observations $\mathbf{Y} = \left(\mathbf{y}_1, \ldots, \mathbf{y}_N\right)$ and fitted values $\hat{\mathbf{P}} = \left(\hat{\mathbf{p}}_1, \ldots, \hat{\mathbf{p}}_N\right)$ is

$$K\left(\mathbf{Y}, \hat{\mathbf{P}}\right) = 2\sum_{i=1}^{N} \left(\ln f_{\mathbf{y}_i}\left(\mathbf{y}_i\right) - \ln f_{\hat{\mathbf{p}}_i}\left(\mathbf{y}_i\right)\right). \tag{13}$$

The difference $K\left(\mathbf{Y},\hat{\mathbf{P}}_{\mathbf{0}}\right)-K\left(\mathbf{Y},\hat{\mathbf{P}}\right)$, where the sub-index $\mathbf{0}$ refers to the null model, is measuring the reduction in deviance achieved by the inclusion of explanatory variables.

Most often the data-generating process is unknown, whereby certain features of the data are to be specified for calculating the deviance. Suppose that the conditional mean and the conditional variance function of $\mathbf{y}$ can be represented by those of a linear exponential probability distribution (denoted f in expression (15)). Then, McCullagh and Nelder [28, Ch. 9] showed that the deviance between observations and fitted values can be computed as

$$-2Q\left(\hat{\mathbf{P}};\mathbf{Y}\right), \tag{14}$$

where

$$Q\left(\hat{\mathbf{P}};\mathbf{Y}\right)=\sum_{i=1}^{N}\left(\ln f_{\hat{\mathbf{p}}_i}\left(\mathbf{y}_i\right)-\ln f_{\mathbf{y}_i}\left(\mathbf{y}_i\right)\right) \tag{15}$$

is the estimated quasi-likelihood for $f_{\mathbf{p}}$ based on data $\mathbf{y}$. Although $\mathbf{y}_i$ might not belong to the support of f, support restrictions are irrelevant for calculating the quasi-likelihood, the important point being the correct specification of the conditional mean and variance functions of the data.

For reasons given in Section 2, expression (4), which is the mean of a (one-trial) Multinomial distribution, seems a reasonable model for the population regression of a vector of relative time shares. In its turn, this adequacy suggests that the covariance matrix specified in (7)-(8), which, except for the presence of σ^2, coincides with that of the Multinomial distribution, could serve as a model for the covariance structure of relative time shares. It is well known, however, that the Multinomial covariance structure presents an important restriction in the fact that all elements outside its main diagonal are negative, which has limited its use in practice. Yet, although the standard time-use data collection instrument permits distinguishing a large number of activities, time-use researchers typically end up analyzing a few aggregates (e.g., the number of activities considered in [11, 13, 39] ranges between two and four), what in my experience increases the likelihood of observing negative unconditional activity correlations. Thus, as the structure of $V(\mathbf{y})$ is partly reflecting that of $V(\mathbf{y}|\mathbf{x})$, the adequacy of the Multinomial covariance structure seems to increase when M is low. Assuming (4) and (7)-(8) are appropriate for the data at hand, the quasi-likelihood is given by

$$Q(\mathbf{P};\mathbf{Y}) = \sum_{i=1}^{N}\left(\sum_{m=1}^{M} y_{im}\left(\mathbf{x}_i\mathbf{b}_m - \ln\left(\sum_{k=1}^{M}\exp(\mathbf{x}_i\mathbf{b}_k)\right)\right) - \sum_{m=1}^{M} y_{im}\ln y_{im}\right), \tag{16}$$

the data total deviance is $-2Q\left(\hat{\mathbf{P}}_{\mathbf{0}};\mathbf{Y}\right)$, and $-2\left(Q\left(\hat{\mathbf{P}}_{\mathbf{0}};\mathbf{Y}\right)-Q\left(\hat{\mathbf{P}};\mathbf{Y}\right)\right)$ is the reduction in deviance achieved by the inclusion of explanatory variables.

An interesting property of LEF models that use the canonical link is that the KL divergence exhibits the Pythagorean property,

$$K\left(\mathbf{Y},\hat{\mathbf{P}}_{\mathbf{0}}\right) = K\left(\mathbf{Y},\hat{\mathbf{P}}\right) + K\left(\hat{\mathbf{P}},\hat{\mathbf{P}}_{\mathbf{0}}\right), \tag{17}$$

(see [17, 37]), whereby the difference $K\left(\mathbf{Y},\hat{\mathbf{P}}_{\mathbf{0}}\right) - K\left(\mathbf{Y},\hat{\mathbf{P}}\right)$ can be interpreted not only as the reduction in deviance due to inclusion of explanatory variables, but also as the deviance explained by the regression model, $K\left(\hat{\mathbf{P}},\hat{\mathbf{P}}_{\mathbf{0}}\right)$. Since the mean function specified in (4) corresponds to the canonical link of the Multinomial distribution, then

$$-2Q\left(\hat{\mathbf{P}}_{\mathbf{0}};\mathbf{Y}\right) = -2\left(Q\left(\hat{\mathbf{P}};\mathbf{Y}\right) + Q\left(\hat{\mathbf{P}}_{\mathbf{0}};\hat{\mathbf{P}}\right)\right), \tag{18}$$

so that the difference $-2\left(Q\left(\hat{\mathbf{P}}_{\mathbf{0}};\mathbf{Y}\right)-Q\left(\hat{\mathbf{P}};\mathbf{Y}\right)\right)$ is the deviance explained by the regression model.

Sometimes time-use observations are better modeled as being dependent within groups but independent among groups, such as when the sample contains more than one observation for each of many individuals or families. In such a case, computing the deviance as shown in (15) is misleading because the quasi-likelihood's shape assumes that observations are independent. If the true variance function of observations belonging to group g, $\mathbf{V}_g$, $g = 1,\ldots,G$, were known, one could derive the associated form for the joint distribution of the repeated measurements. It turns out, however, that it is not necessary to correctly model $\mathbf{V}_g$ to obtain consistent estimates of $\boldsymbol{\beta}$ and $\mathbf{V}\left(\hat{\boldsymbol{\beta}}\right)$, see [24, 44], whereby researchers most often specify a "working" $\mathbf{V}_g$, $\bar{\mathbf{V}}_g$, and estimate $\boldsymbol{\beta}$ by solving the generalized estimating equations (GEEs)

$$\sum_{g=1}^{G} \mathbf{D}'_g \bar{\mathbf{V}}_g^{-1} \mathbf{S}_g = \mathbf{0}, \tag{19}$$

where $\mathbf{D}_g = dE\left[\mathbf{Y}_g \middle| \mathbf{X}_g\right] \big/ d\boldsymbol{\beta}$, $\mathbf{S}_g = \mathbf{Y}_g - E\left[\mathbf{Y}_g \middle| \mathbf{X}_g\right]$, and $\mathbf{Y}_g$ and $\mathbf{X}_g$ denote, respectively, the observations belonging to group g and a matrix with rows $\mathbf{x}_i$, $i \in g$. Since GEEs do not rely on the quasi-likelihood, model comparisons are not based on the likelihood-ratio test but on other asymptotically equivalent procedures.

4 An *R*-Squared Measure of Goodness of Fit for Systems of Time-Demand Equations

The commonly reported goodness-of-fit statistic in the standard linear regression model, R^2, is troublesome when applied to nonlinear contexts, as it can lie outside the $[0,1]$ interval and decrease as explanatory variables are added. For this reason, alternative R^2-type statistics, generally called pseudo-R^2s, have been constructed for particular nonlinear models using a variety of methods. Hauser [18], for instance, proposed a pseudo-R^2 for multinomial regression models calculated using ML statistics, which, among other satisfying properties, lies between 0 and 1 and is non-decreasing as explanatory variables are added. Later on, Cameron and Windmeijer [5] developed a pseudo-R^2 based on the KL divergence for exponential family regression models estimated by ML, of which Hauser's goodness-of-fit statistic is a particular case. In this section, I draw upon [5] to extend Hauser's pseudo-R^2 to be computed using QML statistics, and reinterpret Hauser's pseudo-R^2 in the light of the deviance measure of discrepancy. A possible extension of Cameron and Windmeijer's pseudo-R^2 measure to be computed using QML statistics is left for future research.

Under the conditions that let $-2Q$ to be a measure of deviance, a measure of the proportionate reduction in total deviance achieved by the regression model can be calculated as:

$$R_Q^2 = 1 - Q\left(\hat{\mathbf{P}};\mathbf{Y}\right) \big/ Q\left(\hat{\mathbf{P}}_\mathbf{0};\mathbf{Y}\right). \tag{20}$$

This R_Q^2 has the following properties:

1. R_Q^2 is non-decreasing as explanatory variables are added. Proof: The QMLE maximizes $Q(\mathbf{P};\mathbf{Y})$, which will therefore not decrease as explanatory variables are added, i.e. as constraints on the coefficients are removed.
2. $0 \le R_Q^2 \le 1$. Proof: The lower bound of 0 occurs when the inclusion of explanatory variables leaves the fitted values unchanged, and the upper bound occurs when the model fit is perfect.
3. R_Q^2 is a scalar multiple of the quasi-likelihood ratio (QLR) test statistic for the hypothesis that the coefficients of all the explanatory variables save the constant are 0. Proof: Re-expressing R_Q^2 as $\left(\sum_{i=1}^N \ln f_{\hat{\mathbf{p}}_0}\left(\mathbf{y}_i\right) - \sum_{i=1}^N \ln f_{\hat{\mathbf{p}}_i}\left(\mathbf{y}_i\right)\right) \big/ Q\left(\hat{\mathbf{P}}_\mathbf{0};\mathbf{Y}\right)$, it turns out that $R_Q^2 = \dfrac{\hat{\sigma}_u^2}{-2Q\left(\hat{\mathbf{P}}_\mathbf{0};\mathbf{Y}\right)} QLR$, where $\hat{\sigma}_u^2$ is a consistent estimate of σ^2 calculated as shown in expression (10) using results from unrestricted estimation (see [43, p. 370]).

4. R_Q^2 can be equivalently expressed as

$$R_Q^2 = \frac{Q\left(\hat{\mathbf{P}}_0; \hat{\mathbf{P}}\right)}{Q\left(\hat{\mathbf{P}}_0; \mathbf{Y}\right)}, \tag{21}$$

where $Q\left(\hat{\mathbf{P}}_0; \hat{\mathbf{P}}\right)$ is (up to the factor -2) the estimated deviance between the null and fitted models. Hence, R_Q^2 can also be interpreted as the fraction of deviance explained by the fitted model. Proof: See the discussion surrounding expression (18).

Properties 1 and 2 are standard properties often desired for *R*-squared measures. Property 3 generalizes a similar result for the linear regression model under normality [2] and has the practical intent of avoiding conflicting signals between the ranking of models generated by R_Q^2 and the related statistical test. Property 4 is also desirable as it allows R_Q^2 to be interpreted similarly as the usual R^2 in the linear regression context: either as the proportionate reduction in deviance due to the inclusion of explanatory variables, or as the fraction of deviance explained by the regression model.

5 Application

The involvement of fathers in children's cognitive and non-cognitive development has gained considerable attention in the socio-economic literature (e.g., see [3, 25, 34] and the many references cited therein). Part of this interest has been motivated by the sharp increase in maternal employment over the last decades, which has stimulated research on the consequences of father care for children's development. Moreover, the inclusion of the father's time in the production process of child outcomes has contributed to better understand phenomena such as birth order differences in child results. In order to illustrate the previous methods, I shall study in this section the division of paternal time between formative and non-formative activities in the child's 3-9 years of age when the mother is working.

The data for the analysis are taken from the Spanish Time Use Survey (STUS) 2002-2003, the first full-scale national time use survey conducted in Spain. As is now standard around the world, the time use information in the STUS 2002-2003 was collected by the time diary method: All persons aged 10 years and older in the interviewed households were asked to list their main activity in each 10-minute interval of a complete 24-hours cycle. These activities were then classified by the survey agency into standardized Eurostat codes (see [10, Annex VI]).[4] Respondents also reported whether members of their household under 10 years of age were present during the activity, as well as the age, sex, and relationship to the respondent of each person in the household. This information makes it possible to construct a very accurate measure of the time the father spends with his children under 10. In view of the evidence that parental time investments in children's development become less effective during adolescence [7], this age limitation does not seem very important for this study.

Following [34], I measure formative time as the time the father spends with his children reading, doing homework, doing arts and crafts, doing sport, playing, attending performances and museums, engaging in religious activity, having meals and talking with the children, or providing personal care for the children. The positive relationship between the frequency of activities such as reading, playing, or eating with children and their development is well documented in the literature (e.g., see [6, 34]).

The analysis is limited to fathers residing in two-earner families whose youngest child is between 3 and 9 years of age. I have excluded families whose youngest child is under 3 because some of the above-mentioned formative activities typically involve children older

[4] The STUS development and design followed the guidelines established by Eurostat in 2000 though published in 2004, see [10, 19]. The average number of activity episodes per day, 21.5, indicates diary data of reasonable quality [20].

than infants or toddlers. I also discarded families presenting invalid or missing data in any of the explanatory variables listed below. These restrictions left a sample of 762 fathers, whose use of time on the diary day was classified into two broad activities: formative time (whose relative share of daily time is denoted y_1) and all other activities (y_2). I distinguish only two activities because the data will be also analyzed using one of Wales and Woodland's approaches, which, when $M > 2$, proved difficult to estimate and do not accommodate well usual pseudo-R^2s.[5]

Table 1 presents sample statistics. On average, fathers invest 1.8 hours per day in their children, though the standard deviation (1.8 hours), coefficient of variation (1.0), and percentage of zeros (17.2) reveal substantial variation in time investments. There is some evidence that some of these zeros might be due to the imprecision of the measurement instrument, for if two activities are carried out simultaneously, only time devoted to the main activity is recorded. Thus, for example, 35.1 percent of fathers reporting no formative time

[5] The derived distribution of the observed shares in any of Wales and Woodland's approaches does not follow a linear exponential model, what discards the *R*-squared proposed in [5]. When all M activities are carried out, the density associated with $\mathbf{y}_i$ is given by (or is proportional to) the Multivariate Normal density function (see expressions (7) and (19) in [40]). This fact discards the so-called Relative Gain, $R^2_{RG} = 1 - \frac{\ell_{\max} - \ell_{fit}}{\ell_{\max} - \ell_0}$, where ℓ_0 denotes the value of the log-likelihood function in the intercept-only model, ℓ_{fit} is the value in the fitted model, and $\ell_{\max}$ represents the largest possible value of the log-likelihood: Under normality, $\ell_{\max} \to \infty$ as the determinant of the covariance matrix of $\mathbf{y}$ tends to zero. The related measure $1 - \ell_{fit}/\ell_0$ is not bounded between zero and one because with continuous data the log-likelihood can be negative or positive.

declare to be having meals[6] in the presence of children simultaneously with non-formative main activities (although we cannot know if they are talking with the children). Some zeros, however, might be simply indicating that the father has not invested time in their children. In these cases, it is not possible to distinguish whether the zero is essential (i.e. the result of never investing time in the children), or most probably created by the limited observation period (i.e. the result of not having invested time on the diary day). Since $M = 2$, both the unconditional- and conditional-on-**x** correlations between y_1 and y_2 are negative and equal to -1.

The literature examining the determinants of father care has focused on factors such as the opportunity cost of providing child care, the father's and the mother's availability, and the sex and age of the children living in the household. As regards formative activities, it also seems reasonable to suppose that the father's education may influence the time he invests in his children. For these reasons, the 11 variables whose inclusion in $\mathbf{x}$ is to be assessed are: the father's and the mother's net monthly earnings and work schedules, whether the mother works part-time, the level of schooling attained by the father, the age of the youngest child, the number of children aged 3-9, whether at least one child is a boy, the trimester of interview, and whether the diary pertains to a weekend day. (Below, in a non-nested model selection exercise, we shall discuss the introduction of the father's age as an additional explanatory variable.) Earnings and schooling will be represented by a set of dummy variables whose cardinality is given by the number of answer alternatives in the corresponding survey question. Similar to [15], indicators for nontraditional work hours are constructed from the weekly work schedule module of the STUS 2002-2003. I will consider

[6] Unfortunately, secondary activities are reported with much lower detail than main activities, and it is not possible to distinguish other formative activities beyond having meals.

that a worker works in the evening (respectively, at night) if s/he was on the job at any time between 7PM and 10PM (10PM and 6AM) at least one day of the working week. Additional information on the explanatory variables is given in Table 1.

Table 2 presents the results of a partial analysis of deviance of the use of time in our sample of Spanish fathers. (For table layout simplicity, possible overlaps among the explanatory variables are ignored.) The partial deviance is calculated as the deviance explained by all 11 explanatory variables minus the deviance in the sub-model in which the variable of interest is removed. Table 2 also lists the value of the QLR statistic computed from the difference in the quasi-likelihood with and without the restrictions imposed,

$$QLR = \frac{-2\left(\sum_{i=1}^{N} \ln f_{\hat{\mathbf{p}}_r}\left(\mathbf{y}_i\right) - \sum_{i=1}^{N} \ln f_{\hat{\mathbf{p}}_u}\left(\mathbf{y}_i\right)\right)}{\hat{\sigma}_u^2}, \tag{22}$$

the degrees of freedom of the QLR statistic limiting distribution ($QLR\left(df\right)$), and the *p*-value for testing the statistical significance of each explanatory variable and of the overall model. Under the null, QLR has a χ^2 limiting distribution with df given by the number of restrictions being tested. Since in this study each explanatory term can have associated non-zero coefficients in one activity equation only, $QLR\left(df\right)$ will tally with the number of terms representing the variable whose statistical significance is being evaluated.

The total deviance in the sample amounts to 59.12. When all 11 explanatory factors are included in $\mathbf{x}$, the model is able to explain 11.17 units of this deviance, implying an R_Q^2 of size 0.189. The indicator for weekend day is the major contributor to deviance in Spanish fathers' allocation of time between investments in children's development and all other activities: Its partial deviance represents 46.3 percent of the total deviance. (On average, the fraction of daily time the father invests in his children's development increases 0.045 on weekends, *S.E.*=0.005.) There is also evidence of sizeable effects associated to the age of the youngest child (which is inversely related to the amount of formative time provided by the father) and the father's work schedule (working at night and, especially, in the evening reduce time spent on children's development). The *p*-value for testing the exclusion of each of these variables is well below 0.05, and therefore they contribute significantly to the predictive ability of the model even when all other 10 factors are included in $\mathbf{x}$. On the other hand, very modest effects are associated in general to the other explanatory variables, which considered individually do not serve as significant predictors.

For the sake of comparison, Tables 3 and 4 present the set of model comparisons listed in Table 2 performed under two alternative statistical approaches. In Table 3 the underlying specification is the standard Tobit model, which, when $M = 2$ and the y_m not deleted (in this case y_1) presents no ones, matches the Amemiya-Tobin approach of [40]. Results in Table 4 were obtained by the standard CODA methodology. Firstly, the composition with observed time-use patterns, $\mathbf{y}_i$, was replaced by a new composition $\mathbf{yr}_i = \left(yr_{i1}, yr_{i2}\right)$ presenting no zeros. The lack of information on the nature of the zeros complicated choosing a zero-replacement procedure. For its simplicity, I worked with the non-parametric multiplicative replacement discussed in [27]. I assumed that if two activities were carried out simultaneously in a 10-minute interval, the survey respondent would consider as the main activity the one absorbing most of the 10 minutes. Hence,

$$yr_{im} = \begin{cases} 0.002\phi_{im} & \text{if } y_{im} = 0 \\ y_{im}\left(1 - 0.002 \sum\limits_{k|y_{ik}=0} \phi_{ik}\right) & \text{if } y_{im} > 0 \end{cases} \tag{23}$$

where 0.002 days is the 65 percent of the detection value[7] (5 minutes) and ϕ_{im} is the number of 10-minute intervals in which activity m was carried out secondarily by individual i on the diary day. The problem with this approach is that ϕ_{i1} is unknown (see note 6). Hence, I selected a common ϕ_{i1} by maximizing the adequacy-to-data (as defined next) of a model including the 11 explanatory variables. The selected ϕ_{i1} was $\phi_{i1} = 9$, what suggests that the daily time a father invests in his children could be 45 minutes higher on average. Secondly, the additive log-ratio transformation was applied to $\mathbf{yr}_i$, which is perhaps the most popular transformation in economics, [12] argues, yielding $h_i = \ln\left(yr_{i1}/yr_{i2}\right)$. Finally, a standard analysis of covariance was applied to the transformed data. In Tables 3 and 4, model adequacy is measured with the squared correlation between y_1 and $\hat{y}_1$,[8] and the significance of the explanatory variables is tested with the likelihood ratio (*LR*) statistic.

The Tobit model presents the best fit to data, 0.194, whereas the CODA model presents the poorest, 0.186. The differences with respect to R_Q^2, however, are small. As regards model comparison tests, all conclusions achieved in Table 2 are supported. A very sizable weekend day effect is again evident, as well as significant effects associated to the age

[7] Martín-Fernández et al. [26] have found that this imputation value minimizes the distortion of the data covariance structure when the proportion of zeros is less than 10 percent. Since the proportion of zeros in y_1 is somewhat larger, I have experimented with 0.001 ϕ_{im}. This did not change the main conclusions, although the selected ϕ_{i1} became $\phi_{i1} = 17$.

[8] This *R*-squared, used in [43, p. 529] to assess the adequacy of the Tobit model, lies between 0 and 1, but can decrease as explanatory variables are added. In the CODA context, $\hat{y}_1$ was obtained using the additive logistic transform $\hat{y}_1 = \dfrac{\exp\left(\hat{h}\right)}{1+\exp\left(\hat{h}\right)}$.

of the youngest child and the father's work schedule. (Tobit and CODA estimates for these variables present the same sign as QML estimates. According to the Tobit model, the fraction of daily time invested in children's development increases, on average, 0.038 on weekends, *S.E.*=0.005, whereas the corresponding average partial effect estimated by the CODA model is 0.031, *S.E.*=0.004.) At the 0.05 level, all the other explanatory factors are individually insignificant for explaining Spanish fathers' allocation of time. The non-rejection of the null, however, is established with less confidence in general in the Tobit and CODA approaches. For the latter, I have experimented with $\phi_{i1} = 1$ and $\phi_{i1} = 5$: Although the *R*-squared dropped by construction somewhat (0.169 and 0.184, respectively), model comparison results remained unaltered. Admittedly, a two-activity setup suits particularly well the Multinomial covariance structure (as the conditional-on-**x** correlation between y_1 and y_2 is then negative). Yet, a previous version of this article compared our approach to the CODA model in a 4-activity context, finding again a similar ranking of model comparisons and a slightly better fit to data by our approach.

R_Q^2 can also be used to select among alternative non-nested models, provided that they contain the same number of parameters. Suppose, for example, that we wanted to add each father's age to the set of explanatory variables considered so far, but we wondered whether it is better (in predictive ability terms) to add age in levels or in natural logarithm units. It turns out that when age in levels is added to $\mathbf{x}$, R_Q^2 increases to 0.1903, whereas it becomes 0.1900 when log of age is included in $\mathbf{x}$. The same decision would be adopted in the Tobit model, where the *R*-squared increases to 0.1955 in the former case and to 0.1953 in the latter, and in the CODA context, where the *R*-squared with age added in levels grows to 0.1872, but only achieves 0.1871 with age added in logs. In any of the three models, neither the father's age nor its log achieves statistical significance at the 0.05 level.

6. Conclusion

We have argued that the multinomial logit regression specification and associated quasi-likelihood estimator make up an attractive statistical approach for time-use observations. Within this framework, a variance decomposition exercise can be performed if, in addition, we are willing to specify the conditional variance function of the time-use measurements as that of the Multinomial distribution. Changes in the value of the quasi-likelihood permit then assessing the contribution of an explanatory variable to the total deviance observed in the data. From these quasi-likelihood statistics a measure of fit can be constructed that has the right interpretation at the limits of the unit interval as well as an intuitively appealing interpretation between these limits. An empirical application has illustrated the usefulness of these methods. A partial analysis of deviance revealed that the day of the week, the children's age, and the father's work schedule are the major contributors to deviance of Spanish fathers' allocation of time between children's development activities and all other activities. The R-squared proposed in this paper indicated that age of the father in levels fits better these data than its natural logarithm does.

Acknowledgements

Part of this paper was written during my stay at the University of Alicante, whose hospitality I gratefully appreciate. I wish to thank the anonymous referees, Lola Collado, Daniel Hamermesh, Goeran Kauermann, Angel López, Josep Antoni Martín, Andres Romeu, Jay Stewart, and seminar participants at Alicante, Cartagena, Lueneburg, and at the Departments of Economics and of Computer Science and Applied Mathematics of the University of Girona for helpful comments and encouragement. Financial support from the Instituto Valenciano de Investigaciones Economicas and the Spanish Ministry of Education (ECO2011-29751/ECON) is gratefully acknowledged.

References

[1] Aitchison, J. 1986. *The Statistical Analysis of Compositional Data*. London and New York: Chapman and Hall.

[2] Anderson, T.W. 2003. *An Introduction to Multivariate Statistical Analysis*. 3rd edition. New York: Wiley.

[3] Averett, Susan L., Lisa A. Gennetian, H. Elizabeth Peters. 2005. Paternal child care and children's development. *Journal of Population Economics* 18:391-414.

[4] Bloemen, Hans G., Silvia Pasqua, and Elena G. F. Stancanelli. 2010. An empirical analysis of the time allocation of Italian couples: Are they responsive? *Review of Economics of the Household* 8:345-369.

[5] Cameron, A.C. and F.A.G. Windmeijer. 1997. An R-squared measure of goodness of fit for some common nonlinear regression models. *Journal of Econometrics* 77:329-342.

[6] Del Boca, Daniela, Christopher Flinn, and Matthew Wiswall. 2010. Household choices and child development. Carlo Alberto Notebooks 149. Collegio Carlo Alberto.

[7] Del Boca, Daniela, Chiara Monfardini, Cheti Nicoletti. 2012. Children's and Parents' time-use choices and cognitive development during adolescence. Economic Research Center working paper No. 2012-006, University of Chicago.

[8] Dong, D., B.W. Gould, and H.M. Kaiser. 2004. Food demand in Mexico: an application of the Amemiya-Tobin approach to the estimation of a censored food system. *American Journal of Agricultural Economics* 86(4): 1094-1107.

[9] Efron, B. 1978. The geometry of exponential families. *The Annals of Statistics* 6(2):362-376.

[10] Eurostat. 2004. *Guidelines on harmonized European time use surveys*. Luxembourg: Office for Official Publications of the European Communities.

[11] Freeman, R.B., and R. Schettkat. 2005. Marketization of household production and the EU-US gap in work. *Economic Policy* 20(41):5-50.

[12] Fry, Tim. 2011. Applications in economics. In *Compositional data analysis: Theory and applications*, edited by Vera Pawlowsky-Glahn and Antonella Buccianti. John Wiley and Sons, Ltd.

[13] Gershuny, Jonathan. 2000. *Changing times. Work and leisure in postindustrial society.* New York, Oxford University Press.

[14] Gourieroux, C., A. Monfort, and A. Trognon. 1984. Pseudo Maximum Likelihood methods: theory. *Econometrica* 52(3): 681-700.

[15] Hamermesh, Daniel S. 2002. Timing, togetherness and time windfalls. *Journal of Population Economics* 15:601-623.

[16] Hamermesh, Daniel S., and Gerard A. Pfann. 2005. *The Economics of Time Use*. Elsevier B.V.

[17] Hastie, T. 1987. A closer look at the deviance. *The American Statistician* 41(1):16-20.

[18] Hauser, J.R. 1978. Testing the Accuracy, Usefulness, and Significance of Probabilistic Choice Models: An Information-Theoretic Approach. *Operations Research* 26(3):406-421.

[19] Instituto Nacional de Estadística. 2004. *Encuesta de Empleo del Tiempo 2002-2003 – Proyecto Metodológico y Resultados Nacionales*. Madrid, Instituto Nacional de Estadística.

[20] Juster, F. Thomas. 1985. The validity and quality of time use estimates obtained from recall diaries. In *Time, Goods, and Well-Being*, edited by F. Thomas Juster and Frank P. Stafford. Institute for Social Research, University of Michigan.

[21] Kahneman, Daniel, and Alan B. Krueger. 2006. Developments in the measurement of subjective well-being. *Journal of Economic Perspectives*, 20(1):3-24.

[22] Kalenkoski, Charlene M., David C. Ribar, and Leslie S. Stratton. 2005. Parental child care in single-parent, cohabiting, and married-couple families: time-diary evidence from the United Kingdom. *American Economic Review* 95(2):194-198.

[23] Kimmel, Jean, and Rachel Connelly. 2007. Mothers' time choices. *Journal of Human Resources* 42(3):643-681.

[24] Liang, Kung-Yee, and Scott L. Zeger. 1986. Longitudinal data analysis using generalized linear models. *Biometrika* 73(1):13-22.

[25] Lundberg, Shelly. 2005. Sons, daughters, and parental behaviour. *Oxford Review of Economic Policy* 21(3):340-356.

[26] Martín-Fernández, Josep Antoni, Carles Barceló-Vidal, and Vera Pawlowsky-Glahn. 2003. Dealing with zeros and missing values in compositional data sets using nonparametric imputation. *Mathematical Geology* 35(3):253-278.

[27] Martín-Fernández, Josep Antoni, Javier Palarea-Albaladejo, and Ricardo Antonio Olea. 2011. Dealing with zeros. In *Compositional data analysis: Theory and applications*, edited by Vera Pawlowsky-Glahn and Antonella Buccianti. John Wiley and Sons, Ltd.

[28] McCullagh, P. and J.A. Nelder. 1989. *Generalized Linear Models*. Second edition. Boca Raton (Florida), Chapman & Hall/CRC.

[29] Mullahy, John and Stephanie A. Robert. 2010. No time to lose: Time constraints and physical activity in the production of health. *Review of Economics of the Household* 4:409-432.

[30] Papke, L.E. and J.M. Wooldridge. 1996. Econometric methods for fractional response variables with an application to 401 (K) plan participation rates. *Journal of Applied Econometrics* 11(6): 619-632.

[31] Pawlowsky-Glahn, Vera, and Antonella Buccianti (eds.) 2011. *Compositional data analysis: Theory and applications*. John Wiley and Sons, Ltd.

[32] Pentland, Wendy E., Andrew S. Harvey, M. Powell Lawton, and Mary Ann McColl (eds.) 1999. *Time Use Research in the Social Sciences*. New York, USA: Kluwer Academic Publishers Group.

[33] Presnell, Brett, Scott P. Morrison, and Ramon C. Littell. 1998. Projected multivariate linear models for directional data. *Journal of the American Statistical Association* 93:1068-1077.

[34] Price, Joseph. 2008. Parent-child quality time. Does birth order matter? *The Journal of Human Resources* 43(1):240-265.

[35] Prowse, V. 2009. Estimating labour supply elasticities under rationing: a structural model of time allocation behaviour. *Canadian Journal of Economics* 42(1): 90-112.

[36] Pudney, Stephen. 1989. *Modelling individual choice. The Econometrics of corners, kinks and holes*. Basil Blackwell.

[37] Simon, G. 1973. Additivity of information in exponential family probability laws. *Journal of the American Statistical Association* 68(342):478-482.

[38] Stephens, Michael A. 1982. Use of the von Mises distribution to analyse continuous proportions. *Biometrika* 69(1):197-203.

[39] Sullivan, Oriel. 2010. Changing differences by educational attainment in fathers' domestic labour and child care. *Sociology* 44(4):716-733.

[40] Wales, T.J. and A.D. Woodland. 1983. Estimation of consumer demand systems with binding non-negativity constraints. *Journal of Econometrics* 21: 263-285.

[41] Woodland, A.D. 1979. Stochastic specification and the estimation of share equations. *Journal of Econometrics* 10: 361-383.

[42] Wooldridge, Jeffrey M. 1991. On the application of robust, regression-based diagnostics to models of conditional means and conditional variances. *Journal of Econometrics* 47:5-46.

[43] Wooldridge, Jeffrey M. 2002. *Econometric Analysis of Cross Section and Panel Data*. Cambridge (Massachusetts) and London (England). The MIT Press.

[44] Zeger, Scott L. and Kung-Yee Liang. 1986. Longitudinal data analysis for discrete and continuous outcomes. *Biometrics* 42(1):121-130.

Table 1. Sample descriptive statistics, the Spanish Time Use Survey 2002-2003

Variable	Mean	SD	SD/Mean	Min	Max	% = 0
Formative time (hrs/day)	1.8	1.8	1.0	0	9.7	17.2
Other time	22.2	1.8	0.1	14.3	24	0
Father's age (years)	39.6	5.4		22	61	
Youngest child's age	5.7	2.1		3	9	

Variable (%)	Mean	*Variable (%)*	Mean
Net monthly earnings < €500	3.0	Wife's net monthly earnings < €500	22.4
€500 - €999.99	27.9	€500 - €999.99	40.7
€1000 - €1249.99	29.8	€1000 - €1249.99	14.4
€1250 - €1499.99	14.2	€1250 - €1499.99	8.7
€1500 - €1999.99	14.2	€1500 - €1999.99	9.6
€2000 - €2499.99	5.2	€2000 - €2499.99	2.2
€2500 - €2999.99	2.4	€2500 - €2999.99	1.1
≥ €3000	3.3	≥ €3000	0.9
Illiterate	0.5	Wife works part-time	13.0
1-4 years in school	0.8	Wife works in the evening	31.6
5-7 years in school	8.4	Wife works at night	10.5
8 years in school	29.1	1 kid aged 3-9	75.1
Exactly high school graduate	14.2	2 kids aged 3-9	23.6
Vocational training (2 years)	9.6	3 kids aged 3-9	1.3
Vocational training (4 years)	12.1	At least one son	58.7
College graduate (3 years)	11.4	Diary day is in 1st quarter (2003)	27.2
College graduate (5 years)	12.7	2nd quarter (2003)	28.6
PhD	1.2	3rd quarter (2003)	21.4
Works in the evening	46.6	4th quarter (2002)	22.8
Works at night	17.3	Weekend day (Friday-Sunday)	47.5

Notes: Data are of 762 fathers residing in two-earner families whose youngest child is 3-9 years of age. Formative time is the time the father spends with his children reading, doing homework, doing arts and crafts, doing sport, playing, attending performances and museums, engaging in religious activity, having meals and talking with the children, or providing personal care for the children. Monthly earnings are from the main job. A worker works in the evening (respectively, at night) if s/he was on the job at any time between 7PM and 10PM (10PM and 6AM) at least one day of the working week.

Table 2. Analysis of deviance of Spanish fathers' use of time

Source	Partial deviance	*df*	*QLR*	*QLR*(*df*)	Prob > *QLR*
Model	11.17	36	86.71	36	.000
Earnings	.48	7	3.71	7	.812
Education	1.03	9	8.01	9	.533
Work schedule	1.13	2	8.74	2	.013
Wife's earnings	.65	7	5.06	7	.653
Wife works part-time	.10	1	.75	1	.387
Wife's work schedule	.33	2	2.59	2	.274
Number of children	.14	2	1.11	2	.574
Youngest child's age	1.20	1	9.31	1	.002
At least one son	.11	1	.83	1	.362
Trimester of interview	.09	3	.74	3	.865
Weekend day	5.17	1	40.17	1	.000
Total deviance	59.12	761			
Residual	47.95	725			
R_Q^2	.189				

Notes: Fathers use of time was classified into two broad activities: formative time spent with children and all other activities. See notes to Table 1 for further information. *QLR*: quasi-likelihood ratio statistic, calculated as $\textit{partial deviance}/\hat{\sigma}^2$. $\hat{\sigma}^2$, computed as shown in expression (10) using results from the model with all 11 factors included, equals 0.129.

Table 3. Tobit model comparisons

Source	LR	$LR(df)$	Prob > LR
Model	140.01	36	.000
Earnings	7.83	7	.348
Education	14.70	9	.100
Work schedule	17.71	2	.000
Wife’s earnings	8.75	7	.271
Wife works part-time	.88	1	.349
Wife’s work schedule	3.31	2	.191
Number of children	2.62	2	.269
Youngest child’s age	16.07	1	.000
At least one son	1.70	1	.192
Trimester of interview	1.21	3	.751
Weekend day	60.33	1	.000
R-squared	.194		

Notes: *LR* is the value of the likelihood ratio statistic. *R*-squared is computed as $\widehat{cor}^2\left(y_1, \hat{y}_1\right)$. See notes to Tables 1 and 2 for further information.

Table 4. CODA model comparisons, $\phi_{i1} = 9 \ \forall i$

Source	*LR*	*LR(df)*	Prob > *LR*
Model	153.34	36	.000
Earnings	7.86	7	.345
Education	14.79	9	.097
Work schedule	16.09	2	.000
Wife's earnings	12.54	7	.084
Wife works part-time	1.58	1	.209
Wife's work schedule	3.69	2	.158
Number of children	4.34	2	.114
Youngest child's age	17.20	1	.000
At least one son	2.52	1	.112
Trimester of interview	4.55	3	.208
Weekend day	66.18	1	.000
R-squared	.186		

Notes: *LR* is the value of the likelihood ratio statistic. *R*-squared is computed as $\widehat{cor}^2(y_1, \hat{y}_1)$. See notes to Tables 1 and 2 for further information.