A Cross-Clustered Model of Home-Based Work Participation Frequency During Traditionally Off-Work Hours

Erika Spissu
University of Cagliari - Italy
CRiMM - Dipartimento di Ingegneria del Territorio
Via San Giorgio 12, 09124 Cagliari
Tel: + 39 070 675 6403; Fax: + 39 070 675 6402 E-mail: espissu@unica.it

Naveen Eluru
The University of Texas at Austin
Department of Civil, Architectural and Environmental Engineering
1 University Station, C1761, Austin, TX 78712-0278
Tel: 512-471-4535, Fax: 512-475-8744 Email: naveeneluru@mail.utexas.edu

Ipek N. Sener
The University of Texas at Austin
Department of Civil, Architectural and Environmental Engineering
1 University Station, C1761, Austin, TX 78712-0278
Tel: 512-471-4535, Fax: 512-475-8744 Email: ipek@mail.utexas.edu

Chandra R. Bhat*
The University of Texas at Austin
Department of Civil, Architectural and Environmental Engineering
1 University Station C1761, Austin, TX 78712-0278s
Tel: 512-471-4535, Fax: 512-475-8744 Email: bhat@mail.utexas.edu

and

Italo Meloni
University of Cagliari - Italy
CRiMM - Dipartimento di Ingegneria del Territorio
Via San Giorgio 12, 09124 Cagliari
Tel: + 39 070 675 5268, Fax: + 39 070 675 5261, E-mail: imeloni@unica.it

*****corresponding author

August 2009
Revised November 2009

ABSTRACT

The objective of this paper is to shed light on the determinants of working from home beyond the traditional office-based work hours. Specifically, we examine the frequency of work participation from home for individuals who also have the traditional work pattern of traveling to an out-of-home work place with a fixed number of work hours at the out-of-home work place. The sample for the empirical analysis is drawn from the 2002-2003 Turin Time Use Survey, which was designed and administered by the Italian National Institute of Statistics (ISTAT). From a methodological standpoint, we explicitly recognize both spatial and social clustering effects using a cross-clustered ordered-response structure to analyze the frequency of work participation from home during off-work periods. The model is estimated using the inference technique of composite marginal likelihood (CML), which represents a conceptually, pedagogically, and implementationally simpler procedure relative to traditional frequentist and Bayesian simulation techniques.

Keywords: Social and spatial dependency, composite marginal likelihood estimation, non-traditional work hours, multi-level modeling, cross-cluster analysis.

1. INTRODUCTION

The rapid advances in information and communication technologies or ICTs have substantially altered work patterns across the globe. Several studies have indicated that one consequence of the pervasiveness of the internet is a blurring of the traditional separation between "*work*" and "*non-work*" locations for conducting work (*1*, *2*). A 2008 survey across 2,252 adult Americans reported that 19% increased the amount of time spent working from home because of the availability of the internet (*3*). To provide further evidence of the growing teleworking from home trends, about 15% of U.S. workers worked remotely at home at least once a week in 2006 (*4*), while about 20% of European workers reported working at least a quarter of their working hours from home in 2005 (*5*).

The advances in ICTs are not only blurring work in the context of space (*i.e.*, **where from** work is pursued), but also blurring work in the context of time (*i.e.*, **when** work is pursued). There have been some studies in the social science and work habits literature (see (*6*) for a review) suggesting that, while ICTs provide a convenient means of obtaining and absorbing information almost instantaneously, it has also fed to a "workaholic" culture due to the ability to work virtually anytime with several consequent societal issues such as family time reductions and interruptions. Of course, there has also been a recognition for a long time now in the time-use and activity-based literature of the important potential impacts of ICT-related work patterns on individual time-use and activity-travel patterns (for instance, see *7-11*). In particular, these studies emphasize the importance of understanding work patterns as a precursor to generating and scheduling overall individual and household work and non-work patterns.

The above discussion clearly indicates the role of work patterns in shaping the way humans conduct their day-to-day life in general, and pursue their activity-travel patterns in particular. However, from the standpoint of examining work patterns themselves, the focus of earlier studies has been on work location rather than the temporal dimension of work. This latter dimension is typically considered for the traditional arrangement of individuals who travel out-of-home to work but not for work patterns that entail working partly from home and partly from work. The emphasis of this paper is on the latter kind of work pattern. Specifically, we examine the frequency of work participation from home for individuals who also have the traditional work pattern of traveling to an out-of-home work place with a fixed number of work hours at the out-of-home work place. The data for the analysis is drawn from a time-use survey conducted in Italy, where it is still very rare that employees have the option of telecommuting or of working away from their work place during regular working hours. But, an increasing fraction of Italians are working from home outside traditional work hours, according to a research conducted in Turin, Italy (*12*).

From a methodological standpoint, we use an ordered-response system to model frequency of work participation from home during traditionally off-work hours which explicitly recognize both spatial and social clustering effects using a multi-level structure. This is important since there may be unobserved effects (that is, those effects that cannot be directly captured through explanatory variables) based on spatial grouping effects (for example, individuals residing in a certain neighborhood may be uniformly more likely to work off-hours due to spatial proximity effects) and/or on social grouping effects (for instance, individuals who interact closely with one another in social circles may all be observed to cluster on the propensity to work off-hours from home; note that social grouping does not require any kind of spatial proximity). In such a multi-level clustering context, it is important to recognize and preserve between-cluster heterogeneity [*i.e.*, intrinsic differences across clusters; see (*13*) and (*14*)] because ignoring such heterogeneity, when present, would in general result in mis-estimated standard errors in linear models and (in addition) inconsistent parameter estimation in non-linear models. Besides, one has to consider local cluster-based variations in the relationship between the dependent and independent variables to avoid structural instability, especially in non-linear models. Finally, heterogeneity among aggregate cluster units (neighborhoods or social groups) and heterogeneity among elementary units (individuals) needs to be differentiated. The alternative of ignoring this differentiation and modeling the behavior of interest at a single level invites the pitfalls of either the ecological fallacy when the level of analysis is solely at the aggregate level (*i.e.*, failing to recognize that it is the elementary units which act and not aggregate units) or the atomistic fallacy when the analysis is pursued entirely at the elementary unit level (*i.e.*, missing the context in which elementary units behave).

There has been substantial interest in multi-level analysis in several fields, including education, sociology, medicine, and geography [see (*15*) for a recent review of multi-level models and their applications]. However, the application of the method has been almost exclusively confined to the case of a strictly hierarchical clustering structure. This can be easily handled using a multi-level structure by including a random effects term specific to each cluster, and estimating the parameters of the resulting model using the familiar maximum likelihood estimation [see, for example, (*16*)]. However, the situation changes entirely when elementary units can be classified into more than one higher-level unit (more on this in the next section). The net result of such cross-clustering is that, in general, the dimensionality of integration in the cross random-effects case explodes rapidly, making the likelihood maximization approach ineffective [see (*14*) for details].

In the current paper, we adopt the technique of composite marginal likelihood (CML) estimation, an emerging inference approach in the statistics field, though there has been little to no coverage of this method in econometrics and other fields [see (*17*) and (*18*)]. The CML estimation approach is a simple approach that can be used when the full likelihood function is near impossible or plain infeasible to evaluate due to the underlying complex dependencies, as is the case with econometric models with general cross random effects. The CML approach also represents a conceptually, pedagogically, and implementationally simpler procedure relative to simulation techniques, and also has the advantage of reproducibility of results.

The rest of the paper is organized as follows. Section 2 presents earlier discrete choice studies in the travel demand literature that use a multi-level structure, and positions the current study. Section 3 outlines the econometric methodology. Section 4 presents details of the data and sample characteristics. Section 5 presents the empirical results and, finally, Section 6 concludes the paper.

2. EARLIER CROSS-LEVEL CLUSTERING APPROACHES AND THE CURRENT STUDY

As indicated earlier, strictly hierarchical multi-level analysis has seen substantial application in the literature, especially in the context of linear models. However, the past decade has also seen the application of multi-level analysis in non-linear models in the activity analysis field (*16*, *19*). In Bhat and Zhao (*16*), the clustering is purely spatial and based on residential zone. These authors examine the number of daily shopping stops made by households, while considering spatial clustering effects. In Dugundji and Walker (*19*), the clustering is based on a combination of residential district/post code (to represent spatial clustering effects) and socio-economic grouping (to proxy social interaction effects). These authors examine mode choice to work, while accommodating spatial and social clustering effects. However, both the Bhat and Zhao (BZ) and the Dugundji and Walker (DW) studies adopt strictly hierarchical clustering structures, wherein each individual is assigned to one and only one cluster (and the clusters are mutually exclusive and collectively exhaustive). Such structures lend themselves rather easily to maximum simulated likelihood estimation, since the strictly hierarchical clustering is accommodated through cluster-specific mixing random effects and individuals can be grouped into one of several clusters. The important point is that the dimensionality of integration of the probability expressions appearing in the BZ and DW studies is independent of the number of clusters.

The only earlier study in the travel demand literature that the authors are aware of that captures cross-cluster effects is the one by Bhat (*14*). Bhat, like DW, also models work mode choice, but allows cross-clustering based on residential location and work location. To allow maximum likelihood estimation, Bhat has to use very aggregate spatial definitions of the work location, which reduces the dimensionality of the integration in the likelihood function and allows the use of simulation techniques. However, Bhat's simulation approach is infeasible in the more general case of cross-cluster effects with several clusters in both dimensions, or when the cross-cluster effects are based on clustering in more than two dimensions. The main problem in these more general cases is that the dimensionality of integration is no more independent of the number of clusters in each dimension. To give a sense of the dimensionality, if Bhat had used the same spatial resolution of traffic analysis zones in defining work locations as in defining residential locations (193 traffic analysis zones), the number of dimensions of integration would have been of the order of 193*number of travel modes or 600 dimensions. As importantly, this integration would have to be undertaken in Bhat's study over a conditional likelihood function integrand involving the product of the probabilities of each individual in the entire sample. Consequently, the likelihood maximization involves likelihood evaluations with numerically extremely small values, causing substantial instability problems.[1]

In the current paper, we apply a composite marginal likelihood (CML) approach for cross-clustering in the context of an ordered response structure. Generally speaking, the CML approach, originally proposed by Lindsay (*21*), entails the development of a surrogate likelihood function that involves easy-to-compute, low-dimensional, marginal likelihoods [see (*17, 18*) for extensive reviews and discussions). We implement the CML approach here based on the marginal likelihood of pairs of individuals. The approach is ideally suited for crossed random effects since it entails only bivariate distribution function evaluations, independent of the number of dimensions of clustering or the number of clusters within each dimension [see (*20*) who consider the CML approach for crossed-random effects in generalized linear mixed models]. Further, the CML approach can be applied using simple optimization software for likelihood estimation and is based on a classical frequentist approach. Its basis in the theory of estimating equations [see (*21*, *22*)] ensures that the CML estimator is consistent, unbiased, and asymptotically normally distributed. The CML estimator (theoretically speaking) loses some efficiency relative to traditional maximum likelihood estimation, though this efficiency loss has been showed to be negligible in practice [see (*23*)]. In any case, the CML estimator is perhaps the only practical approach currently to estimate parameters in general cross-random effects contexts.

[1] Note that taking the logarithm of the likelihood function of the entire sample, as is the norm in the maximum likelihood method, offers no benefit whatsoever because the log-likelihood function does not simplify to the sum of the logarithm of the likelihood function of clusters involving fewer individuals than the entire sample. Further, note that even Bayesian techniques are impractical for the case of cross random effects because they require extensive simulation and are time-consuming (*20*). In this regard, both the ML and the Bayesian approach may be viewed as "brute force" simulation techniques that are not straightforward to implement and can create convergence assessment problems.

3. METHODOLOGY
3.1 Model Structure

In the current section, we describe the model structure and estimation methodology in the general context of an ordered-response model with two-dimensional cross-random clustering. In the substantive context of the current paper, the dependent variable in the ordered-response model corresponds to the frequency of work participation from home for individuals who have the traditional work pattern of traveling to an out-of-home work place with a fixed number of work hours at the out-of-home work place. The two-dimensional clustering corresponds to spatial clustering based on the residential location of the individual and social clustering based on the social grouping to which the individual belongs. The specific manner in which the spatial and social clusters are defined and implemented in our empirical analysis is discussed subsequently.

In the usual framework of an ordered-response model, let the underlying latent continuous random propensity z^*_{qij} of individual q in spatial cluster i and social cluster j be related to a vector x_{qij} of relevant explanatory variables as follows:

$$z^*_{qij} = \alpha_{ij} + \beta' x_{qij} + \varepsilon_{qij}\ ,\ \ z_{qij} = k \ \text{ if } \psi_k < z^*_{qij} \le \psi_{k+1}, \tag{1}$$

where α_{ij} is a scalar term associated with spatial group i and social group j, β is a vector of coefficients to be estimated, ε_{qij} is a standard normally distributed random term, z_{qij} is the observed ordinal frequency of working from home during off-work times, k is the index for the ordinal frequency category (k = 1, 2, …, K) and ψ_k is the lower bound threshold for ordinal level k ($\psi_0 < \psi_1 < \psi_2 ... < \psi_K < \psi_{K+1}$; $\psi_0 = -\infty$, $\psi_{K+1} = +\infty$). ε_{qij} is assumed to be independent of the elements in α_{ij}, β and x_{qij}. Also, as formulated above, x_{qij} does not include a constant term. The variance in the scalar term α_{ij} represents intrinsic unobserved heterogeneity across individuals in their propensity to work off-hours from home based on their residential location and social grouping.

Equation (1) represents the micro-level model for individuals. We now allow the scalar term α_{ij} to vary across spatial clusters and social groups in a higher-level macro-model:

$$\alpha_{ij} = \lambda' w_{ij} + u_i + v_j\,, \tag{2}$$

where w_{ij} is a vector of observed variables specific to spatial cluster i or social group j or to the combination of spatial cluster i and social group j, λ is a parameter vector to be estimated, and u_i and v_j are random terms that capture unobserved variations across spatial groups and social groups, respectively, in the propensity of working from home during off-work hours. The latter two error terms are assumed to be realizations from independently and identically normal distributed terms across spatial and social clusters, respectively: $u_i \sim N(0,\sigma^2)$ and $v_j \sim N(0,\eta^2)$. Next, define $\gamma = (\beta', \lambda')'$ and $s_{qij} = (x'_{qij}, w'_{ij})'$. Then, the micro- and macro-models of Equations (1) and (2) can be combined to form:

$$z_{qij}^{*} = \gamma' s_{qij} + u_i + v_j + \varepsilon_{qij}\ ,\ \ z_{qij} = k \ \text{ if } \psi_k < z_{qij}^{*} \le \psi_{k+1}, \tag{3}$$

The usual independence assumption among all error terms is invoked. Note that if σ^2 (variance of u_i) and η^2 (variance of v_j) are equal to zero, then it implies that there are no variations (due to unobserved factors) in the propensity to works off-hours from home across spatial and social clusters, respectively. In this case, the cross-random ordered-response (CROR) model of Equation (3) collapses to the standard ordered-response (SOR) model. The implication is that all unobserved heterogeneity is due to inter-individual differences, and there is no unobserved heterogeneity based on spatial and social clustering. Note also that the propensity specification of Equation (3) generates a covariance pattern among individuals as follows: for two individuals in the same spatial cluster, but not in the same social cluster, the covariance in their propensities to work off-hours is σ^2. For two individuals in the same social cluster, but not in the same spatial cluster, the covariance is η^2. For two individuals in the same spatial and social cluster, the covariance is $\sigma^2 + \eta^2$. Finally, for two individuals not in the same spatial cluster nor in the same social cluster, the covariance is zero.

3.2 Estimation Approach

In the current paper, we use a pairwise marginal likelihood estimation approach, which corresponds to a composite marginal approach based on bivariate margins [see (*18, 24-26*) for the use of the pairwise likelihood approach in the recent past). Each bivariate margin represents the joint probability of the observed frequency of working off-hours from home for a pair of individuals *q* and *h* in the sample. The presence of spatial and social clustering effects leads to covariance effects between the pair of individuals *q* and *h* based on their spatial and social groupings.

In this section, since each individual *q* is uniquely identified with a particular spatial cluster *i* and a particular social cluster *j*, it is convenient from a presentation standpoint to suppress the indices *i* and *j*. Thus, we will use the notation z_q for z_{qij}, and s_q for s_{qij}. Also, let d_q be the actual observed ordinal frequency of working from home during off-work hours for individual *q*. The pairwise marginal likelihood function may then be written, after defining $\delta = (\gamma', \sigma, \eta)'$, as:

$$\begin{aligned}
L_{CML}(\delta) &= \prod_{q=1}^{Q-1} \prod_{h=q+1}^{Q} \left[P(z_q = d_q, z_h = d_h) \right], \\
P(z_q = d_q, z_h = d_h) &= \Phi_2\left(u_{(d_q)}, u_{(d_h)}, \theta_{qh}\right) - \Phi_2\left(u_{(d_{q-1})}, u_{(d_h)}, \theta_{qh}\right) \\
&\quad - \Phi_2\left(u_{(d_q)}, u_{(d_{h-1})}, \theta_{qh}\right) + \Phi_2\left(u_{(d_{q-1})}, u_{(d_{h-1})}, \theta_{qh}\right), \\
u_{(d_q)} &= \left(\frac{\psi_{(d_q)} - \gamma' s_q}{\sqrt{1+\sigma^2+\eta^2}} \right) \text{ and } \theta_{qh} = \frac{G_{qh}\sigma^2 + R_{qh}\eta^2}{1+\sigma^2+\eta^2}.
\end{aligned} \tag{4}$$

In the above expression, $G_{qh} = 1$ if *q* and *h* are in same spatial cluster and $G_{qh} = 0$ otherwise. Similarly, $R_{qh} = 1$ if *q* and *h* are in same social cluster and $R_{qh} = 0$ otherwise. The bivariate probability expressions in the pairwise marginal likelihood function of Equation (4) are straightforward to compute, since they only entail four bivariate standard normal expressions. The pairwise marginal likelihood function comprises $Q(Q-1)/2$ pairs of bivariate probability computations, which can itself become quite time consuming. Fortunately, the individuals that have no spatial and no social interdependencies can be pre-identified. Our coding exploits this situation to enable the relatively fast maximization of the logarithm of the pairwise marginal likelihood function.

The CML estimator obtained by maximizing the logarithm of the function in Equation (4) with respect to the γ, σ, and η parameters is consistent and asymptotically normal distributed with the asymptotic variance matrix given by Godambe's (*27*) sandwich information matrix:

$$F(\delta)=\frac{1}{Q}\left[H(\delta)\right]^{-1}J(\delta)\left[H(\delta)\right]^{-1}, \tag{5}$$

where

$$H(\delta)=E\left[-\frac{\partial^2 \log L_{CML}(\delta)}{\partial\delta\partial\delta'}\right] \text{ and}$$
$$J(\delta)=E\left[\left(\frac{\partial \log L_{CML}(\delta)}{\partial\delta}\right)\left(\frac{\partial \log L_{CML}(\delta)}{\partial\delta'}\right)\right].$$

The "bread" matrix $H(\delta)$ of Equation (5) can be estimated in a straightforward manner using the Hessian of the negative of $\log L_{CML}(\delta)$, evaluated at the CML estimate $\hat{\delta}$. This is because the information identity remains valid for each pairwise term forming the composite marginal likelihood. Thus, $H(\delta)$ can be estimated as:

$$\hat{H}(\hat{\delta})=\left(-\left[\sum_{q=1}^{Q-1}\sum_{h=q+1}^{Q}\frac{\partial^2 \log L_{CML,qh}(\hat{\delta})}{\partial\hat{\delta}\partial\hat{\delta}'}\right]\right), \tag{6}$$

where

$$L_{CML,qh}(\hat{\delta})=[P(z_q=d_q, z_h=d_h)] \;\; |\,\hat{\delta}.$$

However, the estimation of the "vegetable" matrix $J(\delta)$ is more difficult, since $\frac{\partial \log L_{CML}(\delta)}{\partial\delta}$ vanishes when evaluated at the CML estimate $\hat{\delta}$. Further, one cannot estimate $J(\delta)$ as the sampling variance of individual contributions to the composite score function because of the underlying spatial and social dependence in observations. In addition, the non-decaying correlation pattern of the current framework does not permit the use of the windows resampling procedure of Heagerty and Lumley (*28*) to estimate $J(\delta)$ as in Bhat *et al.*, (*18*). Hence we resort to pure Monte Carlo computation to estimate $J(\delta)$ (see *18*). In this approach, we generate B data sets $Z^1, Z^2, ..., Z^B$, where Z^b (b = 1, 2, …, B) is a vector of one possible realization for $(z_1, z_2, ..., z_Q)$ for the exogenous variable vector $S = (s_1, s_2, ..., s_Q)$ (under the assumed model with $\delta = \hat{\delta}(S)$). Once these datasets are generated, the estimate of $J(\delta)$ is given by:

$$J(\delta) = \frac{1}{B}\sum_{b=1}^{B}\left[\left(\frac{\partial \log L_{CML}(\delta)}{\partial \delta}\right)\Bigg|_{Z^b}\left(\frac{\partial \log L_{CML}(\delta)}{\partial \delta'}\right)\Bigg|_{Z^b}\right]$$

The above computation is not very demanding because the model in Equation (1) can be generated in a straight-forward manner. We tested various values of B for the stability in the estimate of $J(\delta)$, and found that a value of B = 1000 was much more than adequate for reasonable accuracy.

4. DATA AND SPATIAL-SOCIAL CLUSTER DEFINITIONS

4.1 Data Sources and Sample Used

The primary source of data used in the paper is the 2002/2003 Turin Time Use Survey, which was designed/administered by the Italian National Institute of Statistics (ISTAT) and sponsored by the Turin Town Council and 14 neighboring town councils (Baldissero Torinese, Beinasco, Borgaro Torinese, Collegno, Grugliasco, Moncalieri, Nichelino, Orbassano, Pecetto Torinese, Pino Torinese, Rivoli, San Mauro Torinese, Settimo Torinese, Venaria Reale). The survey collected a daily activity time-use diary from each of 4537 household members aged 3 years and older from 1830 households [see (*29*) for details of the survey design and administration procedures]. Detailed work characteristics and demographic information was also obtained from all surveyed households and individuals, including residence location in one of the 24 municipalities in the Turin area.[2]

In addition to the main time-use survey data, the 2001 Census of population (*30*) and the 2001 Census of Industry and Services (*31*) data sets were used to obtain built environment variables including housing type measures, number of commercial, industrial and service units, population density, household density, and employment density.

The final sample used for the estimation includes 2042 individuals aged 14 or more. Since the focus of the study is on work participation during traditionally off-work hours, only workers ("employed" individuals as identified in the survey) were considered in the analysis. The survey specifically asked workers to provide their off-work hours participation frequency from home in ordered-response categories, which serves as the dependent variable for the model. These categories and the corresponding distribution among workers, are as follows: (1) "never": 1597 (78.2%), (2) "few times a month": 173 (8.5%), (3) "few times a week": 182 workers (8.9%), and (4) "everyday" 90 (4.4%). These sample shares of individuals working during off-hours at home, as obtained from the Turin area, are similar to corresponding figures for the entire of Italy (*29*).

4.2 Spatial-Social Cluster Definitions

As discussed earlier, this study incorporates both spatial and social clustering effects. An issue in introducing such effects is that the analyst has to define the spatial unit and the social unit for

[2] The information regarding household residence within the ten districts of Turin city, while collected in the time use survey, was missing in the data. However, we had access to the set of sampled households (with household/individual characteristics) residing in each municipality (label this the location data set). As a means to identify each sampled household in the time-use survey with a specific municipality in the location data set, we used a probabilistic linking procedure based on individual/household demographics. A household in the time-use survey was linked to a household in the location data set only if the probability of a true match was 0.95 or higher. The details of this linking procedure are available upon request from the authors.

accommodating the clustering effects. In the context of the current study, we used the finest spatial resolution to which individual residential locations could be identified with, which corresponded with locations in 24 municipalities (14 municipalities outside Turin and 10 municipalities inside Turin). Two different spatial clustering schemes were considered based on (1) whether or not workers were resident in the same zone, and (2) whether or not workers were resident in the same or immediately adjoining zone. The second spatial clustering scheme is more expansive than the first.

A number of social clustering schemes may be defined based on the social circle in which each worker moves as part of her/his daily life. In the social science literature, various different characterizations of social interaction and social proximity have been proposed [see, for example, (*32*)] based on direct relationships (such as with family members, work colleagues, and friends), indirect relationships (with individuals living in the same environment), and cultural relationships (with individuals of the same social status, race, and ethnicity). In the transportation literature, social interactions/proximity effects have been examined in one of the following two broad ways: (1) Using social network geographies (*33*), and (2) Using the egocentric approach (*34*). The former approach focuses on identifying all the possible social connections that each individual has through a survey or other data collection methods. The latter approach, on the other hand, clusters individuals in social groups based on their demographics and/or attitudes toward joint activities. In the current analysis, we do not have any explicit information on the social circle of each individual in the sample, but we do have rich demographic information on each individual. Thus, our definition of social interactions is based on an egocentric approach as in Dugundji and Walker (*19*), where we consider social interactions to be particularly strong among individuals who share certain common demographic characteristics.[3] The demographic groupings considered in the empirical analysis included the following key variables (and also groupings based on various possible combinations of these key variables): (1) whether or not workers share the same marital status (never married, currently married, or separated/divorced/widowed), and (2) whether or not workers have children (including examination of children in different age groups).

An important point to note here. Regardless of the definitions considered for the social clustering effect, there is always a diversity of individuals in each municipality based on social grouping. Thus, it is not possible to have independent and mutually exclusive clusterings of individuals in the sample based on both spatial and social clustering. That is, the cross-random effects lead to a global interaction network over the entire sample, leading to a kind of a giant cluster with dependencies between each pair of workers.

5. MODEL ESTIMATION RESULTS

5.1 Variable Specification and Cluster Definitions

Several different exogenous variables, functional forms of variables, and variable interactions were considered in the model specifications. The exogenous variables included (1) individual demographics (age, gender, *etc.*), (2) individual work-related characteristics (work schedule, number of jobs, *etc.*), (3) household characteristics (household composition, household income, *etc.*), and (4) built environment measures of the individual's residence location (in terms of work-related characteristics, work occupation status was collected in a very ambiguous manner that did not allow us to use it as an exogenous variable; including occupation status in the analysis is important in further research). The final specification was based on a systematic process of removing statistically

---

[3] Note, however, that the methodology proposed in this paper is generic, and can be applied with any approach that identifies the social "network" of an individual.

insignificant variables and combining variables when their effects were not significantly different. The specification process was also guided by prior research and intuitiveness/parsimony considerations. We should also note here that, for the continuous variables in the data (such as age, work hours at the out-of-home work location, and income), we tested alternative functional forms that included a linear form, a spline (or piece-wise linear) form, and dummy variables for different ranges. For all the continuous variables, the use of dummy variables provided the best results and is employed in the final specification (as we will note later, income, however, did not turn out to be statistically significant). Different threshold values to define the dummy variables were tested, and the ones that provided the best fit were used.

In addition to several variable specifications, we also considered various spatial and social clustering schemes through the specification of the G_{qh} and R_{qh} dummy variables, as discussed in Section 4.2. The best specification was obtained with a spatial clustering scheme based on "whether or not two workers reside in the same Municipality" and the social clustering scheme based on whether or not workers share the same marital status (never married, married, or separated/divorced/widowed). In the discussions below, we present only the results of the final variable specification with the best spatial/social clustering scheme to keep the discussions streamlined.

5.2 Estimation Results

Table 1 presents the model estimation results. The base category is listed for each discrete exogenous variable in parenthesis. The coefficients in the table indicate the effects of variables on the latent propensity of work participation from home during off-work hours (that is, they represent elements of the γ vector as defined in Section 3.1). The reader will note that, while many different functional forms were tested for several variables, the final specification includes only dummy variables. Thus, the magnitudes of the coefficients also provide an estimate of the importance of the variables in influencing off-work participation propensities and participation probabilities.

5.2.1 Individual Demographics

The effects of individual demographics indicate that young individuals (<45 years of age) are less likely than individuals over 45 years of age to work from home during off-work hours. This is particularly the case for the very young individuals (≤30 years of age). This result may be a reflection of the generally wider social networks of young individuals (see *35*) and the generally higher participation rates of younger individuals in out-of-home discretionary and maintenance activities (see *36-39*). Such tendencies would reduce the time available for, or the inclination to, work at home outside standard work hours. Further, the results point to the increased home-based work during off-work times among more educated individuals relative to less educated individuals, perhaps because of demanding and high-status work positions.

5.2.2 Work Characteristics

Among the work-related characteristics, the results suggest that individuals who work more than 40 hours a week at their out-of-home work place are also more likely to work during off-work hours, presumably because of overall "workaholic" tendencies. The work schedule variables show that individuals who work in the evening and on Saturdays are more likely to bring work home compared to other workers, while those who work in the night are less likely to bring work home relative to those who work at any other time. In addition, individuals with a second job are more likely to bring work home, presumably due to time pressures and the need to juggle multiple things at multiple jobs. In addition, self employed individuals are more likely to work during off-work times relative to those who are not self-employed.[4] This is not surprising, since self-employed individuals have a personal vested incentive to work harder and longer hours to ensure that their personal venture is successful.

5.2.3 Household Characteristics

Among the household characteristics, the results show a lower tendency to bring work home among individuals who live in households with young adults aged 18 to 24 years (*i.e.*, driving license age adults) compared to individuals living in households with young or no children. While this result needs further exploration in future studies, one reason may be because of more social interaction within the family when there are multiple adults in the household (relative to when there are young children or no children). As expected, internet access at home increases the tendency to work from home during off-work hours (*6*). Of course, this raises the issue of whether it is the presence of internet access that makes individuals work more from home during off-work times, or whether the need to work during off-work time motivates individuals to set an up internet connection at home. While both of these two directions of effects are potentially possible (and may apply to different sets of individuals), internet access is becoming a way of life even for maintaining social connections and family-connectedness (*3*). Thus, it is likely that setting up an internet connection at home is not solely influenced by a desire to work from home during off-work times, but, once available, internet access "sucks" people into work-related activities. Finally, the presence of a house-helper has a positive impact on the propensity to work from home during off-work times. The fact that the individual's household contracts out household tasks is potentially a sign of work spillover effects at home.

5.2.4 Built Environment Attributes of Residence Location

As indicated earlier in the paper, we considered several built environment measures associated with the municipality of each individual's residence in our model specifications. However, none of these came out to be statistically significant. Of course, it is possible that this is because of using rather aggregate spatial units to compute built environment measures. Future research efforts should explore finer resolutions of space for computing built environment measures.

5.2.5 Threshold Parameters

The thresholds do not have substantial interpretations. They simply serve to translate the underlying latent propensity to work off-hours to the ordinal categories of working off-hours.

5.3 Spatial and Social Dependence Effects

As discussed in the previous sections, the cross-random ordered-response (CROR) model of this paper accommodates spatial dependence among workers residing in the same municipality and social dependence among workers based on marital status. The variance terms σ^2 (variance of u_i in Equation (3)) and η^2 (variance of v_j in Equation (3)) provide the magnitude of this dependence, as indicated in Section 3.1. Table 1 indicates the estimated standard deviations σ and η, both of which

[4] In Italy, it is the norm even for self-employed individuals to have an out-of-home office location, with set times of work at the out-of-home office.

are statistically significant at the 0.05 level of significance (note that the standard deviation terms should be positive, and so a one-tailed t-test is warranted). The spatial clustering effect is particularly strong and highly statistically significant. In terms of the CML log-likelihood value at convergence, the CROR value is -2698809 compared to the SOR value of -2700118. The composite marginal likelihood ratio test (CLRT) statistic, computed as twice the difference in the composite marginal log-likelihood values, yields a value of 2618. However, this CLRT statistics does not have the standard chi-squared asymptotic distribution under the null hypothesis as in the case of the maximum likelihood inference procedure. While one can use a bootstrapping approach to obtain the precise distribution of the CLRT statistic, several bootstrapping runs are needed, which becomes cumbersome. In any case, the t-statistics on the σ and η parameter estimates are statistically significant, indicates the data fit superiority of the CROR model over the SOR model. Further, as in the usual likelihood approach, one may compute an adjusted rho-bar squared value $\bar{\rho}_c^2$ in the composite marginal likelihood approach for the CROR model and the SOR models as $\bar{\rho}_c^2 = 1 - [(\log L_{CML}(\hat{\delta}) - H) / \log L_{CML}(C)]$, where $\log L_{CML}(\hat{\delta})$ is the composite marginal log-likelihood at convergence, H is the number of model parameters excluding the thresholds, and $\log L_{CML}(C)$ is the log-likelihood with only thresholds in the model. The value of $\bar{\rho}_c^2$ for the CROR model is 0.142, while that for the SOR model is 0.141.

5.4 Aggregate Elasticity Effects

The parameters on the exogenous variables in Table 1 do provide a sense of the relative magnitudes of effects of variables on the propensity to work from home during off-work hours. This is because all variables in the model are dummy variables. The results indicate that the education variables and having a second job have the highest impact, followed by presence of a house-helper, young age (age between 14 and 30 years), work schedule characteristics, self-employment status, and hours of work at the out-of-home work location. Interestingly, internet access does not have as substantial an effect as the other variables.

The coefficients in Table 1, while providing a sense of relative magnitudes of the effect of variables on the propensity to work from home during off-work hours, do not provide the absolute magnitude of effects of variables on the probability of each frequency level of working from home during off-work periods. To obtain such absolute magnitude effects, we compute the aggregate level "elasticity effects" of variables. In particular, we change the value of each dummy explanatory variable to one for the subsample of observations for which the variable takes a value of zero and to zero for the subsample of observations for which the variable takes a value of one. We then sum the shifts in expected aggregate shares of each frequency level of working from home in the two subsamples after reversing the sign of the shifts in the second subsample, and compute an effective percentage change in the expected aggregate share of teenagers participating in each number of activity episodes due to a change in the dummy variable from 0 to 1.

The elasticity effects are presented in Table 2 for each frequency level of working from home during off-work hours. The numbers in the table may be interpreted as the percentage change in the probability of working from home due to a change in the variable from 0 to 1. For instance, the first number in Table 2 indicates that the probability of a young individual (14-30 years of age) never working from home is 11.55% higher than the probability of an individual older than 45 years never bringing work home, other characteristics being equal. The results in Table 2 confirm the relative magnitude effects of variables discussed earlier. Indeed, it is interesting to note that, at least in Italy, having internet access is "trumped up" by other factors in the effect on working from home beyond the usual workday at the office. Studies of work-life balance have focused recently on how technology impacts home-based work and work-life balance [see, for example, (*2*, *40*)], which is clearly an important and valuable direction of research. However, it appears from our results that work characteristics and individual/household demographics play a much more important role in determining who works from home during off-work hours and who does not. In particular, those with high education levels, work multiple jobs, and are self-employed are particularly predisposed to working from home.

6. CONCLUSIONS

The rapid advances in telecommunications and information technologies (such as computers, and smart phones) have altered work patterns across the globe. These advances are not only blurring work in the context of space (*i.e.*, **where from** work is pursued), but also blurring work in the context of time (*i.e.*, **when** work is pursued). However, the emphasis in the work-based research literature has been on the spatial context (*i.e.*, where from work is conducted, such as telecommuting from home or working at a satellite work location or working at the regular work location), rather on the temporal context of work. This latter dimension is typically considered for the traditional arrangement of individuals who travel out-of-home to work (in terms of work start time at the traditional work place and work end time at the traditional work place), but not for work patterns that entail working partly from home and partly from work on a workday, or for work patterns that involve working from home beyond the traditional office-based work hours. The emphasis of this paper is on understanding the determinants of working from home beyond the traditional office-based work hours. Specifically, we examine the frequency of work participation from home for individuals who also have the traditional work pattern of traveling to an out-of-home work place with a fixed number of work hours at the out-of-home work place. The sample for the empirical analysis is drawn the 2002-2003 Turin Time Use Survey, which was designed and administered by the Italian National Institute of Statistics (ISTAT).

From a methodological standpoint, we explicitly recognize both spatial and social clustering effects using a cross-clustered ordered-response structure. This is important since there may be unobserved effects (that is, those effects that cannot be directly captured through explanatory variables) based on spatial grouping effects and/or on social grouping effects. The net result of such cross-clustering is that, in general, the dimensionality of integration in the cross random-effects case explodes rapidly, making the likelihood maximization approach ineffective if not infeasible. In the current paper, we adopt the technique of composite marginal likelihood (CML) estimation, which represents a conceptually, pedagogically, and implementationally simpler procedure relative to simulation techniques, and also has the advantage of reproducibility of results.

The results from our analysis indicate the statistically significant presence of spatial and social clustering in the underlying propensity to work from home during off-work hours. Ignoring such clustering can, and in general, will lead to inconsistent model parameter results as well as poorer data fit. The results also demonstrate the practical flexibility of the CML approach in situations where traditional maximum likelihood techniques or Bayesian techniques are either practically infeasible or ineffective. Several variables were considered in the model specifications as determinants of the frequency of home-based work during traditionally off-work hours. The results show that age, education level, number of hours of work, work schedule, number of jobs, whether self-employed or not, presence of young adults in the household, internet access at home, and presence of a house-helper all have statistically significant effects. Among these, education level and number of jobs have the most dominant impacts, while internet access from home has a relatively minor impact. None of the built environment variables turned out to impact working from home during off-work periods, though this result needs further exploration in future studies using a fine resolution of space to compute built environment measures.

The empirical results have implications for individual time-use and activity-travel patterns, especially in the context of how home-based work during the normal off-work periods may alter the generation and scheduling of daily household work and non-work patterns. In this regard, the model system developed in this paper can constitute one element of a larger-scale activity-based travel demand modeling system that uses demographic and work characteristics (at the out-of-home workplace) to predict complete daily individual activity-travel patterns. The results from the current analysis should also be useful for family and social science research, since research in those areas has suggested that working from home has impacts on individuals' satisfaction levels with family life, social interactions, and leisure participation. For instance, spending more time working at home can translate to less time with family, less time eating meals together with family, less time socializing with neighbors and friends, and also less time pursuing hobbies and other recreational pursuits.

Overall, examining work-related activities at home, and identifying those who pursue such work patterns, is important not only from the standpoint of predicting activity-travel patterns, but also because of its far-reaching impacts on the way we live and spend time. This is a relatively under-researched area in the literature, and it is hoped that this research will spur more focus on not only the "where" of work patterns, but also the "when" of work patterns. In this regard, while the Italian data set used in the current analysis is appropriate in many ways, it is also limited in the exogenous variables collected and available to predict the frequency of working from home during traditionally off-work times. Future efforts should consider a more comprehensive set of variables.

ACKNOWLEDGEMENTS
The authors acknowledge the helpful comments of three anonymous reviewers on an earlier version of the paper. The authors are grateful to the Turin Town Council (Office of Statistics) for providing the Time use Survey (2002-2003) data conducted in Turin. The corresponding author would like to acknowledge the support of an International Visiting Professorship from the University of Cagliari. SWUTC funding to UT research team is also gratefully acknowledged. Finally, the authors are grateful to Lisa Macias for her help in formatting this document.

REFERENCES

1. Sall, E. A., and C. R. Bhat. An analysis of weekend work activity patterns in the San Francisco Bay Area. *Transportation*, Vol. 34, 2007, pp. 161–175.
2. Valcour, P., and L. W. Hunter. Technology, Organizations, and Work-Life Integration. In E. E. Kossek and S. J. Lambert (eds), *Work and Life Integration: Organizational, Cultural, and Individual Perspectives*, Erlbaum Associates, Mahwah NJ, 2005, pp. 61–84.
3. Kennedy, T. L. M., A. Smith, A. T. Wells, and B. Wellman. Networked families. Pew Internet & American Life Project, 2008. http://www.pewinternet.org/.
4. WorldatWork. Telework Trendlines for 2006. February 2007. http://www.worldatwork.org/.
5. EWCS. Fourth European Working Conditions Survey. European Foundation for the Improvement of Living and Working Conditions, 2007. http://www.eurofound.europa.eu/.
6. Kaufman-Scarborough, C. Time Use and the Impact of Technology Examining Workspaces in the Home. *Time & Society*, Vol. 15, No. 1, 2006, pp. 57-80.
7. Golob, T.F., and A. C. Regan. Impacts of information technology on personal travel and commercial vehicle operations: research challenges and opportunities, *Transportation Research Part C*, Vol. 9, No. 2, 2001, pp. 87-121.
8. Bailey, D. E., and N. B. Kurland. A Review of Telework Research: Findings, new directions, and lessons for the study of modern work. *Journal of Organizational Behavior*, Vol. 23, No. 4, 2002, pp. 383-400.
9. Srinivasan, S., and C. R. Bhat. An Exploratory Analysis of Joint-Activity Participation Characteristics Using the American Time Use Survey. *Transportation*, Vol. 35, No. 3, 2008, pp. 301-328.
10. Mokhtarian, P.L., I. Salomon, and S. L. Handy. The Impacts of ICT on Leisure Activities and Travel: A Conceptual Exploration. *Transportation*, Vol. 33, No. 3, 2006, pp. 263-289.
11. Pinjari, A. R., and C. R. Bhat. A Multiple Discrete-Continuous Nested Extreme Value (MDCNEV) Model: Formulation and Application to Non-Worker Activity Time-Use and Timing Behavior on Weekdays. *Transportation Research Part B*, Vol. 44, No. 4, 2008, pp. 562-583.
12. Belloni, A. M. *Andare a tempo, Il caso di Torino: una ricerca sui tempi della città*, Franco Angeli (in Italian), 2007.
13. Jones, K., and C. Duncan. People and places: the multilevel model as a general framework for the quantitative analysis of geographical data. In *Spatial Analysis: Modelling in a GIS Environment*, P. Longley and M. Batty (eds.), GeoInformation International, Cambridge, 1996, 79-104.
14. Bhat, C. R. A Multi-Level Cross-Classified Model for Discrete Response Variables. *Transportation Research Part B*, Vol. 34, No. 7, 2000, pp. 567-582.
15. Reise, S. P., and N. Duan (eds). *Multilevel Modeling: Methodological Advances, Issues, and Applications*, Lawrence Erlbaum, Mahwah, NJ, 2003.
16. Bhat, C. R., and H. Zhao. The Spatial Analysis of Activity Stop Generation, *Transportation Research Part B*, Vol. 36, No. 6, 2002, pp. 557-575.
17. Varin, C. On composite marginal likelihoods. *Advances in Statistical Analysis*, Vol. 92, No. 1, 2008, pp. 1-18.
18. Bhat, C. R., I. N. Sener, and N. Eluru. A Flexible Spatially Dependent Discrete Choice Model: Formulation and Application to Teenagers' Weekday Recreational Activity Participation. *Transportation Research Part B*, forthcoming.

19. Dugundji, E., and J. Walker. Discrete Choice with Social and Spatial Network Interdependencies An Empirical Example Using Mixed Generalized Extreme Value Models with Field and Panel Effects. In *Transportation Research Record: Journal of the Transportation Research Board*, No. 1921, Transportation Research Board of the National Academies, Washington, D.C., 2005, pp. 70-78.
20. Bellio, R., and C. Varin. A Pairwise Likelihood Approach to Generalized Linear Models with Crossed Random Effects. *Statistical Modelling*, Vol. 5, No. 3, 2005, pp. 217-227.
21. Lindsay, B. G. Composite likelihood methods. *Contemporary Mathematics*, Vol. 80, 1988, pp. 221-239.
22. Cox, D., and N. Reid. A Note on Pseudolikelihood Constructed from Marginal Densities. *Biometrika*, Vol. 92, 2004, pp. 729-737.
23. Varin, C., and P. Vidoni. Pairwise Likelihood Inference for General State Space Models. *Econometric Reviews*, Vol. 28, No. 1, 2009, pp. 170-185.
24. Varin, C., and C. Czado. A Mixed Autoregressive Probit Model for Ordinal Longitudinal Data, 2009. Available at: http://www-m4.ma.tum.de/Papers/Czado/varin_czado_revision7.pdf
25. Apanasovich, T. V., D. Ruppert, J. R. Lupton, N. Popovic, N. D. Turner, R. S. Chapkin, and R. J. Carroll. Aberrant crypt foci and semiparametric modelling of correlated binary data. *Biometrics*, Vol. 64, No. 2, 2008, pp. 490-500.
26. Ferdous, N., N. Eluru, C. R. Bhat, and I. Meloni, A Multivariate Ordered Response Model System for Adults' Weekday Activity Episode Generation by Activity Purpose and Social Context. *Transportation Research Part B*, forthcoming.
27. Godambe, V. An Optimum Property of Regular Maximum Likelihood Equation. *Annals of Mathematical Statistics*, Vol. 31, 1960, pp. 1208-1211.
28. Heagerty, P. J., and T. Lumley. Window Subsampling of Estimating Functions with Application to Regression Models. *Journal of the American Statistical Association*, Vol. 95, 2000, pp. 197-211.
29. Istat. Multipurpose "*Time Use*" Survey, 2006, http://www.istat.it/.
30. Istat. Census of population, 2001, http://www.istat.it/.
31. Istat. Census of Industry and Services, 2001, http://www.istat.it/.
32. Agrawal, A., D. Kapur, and J. McHale. How do spatial and social proximity influence knowledge flows? Evidence from patent data. *Journal of Urban Economics*, Vol. 64, No. 2, 2008, pp. 258-269.
33. Axhausen, K. W. Social networks and travel: Some hypotheses. In *Social Dimension of Sustainable Transport: Transatlantic Perspectives*, K. P. Donaghy, S. Poppelreuter, and G. Rudinger (eds.), Ashgate, Aldershot, U.K., 2005, pp. 90-108.
34. Carrasco, J. A., B. Hogan, B. Wellman, and E. J. Miller. Collecting social network data to study social activity-travel behavior: an egocentric approach. *Environment and Planning B: Planning and Design*, Vol. 35, 2008, pp. 961-980.
35. Frei, A. and K. W. Axhausen. Size and structure of social network geographies. *Arbeitsberichte Verkehrs und Raumplanung*, 439, IVT, ETH Zürich, Zürich, 2007.
36. Chen, C., and P. Mokhtarian. Tradeoffs between Time Allocations to Maintenance Activities/Travel and Discretionary Activities/Travel. *Transportation*, Vol. 33, No. 3, 2006, pp. 223-240.

37. Bhat, C. R. A Multiple Discrete-Continuous Extreme Value Model: Formulation and Application to Discretionary Time-Use Decisions. *Transportation Research Part B*, Vol. 39, No. 8, 2005, pp. 679-707.
38. Meloni, I., M. Bez, and E. Spissu. Random Utility-Based Microeconomic Model for Discretionary Activity Time Allocation. In *Transportation Research Record: Journal of the Transportation Research Board*, No. 2054, Transportation Research Board of the National Academies, Washington, D.C., 2008, pp. 74-82.
39. Srinivasan, K. K., and S. R. Athuru. Analysis of Within-Household Effects and Between Household Differences in Maintenance Activity Allocation. *Transportation*, Vol. 32, No. 5, 2005, pp. 495-521.
40. Lewis, S. and C. L. Cooper. The Work-Family Research Agenda in Changing Contexts. *Journal of Occupational Health Psychology*, Vol. 4, 1999, pp. 382-93.

LIST OF TABLES

TABLE 1 CROR Estimation Results

	CROR	
	Estimates	Est./s.e.
Individual characteristics		
Age (Age > 45 is the base)		
Age 14 – 30	-0.410	-3.87
Age 31 – 45	-0.202	-2.52
Education (less than 2-3 years of high school is the base)		
High-level education (Bachelor degree or higher)	1.146	11.52
Medium-level education (between 4-5 years of high school and a Bachelor's degree)	0.673	8.46
Work characteristics		
Work-Hours per Week (40 hours and lower is the base)		
41 hours and more	0.314	4.64
Work schedule involves working in the:		
Evening	0.207	2.38
Night	-0.340	-2.60
Saturday	0.383	4.92
Having a second job	0.555	3.66
Occupation (Not Self-employed is the base)		
Self-employed	0.331	3.91
Household characteristics		
Presence of young adults 18-24 years (Presence of young children/no children is the base)	-0.230	-2.02
Availability of Internet access at home	0.147	2.21
Presence of a house-helper	0.392	3.90
Thresh01	1.290	9.02
Thresh02	1.712	11.74
Thresh03	2.454	16.31
Correlation terms		
Spatial correlation (σ)	0.114	2.71
Social correlation (η)	0.061	1.75
Mean Log-likelihood	-1321.650	

TABLE 2 Aggregate Elasticity Effects (% of variation)

	Propensity to work from home during off-work hours			
	Never	**Few times at month**	**Few times at week**	**Everyday**
Individual characteristics				
Age (Age > 45 is the base)				
Age 14 – 30	11.55	-32.41	-43.40	-58.73
Age 31 – 45	6.10	-15.90	-22.63	-34.04
Education (less than 2-3 years of high school is the base)				
High-level education (Bachelor degree or higher)	-44.18	82.03	160.00	321.52
Medium-level education (between 4-5 years of high school and a Bachelor's degree)	-20.83	52.09	74.57	126.38
Work characteristics				
Work-Hours per Week (40 hours and lower is the base)				
41 hours and more	-10.16	26.19	38.23	56.22
Work schedule involves working in the:				
Evening	-6.55	16.72	24.46	36.94
Night	9.38	-26.56	-34.96	-47.90
Saturday	-11.66	31.57	43.89	61.61
Having a second job	-19.68	42.11	71.46	131.58
Occupation (Not Self-employed is the base)				
Self-employed	-10.87	27.40	40.77	61.71
Household characteristics				
Presence of young adults 18-24 years (presence of younger children/no children is the base)	6.61	-18.35	-24.68	-34.36
Availability of Internet access at home	-4.48	11.97	16.76	24.20
Presence of a house-helper	-13.42	31.46	49.99	81.71