A Copula-Based Sample Selection Model of Telecommuting Choice and Frequency

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

and

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

*corresponding author

ABSTRACT

The objective of this study is to contribute to the telecommuting literature by jointly examining the propensity and frequency of workers to telecommute, using a rich set of individual demographics, work-related and industry characteristics, household demographics, and commute trip/work location characteristics. The data are drawn from the Chicago Regional Household Travel Inventory, collected between 2007 and 2008. From a methodological standpoint, the current study adopts a copula approach that allows the testing of several types of dependency structures between the telecommuting choice and frequency behavioural processes. To our knowledge, this is the first formulation and application in the econometric literature of a copula approach for the case of a binary self-selection mechanism with an ordered-response outcome.

The results clearly indicate that telecommuting choice and the frequency of telecommuting are governed by quite different underlying behavioral processes. For instance, women are less likely to telecommute relative to men, though there is no statistically significant difference in telecommuting frequency propensity between men and women telecommuters. Similarly, full-time employed individuals (≥ 30 hours per week) are more likely to have a telecommuting arrangement than those working part-time (< 30 hours per week). However, among those who telecommute, full-time employed individuals telecommute less frequently than part-time employed individuals. Further, the results suggest that the analyst risks the danger of incorrect conclusions regarding dependency in the telecommuting choice and frequency behavioral processes, as well as inconsistent and inefficient parameter estimates, by imposing incorrect dependency structures or assuming independence between the two behavioral processes. Overall, the empirical results indicate the important effects of several demographic and work-related variables on telecommuting choice and frequency, with implications for transportation planning and transportation policy analysis.

Keywords: Telecommuting choice, telecommuting frequency, copula approach, revealed preference analysis, sample selection models, ordered-response structure

1. INTRODUCTION

There is evidence of increasing telecommuting adoption over the past several years in the U.S. As indicated by Mokhtarian *et al.* (2005), quantifying the magnitude of telecommuting is a challenge, simply because telecommuting is a multifaceted and complex phenomenon. The situation can be exacerbated by the absence of a clear definition of telecommuting when statistics regarding telecommuting are presented. However, in a recent study conducted by World at Work (2009), the number of employee telecommuters (defined as regular full-time or part-time employees who do not work on a contract basis and are not self-employed, and "who work at home or at a remote location at least one day per month during normal business hours") was estimated to have climbed from 9.9 million in 2005 to 12.4 million in 2006 to 17.2 million in 2008.[1] However, even as the number of employee telecommuters appears to have increased, the World at Work study reports that the share of teleworkers working from home every day in the week has declined from 51% in 2006 to 40% in 2008.[2] Further, the share of teleworkers working from home once every week has also dropped, according to the report, from 77% to 72%, leading the report to conclude that "occasional telecommuting is on the rise". These differing and opposite trends in telecommuting adoption and the intensity of adoption (or telecommuting frequency), in conjunction with the potential benefits of telecommuting to the economy and the environment, has led to an increased interest in understanding the underlying processes determining telecommuting choice (or adoption) and telecommuting frequency. The current study contributes to such an understanding by modeling telecommuting choice and telecommuting frequency jointly. The sample used in the analysis is drawn from the 2008 Chicago Regional Household Travel Inventory (see CRHTI, 2009), and offers the opportunity to study telecommuting behavior using a very recent revealed preference survey.

The rest of the paper is structured as follows. Section 2 presents a brief overview of the earlier literature on telecommuting and positions the current study within this broader context.

---

[1] It is unclear in the study by World at Work (2009) if a "regular full-time or part-time employee" is necessarily one who has a regular workplace location outside home. While this appears to be the intent of the definition of "regular employee", this is unclear.

[2] In computing this share, the study considers both employee telecommuters (as defined earlier) as well as contract telecommuters ("individuals who work on a contract basis for an employer or are self-employed, and who work at home or at a remote location at least one day per month during normal business hours"). The share is not provided separately for employee telecommuters. Hence, we use the label "teleworkers working from home" when presenting the share statistics. However, given that there is likely to be less fluctuation over time in the number of days of teleworking from home among those who are self-employed or home-based workers, it is not unreasonable to assume that the decrease in share applies to employee telecommuters.

Section 3 describes the data collection procedures as well as the sample used in the analysis. Section 4 outlines the modeling methodology employed for the empirical analysis of the current study. Section 5 presents the empirical results. Finally, Section 6 summarizes important findings from the study and concludes the paper.

2. OVERVIEW OF EARLIER STUDIES AND CURRENT PAPER

In this section, we provide an overview of earlier telecommuting studies to demonstrate the level of interest in the topic and the types of analyses that have been conducted. The intent of the discussion is not to provide an extensive review of the literature, but rather to present important trends in the study of telecommuting (see Tang *et al.*, 2008 and Walls and Safirova, 2004 for detailed reviews on the subject).

The quantitative studies of telecommuting may be broadly classified into two categories: (1) Quantitative studies using stated-preference survey data, and (2) Quantitative studies using revealed-preference survey data. The first category of studies, which also, in general, appeared earlier in time than the second category of studies, was based on stated preference surveys ostensibly because the penetration rate of telecommuting in the worker population until the mid-1990s was not adequate to support quantitative modeling using revealed preference data (Mannering and Mokhtarian, 1995). For instance, Bernardino *et al.* (1993) and Yen and Mahmassani (1994) used ordered response frameworks to model the stated telecommuting willingness of individuals, while Sullivan *et al.* (1993) estimated a multinomial logit model (rather than an ordered-response model) to analyze stated telecommuting choice and participation frequency. The above studies, while providing useful insights regarding the stated preferences of individuals to adopt telecommuting, do not adequately examine the actual individual choices/constraints that influence telecommuting adoption and frequency. As a result, they are likely to be of limited value for informing the development of policy strategies (Mokhtarian and Salomon, 1996a).

The earliest published research effort in the second group of revealed preference studies appears to be the one by Olszewski and Mokhtarian (1994). The emphasis of this study was solely on the telecommuting frequency dimension among telecommuters; the study did not examine the choice to telecommute. The results from the study indicated statistically insignificant effects of age, gender, number of children in the household, and commute distance on telecommuting frequency, though some of these results may simply be an artifact of the limited sample size in the analysis.

Subsequent to the Olszewski and Mokhtarian study, Mannering and Mokhtarian (1995) employed a multinomial logit model based on revealed preference data with three possible alternatives: "never telecommute", "infrequently telecommute", and "frequently telecommute". However, the study was limited by the small percentage of (frequent) telecommuters within the survey sample. Several other revealed preference studies have also focused on the choice of telecommuting, occasionally with some representation of frequency in the broad manner of Mannering and Mokhtarian (1995) (see, for instance, Bernardino and Ben-Akiva, 1996, Mokhtarian and Salomon, 1996b, and Mokhtarian and Salomon, 1997). Another revealed preference study with a more national focus (rather than the regional focus of the studies just mentioned) is the one by Drucker and Khattak (2000).

Unlike the revealed preference studies discussed above that were undertaken in the 1990s, the past few years have seen more research with revealed preference data focusing on both the telecommuting choice as well as a measure of frequency that includes a time frame of reference (such as once a month, once a week, 2-3 times a week, and 4-5 times a week) as opposed to previous broad characterizations as "infrequently" or "frequently" telecommute. Some of these studies also explicitly recognize that the telecommuting choice decision (*i.e.*, whether to telecommute at all or not) and the frequency of telecommuting may be governed by quite different underlying behavioral processes rather than being governed by a single behavioral process. For instance, Popuri and Bhat (2003) were the first to jointly model the distinct choice and frequency decisions, and showed that failure to accommodate this correlation can lead to inconsistent parameter estimates. However, their data set does not have job-related characteristics (such as industry and occupation categories) that may significantly influence telecommuting. In this regard, Walls *et al.* (2007) examined both the choice and frequency decisions of telecommuting using an extensive set of job-related factors and found substantial influences of these work-related factors. In their study, Walls *et al.* considered the correlation in unobserved factors in the choice and frequency decisions by including a Heckman's (1979) correction term in the frequency model after estimating the telecommuting binary choice model parameter estimates. They found this correction term to be statistically insignificant, and so estimated independent models of choice and frequency. However, the textbook Heckman's correction term is valid only for a continuous outcome equation, and not for the ordered response outcome of frequency that Walls *et al.* (2007) employ. The appropriate procedure for the normally distributed underlying processes of choice and frequency that Walls *et al.* assume would be the joint estimation technique of Popuri and Bhat (2003). Finally, Tang *et al.* (2008) examined the effect of

objective residential neighborhood built environment factors, as well as subjective perceptions of these factors, on both the adoption and frequency of telecommuting, using a single multinomial logit model (MNL). One limitation of their study is that they considered very few individual/household demographic variables, and no work-related variables (other than commute time).

Overall, the above discussion illustrates the substantial recent interest in jointly analyzing the choice and frequency of telecommuting. The objective of this study is to contribute to this telecommuting literature in several important ways. First, the sample used in this study includes the revealed preference survey responses of 9624 workers from the Chicago region. The sample comprises 1534 telecommuters, which constitutes the largest number of telecommuters in any study so far that we are aware of. The large sample of telecommuters should aid in comprehensively and rigorously "teasing out" the factors that influence the telecommuting adoption and frequency decisions. In fact, the richness of the data allows us to incorporate a variety of variables, including individual demographics, work-related and industry characteristics, household demographics, and commute trip/work location characteristics. Second, the data sample is obtained from the recently completed 2008 Chicago Regional Household Travel Inventory (CRHTI), thus providing us with the ability to develop a very current perspective of the process driving telecommuting decisions (at least in the Chicago region). In contrast, even the recent studies by Walls *et al.* (2007) and Tang *et al.* (2008) have used relatively dated data from 2002 and 2003, respectively. Third, the survey reduces the ambiguity in the difference between home-based telecommuting and operation of a home-based business by removing individuals who indicated that they were self-employed and worked primarily from home. Thus, the sample of workers considered in the current analysis includes only those who stated expressly that their primary/main work location is a location outside home that they travel to

routinely.[3] Finally, from a methodological perspective, we jointly model the choice and frequency of telecommuting rather than independently modeling the two decisions. The failure to capture the jointness among these two inter-related choices can lead to inconsistent parameter estimates and misinformed policy actions, as discussed in Popuri and Bhat (2003). However, we go one step beyond the methodological approach of Popuri and Bhat by using a flexible copula-based approach to characterize the dependency between the error terms in the telecommuting choice and frequency equations. The copula approach allows the testing of several types of dependence structures rather than pre-imposing the very restrictive bivariate normal distribution assumption of Popuri and Bhat.

3. DATA AND SAMPLE DESCRIPTION

3.1. Data Sources

The data used in this study are drawn from the 2008 Chicago Regional Household Travel Inventory (CRHTI), which was sponsored by the Chicago Metropolitan Agency for Planning (CMAP), the Illinois Department of Transportation (IDOT), the Northwestern Indiana Regional Planning Commission, and the Indiana Department of Transportation. The survey was administered in both English and Spanish using standard postal mail-based survey methods and computer-aided telephone

---

[3] The Chicago survey asks the following question if and only if the respondent stated expressly that her/his primary/main work location is a location outside home: "*Does your employer allow you to work from home for pay on a regular basis? This would be in place of driving to a regular work location, something that is commonly referred to as -telework.-*" We will assume here that respondents mentally replaced the word "driving" with "traveling" in the question above, so that an individual whose employer allows working from home for pay and who travels by bicycle or transit or walk would also have answered positively to the question. If an individual answered affirmatively to the above question, the person was asked the following question: "*About how often do you work at home instead of traveling to your usual workplace?*". This telecommuting frequency question was not asked if a respondent answered negatively to the first question. Interestingly, 97% of those who answered positively to the first question indicated that they telecommuted at least once in the year (and the 3% of those who answered positively to the first question, but returned a "don't know or "refused" response to the frequency question, had to be removed from the analysis anyway because several demographic and work-related variables were missing for these individuals). Thus, effectively speaking, those who worked from home at least once a year based on the telecommuting frequency question were characterized as telecommuters, while others were considered as non-telecommuters. Of course, an issue with this classification is that there may be those occasional telecommuters even in the pool of individuals whose employers do not allow telecommuting on a regular basis. Since such individuals were not presented with the telecommuting frequency question, they necessarily are classified as non-telecommuters. However, it is also very likely that the frequency levels of telecommuting for such individuals will be rather low (such as that informal arrangement to work from home on the day that a child is sick). For frequency levels of once a week or even once a month, a person would most likely have to be with an employer that allows regular telecommuting. Thus, the potential misclassification in our approach may not be substantial. In fact, the misclassification here may not be any more than the misclassification in some earlier studies that consider those who do not telecommute more frequently than once a month as non-telecommuters. Of course, all this is somewhat speculative, and does not detract from the fact that it would have been nice to present the frequency question to every respondent (regardless of whether or not the employer allowed telecommuting) as well as to have a response category of "never" for the frequency question.

interview (CATI) technology through Travel Tracker Survey to facilitate the organization and storage of the data. A dual sampling frame approach was used, with one sampling frame being the list of land-line telephone numbers in the study area and the other being an address-based frame of all residential addresses that receive U.S. postal mail. Further details of the survey design and implementation methods are available in NuStats (2008).

The survey was conducted expressly to inform the development of regional travel demand models for the Chicago region. It involved the collection of activity and travel information for all household members (regardless of age) during a randomly assigned 1-day or 2-day period (the 1-day period sample focused only on weekdays, while the 2-day period sample targeted two consecutive days including the Sunday/Monday and Friday/Saturday pairs but not the Saturday/Sunday pair). The final sample included information from 14,315 households.

3.2. Sample Formation and Description

The data assembly process involved several steps. First, since the focus of the study is on telecommuting, only employed individuals whose primary/main work location is a location outside home were selected from the overall sample. Second, two specific dimensions of each employed individual's work pattern were considered for the current analysis: (1) Telecommuting choice (whether or not person telecommutes – see footnote 3), and (2) Telecommuting frequency (obtained in one of the five categories of "once a year", "a few times a year", "once a month or more", "once a week or more", and "almost every day"). In the current analysis, we use a binary model for the telecommuting choice component and a five-point ordered-response model for the telecommuting frequency component. Finally, several screening and consistency checks were undertaken to obtain the final sample of 9624 employees.

The data sample for analysis includes 1534 telecommuters (15.9% of the overall sample). This telecommuting percentage is similar to that found in Popuri and Bhat (2003) in the New York City area, though it is lesser than the 25% or so telecommuting percentages reported in Walls *et al.* (2007) and Tang *et al.* (2008). This lower percentage in our study is potentially because we are better able to distinguish between telecommuters and home-based business (HBB) workers (*i.e.*, those who work out of home). Tang *et al.* acknowledge that their characterization of telecommuters is likely to be a mix of actual telecommuters and HBB workers. In terms of telecommuting frequency, the split in the sample of telecommuters is as follows: 36 (2.4%) telecommute once a

year, 194 (12.6%) telecommute a few times a year, 461 (30.1%) telecommute once or more per month, 649 (42.3%) telecommute once or more per week, and 194 (12.6%) telecommute almost every day. As expected, most of those who telecommute do so at least once a month.[4]

4. METHODOLOGY

4.1. Model Structure

In our empirical analysis, there are two dependent variables - telecommuting choice, modeled using a binary choice structure, and telecommuting frequency, modeled using an ordered-response structure. These two dependent variables are jointly analyzed using a copula approach that enables flexible dependency in the latent propensities underlying the choice and frequency dimensions. Mathematically, the model system is as follows:

$$t_q^* = \beta' x_q + v_q, \ t_q = 1 \text{ if } t_q^* > 0 \text{ and } t_q = 0 \text{ if } t_q^* < 0$$

$$s_q^* = \gamma' z_q + \eta_q, \ s_q = k \text{ if } \delta_{k-1} < s_q^* < \delta_k, \ k = 1, 2, ..., K, \ s_q \text{ observed only if } t_q^* > 0, \tag{1}$$

---

[4] Note that, in choice modeling, the exogenous sample maximum likelihood (ESML) procedure (*i.e.*, the usual maximum likelihood procedure based on a strictly random sample) is entirely appropriate to other samples as long as the dependent variable proportions in the sample match up to the corresponding population proportions. Whether the sample is also representative of the population on the exogenous variables or not is irrelevant. The reader is referred to Manski and Lerman (1977) and Cosslett (1981) for further details. In the context of the current paper, whether or not the relationship extracted from the sample is representative of the population relationship hinges on how closely the telecommuting choice/frequency distribution in the sample is representative of the telecommuting choice/frequency distribution in the population. Further, because of the binary nature of the choice variable, the parameters on all exogenous variables (except the constant) in the binary telecommuting choice equation are consistent, even if the sample share of telecommuters does not correspond to the population share of telecommuters. Of course, to obtain consistent estimates of the parameters in the telecommuting frequency equation, as well as the dependency in unobserved factors between the choice and frequency equations, one needs to have reasonably representative shares in each telecommuting frequency category. In this study, we examined both the unweighted telecommuting frequency shares from the data sample of 9624 employees and the weighted shares based on weighting the original Chicago data sample at the household level using the 2005-2007 American Community Survey (ACS) data controls (see Frank, 2009 for details). These latter weighted shares of telecommuting frequency (among those who telecommute) came out to be as follows: 10.4% telecommute once a year, 8.6% telecommute a few times a year, 30.8% telecommute once or more per month, 23.9% telecommute once or more per week, and 26.3% telecommute almost every day. After examining the unweighted and weighted shares, we decided to use the unweighted shares because they seemed more likely to be representative of population frequency shares, especially the split between telecommuting once or more per week and telecommuting almost every day. In particular, the unweighted shares are closer to the frequency shares from the 2005 U.S. Census Bureau Survey of Income and Program Participation (SIPP) data (http://www.census.gov/population/www/socdemo/workathome/2005%20Table %208.xls, accessed March 7, 2010), which indicates that a majority of individuals who work at home at least once a week (but also work outside home during the week) are likely to work from home 1-2 days in the week rather than almost every day in the week. The unweighted shares are also more consistent with an overall telecommuting frequency of about 1.2 days per week among telecommuters, obtained in Tang *et al.* (2008) and several other studies.

where q is an index for individuals, k is an index for frequency level, t_q is an observed binary variable indicating whether or not person q chooses to telecommute ($t_q = 1$ if person q telecommutes, 0 otherwise), t_q^* is an underlying continuous variable related to the observed binary variable t_q as shown above, s_q is an observed ordinal variable representing the frequency of telecommuting if individual q telecommutes, s_q^* is a latent continuous variable representing the propensity underlying the telecommuting frequency decision, the δ_k terms represent thresholds that relate s_q^* to the observed variable s_q in the usual ordered-response structure ($\delta_0 = -\infty, \delta_K = \infty; -\infty < \delta_1 < \delta_2 < \ldots < \delta_{K-1} < \infty$), x_q and z_q are vectors of explanatory variables (as written in Equation (1), x_q includes a constant, but z_q does not), β and γ are vectors of parameters to be estimated, and v_q and η_q are random error terms, which may take any parametric distribution. In the current study, we examine both logistic and normal marginal distributions for these error terms, and choose the distribution that provides the best data fit. The error terms v_q are assumed to be independent and identically distributed (IID) across individuals q, and the error terms η_q are also assumed to be IID across individuals q. Further, for the logistic case, a standard logistic distribution is used for the error terms, while, for the normal case, a standard normal distribution is used for the error terms (these standardizations are innocuous normalizations needed for econometric identification). For presentation ease, let the marginal distribution of v_q be $F(.)$ and the marginal distribution of η_q be $G(.)$.[5]

   With the notational preliminaries above, the probability that individual q does not telecommute is given by:

$$\Pr[t_q = 0] = \Pr[v_q < -\beta' x_q] = F(-\beta' x_q). \tag{2}$$

The probability that the individual q telecommutes and does so at a frequency level k ($k = 1, 2, \ldots, K$) can be written from Equation (1) as:

---

[5] Thus, in the context of the current analysis, $F(.)$ may be the standard logistic cumulative distribution function or the standard normal distribution function. The same is the case with $G(.)$. Note that, in the copula approach we use, it is not necessary that both $F(.)$ and $G(.)$ should be simultaneously logistic (logistic-logistic) or simultaneously normal (normal-normal). Rather, we can also test the normal-logistic and logistic-normal pairings.

$$\Pr[t_q = 1, s_q = k] = \Pr[v_q > -\beta'x_q, \delta_{k-1} - \gamma'z_q < \eta_q < \delta_k - \gamma'z_q]$$

$$= \Pr[v_q > -\beta'x_q, \eta_q < \delta_k - \gamma'z_q] - \Pr[v_q > -\beta'x_q, \eta_q < \delta_{k-1} - \gamma'z_q]$$

$$= \Pr[\eta_q < \delta_k - \gamma'z_q] - \Pr[v_q < -\beta'x_q, \eta_q < \delta_k - \gamma'z_q] - \left(\Pr[\eta_q < \delta_{k-1} - \gamma'z_q] - \Pr[v_q < -\beta'x_q, \eta_q < \delta_{k-1} - \gamma'z_q]\right)$$

$$= G\left(\delta_k - \gamma'z_q\right) - \Pr[v_q < -\beta'x_q, \eta_q < \delta_k - \gamma'z_q] - \left(G\left(\delta_{k-1} - \gamma'z_q\right) - \Pr[v_q < -\beta'x_q, \eta_q < \delta_{k-1} - \gamma'z_q]\right)$$

$$(3)$$

The above joint probability depends upon the dependence structure between the random variables v_q and η_q. As highlighted before, the incorporation of the dependency effects can be greatly facilitated by using a copula approach for modeling joint distributions. The copula approach does not need the *a priori* specification of the functional form of the dependence surface. Indeed, we can test different functional forms, and select the one that empirically fits the data best rather than pre-imposing the very restrictive, but commonly used, bivariate normal distribution assumption. To our knowledge, we are the first to formulate and estimate a copula-based model for the case of a binary self-selection model with an ordinal outcome equation.

In the specific context of the current study, a joint bivariate distribution function of the random variables v_q [with the marginal distribution $F(.)$] and η_q [with the marginal distribution $G(.)$] may be generated as follows (see Sklar, 1973):

$$J(v,\eta) = \Pr(v_q < v, \eta_q < \eta) = \Pr[U_1 < F(v), U_2 < G(\eta)] = C_\theta[u_1 = F(v), u_2 = G(\eta)], \tag{4}$$

where C_θ is a copula function and θ is a dependency parameter (assumed to be scalar), together characterizing the dependency between v_q and η_q. A rich set of bivariate copulas $C_\theta(u_1, u_2)$ are available to generate the dependence between the random variables v_q and η_q, including the Gaussian copula (*i.e.* the bivariate normal dependency structure), the Farlie-Gumbel-Morgenstern (FGM) copula, and the Archimedean class of copulas (including the Clayton, Gumbel, Frank, and Joe copulas). For given functional forms of the margins, the precise bivariate dependence profile between the variables v_q and η_q is a function of the copula $C_\theta(u_1, u_2)$ used, and the dependence parameter θ (see Bhat and Eluru, 2009 and Bhat and Sener, 2009 for discussions of copula-based

approaches). But, regardless of the margins, the overall nature of the dependence between v_q and η_q is determined by the copula function.

4.2. Model Estimation

The parameters to be estimated in the joint binary choice-ordered response model (that is, telecommuting choice-telecommuting frequency models) include the β vector, the $(K-1)$ δ_k parameters $(\delta_0 = -\infty, \delta_K = \infty; -\infty < \delta_1 < \delta_2 < \ldots < \delta_{K-1} < \infty)$, the vector γ, and the dependency parameter θ.

The probability that an individual q telecommutes and does so at a frequency level k ($k = 1$, 2, …, K) can be obtained from Equation (3) as follows:

$$\Pr[t_q = 1, s_q = k] = G(\delta_k - \gamma' z_q) - G(\delta_{k-1} - \gamma' z_q) - \left[C_\theta(u_{q1}, u_{q,k,2}) - C_\theta(u_{q1}, u_{q,k-1,2}) \right], \tag{5}$$

where $u_{q1} = F(-\beta' x_q)$, $u_{q,k,2} = G(\delta_k - \gamma' z_q)$, and $u_{q,k-1,2} = G(\delta_{k-1} - \gamma' z_q)$

Next, let $I[.]$ be an indicator function taking the value of unity if the expression in parenthesis is true and 0 otherwise. Also, define a set of dummy variables M_{qk} as below:

$$M_{qk} = I[t_q = 1] \times I[s_q = k]. \tag{6}$$

Then, the log likelihood function for the copula model takes the following form:

$$\log L = \sum_{q=1}^{Q} \left(I[t_q = 0] \times \log[\Pr(t_q = 0)] + \sum_{k=1}^{K} M_{qk} \log[\Pr(t_q = 1, s_q = k)] \right). \tag{7}$$

All the parameters in the model are estimated by maximizing the log-likelihood function in Equation (7). The model estimation was pursued using the GAUSS matrix programming language. The asymptotic covariance matrix of the estimator is obtained in the usual way as the inverse of the information matrix, which is itself estimated numerically as the negative of the second derivatives matrix of the log-likelihood function at the convergent values.

5. MODEL RESULTS

5.1. Variable Specification

Several variable specifications and functional forms were considered in the model. The final model specification was based on intuitive considerations, insights from previous literature, parsimony in specification, and statistical fit/significance considerations. The final specification includes some variables that are not highly statistically significant, but which are included because of their intuitive effects and potential to guide future research and survey efforts in the field.

5.2. Model Specification and Data Fit

The empirical analysis involved estimating models with two different univariate (*i.e.*, marginal) distribution assumptions (normal and logistic) for the error terms v_q and η_q, and seven different copula structures (independence, Gaussian, FGM, Clayton, Gumbel, Frank, and Joe).[6] As discussed in Section 4, in the copula approach, there is no need to assume that the marginal distributions of the v_q and η_q error terms are simultaneously normal (normal-normal) or logistic (logistic-logistic); instead v_q and η_q terms can have a normal-logistic or logistic-normal distribution. We examined all these four possible combinations for the error terms v_q and η_q, as well as the seven different copula structures. The result is 24 copula models with dependency, and four independence copula models.

The Bayesian Information Criterion (BIC) is employed to select the best copula model among the 24 competing non-nested copula models with dependency (see Quinn, 2007, Genius and Strazzera, 2008, Trivedi and Zimmer, 2007, page 65). The copula model that results in the lowest BIC value is the preferred one. But, if all the competing models have the same exogenous variables and a single copula dependence parameter θ, as in the current empirical case, the BIC information selection procedure measure is equivalent to selection based on the largest value of the log-likelihood function at convergence. Based on the BIC, the Normal-Normal Frank (NNF) model

---

[6] Due to space considerations, we are unable to provide additional details on the structures of different copula types. Interested readers are referred to Bhat and Eluru (2009). Also, note that the independence copula corresponds to $C_\theta(u_1, u_2) = u_1 u_2$.

provided the best data fit, with a corresponding Kendall's measure of dependency of +0.202.[7] The positive dependency measure obtained in the Frank copula relating the v_q and η_q terms is intuitive, indicating that unobserved factors (such as feeling more productive working from home or preferring to work without others around) that increase an employee's propensity to telecommute also increase the employee's inclination to telecommute frequently. Similarly, unobserved factors (such as a social preference to work with other co-workers; for example, see Páez and Scott, 2007) that decrease an employee's propensity to telecommute may also decrease the employee's inclination to telecommute frequently.[8]

In terms of the copula dependency surface, the Frank copula, like the Gaussian copula, allows for positive and negative dependence, is comprehensive in its coverage, is radially symmetric in its dependence structure, and imposes the assumption of asymptotic independence. However, as shown visually in Bhat and Eluru (2009), the dependence surface of Frank's copula shows very strong central dependency (stronger than the Gaussian copula) and very weak tail dependence (weaker than the Gaussian copula). In the current empirical context, this means that, due to unobserved factors, individuals are likely to be substantially clustered around the medium-medium levels of the two-dimensional (latent) telecommuting propensity-frequency inclination spectrum, and less so at the low-low end or the high-high end of the spectrum. The central clustering tendency and the lower dependency at the extremes are much more pronounced than that implied by the Gaussian copula.

---

[7] The Frank copula function (Frank, 1979) is given by: $C_\theta(u_1, u_2) = -\frac{1}{\theta} \ln\left(1 + \frac{(e^{-\theta u_1} - 1)(e^{-\theta u_2} - 1)}{e^{-\theta} - 1}\right)$, $-\infty < \theta < \infty$.

The Kendall's measure of dependency (τ) transforms the copula dependency parameter (θ) into a number between -1 and 1 (see Bhat and Eluru, 2009). For the Frank copula, $\tau = 1 - \frac{4}{\theta}\left[1 - \frac{1}{\theta}\int_{t=0}^{\theta} \frac{t}{e^t - 1} dt\right]$. Independence is attained in Frank's copula as $\theta \to 0$.

[8] A data-based limitation of the current study is that it does not consider the potential jointnesss in decision-making among individuals in their choice of telecommuting (or not telecommuting). Accommodating such social interaction effects in telecommuting decisions is an important avenue for further research (see Páez et al., 2008 for some initial work along these lines).

The likelihood value at convergence of the Normal-Normal Frank (NNF) copula is -5123.21.[9] Among the four independence copula models, the Normal-Normal Independence (NNI) model provided the best data fit, with a likelihood value of -5125.73. Since both the NNF and the NNI models have the same margins for both v_q and η_q, they can be compared using a likelihood ratio test (the NNI model, which is equivalent to independent models of telecommuting choice and frequency, is obtained by restricting the dependence parameter in the NNF model to zero). The chi-squared test statistic is 5.04, strongly rejecting the null hypothesis of independence between the telecommuting choice and frequency equations at the 0.025 level of significance for one degree of freedom. Interestingly, the log-likelihood value at convergence for the classic textbook structure (see Lee, 1983) that assumes a normal-normal Gaussian (NNG) model structure is -5124.89, with a corresponding Kendall's measure of dependency of 0.098.[10] The likelihood ratio statistic for the test between the NNG and NNI models is only 1.68. Thus, one is unable to reject the null hypothesis of independence between telecommuting choice and frequency at the usual levels of significance used in hypothesis testing. The implication is clear. One can get inappropriate results regarding the dependency between two random variables just because of the imposition of a specific parametric form for the dependency. In the current empirical context, using the typical bivariate normal distributional assumption between the telecommuting choice and frequency equations provides the incorrect result that there is no statistically significant dependency. Intuitively, this is because a reasonably significant positive dependency in the NNG model implies a dependency level at the edges that is higher than that reflected in the actual data. To compensate for this, the NNG model estimates a low and statistically insignificant dependency level. On the other hand, the NNF model better replicates the positive dependency surface relationship between the latent telecommuting propensity and frequency inclination variables, both at the center as well as at the edges.

---

[9] Note that the log-likelihood value at sample shares for the no-telecommuting and the five categories of telecommuting is -6271.56 (this is equivalent to the log-likelihood for the model that has only a constant in the telecommuting choice equation, only the thresholds in the telecommuting frequency equation, and assumes independence between the choice and frequency equations). The likelihood ratio test for the NNF model comparison with this sample shares model is 2296.70, which is larger than the corresponding chi-squared table value with 46 degrees of freedom (corresponding to the number of non-constant and non-threshold parameters in the final model) at any reasonable level of statistical significance. This clearly indicates the value of the model developed here.

[10] For the Gaussian copula, $\tau = \dfrac{2}{\pi}\arcsin(\theta)$. Independence is attained in Gaussian's copula when $\theta = 0$.

5.3. Estimation Results

To conserve on space, we only present the results for the best NNF model.[11] The results are presented in Table 1, where the first main number column is used to present descriptive aggregate statistics of each variable in the group of non-telecommuters and in the group of telecommuters. Thus, the entry "4080 (50.4)" for the "Female" variable in the non-telecommuter column indicates that 4080 of the 8090 non-telecommuters are female, which corresponds to 50.4% of the non-telecommuter sample. The estimated coefficients from the model are presented in the second main number column of Table 1, and will be the focus of discussion in this and the next few sections. The elasticity effects of variables are presented in the final column of the table, and are discussed in Section 5.4.

The highly significant negative constant in the binary telecommuting choice model is, in part, a reflection of the large share of non-telecommuters in the sample, although it also serves as an overall adjustor term to fit the data best given the exogenous variables (if the exogenous variables were all purely dummy variables, the constant may be viewed as the sample share indicator in the base segment formed by the combination of the base context for each dummy variable; however, this interpretation does not hold in the presence of ordinal and continuous exogenous variables). The thresholds at the top of Table 1 for the ordered-response frequency model do not have any substantive interpretations. They simply serve the purpose of mapping the latent propensity into the observed frequency levels. Unlike the binary telecommuting choice model, we did not include a separate constant term in the ordered-response telecommuting frequency model because all four threshold parameters plus a constant cannot be separately identified (*i.e.*, one of them is redundant). Also note that, for dummy exogenous variables, the category that does not appear in the table is the base category. This base category is explicitly identified in the text discussion below.

5.3.1. Individual Demographics

The first set of exogenous variables corresponds to individual demographics. The effect of the "female" variable indicates that women have a lower propensity to telecommute compared to men. We also examined gender interaction effects with the presence of children, but found that the gender

---

[11] The estimates from the other copula models and the independent model were, as one would expect, different from those obtained from the NNF model. Further, the standard errors of the telecommuting frequency model estimates were, in general, smaller than those from the other models, indicating efficiency benefits as well from using the NNF structure.

difference was not affected by the presence of children in the household (however, there was a "presence of children" effect independent of gender, which is discussed under household demographics). The age-related effects suggest a lower propensity among young adults less than 30 years of age (relative to their older peers) to telecommute, and this age-based difference is particularly strong for women relative to men. Men over 30 years of age are the most likely to telecommute relative to their younger counterparts and their (younger and older) female peers, while women under 30 years of age are the least likely to telecommute. Also, in the group of telecommuters, young women have a lower propensity to telecommute frequently relative to their older counterparts and their (younger and older) male peers (see the negative sign on the "female*age less than 30 years of age" variable in the telecommuting frequency column). These gender/age effects are consistent with the findings in the literature (see, for instance, Mannering and Mokhtarian, 1995, Drucker and Khattak, 2000, Popuri and Bhat, 2003, Mokhtarian and Meenakshisundaram, 2002, and Walls *et al.,* 2007). As suggested also by Tang *et al.* (2008), the lower telecommuting propensity of women relative to men may be because men (and men over 30 years of age in particular) occupy jobs with "more autonomy and bargaining power", as well as jobs that need telecommunications expertise. The general finding that individuals above 30 years of age are more likely to telecommute than those younger than 30 years may be attributable to older, experienced, employees being more able to exercise personal choices regarding work arrangements. This result may also be a reflection of the higher social value that younger employees place (relative to their older peers) on being around others at the work place. Education is clearly a very important factor that positively influences the choice of telecommuting and the frequency of telecommuting, another recurring finding in the literature (the base category for the education variables in the table corresponds to an education level below a bachelor's degree). Finally, the results show a positive propensity to telecommute among employees with a driver's license. This is a result also obtained in Drucker and Khattak (2000), but needs further exploration to analyze the underlying reasons.

5.3.2. Work-Related and Industry Characteristics

Full-time employed individuals ($\geq$ 30 hours per week) are more likely to have a telecommuting arrangement than those working part-time (< 30 hours per week). It may be argued that employers are in general less willing to allow part-time employees to telecommute (because these individuals are already showing up to work only partly in the week). However, among those who telecommute,

the results in the table reveal that full-time employed individuals telecommute less frequently than part-time employed individuals. It is possible that full-time employed individuals have more obligations to be at work frequently (see also Tang *et al.*, 2008 and Yeraguntla and Bhat, 2005).

Individuals with flexible work schedules are more likely (than individuals with no work schedule flexibility) to telecommute and telecommute frequently. One would anticipate that individuals who want work flexibility will look for jobs that provide them both temporal flexibility (as captured in the work schedule flexibility variables) as well as spatial flexibility (*i.e.,* telecommuting options). Thus, the positive association between work schedule flexibility and telecommuting propensity/frequency is to be expected. The propensity of telecommuting also increases with an increase in the number of jobs, presumably a reflection of trying to manage time more efficiently by working at home and saving work-related travel time to multiple work locations.

An important empirical contribution of the current study is the variety of industry types incorporated in the models. The base for introducing the industry dummy variables in our specification includes manufacturing, transportation, retail and other industries (for ease, we will refer to the base category as MATRE). We chose these categories as the base since it is quite likely that individuals in these industries will need to travel every day to their work location.[12] The results indicate statistically significant differences among individuals in different industries in their telecommuting propensity and frequency. Workers in the communications industry are more likely to telecommute and to do so frequently relative to those in the MATRE category. Further, employees in service-related industries, in general, also have a higher telecommuting propensity than the MATRE industry category, sometimes also reinforced by higher telecommuting frequency. The only exceptions are for employees in educational services, and health care or social assistance. This is indeed quite expected, since the jobs of workers in these latter two service professions naturally require face-to-face interactions with students and those who need health care/social assistance, respectively. But among those who are able to telecommute in these two professions, the frequency of telecommuting is higher than in the MATRE industries. Finally, individuals working for the government are the least likely to telecommute. Individuals working for the government may need to

---

[12] While it is possible that individuals in the same occupation may have the same order of telecommuting propensities even if in different industries, we were unable to include occupation information because this information was collected in the survey in an open-ended text form. In contrast, the industry information was collected in closed categories and was more useable. However, these industry effects should be viewed with some caution, and only as broad characterizations of the type and mix of jobs within any given industry.

be involved in quick coordination/organization responses in uncertain conditions, and are likely to participate in interactive knowledge and information based activities. Such work-related characteristics and activities are likely to be facilitated by face-to-face contact and interactions with colleagues and others (see Storper and Venables, 2004). Further, it is possible that government employees may not be able to work from home because of the need to work with sensitive information that can be accessed only in their secure work location environment, or because their immediate management team subtly (and not so subtly) discourages telecommuting to prevent the risk of the department being perceived in a stereotypic manner as a malingerer bureaucracy, or because government agencies tend to have more traditional managerial styles that discourage telecommuting.

5.3.3. Household Demographics

The results of household demographics show that individuals in households with children (15 years or younger) are more likely to telecommute than those in households with no children, presumably because of child-care and child-escort responsibilities. Also, individuals in households with more workers have a higher preference to adopt telecommuting and to telecommute frequently relative to households with fewer workers. As household income increases, individuals are significantly more likely to telecommute, a finding that is consistent with many previous studies (see Mannering and Mokhtarian, 1995, Bernardino and Ben-Akiva, 1996, and Popuri and Bhat, 2003). This may be attributed to more control over work location-related and work timing choices as one "climbs the work ladder". The effect of the "number of household vehicles" variable is interesting, and suggests a lower telecommuting propensity and frequency among individuals in households with more vehicles. This result differs from those of Drucker and Khattak (2000) and Popuri and Bhat (2003). However, it may simply be a reflection of individuals who telecommute choosing to own fewer vehicles. Future studies should examine the potential jointness in the choices of the number of vehicles and telecommuting.

5.3.4. Commute-Trip/Work Location Characteristics

A general caveat regarding the effect of commute trip/work location characteristics on telecommuting choice and frequency. All of these attributes are potentially endogenous to the choice/frequency of telecommuting, although almost all earlier studies, like the current study, have

considered such variables as exogenous to the choice of telecommuting. More broadly speaking, an argument could be made that all work-related decisions (including telecommuting, work schedule flexibility, full time versus part-time, and perhaps even industry type) and residential location choice decisions should be modeled in one single joint model system that also implicitly determines the choice of a work location and commute trip attributes. But, in the process of practical modeling, the analyst needs to make informed judgments and assumptions regarding what may be considered exogenous variables. We suggest that an area of future research should be to comprehensively examine the various choices surrounding work characteristics, residential location decisions, and telecommuting, to provide meaningful guidance regarding which variables may be considered more endogenous than others (see, for instance, the studies by Ellen and Hempstead, 2002, and Ory and Mokhtarian, 2006, which begin to address this issue; the suggestions from these studies is that individuals tend to make their work/home location choices prior to decisions on telecommuting). In the rest of this section, we discuss the effects of commute trip/work-related characteristics on telecommuting choice/frequency, though the caveat just discussed about the potential endogeneity of these characteristics should be kept in mind.

The first variable under commute trip/work-related characteristics corresponds to the direct one-way home-to-work commute distance of employees. As expected, individuals whose (one-way) commute distance is longer than 25 miles are more likely to telecommute (and telecommute with high frequency) compared to individuals with a (one-way) commute distance less than 25 miles (see also Mokhtarian and Meenakshisundaram, 2002). The next variable suggests that the frequency of telecommuting decreases if the roadway type normally traveled on to work includes an expressway, probably due to less-stressful driving conditions on expressways than on other arterial streets. The positive influence of non-car modes of travel (walk/bicycle/transit) to work is consistent with Tang *et al.*'s (2008) finding that individuals with pro-bike and pro-transit views have a higher propensity to telecommute relative to others. One explanation is that individuals who bicycle/walk/use transit to reach work are environmentally conscious, and see telecommuting as another means to reduce auto travel. Next, vehicle availability for work positively influences the frequency of telecommuting. Although this effect is consistent with most of the literature in the field, the reason for this positive relationship needs further exploration in future studies. Finally, individuals who make several non-work trips on the workday are more likely to telecommute, while those who have to pay to park at work have a higher frequency of telecommuting than those who do not have to pay to park (we also

examined the effects of the amount of any parking costs and tolls usually paid on the commute trip, but both of these policy-relevant variables did not turn out to be statistically significant even at the 0.15 level of significance).

5.4. Elasticity Effects

The parameters on the exogenous variables in the second main column of Table 1 do not directly provide a sense of the absolute magnitude of the effects of variables. To obtain such order-of-magnitude effects, we have chosen to assign cardinal values to each of the ordinal levels of telecommuting frequency, and then compute the elasticity effects of exogenous variables on the expected total number of days per month of telecommuting. The cardinal value assignments for the telecommuting ordinal frequency levels in the model are as follows: (1) telecommuting once a year ($k = 1$ in the notation of Section 4.1): $1/12 = 0.083$ telecommuting days per month, (2) a few times a year ($k = 2$): $4/12 = 0.333$ telecommuting days per month, (3) once a month or more ($k = 3$): $12/12 = 1$ telecommuting day per month, (4) once a week or more ($k = 4$): $4*12/12 = 4$ telecommuting days per month, (5) almost every day ($k = 5$): $22*12/12 = 22$ telecommuting days per month. With these assignments, the expected value of the number of telecommuting days per month for individual q (d_q) using Equation (3) in Section 4.1 is:

$$E(d_q) = \sum_{k=1}^{5} c_k \times \Pr[t_q = 1, s_q = k], \tag{8}$$

where c_k is the cardinal value assignment corresponding to telecommuting ordinal frequency level k. Note that the expected value above is a function of variables in both the vectors x_q and z_q (see Equation (3)). If there are common variables in x_q and z_q (such as age, employment level, and industry characteristics in our empirical specification), these variables will impact the expected value of the number of telecommuting days per month both through the telecommuting choice binary model and the telecommuting frequency ordered response model.

To compute the aggregate-level "elasticity" effect of a dummy exogenous variable, we change the value of the 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 the expected aggregate number of telecommuting days per month in the two subsamples after reversing the sign of the shifts in the second subsample, and

compute the effective percentage change in the expected total number of telecommuting days per month across all individuals in the sample due to a change in the dummy variable from 0 to 1. To compute the aggregate level "elasticity" effect of an ordinal variable, we increase the value of the variable by 1 and compute the percentage change in the expected total number of telecommuting days per month across all individuals in the sample.

The final column of Table 1 provides the elasticity effects. The first entry in the table indicates that the number of telecommuting days per month for women over the age of 30 years is, on average, about 6.64% less than the number of telecommuting days per month for men over the age of 30 years. Other entries may be similarly interpreted.[13] The results reveal that employees with flexible work schedules (especially if they are fully flexible) and employees working in real-estate, rental or leasing services are substantially more likely to telecommute frequently than those with no work schedule flexibility and employees in the MATRE (manufacturing, transportation, retail and other) industry category, respectively. These variables have the highest impacts on the number of days of telecommuting per month. Other variables with substantial positive impacts include being in industries related to communications (relative to being in the MATRE industries category), one-way commute distance, being a male under the age of 30 years (relative to being a female under 30 years), being in industries related to management of companies or enterprises (relative to being in the MATRE industries category), having to pay to park at work (relative to free parking at work), holding a graduate degree (relative to an education level lower than an undergraduate degree), and using a non-motorized mode to get to work (relative to the use of a motorized personal mode to get to work). For all the variables identified above (except for the "pay to park at work" variable), the high positive impact is because these variables positively influence both the choice and frequency model components of telecommuting. Further, the magnitudes of the estimated parameters on these variables in each model component are quite high relative to the estimated parameters on other variables. For the "pay to park at work" variable, the net effect on number of telecommuting days per month is quite substantial (even though it does not affect the telecommuting choice component) because it has a high positive effect in the frequency component of the model system. Finally, the results show that being a full-time employee (relative to being a part-time employee), the number of

---

[13] The elasticity effect in the row corresponding to "female × age less than 30 years" in Table 1 is computed to provide the average percentage difference, between a women less than 30 years and a man less than 30 years, in the expected number of telecommuting days per month (the net coefficient used in this computation is –0.071 – 0.248 = –0.319).

non-work trips on the work-day, the number of jobs held, and being a female over the age of 30 years (relative to being a male over 30 years) have a much smaller impact on the number of telecommuting days per month relative to other explanatory variables.

6. CONCLUSIONS AND IMPLICATIONS

In the current paper, we contribute to the existing telecommuting literature by jointly analyzing the choice and the frequency of telecommuting, using data from the 2007/2008 Chicago Regional Household Travel Inventory (CRHTI). The empirical results indicate the important effects of several demographic and work-related variables. First, the results clearly indicate that telecommuting choice and the frequency of telecommuting may be governed by quite different underlying behavioral processes rather than being governed by a single behavioral process. In particular, the determinant factors of choice and frequency can be different. Thus, according to our results, gender and presence of children in the household impacts the telecommuting choice decision, but not the frequency dimension. Further, a factor that has a particular direction of effect on telecommuting choice may have the opposite effect on frequency. For instance, our results indicate that full-time employment is positively associated with the choice of telecommuting, but negatively associated with the frequency of telecommuting. Second, unobserved factors that predispose an individual to choose to telecommute also increase the individual's telecommuting frequency. But the results also emphasize that pre-imposing a specific dependency structure between the telecommuting choice and frequency decisions can lead to inappropriate conclusions regarding the presence and extent of dependency. In the current paper, we found that using the typical bivariate normal distribution assumption between the telecommuting choice and frequency equations provides the incorrect conclusion of no statistically significant dependency, while using the Frank copula indicates the clear presence of dependency. Further, the influence of exogenous variables from models assuming different dependency structures are different from one another, and the standard errors of the telecommuting frequency model estimates were, in general, smaller from the best-fit Frank copula structure than those from other structures. Overall, one risks the danger of incorrect conclusions regarding dependency in the telecommuting choice and frequency behavioral processes, as well as inconsistent and inefficient parameter estimates, by imposing incorrect dependency structures. It behooves the analyst to empirically test alternative profiles of dependency (*i.e.*, copulas) and select the most appropriate one. Third, work schedule flexibility and industry type are important determinants of

telecommuting choice and frequency. In particular, workers whose schedules are fully flexible and who are in the real estate, rental, or leasing industries are much more likely to telecommute than their peers. Fourth, several factors related to the commute trip and work location influence telecommuting choice and frequency. For instance, our results suggest that individuals who have to pay to park at the work place are more frequent telecommuters than those who do not have to pay to park. Also, those who usually bicycle, walk, or use transit to reach their work place are also more likely to telecommute. Of course, these commute mode choice decisions may be related to built environment attributes at the residence end and/or at the work end, so they may be proxying for built environment effects. Future studies would benefit from the consideration of a comprehensive set of built environment variables, in addition to the many categories of variables included here. As indicated earlier, it would also be helpful to examine the many choices surrounding work characteristics, residential location decisions, and telecommuting to provide meaningful guidance regarding which variables may be considered endogenous and which exogenous in telecommuting choice/frequency modeling.

The empirical results have implications for transportation planning analysis, especially because of the projected changes in demographic and employment-related variables (such as age, households with and without children, and work characteristics) in the U.S. population. The models estimated in this paper can be used to assess the impacts of these changes. The model results can also be used to target specific employee groups, and employer groups based on industry sector, to increase the extent of telecommuting. Companies can use the results to predict how many employees would show up to work on any given workday, which may help plan for office space and parking space. Companies and planning agencies can also evaluate the effects of imposing parking fees at the work place. Finally, the predictions from the model system developed in this paper can feed into larger-scale activity-based travel demand modeling systems that use work-related decisions of individuals as a "peg" around which to schedule other activities and travel.

REFERENCES

Bernardino AT, Ben-Akiva M, 1996, "Modeling the adoption of telecommuting: comprehensive framework" *Transportation Research Record* **1552** 161-170.

Bernardino AT, Ben-Akiva M, Salomon I, 1993, "Stated preference approach to modeling the adoption of telecommuting" *Transportation Research Record* **1413** 22-30.

Bhat CR, Eluru N, 2009, "A copula-based approach to accommodate residential self-selection effects in travel behavior modeling" *Transportation Research Part B* **43**(7) 749-765.

Bhat CR, Sener IN, 2009, "A copula-based closed-form binary logit choice model for accommodating spatial correlation across observational units" *Journal of Geographical Systems* **11**(3) 243-272.

Chicago Regional Household Travel Inventory (CRHTI), 2009, Chicago Metropolitan Agency for Planning, http://www.cmap.illinois.gov/TravelTrackerData.aspx.

Cosslett S, 1981, "Maximum likelihood estimator for choice based samples" *Econometrica* **49**(5) 1289-1316.

Drucker J, Khattak AJ, 2000, "Propensity to work from home: modeling results from the 1995 nationwide personal transportation survey" *Transportation Research Record* **1706** 108-117.

Ellen IG, Hempstead K, 2002, "Telecommuting and the demand for urban living: a preliminary look at white-collar workers" *Urban Studies* **39**(4) 749-766.

Frank MJ, 1979, "On the simultaneous associativity of F(x, y) and x + y - F(x, y)" *Aequationes Mathematicae* **19**(1) 194-226.

Frank P, 2009, "Weighting the Chicago regional household travel inventory survey" Chicago Metropolitan Agency for Planning, http://www.cmap.illinois.gov/uploadedFiles/regional_data/TravelTrackerSurvey/TravelTrackerWeighting.pdf.

Genius M, Strazzera E, 2008, "Applying the copula approach to sample selection modeling" *Applied Economics* **40**(11) 1443-1455.

Heckman JJ, 1979, "Sample selection bias as a specification error" *Econometrica* **47**(1) 153-161.

Lee LF, 1983, "Generalized econometric models with selectivity" *Econometrica* **51**(2) 507-512.

Manski CF, Lerman SR, 1977, "The estimation of choice probabilities from choice based samples" *Econometrica* **45**(8) 1977-1988.

Mannering JS, Mokhtarian PL, 1995, "Modeling the choice of telecommuting frequency in California: an exploratory analysis" *Technological Forecasting and Social Change* **49**(1) 49-73.

Mokhtarian PL, Meenakshisundaram R, 2002, "Patterns of telecommuting engagement and frequency" *Prometheus* **20**(1) 21-37.

Mokhtarian PL, Salomon I, 1996a, "Modeling the choice of telecommuting: 2. A case of the preferred impossible alternative" *Environment and Planning A* **28**(10) 1859-1876.

Mokhtarian PL, Salomon I, 1996b, "Modeling the choice of telecommuting: 3. Identifying the choice set and estimating binary choice models for technology-based alternatives" *Environment and Planning A* **28**(10) 1877-1894.

Mokhtarian PL, Salomon I, 1997, "Modeling the desire to telecommute: the importance of attitudinal factors in behavioral models" *Transportation Research Part A* **31**(1) 35-50.

Mokhtarian PL, Salomon I, Choo S, 2005, "Measuring the measurable: why can't we agree on the number of telecommuters in the U.S.?" *Quality & Quantity* **39**(4) 423-452.

NuStats, 2008, Chicago Regional Household Travel Inventory Draft Final Report. Prepared for Chicago Metropolitan Agency for Planning, Available at: http://www.cmap.illinois.gov/TravelTrackerData.aspx.

Olszewski P, Mokhtarian PL, 1994, "Telecommuting frequency and impacts for State of California employees" *Technological Forecasting and Social Change* **45**(3) 275-286.

Ory DT, Mokhtarian PL, 2006, "Which came first, the telecommuting or the residential relocation? An empirical analysis of causality" *Urban Geography* **27**(7) 590-609.

Páez A, Scott DM, 2007, "Social influence on travel behavior: a simulation example of the decision to telecommute" *Environment and Planning A* **39**(3) 647-665.

Páez A, Scott DM, Volz E, 2008, "A discrete-choice approach to modeling social influence on individual decision making" *Environment and Planning B* **35**(6) 1055-1069.

Popuri YD, Bhat CR, 2003, "On modeling choice and frequency of home-based telecommuting" *Transportation Research Record* **1858** 55-60.

Quinn C, 2007, "The health-economic applications of copulas: methods in applied econometric research" Health, Econometrics and Data Group (HEDG) Working Paper 07/22, Department of Economics, University of York.

Sklar A, 1973, "Random variables, joint distribution functions, and copulas" *Kybernetika* **9** 449-460.

Storper M, Venables AJ, 2004, "Buzz: face-to face contact and the urban economy" *Journal of Economic Geography* **4**(4) 351-370.

Sullivan MA, Mahmassani HS, Yen J-R, 1993, "Choice model of employee participation in telecommuting under a cost-neutral scenario" *Transportation Research Record* **1413** 31-41.

Tang W, Mokhtarian PL, Handy S, 2008, "The role of neighborhood characteristics in the adoption and frequency of working at home: empirical evidence from Northern California", Institute of Transportation Studies, UC Davis, http://repositories.cdlib.org/itsdavis/UCD-ITS-RR-08-21.

Trivedi PK, Zimmer DM, 2007, "Copula modeling: an introduction for practitioners" *Foundations and Trends in Econometrics* 1(1), Now Publishers.

Walls M, Safirova E, 2004, "A review of the literature on telecommuting and its implications for vehicle travel and emissions" Discussion Paper 04-44, December, Resources for the Future, Washington DC.

Walls M, Safirova E, Jiang Y, 2007, "What drives telecommuting? Relative impact of worker demographics, employer characteristics, and job types" *Transportation Research Record* **2010** 111-120.

World at Work, 2009, "Telework Trendlines 2009: A Survey Brief by World at Work", February, http://www.workingfromanywhere.org/news/Trendlines_2009.pdf.

Yen J-R, Mahmassani HS, 1994, "The telecommuting adoption process: conceptual framework and model development" Report 60055-1, prepared for the Southwest Region University Transportation Center, Center for Transportation Research, The University of Texas at Austin, Austin, Texas.

Yeraguntla A, Bhat, CR, 2005, "Classification taxonomy and empirical analysis of work arrangements" *Transportation Research Record* **1926** 233-241.

Table 1. Estimation Results for Employees' Telecommuting Choice and Frequency Models

Explanatory Variables	Univariate Statistics		Estimation Results for Joint Model - *NNF*				Elasticity Effects for Expected Number of Telecommuting Days per Month
	Non-Telecommuter # (%) or mean	Telecommuter # (%) or mean	Telecommuting Choice Model		Telecommuting Frequency Model		
			Estimate	t-stat	Estimate	t-stat	
Constant	-	-	-3.039	-21.16	-	-	-
Threshold 1	-	-	-	-	-1.518	-5.02	-
Threshold 2	-	-	-	-	-0.461*	-1.55	-
Threshold 3	-	-	-	-	0.496*	1.86	-
Threshold 4	-	-	-	-	1.791	7.83	-
Individual Demographics							
Female	4080 (50.4)	640 (41.7)	-0.071*	-1.79	-	-	-6.64
Age less than 30 years	1206 (14.9)	88 (5.7)	-0.103*	-1.17	-	-	-9.47
Female × Age less than 30 years	633 (7.8)	32 (2.1)	-0.248	-1.99	-0.604	-3.60	-64.33
Education: Bachelor's or Undergraduate degree	2290 (28.3)	544 (35.4)	0.316	6.54	-	-	30.13
Education: Graduate degree	2215 (27.4)	685 (44.7)	0.430	8.55	0.117*	1.94	55.31
Driver license	7664 (94.7)	1512 (98.6)	0.314	2.75	-	-	27.58
Work-Related and Occupation Characteristics							
Full-time employment (>30 hours/week)	6472 (80.0)	1320 (86.0)	0.262	4.92	-0.243	-2.83	0.77
Partially flexible	3594 (44.4)	807 (52.6)	0.955	17.75	0.224*	1.94	106.10
Fully flexible	895 (11.1)	633 (41.3)	1.694	28.07	0.657	4.96	351.75
Number of jobs	1.13	1.16	0.051*	1.42	-	-	4.80
Industry							
Communications	341 (4.2)	142 (9.3)	0.468	6.40	0.242	2.29	86.00
Service-based							
Service – Finance and insurance	5443 (6.7)	188 (12.3)	0.223	3.33	-	-	21.55
Service – Real estate, rental, or leasing	78 (1.0)	51 (3.3)	0.606	4.58	0.619	3.74	192.26
Service – Professional, scientific, or technical service	858 (10.6)	310 (20.2)	0.314	5.86	-	-	30.68
Service – Management of companies, or enterprises	162 (2.0)	52 (3.4)	0.272	2.51	0.239*	1.58	60.78
Service – Arts, entertainment, or recreation	214 (2.6)	28 (1.8)	-	-	0.277*	1.43	31.82
Service – Educational services	1159 (14.3)	147 (9.6)	-0.088*	-1.25	0.236	2.38	16.53
Service – Health care or social assistance	1046 (12.9)	154 (10.0)	-0.111*	-1.77	0.282	2.99	19.17
Government	759 (9.4)	86 (5.6)	-0.155	-1.99	-	-	-14.09
Household Demographics							
Presence of children less than or equal to 15 years	2646 (32.7)	312 (39.9)	0.101	2.67	-	-	9.49
Number of household workers	1.91	1.86	0.054*	1.89	0.081*	1.81	14.24
Household income between 75K-100K	1695 (21.0)	294 (19.2)	0.286	5.31	-	-	27.25
Household income greater than 100K	2904 (35.9)	935 (61.0)	0.466	9.75	-	-	44.99
Number of household vehicles	2.03	1.99	-0.046*	-1.90	-0.046*	-1.40	-8.74
Commute-Trip/Work Location Characteristics							
One-way commute distance more than 25 miles	1962 (36.6)	743 (48.4)	0.233	6.09	0.400	7.23	65.26
Commute trip made on an expressway	56 2 (6.9)	86 (5.6)	-	-	-0.231	-2.11	-22.12
Walk/bike to work	372 (4.6)	98 (6.4)	0.195	2.34	0.256	2.20	52.88
Transit to work	1266 (15.6)	352 (22.9)	0.209	4.34	-	-	20.04
Vehicle available for work	1592 (19.7)	516 (33.6)	-	-	0.274	4.45	29.89
Number of non-work trips on the work-day	1.91	2.40	0.018	2.15	-	-	1.68
Pay to park at work?	125 (1.5)	26 (1.7)	-	-	0.463	2.16	56.43
Number of Observations			9624				
Dependency parameter estimate (t-stat)	-		1.880 (2.33)				-
Log-likelihood at sample shares			-6271.56				
Log-likelihood at convergence			-5123.21				

* The significance level of these parameters is lower than 0.05, as can be observed from the t-statistic.