BLS WORKING PAPERS

U.S. Department of Labor
U.S. Bureau of Labor Statistics
Office of Employment and Unemployment Statistics

Tobit or Not Tobit?

Jay Stewart, U.S. Bureau of Labor Statistics

Working Paper 432
November 2009

Tobit or Not Tobit?

October 2009

Jay Stewart*
U.S. Bureau of Labor Statistics
2 Massachusetts Ave., NE
Washington, DC 20212
USA
Stewart.Jay@bls.gov
(202) 691-7376

* Any views expressed here are mine and do not necessarily reflect those of the Bureau of Labor Statistics. I thank Rachel Connelly, Lewis Davis, Matt Dey, Tim Erickson, Harley Frazis, Dan Hamermesh, Sabrina Pabilonia, and participants of seminars at the Bureau of Labor Statistics and Union College for their helpful comments and suggestions.

Tobit or Not Tobit?

Abstract

Time-use surveys collect very detailed information about individuals' activities over a short period of time, typically one day. As a result, a large fraction of observations have values of zero for the time spent in many activities, even for individuals who do the activity on a regular basis. For example, it is safe to assume that all parents do at least some childcare, but a relatively large fraction report no time spent in childcare on their diary day. Because of the large number of zeros Tobit would seem to be the natural approach. However, it is important to recognize that the zeros in time-use data arise from a mismatch between the reference period of the data (the diary day) and the period of interest, which is typically much longer. Thus it is not clear that Tobit is appropriate.

In this study, I examine the bias associated with alternative estimation procedures for estimating the marginal effects of covariates on time use. I begin by adapting the infrequency of purchase model, which is typically used to analyze expenditures, to time-diary data and showing that OLS estimates are unbiased. Next, using simulated data, I examine the bias associated with three procedures that are commonly used to analyze time-diary data—Tobit, the Cragg (1971) two-part model, and OLS—under a number of alternative assumptions about the data-generating process. I find that the estimated marginal effects from Tobits are biased and that the extent of the bias varies with the fraction of zero-value observations. The two-part model performs significantly better, but generates biased estimated in certain circumstances. Only OLS generates unbiased estimates in all of the simulations considered here.

Jay Stewart
U.S. Bureau of Labor Statistics
2 Massachusetts Ave., NE
Washington, DC 20212
USA
Stewart.Jay@bls.gov
(202) 691-7376

Introduction

With the introduction of the American Time Use Survey (ATUS), there has been a renewed interest in research using time-diary data. One feature of these data is that a large fraction of individuals have zero values for the time spent in many activities. So far there has been no general agreement on the correct approach to dealing with these zero-value observations. Researchers have used OLS (Bonke, 1992; and Frazis and Stewart, forthcoming), a two-part model similar to the one proposed by Cragg (1971) (Cawley and Liu, 2007), and Tobit (Souza-Poza, Schmid, and Widmer, 2001; Kalenkoski, Ribar, and Stratton, 2005; Kimmel and Connelly, 2007). Some authors report estimates from more than on estimation procedure (Hamermesh, 2009; Price, 2008). However, Tobit has been the predominant approach in more-recent studies. The Tobit model would seem to be a sensible approach, because it was developed specifically for situations where the dependent variable is truncated at zero or some other cutoff.

The standard discussion of the Tobit model (Tobin, 1958) assumes that there is a latent variable (for example, desired expenditures) underlying the observed dependent variable (actual expenditures). The two are equal when the latent variable is greater than zero, but the observed variable is zero when the latent variable is negative. In economic models, this corresponds to a corner solution in the utility maximization program where the individual's optimal value of the dependent variable is negative but nonnegativity constraints force the value to be zero. It is well-known that, under these assumptions, OLS parameter estimates are downward biased and inconsistent while Tobit estimates are consistent and asymptotically normal (Amemiya, 1973).[1]

This interpretation is not generally appropriate for time-diary data, because a zero does not necessarily imply that the individual never does the activity. The fixed costs of engaging in

[1] This assumes that the distribution of errors is normally distributed.

an activity may lead the individual to do the activity on some days but not on others, as can random events such as illness or a change of schedule.[2] Important examples of activities that are not done every day include time spent working by the employed, time spent looking for work by the unemployed, and time spent in childcare by parents. In these cases, zeros arise because the reference period of the data (the diary day) is shorter than the period of interest (i.e., the period over which decisions are made). In this sense, time-diary data are very similar to expenditure data. For example, expenditures on clothing are often zero in expenditure diaries, but casual empiricism suggests that virtually everybody purchases clothing.

The infrequency of purchase model (IPM) was specifically developed to address the mismatch between the period of interest and the reference period of the data. It has been shown that using OLS to estimate the effect of income on consumption in an IPM framework results in biased and inconsistent parameter estimates (Keen, 1986). Alternatives to OLS include an instrumental variables estimator proposed by Keen (1986) and a two-part model along the lines of Cragg (1971), which generalizes the Tobit model (Blundell and Meghir, 1987).[3]

Along these lines, there is an alternative interpretation of the Tobit model that does not assume there is a latent variable that takes on negative values. Instead, it only assumes that there is information conveyed in the probability that the dependent variable is equal to zero (see McDonald and Moffitt 1980). But even if this interpretation is correct Tobit still may not be appropriate because, as Cragg (1971) points out in the context of estimating expenditure models,

---

[2] The standard household production model can be modified to incorporate timing (for example, see Stewart, forthcoming). In this type of model, it can easily be shown that if the daily fixed cost of engaging in an activity is sufficiently high, the individual will not engage in the activity every day.

[3] Cragg (1971) proposes a double-hurdle model, where the first hurdle is the decision to ever spend money on the good. Since I am restricting my attention to situations where this decision is taken as given, the double-hurdle model reduces to a two-part model. In the first part of the two-part model, a probit is estimated over all observations to determine the probability that individuals purchase the good during the reference period. In the second part, an OLS regression is estimated over the non-zero-value observations. The estimated average probability from the probit is combined with the coefficients from the OLS regression to arrive at unconditional marginal effects.

it assumes that the process that determines whether a person purchases a good is the same as the one that determines the amount spent on that good.

In the analysis that follows, I examine the appropriateness of alternative procedures for estimating the effects of covariates on the average amount of time spent in an activity when there are zero-value observations in the data. I begin by adapting the IPM to time-diary data and showing that, in this context, OLS is unbiased. Next, I generate simulated time-use data using the IPM framework and estimate the effect of covariates on time use using OLS, Tobit, and the two-part model. I compute the bias associated with each procedure, and examine how the bias and mean squared error (MSE) vary with the fraction of zero-value observations in the data.

Adapting the Infrequency of Purchase Model to Time-Diary Data

Using the notation from Keen (1986), expenditures on good *k* are equal to:

$$(1) \qquad e_{hk} = \frac{w_{hk}\{\bar{c}_{hk} + u_{hk}\}}{p_{hk}},\ k = 1,\ldots,N,$$

where e_{hk} and $\bar{c}_{hk}$ denote expenditures and consumption of good *k* by household *h*, p_{hk} denotes the probability that good *k* is purchased during the reference period, w_{hk} is an indicator that equals 1 if household *h* is observed purchasing good *k* during the reference period, and u_{hk} is a random term (where $E(u_{hk}) = 0$) that captures variation in the amount spent on good *k*. Note that u_{hk} is constrained such that $u_{hk} \geq -\bar{c}_{hk}$ (so that expenditures are always non-negative), and the two random terms w_{hk} and u_{hk} are assumed to be independently distributed.

The terms in equation (1) have natural interpretations in the context of time-diary data. Individuals determine how much time they wish to spend in each activity over some period of time, such as a month, and then allocate that time to individual days. Using the notation of the IPM, $\bar{c}_{hk}$ is the amount of time that the individual spends in activity *k* each month (expressed as

a daily average) and e_{hk} is the observed amount of time spent doing activity k on the diary day. The remaining terms in equation (1) have analogous interpretations—p_{hk} is the probability that the individual does activity k on any given diary day, $w_{hk} = 1$ if the individual engaged in activity k on the diary day, and u_{hk} is a random term that captures day-to-day variation in the amount of time spent in activity k. If $\bar{c}_{hk}$ is a linear function of a set of covariates (to keep things simple, I consider the one-covariate case), then:

$$(2) \qquad \bar{c}_{hk} = \alpha_k + \beta_k x_h ,$$

where x_h is a covariate that is thought to influence time spent in activity k. Combining equations (1) and (2) yields:

$$(3) \qquad \begin{aligned} e_{hk} &= \alpha_k + \beta_k x_h + \{(w_{hk} - p_{hk})\bar{c}_{hk} + w_{hk}u_{hk}\} / p_{hk} \\ &= \alpha_k + \beta_k x_h + \eta_{hk}, \end{aligned}$$

which can be estimated using OLS.

Expressing e_{hk}, $\bar{c}_{hk}$, and x_{hk} as deviations from their respective means (and using the "dot" notation), the estimated coefficient, $\hat{\beta}_k$, is given by:

$$\hat{\beta}_k = \frac{\sum_h \dot{x}_h \dot{e}_{hk}}{\sum_h \dot{x}_h^2} = \frac{\sum_h \dot{x}_h (\beta_k \dot{x}_h + \dot{\eta}_{hk})}{\sum_h \dot{x}_h^2},$$

where $\dot{\eta}_{hk} = \{(w_{hk} - p_{hk})\dot{\bar{c}}_{hk} + w_{hk}u_{hk}\} / p_{hk}$. Arranging terms and taking expectations, we have:

$$E(\hat{\beta}_k) = \beta_k + E\left(\frac{\sum_h \dot{x}_h \dot{w}_{hk} \dot{\bar{c}}_{hk}}{p_{hk} \sum_h \dot{x}_h^2}\right) + E\left(\frac{\sum_h w_{hk} u_{hk}}{p_{hk} \sum_h \dot{x}_h^2}\right),$$

where $\dot{w}_{hk} = (w_{hk} - p_{hk})$. Substituting equation (2) for $\dot{\bar{c}}_{hk}$ yields:

$$E(\hat{\beta}_k) = \beta_k + E\left(\frac{\sum_h \beta_k \dot{x}_h^2 \dot{w}_{hk}}{p_{hk} \sum_h \dot{x}_h^2}\right) + E\left(\frac{\sum_h w_{hk} u_{hk}}{p_{hk} \sum_h \dot{x}_h^2}\right).$$

Given the assumption that $E(u_{hk}w_{hk}) = E(u_{hk}) = 0$, the third term is equal to zero. The second term is also equal to zero as long as $E(\dot{x}_h^2\dot{w}_{hk}) = 0$. Because $\dot{w}_{hk}$ is the deviation of w_{hk} around its mean value of p_{hk}, $E(\dot{x}_h^2\dot{w}_{hk}) = 0$ even if p_{hk} is a function of x_h. Thus, $E(\hat{\beta}_k) = \beta_k$ and estimating equation (3) using OLS will generate unbiased estimates of β_k.

The rest of the paper is devoted to comparing three alternative estimation procedures that have been used with time-diary data: OLS, Tobit, and Cragg's two-part model. I construct a simulated sample using the data-generating process described above, and then use these three models to estimate parameters under alternative assumptions about the fraction of zero-value observations.

Construction of the Simulated Data

To construct the sample for the simulations, I started by assuming that all individuals are "doers" (i.e., that they do activity *k* for at least a few minutes every month). To allow for random variation (due to unobserved factors) in the amount of time spent in activity *k* and to make the simulation more consistent with the assumptions of the Tobit model, I modified the adapted IPM slightly by adding a normally-distributed error term, θ_{hk}, to equation (2). For example, if activity *k* is childcare then θ_{hk} might be large and positive for the month if a child stayed home from school for a few days with the flu and required additional care. Thus equation (2) becomes:

$$(2') \qquad \bar{c}_{hk} = X_h \mathrm{B}_k + \theta_{hk},$$

where X_h is a vector of covariates (including an intercept), B_k is a vector of coefficients, and the error $\theta_{hk} \sim N(0, \sigma^2_{\theta_{hk}})$ and is uncorrelated with w_{hk} or u_{hk}. Assuming three covariates and dropping activity subscripts to reduce clutter, equation (2′) becomes:

$$(2'') \qquad \bar{c}_h = \alpha + \beta_1 x_{h1} + \beta_2 x_{h2} + \beta_3 x_{h3} + \theta_h ,$$

where $\alpha = 10$, $\beta_1 = 1.5$, $\beta_2 = -3$, and $\beta_3 = 2$. The data for the x_i and θ_h were generated using a random number generator and are distributed as follows: $x_1 \sim U[1,4]$, $x_2 \sim U[-2,3]$, x_3 is Bernoulli with $Prob(x_3 = 1) = 0.5$, and $\theta_h \sim N(0,1)$.[4] The assumption that $\alpha = 10$ ensures that $\bar{c}_h > 0$ for nearly all respondents. The sample size for each simulation was 50,000 observations, minus the small number of "respondents" who were dropped because $\bar{c}_h \leq 0$.

The next step was to generate daily values for time spent in activity *k*. For each individual in the sample, I created 28 days of data and generated the amount of time spent in activity *k* each day. Combining equations (1) and (2′), the amount of time spent doing activity *k* on day *d* is given by:

$$(4) \qquad e_{hd} = \frac{w_{hd}(\bar{c}_h + u_{hd})}{p_h} = \frac{w_{hd} X_h \mathrm{B}}{p_h} + \frac{w_{hd}\theta_h}{p_h} + \frac{w_{hd} u_{hd}}{p_h} .$$

Taking expectations verifies that:

$$E(e_{hd}) = X_h \mathrm{B} .$$

I implemented equation (4) as follows. First, I set $e_{hd} = \bar{c}_h \times \delta$, where $\delta \sim U[0,1]$ on weekdays and $\delta \sim U[0,2]$ on weekends. Note that this implies individuals spend more time in activity *k* on weekends and that, by construction, $e_{hd} > 0$ for all days. To generate zero observations, I sorted the days for each individual by e_{hd}, and set $e_{hd} = 0$ for the T_h days with the lowest values for e_{hd}. Thus, $p_h = T_h/28$. The values of e_{hd} for the remaining $(28 - T_h)$ days are inflated proportionately so that $\sum_d e_{hd} = 28 \times \bar{c}_h$. Note that this last step also preserves the normality of θ_h.

[4] The results are not sensitive to the variance of θ_h. I ran several sets of simulations with $\theta_h \sim N(0,2)$, and got nearly identical results.

I ran seven sets of simulations, each of which used a different algorithm to determine the relationship between T_h and the variables in the model (see the Appendix for a description of the algorithms). All of the algorithms have a random component so that the fraction of zero observations varies across individuals. Noting that $E(w_h) = p_h$ and letting $\overline{w} = \sum_h w_{hd} \big/ N$, where N is the final sample size, the fraction of zero observations is $\bar{q} = (1 - \overline{w})$. The relationships between q_h $(= 1 - p_h)$ and the variables in the model are:

(1) q_h is unrelated to the value of $\bar{c}_h$ or any of the x_i.

(2) q_h is negatively related to the value of $\bar{c}_h$.

(3) q_h is negatively related to the value of x_1.

(4) q_h is positively related to x_2.

(5) q_h is negatively related to the value of x_3.

The first set of simulations, while not very realistic, provides a useful baseline. I present three sets of simulations for (2), because this would seem to be the most likely case. Cases (3) – (5) cover situations where one of the covariates affects q_h directly rather than indirectly through their effects on $\bar{c}_h$. For each x_i, I ran two sets of simulations—one where q_h is a positive function of x_i and one where the relationship is negative—but I only report the simulations that resulted in a negative correlation between q_h and $\bar{c}_h$.[5] For each set of simulations, I varied the values of the T_h so that the fraction of zero-value observations ranged between 0 and 0.9.[6]

[5] The other simulation results are available from the author on request.
[6] To estimate the two-part model, it was necessary to truncate the range to between 0.005 and 0.9.

Simulation Results

Once the data were generated, I randomly selected one day for each individual in the sample, and estimated the βs using OLS, Tobit, and the two-part model. I report the estimated coefficients from OLS and unconditional marginal effects for the Tobit and two-part models.[7] For OLS and Tobit, I simply estimated the simulation version of equation (3) over all observations in the sample:

$$(5) \qquad e_{hd} = \alpha + \beta_1 x_{h1} + \beta_2 x_{h2} + \beta_3 x_{h3} + \varepsilon_{hd}.$$

The unconditional marginal effects for the Tobit model were computed as:

$$\frac{\partial E(e_{hd} \mid \mathbf{x})}{\partial x_{hi}} = \hat{\beta}_i^T \Phi\left(\frac{\hat{\alpha}^T + \sum_{j=1}^{3} \hat{\beta}_j^T \bar{x}_{hj}}{\hat{\sigma}^T} \right)$$

using the **mfx** command in STATA,[8] where the *T* superscript indicates the Tobit coefficients. For the two-part model, I estimated:

$w_{hd} = \gamma_0 + \gamma_1 x_{h1} + \gamma_2 x_{h2} + \gamma_3 x_{h3} + \varepsilon_{hd}$ over all observations using probit, and

$e_{hd} = \alpha^{2P} + \beta_1^{2P} x_{h1} + \beta_2^{2P} x_{h2} + \beta_3^{2P} x_{h3} + \varepsilon_{hd}^{2P}$ over observations for which $e_{hd} > 0$ using OLS.

The marginal effects were computed as:

$$\frac{\partial E(e_{hd} \mid \mathbf{x})}{\partial x_{hi}} = \frac{\partial \Phi\left(\frac{\hat{\gamma}_0 + \sum_{j=1}^{3} \hat{\gamma}_j \bar{x}_{hj}}{\hat{\sigma}} \right)}{\partial x_{hi}} \times \left[\hat{\alpha}^{2P} + \sum_{j=1}^{3} \hat{\beta}_j^{2PT} \bar{x}_{hj|e_{hd}>0} \right] + \Phi\left(\frac{\hat{\gamma}_0 + \sum_{j=1}^{3} \hat{\gamma}_j \bar{x}_{hj}}{\hat{\sigma}} \right) \times \hat{\beta}_i^{2P}.$$

Figures 1-5 show the simulation results. The three panels in each figure correspond to the three procedures, and show the bias in the estimated marginal effects for the covariates,

---

[7] I also examined the bias associated with using Tobit *coefficients*, rather than marginal effects. The coefficients generally overestimated the true parameters, with the bias increasing sharply as $\bar{q}$. increases. I do not report the coefficients, because they are rarely reported in time-use research.

[8] Note that for x_3, the Bernoulli-distributed covariate, I used the STATA default of computing the marginal effect as the effect of a discrete jump between 0 and 1.

expressed as a percentage of the true parameter values and graphed against $\bar{q}$, the fraction of zero observations. I computed the bias as $(\widehat{ME}_i - \beta_i)/\beta_i$ so that a positive value indicates that the magnitude of β_i has been overestimated, while a negative value indicates that the magnitude has been underestimated. A wrong-signed coefficient would cause the bias to be less than −1.

Figure 1 shows the baseline set of simulations, where q_h is independent of any of the variables in the model. The Tobit marginal effects underestimate the true effects, and the magnitude of this bias increases with $\bar{q}$. The bias is large (about 30 percent) when $\bar{q}$ = 0.4, and grows to over 50 percent when $\bar{q}$ > 0.8. In contrast to the Tobit model, both the OLS and the two-part model generate estimates that are unbiased, even as $\bar{q}$ becomes large. What does happen is that as $\bar{q}$ increases (greater than about 0.7), the variability of these estimates becomes quite large. For x_1 and x_3 the parameter estimates are off by over 30 percent in a few cases, while for x_2 the parameter estimates are never off by more than 10 percent. Even so, these extreme estimates still have smaller bias than the average Tobit estimates.

Table 1 shows the MSE of the estimated coefficients and marginal effects for different ranges of $\bar{q}$, where each panel corresponds to a different figure. We can see that, for all three procedures, the MSE increases as $\bar{q}$ increases. Despite the considerable variability in the OLS and two-part model estimates, the small bias in both procedures results in MSEs that are less than 5 percent of the Tobit estimates' MSE.

As noted above, it is more realistic to assume that individuals who spend more time per month doing an activity are less likely to report zero time spent in the activity on their diary day. The simulation results in Figures 2a, 2b, and 2c show cases where q_h is negatively related to $\bar{c}_h$. The results in the three figures are fairly similar to those in Figure 1. As in Figure 1, OLS

generates unbiased estimates, with the variability of these estimates increasing as $\bar{q}$ increases, while Tobit marginal effects underestimate true parameter values. There is a slight bias in the two-part model (less than 5 percent) over some values of $\bar{q}$. However, the main differences between these figures and Figure 1 are that the Tobit marginal effects are closer to the true values, and that the magnitude of the bias does not increase as rapidly with $\bar{q}$. In Figure 2a, the bias is smaller than in Figure 1, but is still quite large. In Figures 2b and 2c, the marginal effects estimates are mostly within about 10 percent of the true parameter values for values of $\bar{q} < 0.7$, and are fairly close to the lower bounds of the estimates from OLS and the two-part model.[9]

The MSEs that correspond to Figures 2a, 2b, and 2c, exhibit the same pattern as those corresponding to Figure 1, with MSEs increasing as $\bar{q}$ increases and Tobit marginal effects having larger MSEs. But in these simulations, the difference between Tobit and the other two procedures has narrowed. For larger values of $\bar{q}$ Tobit MSEs are smaller than those corresponding to Figure 1, while MSEs for OLS and the two-part model are about the same as in Figure 1 or slightly larger. However, Tobit MSEs are still considerably larger than those of the other two procedures.

The performance of Tobit and the two-part model deteriorates in Figures 3-5, where q_h is a direct function of one of the covariates. Tobit marginal effects are still downward biased, except for the marginal effects on the covariate that directly affects q_h. For example, the bias in the marginal effect of x_1 is positive and increases rapidly with $\bar{q}$ when q_h is a positive function

[9] I also ran a set of simulations where q_h was positively related to c_{hk}. Both OLS and the two-part model generated unbiased estimates until $\bar{q}$ reached the 0.75 to 0.80 range, at which point the estimates became downward biased for all three coefficients. Tobit marginal effects were downward biased, with the bias being quite large. These results are available from the author on request.

of x_1.[10] The two-part model performs better than Tobit, but some coefficients are biased, with the pattern of bias depending on which covariate directly affects q_h. In Figure 3, where q_h is a function of x_1, we see that the marginal effect of x_1 is unbiased for $\bar{q} < 0.6$. For larger values of $\bar{q}$, the bias in the marginal effect is positive and increasing, while the bias in the marginal effects of x_2 and x_3 are negative and decreasing. This pattern is similar to what was observed for Tobit, but less extreme. In Figure 4, the bias does not appear to be particularly severe. But in Figure 5, where q_h is a function of x_3, the marginal effect of x_3 is downward biased for $\bar{q} > 0.4$, although the marginal effects of the other two covariates are still unbiased for all values of $\bar{q}$.

In contrast to the Tobit and two-part models, nothing changes when q_h is a direct function of one of the covariates. OLS coefficients are still unbiased, and the variation of these coefficients still increases as $\bar{q}$ becomes large.

Turning back to the MSEs in Table 1, we see that OLS and the two-part model are fairly close in most cases. In Figure 3, the MSEs for OLS and the two-part model are nearly identical for all three covariates until $\bar{q} > 0.6$, with the differences becoming quite large for $\bar{q} > 0.8$. In Figure 5, the only difference is for x_3. Contrary to the other sets of simulations, the MSE for x_3 is largest when $\bar{q}$ is in the 0.6 - 0.8 range, and then becomes smaller as $\bar{q}$ approaches 0.9.

Discussion and Conclusions

The simulation results clearly show that marginal effects from the Tobit model are biased, that the bias is often large, and that the extent of the bias increases as the fraction of zero observations increases. It seems likely that one of the main reasons for this poor performance is

[10] In the opposite case, where the resulting correlation between the covariates and $\bar{q}$ results in a positive correlation between $\bar{q}$ and the time spent in the activity, all three coefficients are downward biased with the bias increasing as $\bar{q}$ increases.

that the Tobit model assumes that the process that determines whether an individual engages in an activity is the same one that governs how much time is spent in that activity. This explanation is consistent with the findings of Daunfeldt and Hellstrom (2001) who, in their study of time spent in household production activities, reject the Tobit model in favor of the two-part model. My simulations confirm that the two-part model performs better than Tobit. As long as the probability of doing the activity on a given day does not depend on any of the covariates, the two-part model generates estimated marginal effects that are unbiased and invariant to the fraction of zeros in the data. However, if the probability of doing the activity on any given day is a function of one of the covariates, the two-part model behaves unpredictably. This is unfortunate, because a potential advantage of the two-part model is the ability to decompose the marginal effects to examine the effects of covariates on incidence and intensity.

In contrast to the two models that were specifically designed to address the problem of zero observations, OLS estimates are unbiased and robust to a number of assumptions about the relationship between the variables in the model and the probability of doing the activity. Both OLS and the two-part model outperform Tobit in all simulations. There is virtually no difference between OLS and the two-part model, except in cases where the probability of doing the activity is a function of one of the covariates—in these cases, OLS outperforms the two-part model.

There are two issues that I did not address in this study: standard errors and what happens if it is not possible to identify doers. The presence of zeros in the data are likely to affect standard errors by introducing heteroskedasticity into the residual. However, using robust methods to compute standard errors should address this problem.

The second issue is more serious. If it is not possible to identify doers, then none of the three procedures performs particularly well. I ran two sets of simulations where a fraction of the

sample included non-doers as well as doers. The simulated data were constructed as described earlier, except that the intercept was adjusted downward to generate more zeros.[11] This is essentially the Tobit assumption. As in the simulations described above, the Tobit model generates downward biased estimates and the bias increases as the fraction of zero observations (including non-doers) increases. The estimated marginal effects from the two-part model have a large bias when the total fraction of zero observations is only slightly larger than the fraction of "true" zero observations, and the bias decreases as the fraction of zero observations increases (as true zeros become a smaller fraction of all zero observations). It appears that it is the mixture of true zeros and reference-period-mismatch zeros that leads to biased estimates. Finally, and not surprisingly, OLS is downward biased. However, the magnitude of the bias is invariant to the fraction of zero observations in the data.

It is worth reiterating that, when the sample is restricted to doers, the reason for zeros in time-diary data is the mismatch between the length of the reference period (the diary day) and the period over which decisions are made. Thus, it stands to reason that lengthening the reference period should reduce the fraction of zero observations. A study by Foster and Kalenkoski (2008) examines how the diary window length affects OLS and Tobit estimates. Their findings are generally consistent with my results. They find that Tobit marginal effects are smaller than OLS estimates but that the difference is not that large. This is consistent with my finding that Tobit marginal effects are downward biased while OLS estimates are unbiased, and that the bias associated with Tobit marginal effects is small as long as the fraction of zero observations is not too large. My calculations based on their Table 2 indicate that the fraction of zeros is between 0.35 and 0.39 for the 48-hour window and between 0.38 and 0.43 for the 24-

[11] In the first set the constant was set to 1 (vs. 10), which resulted in 22 percent of the sample being non-doers. The corresponding numbers in the second set were −2 and 42 percent.

hour window. In this range, I find that Tobit marginal effects are fairly close to OLS estimates. They also argue that Tobit marginal effects are more sensitive to window length than OLS estimates, but the differences between the two sets of coefficients do not appear to be statistically significant. This is not too surprising given that the difference in the fraction of zeros is not that different for the two window lengths.

Given the robustness of OLS to alternative assumptions about the data-generating process and the ease of estimating OLS, it is hard to recommend either Tobit or the two-part model. If the researcher is interested in the likelihood of engaging in the activity on a given day the two-part model, though unpredictable, outperforms Tobit. But for most policy-related questions, it is only necessary to know how certain covariates affect the average amount of time spent in an activity and the added information about the probability of engaging in the activity on a given day adds little.

Appendix: Algorithms for generating zero-value observations

For each respondent, observations are sorted by e_{hd} and assigned a rank, $R_h(e_{hd})$. The lowest value of e_{hd} is ranked 1, the second lowest is ranked 2, etc. Values of e_{hd} are set to zero if $R_h(e_{hd}) \leq T_h$, where T_h is determined as follows:

Figure 1: The number of zero days is unrelated to the value of $\bar{c}_h$ or any of the x_i.

$$T_h = \text{round}(U(0,1) \times \tau)$$

Figure 2: The number of zero days is negatively related to the value of $\bar{c}_h$.

$$T_h = \text{round}(U(0,1) \times \tau) - \text{round}(U(0,1) \times \bar{c}_h)$$

$$T_h = \text{round}(2 \times U(0,1) \times \tau) - \text{round}(3 \times U(0,1) \times \bar{c}_h)$$

$$T_h = \text{round}(3 \times U(0,1) \times \tau) - \text{round}(3 \times U(0,1) \times \bar{c}_h)$$

Figure 3: The number of zero days is negatively related to the value of x_1.

$$T_h = \text{round}((4 - x_{h1}) \times U(0,1) \times \tau)$$

Figure 4: The number of zero days is positively related to the value of x_2.

$$T_h = \text{round}(0.5 \times (3 + x_{h2}) \times U(0,1) \times \tau)$$

Figure 5: The number of zero days is negatively related to the value of x_3.

$$T_h = \text{round}((2 - x_{h3}) \times U(0,1) \times \tau)$$

In each set of simulations, the parameter τ was initially set to 0 set to 0 and incremented by 1 in each subsequent simulation until the percent of zero-value observations in the entire sample reached 90 percent (so that $\bar{q}$ ranged from 0 to 0.9). Note that values of $T_h < 0$ are treated the same as values of 0.

References

Amemiya, Takeshi (1973) "Regression Analysis when the Dependent Variable is Truncated Normal." *Econometrica* 41(6), November 1973), pp. 997-1016.

Blundell, Richard and Costas Meghir (1987) "Bivariate Alternatives to the Tobit Model." *Journal of Econometrics* 34, pp. 179-200.

Cragg, John G. (1971) "Some Statistical Models for Limited Dependent Variables with Application to the Demand for Durable Goods." *Econometrica* 39(5), Sept. 1971, pp. 829-844.

Daunfeldt, Sven-Olov and Jörgen Hellström (2007) "Intra-household Allocation of Time to Household Production Activities: Evidence from Swedish Household Data." *Labour* 21(2), pp. 189-207.

Flood, Lennart and Urban Grasjo (2001) "A Monte Carlo Simulation Study of Tobit Models." *Applied Economics Letters* 8, pp. 581-584.

Foster, Gigi and Charlene M. Kalenkoski (2008) "Tobit or OLS? An Empirical Evaluation Under Different Diary Window Lengths." Unpublished manuscript, Ohio University.

Frazis, Harley and Jay Stewart (2009) "How Does Household Production Affect Measured Income Inequality?" forthcoming in *Journal of Population Economics*.

Hamermesh, Daniel S. (2009) "Grazing and Making Fat: Determinants and Effects." NBER Working Paper No. 15277, August 2009.

Keen, Michael (1986) "Zero Expenditures and the Estimation of Engel Curves." *Journal of Applied Econometrics* 1(3), July 1986, pp. 277-286.

Kimmel, Jean and Rachel Connelly (2007) "Mothers' Time Choices: Caregovomg. Leisure, Home Production, and Paid Work." *Journal of Human Resources* 42(3), Summer 2007, pp. 643-681.

Kalenkoski, Charlene, David Ribar, and Leslie Stratton (2005) "Parental Childcare in Single-Parent, Cohabiting, and Married-Couple Families: Time-Diary Evidence from the United Kingdom." *American Economic Review Papers and Proceedings* 95(2), May 2005, pp. 194-198.

McDonald, John and Robert Moffitt (1980) "The Uses of Tobit Analysis." *The Review of Economics and Statistics* 62(2), May 1980, pp. 318-321.

Price, Joseph (2008) "Parent-Child Quality Time: Does Birth Order Matter?" *Journal of Human Resources* 43(1), pp. 240-265.

Souza-Poza, Alfonso, Hans Schmid, and Rolf Widmer (2001) "The Allocation and Value of Time Assigned to Housework and Child-care: An Analysis for Switzerland." Journal of Population Economics 14, pp. 599-618.

Stewart, Jay (2009) "The Timing of Maternal Work and Time with Children." Unpublished manuscript, Buearu of Labor Statistics.

Tobin, James (1958) "Estimation of Relationships for Limited Dependent Variables." *Econometrica* 26(1), Jan. 1958, pp. 24-36.

Table 1: Mean Squared Error for Alternative Procedures by Fraction of Zero-Value Observations

Corresponding Figure and Ranges for Fraction of Zero Observations		x_1			x_2			x_3		
		OLS	Tobit	2-Part Model	OLS	Tobit	2-Part Model	OLS	Tobit	2-Part Model
1	$q \le 0.2$	0.00	0.04	0.02	0.00	0.13	0.07	0.02	0.11	0.08
	$0.2 < q \le 0.4$	0.00	0.11	0.00	0.00	0.45	0.00	0.01	0.25	0.01
	$0.4 < q \le 0.6$	0.01	0.39	0.01	0.00	1.20	0.00	0.03	0.65	0.02
	$0.6 < q \le 0.8$	0.02	0.69	0.02	0.01	2.74	0.01	0.07	1.18	0.06
	$q > 0.8$	0.04	1.10	0.03	0.02	4.26	0.02	0.11	2.01	0.10
2a	$q \le 0.2$	0.00	0.01	0.00	0.00	0.05	0.00	0.01	0.03	0.01
	$0.2 < q \le 0.4$	0.00	0.03	0.00	0.00	0.14	0.00	0.01	0.07	0.02
	$0.4 < q \le 0.6$	0.01	0.11	0.01	0.00	0.36	0.01	0.03	0.19	0.03
	$0.6 < q \le 0.8$	0.03	0.27	0.02	0.01	0.98	0.01	0.08	0.47	0.07
	$q > 0.8$	0.07	0.46	0.07	0.02	1.71	0.01	0.15	0.82	0.15
2b	$q \le 0.2$	0.00	0.01	0.01	0.00	0.05	0.01	0.01	0.03	0.01
	$0.2 < q \le 0.4$	0.00	0.00	0.01	0.00	0.02	0.01	0.01	0.01	0.03
	$0.4 < q \le 0.6$	0.01	0.02	0.01	0.00	0.03	0.01	0.04	0.07	0.03
	$0.6 < q \le 0.8$	0.03	0.08	0.03	0.01	0.18	0.02	0.13	0.16	0.13
	$q > 0.8$	0.05	0.17	0.05	0.01	0.50	0.02	0.10	0.24	0.09
2c	$q \le 0.2$	0.00	0.01	0.01	0.00	0.05	0.00	0.01	0.02	0.01
	$0.2 < q \le 0.4$	0.00	0.01	0.01	0.00	0.01	0.02	0.02	0.02	0.03
	$0.4 < q \le 0.6$	0.01	0.02	0.02	0.01	0.04	0.02	0.04	0.06	0.03
	$0.6 < q \le 0.8$	0.02	0.07	0.02	0.01	0.16	0.01	0.05	0.13	0.06
	$q > 0.8$	0.05	0.16	0.05	0.02	0.47	0.02	0.18	0.34	0.17
3	$q \le 0.2$	0.00	0.08	0.00	0.00	0.13	0.00	0.02	0.12	0.02
	$0.2 < q \le 0.4$	0.01	2.63	0.01	0.00	0.56	0.01	0.01	0.29	0.02
	$0.4 < q \le 0.6$	0.02	17.59	0.01	0.00	1.50	0.01	0.03	0.70	0.03
	$0.6 < q \le 0.8$	0.04	64.10	0.11	0.02	3.30	0.26	0.11	1.34	0.13
	$q > 0.8$	0.06	85.71	0.65	0.03	4.83	0.66	0.26	1.96	0.31
4	$q \le 0.2$	0.00	0.04	0.00	0.00	0.04	0.00	0.03	0.12	0.03
	$0.2 < q \le 0.4$	0.01	0.12	0.01	0.00	0.02	0.01	0.01	0.28	0.01
	$0.4 < q \le 0.6$	0.01	0.33	0.01	0.00	0.45	0.04	0.02	0.61	0.03
	$0.6 < q \le 0.8$	0.03	0.78	0.04	0.00	1.44	0.01	0.07	1.42	0.09
	$q > 0.8$	0.04	1.06	0.04	0.02	1.35	0.01	0.12	1.92	0.10
5	$q \le 0.2$	0.00	0.04	0.00	0.00	0.13	0.00	0.02	0.07	0.03
	$0.2 < q \le 0.4$	0.01	0.13	0.01	0.00	0.48	0.00	0.02	1.76	0.03
	$0.4 < q \le 0.6$	0.01	0.35	0.01	0.00	1.27	0.00	0.06	15.16	0.43
	$0.6 < q \le 0.8$	0.03	0.79	0.03	0.01	2.83	0.01	0.06	29.07	0.53
	$q > 0.8$	0.04	1.11	0.04	0.02	4.38	0.03	0.14	33.51	0.24

Figure 1: The Fraction of Zero Observations is Independent of Variables in the Model

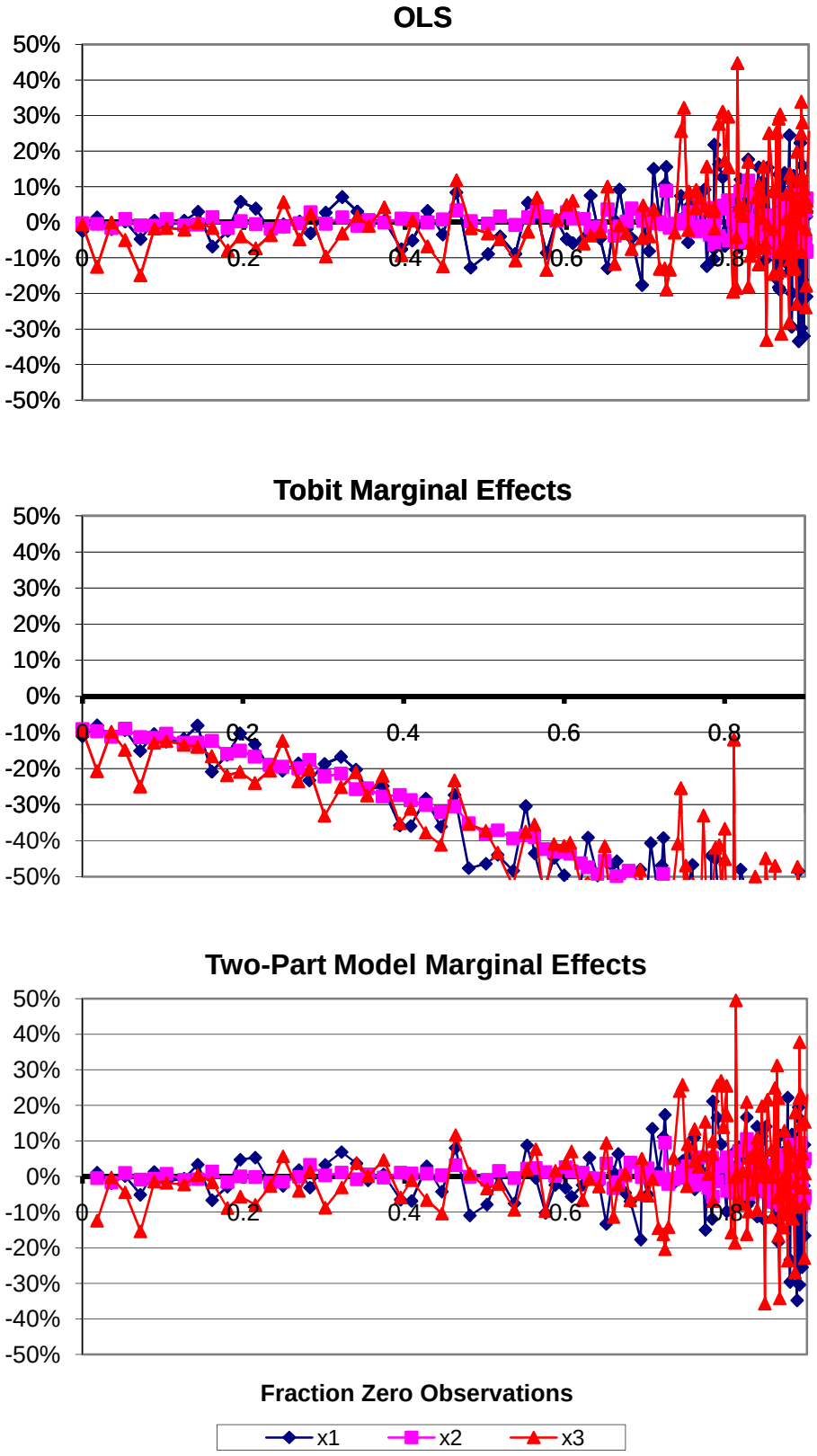

Figure 2a: The Fraction of Zero Observations is Negatively Related to the Amount of Time Spent in the Activity

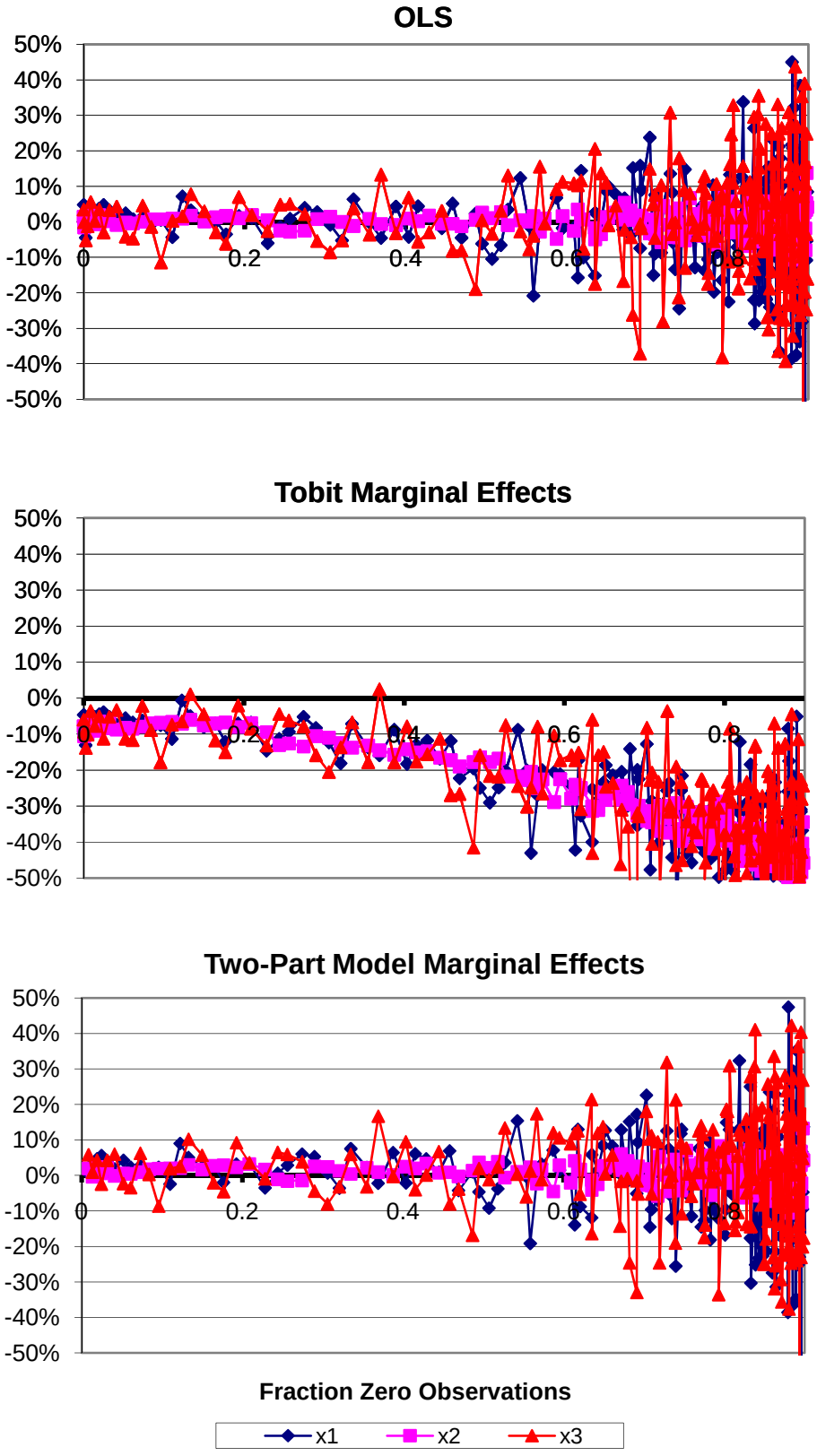

Figure 2b: The Fraction of Zero Observations is Negatively Related to the Amount of Time Spent in the Activity

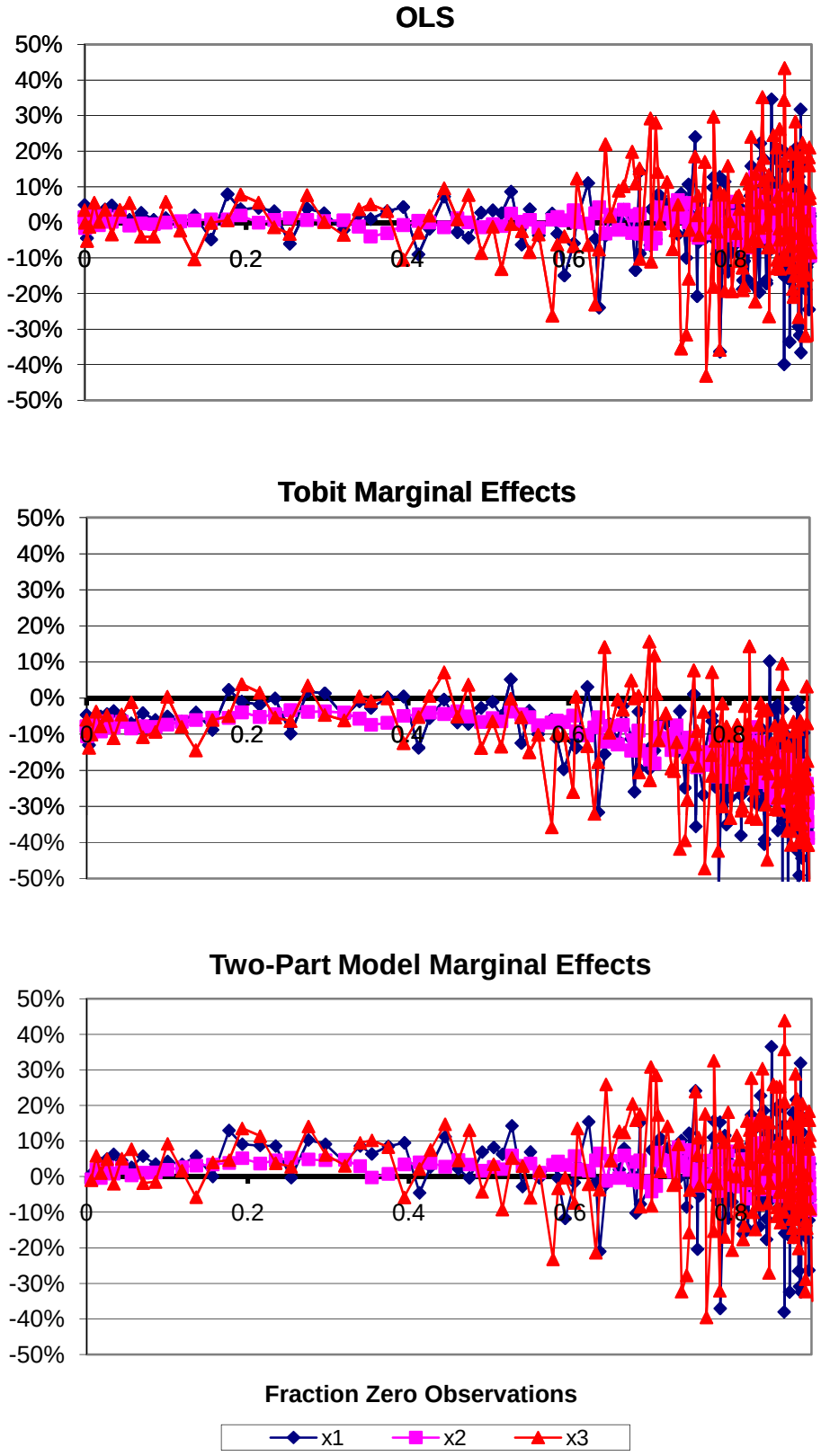

Figure 2c: The Fraction of Zero Observations is Negatively Related to the Amount of Time Spent in the Activity

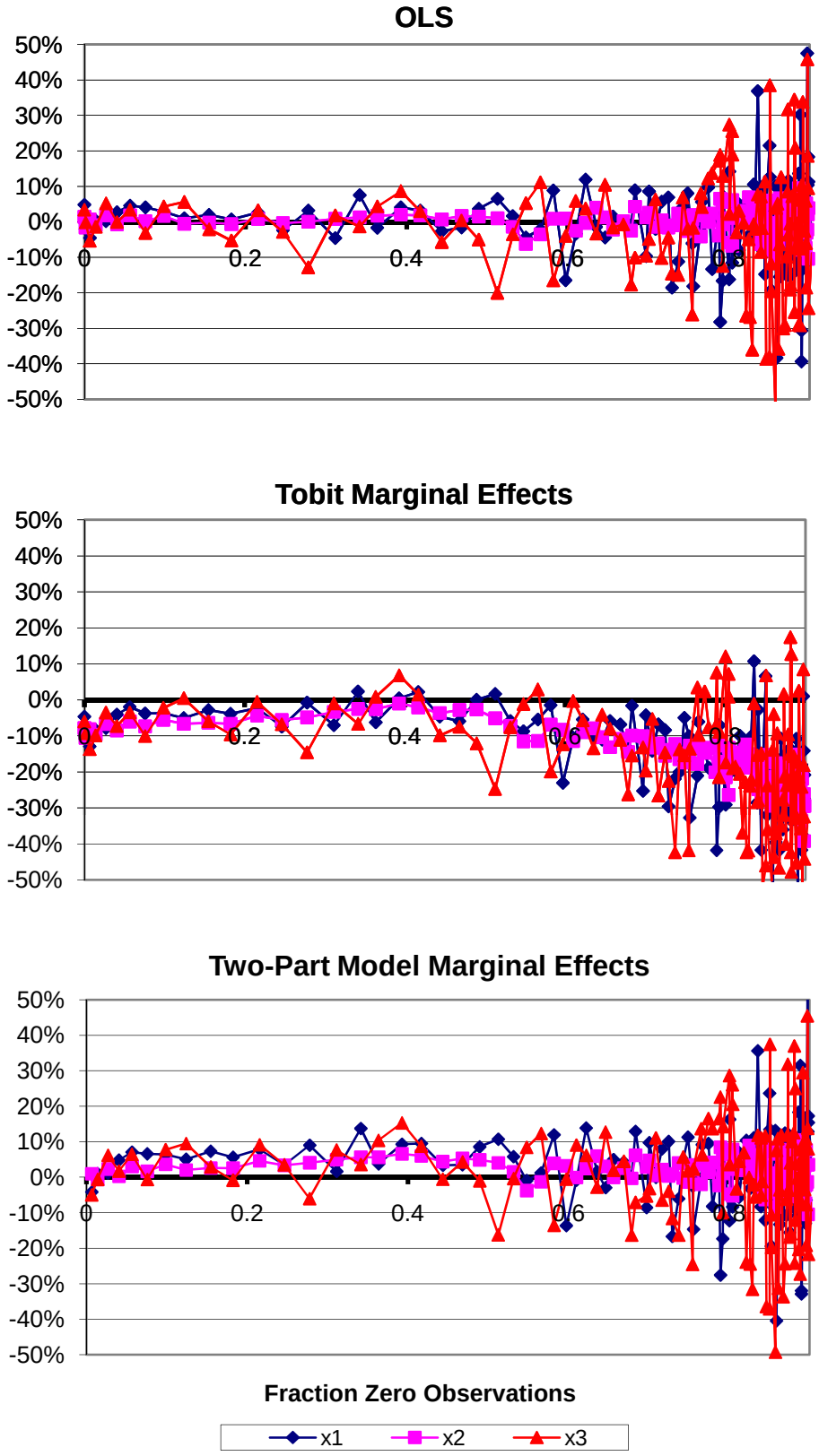

Figure 3: The Fraction of Zero Observations is Negatively Related to x_1

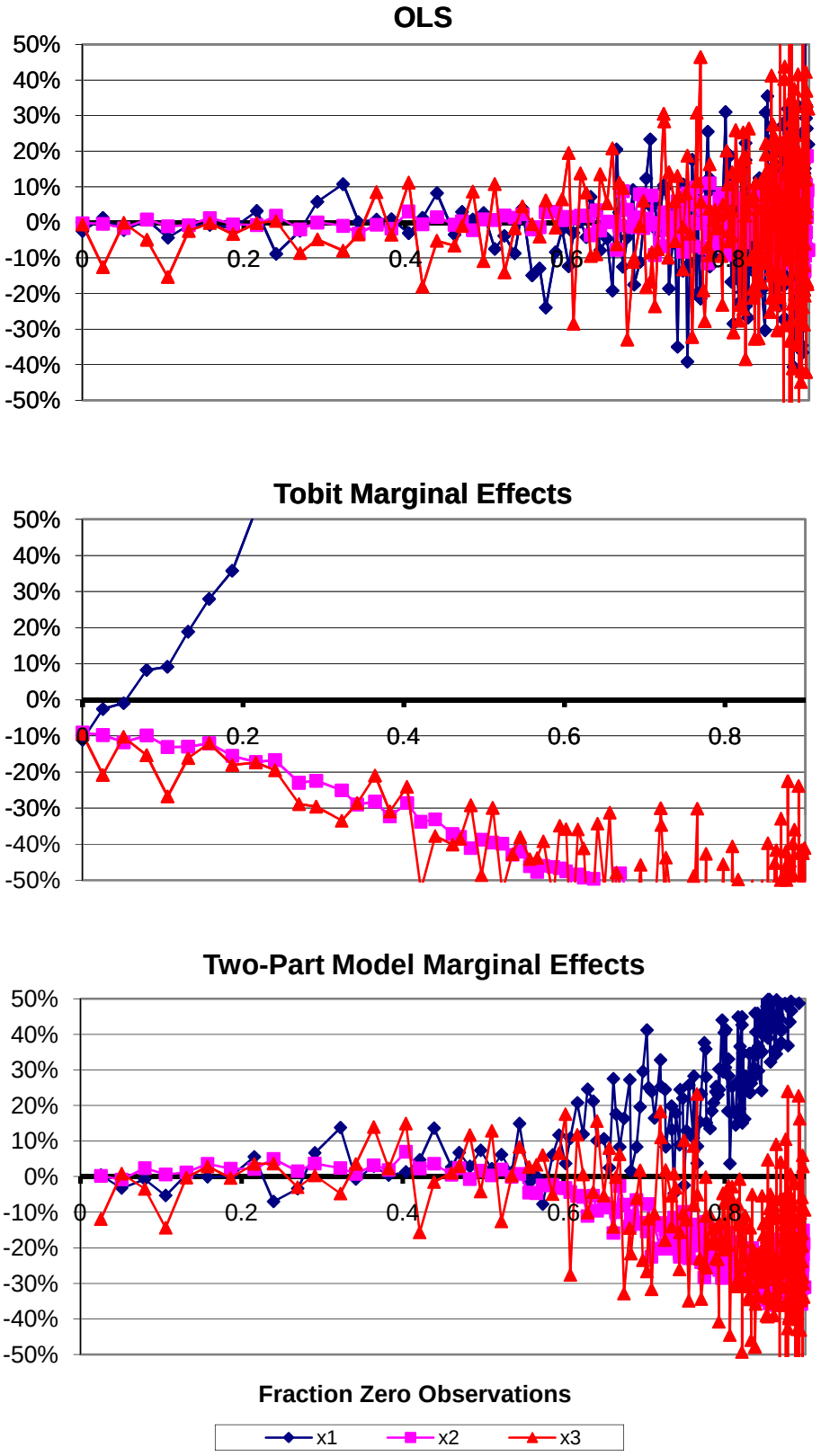

Figure 4: The Fraction of Zero Observations is Positively Related to x_2

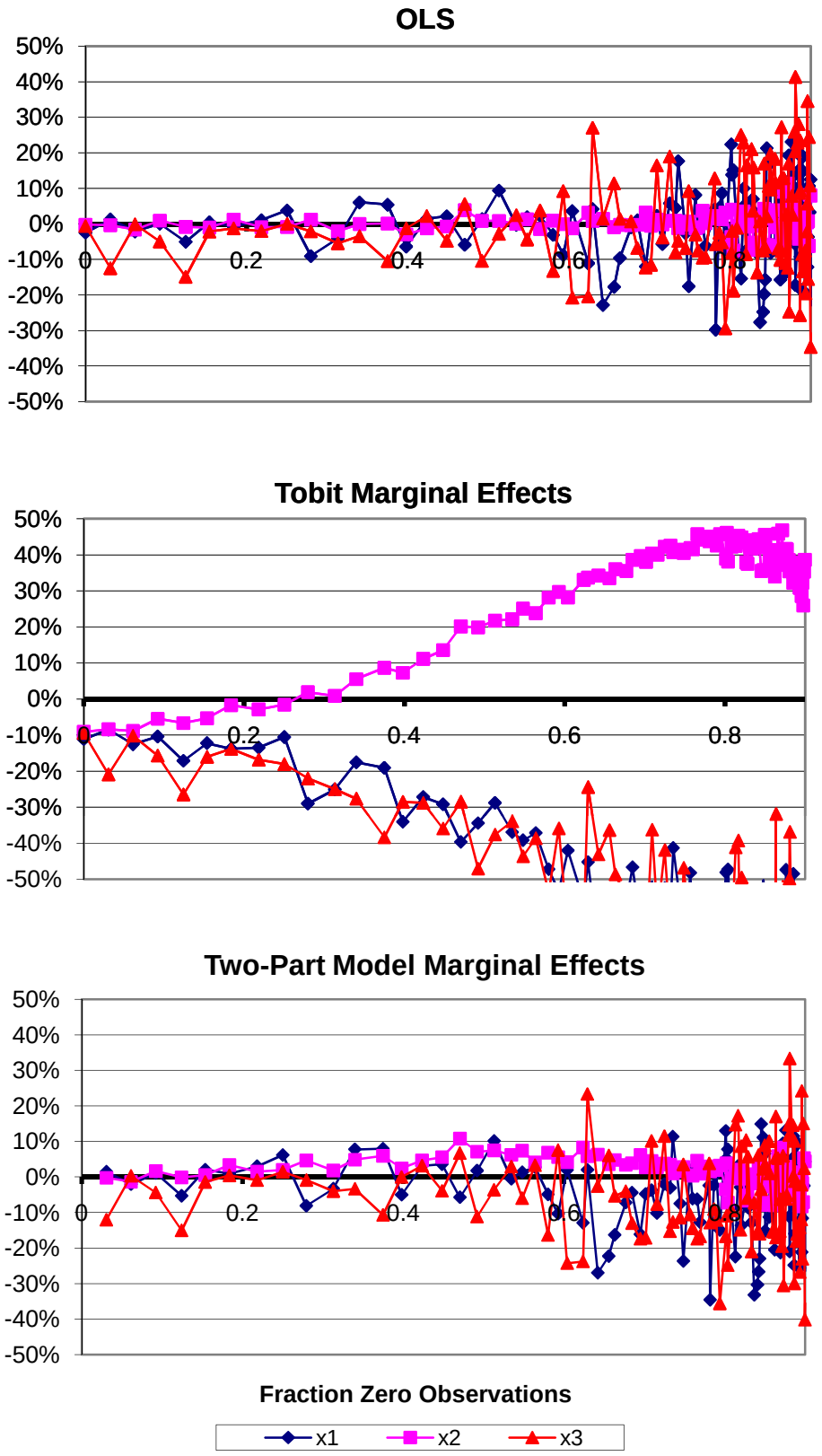

Figure 5: The Fraction of Zero Observations is Negatively Related to x_3