

Momanyi, Kevin

**Working Paper**

## Telecare and Unplanned Hospitalization in Scotland: Evidence From Linked Survey and Administrative Data

*Suggested Citation:* Momanyi, Kevin (2017) : Telecare and Unplanned Hospitalization in Scotland: Evidence From Linked Survey and Administrative Data, ZBW - Leibniz Information Centre for Economics, Kiel, Hamburg

This Version is available at:

<https://hdl.handle.net/10419/256933>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# Telecare and Unplanned Hospitalization in Scotland: Evidence From Linked Survey and Administrative Data

Kevin Momanyi \*

*University of Aberdeen*

---

## Abstract

In the face of increasing health care costs, policy makers in Scotland are interested in coming up with innovative ways of reducing unplanned hospital admissions. This paper investigates whether the use of telecare devices could be one such way. Unlike the previous studies looking into the same issue, we link the Scottish Homecare Census data to three other information sources-including the Scottish Morbidity Records-and estimate the treatment effect using time series analysis. We also employ an estimation strategy that controls for various methodological challenges associated with the analysis of observational data. A key highlight of our investigation is that the type of telecare device matters in as far as reducing unplanned hospitalization is concerned. For instance, when we conduct an analysis of the effect of using a community alarm, we find community alarm users to have a lower likelihood of being hospitalized as emergency cases than non-users, holding other factors constant. When we, however, consider the other telecare devices, the results show the opposite effect.

*Key words:* Unplanned hospital admissions, telecare, time series analysis.

---

*Preprint submitted to the University of Aberdeen, 24-Month Assessment Report*

## 1 Introduction

The provision of better health care services and advances in medical technology have led to improved life expectancy rates for many countries worldwide (National Records of Scotland, 2015). These gains, however, have also been accompanied by an increased demand for social care, multi-morbidity (Barnett et al., 2012) and unplanned hospitalization (Wolff, Starfield and Anderson, 2002). In Scotland, for example, the number of emergency cases has been on the rise over the recent years and there is therefore need to find out how this trend can be reversed. This study investigates one such way by determining the relationship between the use of telecare devices i.e. community alarms and automated sensors, and unplanned hospitalization.

While there are a number of studies that investigate the determinants of unplanned admission to hospital (see for example, Shelton, 2000, Santolino, 2011, Verdaguer, 2003, Cannolly, 2006, Billings, 2006, Lyon et al., 2007, Coburn, 2002), we are not aware of any study in Scotland that considers telecare use as one of the potential predictors. Yet telecare could reduce the likelihood of hospitalization by enabling the users to summon for help immediately or before an adverse event happens (Steventon et al., 2013). As far as we know, the majority of the existing studies looking into the impact of telecare are randomized controlled trials and quasi-experiments with relatively small samples (see for example, Akematsu & Tsuji, 2012a, Akematsu & Tsuji, 2012b, Akematsu, 2013, Henderson et al., 2014, Hirani et al., 2013, Steventon et al., 2013, Brownsell, 2008). We therefore contribute to the literature by conducting our study using routinely collected large scale non-experimental data.

---

\* Corresponding author

*Email address:* kevzy44@gmail.com; *Date:* October 2017 (Kevin Momanyi).

The specific objectives of the paper are (i) to investigate the relationship between the use of telecare devices and admission to hospital in the emergency department; (ii) to identify the predictors of emergency admission, and (iii) to report the treatment effects for telecare. The next section of the paper describes the data sources. Section 3 discusses the methodology. Section 4 presents the results, while Section 5 concludes.

## **2 The data**

As pointed out in Section 1, we use large scale non-experimental data to address our study objectives. The data is obtained from four different sources i.e. the Scottish Homecare Census for 2011, the Scottish Morbidity Records (SMRs) for the 2010/2011 financial year, Self-Directed Support data (SDS) for 2011 and prescribing data from the beginning of April 2010 to the end of March 2011. Our covariates of interest are contained in the Scottish Morbidity Records and the Homecare Census.

The Scottish Morbidity Records contain episode level data on both acute hospital admissions and admissions to psychiatric care. We use the information contained in these records to compute the outcome measure. The Scottish Homecare Census contains information on telecare use for 27,800 homecare clients who were residents in one of the following five local council areas: South Ayrshire, Stirling, the City of Edinburgh, South Lanarkshire and Clackmannanshire. Among the variables contained in the dataset are ethnicity, a variable for the use of telecare devices and a variable for client group (indicating whether or not a particular individual had a mental health problem including dementia, a learning disability or a physical disability).

In this paper, we generate weekly time series data with repeated cross-sections using the dates of admission in the Scottish Morbidity Records; and by so doing, we extend the experimental studies in the literature that only investigate changes in outcomes at specific points in the follow up period. The use of a pooled dataset also has a distinct advantage over the other observational study designs in that it allows us to analyze the trend in unplanned hospitalization over the period of analysis and obviate various methodological issues.

### **3 Methodology**

#### *3.1 Modeling framework*

Following the Andersen’s Behavioral Model of Health Services Use (Andersen, 1995), we can develop a modeling framework that shows the relationship between telecare use and unplanned admission to hospital. According to the framework, the predictors of unplanned hospitalization can be broadly classified into three categories i.e. ‘predisposing’, ‘enabling’ and ‘need’ factors. The ‘predisposing’ factors are those factors that increase the likelihood of unplanned admission to hospital e.g. old age. The ‘enabling’ factors are those factors that enable individuals to change their patterns of service use. The ‘need’ factors are those factors that are related to individuals’ care needs e.g. comorbidity.

In this paper, we conceptualize the use of telecare as an ‘enabling’ factor. We assume that individuals maximize utility by using telecare devices so as to reduce the probability of unplanned hospitalization. We further assume that individuals also maximize their health status in pursuit of welfare gains. The

utility function for a particular individual is thus given by:

$$U = U(X, T, H) \tag{1}$$

where  $U$  is the utility function for a particular individual;  $X$  is a vector of ‘predisposing’ and ‘need’ factors;  $T$  denotes the use of telecare/community alarm, and  $H$  is the individual’s health status.

The health status of the individual is, in turn, expressed as a function of the use of telecare, other observable health inputs,  $D$ , and unobservable biological endowments,  $\mu$ , as follows:

$$H = H(T, D, \mu) \tag{2}$$

### 3.2 *Estimation issues*

Due to the intricate nature of the health production function shown in Equation 2 in that it contains an unobservable component as one of its arguments and is also part of an individual’s utility maximizing behavior, the empirical version of the utility function could suffer from endogeneity and unobserved heterogeneity. Endogeneity in this case will be brought about by the correlation of the variable for telecare use with the unobservable factors in Equation 2, whereas unobserved heterogeneity will emerge if the unobservable factors that bring about endogeneity also lead to systematic differences in the probability of unplanned admission to hospital between telecare/community alarm users and non-users.

We are also likely to encounter some challenges in the course of estimating the treatment effect since the study uses observational data. Firstly, there could

be instances where the outcome measure is missing for some homecare clients in the dataset thus leading to sample selection bias. Secondly, the fact that we use the Scottish Morbidity Records to generate time series data yet the Homecare Census data is cross-sectional implies that the variables contained therein could be endogenous due to measurement error.

In this paper, we control for potential sample selection bias using the Amemiya-Heckman approach (Amemiya, 1980). This technique is implemented in three steps. We first obtain the predicted probabilities of a probit model relating inclusion into the study sample to several exogenous covariates and some instruments. We then express the correction term for potential sample selection bias as the ratio of the probability density function (pdf) to the cumulative distribution function (cdf) for each observation in the dataset. The pdf and cdf are constructed from the predicted probabilities obtained in the first step. The last step entails including the quotient computed in the second step as an additional explanatory variable in our empirical model of interest. A statistically significant coefficient for this variable is indicative of sample selection bias.

The problem of potential endogeneity and potential unobserved heterogeneity is controlled for using a two-stage estimation procedure (See for example, Terza, Basu & Rathouz, 2008). In the first stage, we estimate a reduced form model of the use of telecare using a probit model. We then include, in our empirical model, the residuals of the model as well as an interaction term of these residuals with the treatment variable as controls for potential endogeneity and potential unobserved heterogeneity respectively. Since the study covariates in this paper are generated as time series data, we also add a variable for time trend that controls for unexplained variations in the probability of unplanned

hospitalization over the 2010/2011 financial year.

### *3.3 Identification strategy*

Our objective in this paper is to determine the effect of the use of telecare devices on unplanned hospitalization by estimating the utility function in Equation 1. In order to do so, however, we need to properly identify the arguments of the utility function. The discussion in the previous sub-section suggests that the variable for telecare is potentially endogenous and a selection issue could arise due to missingness. Further, given that the proposed corrective measures entail solving a reduced form model of telecare use and a sample selection equation, we need two instrumental variables for exact identification (Pokropek, 2016; Greene, 2002). These instruments, however, ought to be correlated with the variable for telecare use but should not be determined in the utility function defined in Equation 1 (Murray, 2006; Wooldridge, 2002).

The first instrument is a variable for the proportion of telecare users in each of the local council areas mentioned in Section 2, whereas the second is the ‘Scottish Index of Multiple Deprivation’. We expect these variables to be correlated with the likelihood of telecare use but not endogenously determined in Equation 1 since they are aggregated measures at the population level.

We determine whether the chosen instruments are correlated with the treatment variable by testing for their statistical significance in the reduced form model of telecare use. The instruments are said to be relevant if the coefficients are significantly different from zero.



### 3.4 Econometric specification

We estimate the utility function in Equation 1 using generalized linear models. These models generalize the linear regression model by allowing the outcome variable to be related to the independent variables via a link function. Accordingly, they encompass several other models including binary regression models, count models, multinomial models and ordinal models (see McCullagh, 1999 for a comprehensive discussion of generalized linear models). In this paper, the dependent variable,  $Y_t$ , is defined as a nominal variable and we thus specify the empirical model as a multinomial model.

$$Y_t = \begin{cases} 1 & \text{for admission as a daycase at time } t, \\ 2 & \text{for admission as an elective inpatient at time } t, \\ 3 & \text{for admission as an emergency case at time } t. \end{cases} \quad (3)$$

Following the Amemiya-Heckman correction for potential sample selection bias and the approach by Terza, Basu and Rathouz (2008), the empirical model can be thought of as comprising the following three equations:

$$Pr(T = 1|F, G, t) = \Phi(\delta_0 + \delta_1 F + \delta_2 G + \delta_3 t) \quad (4)$$

$$Pr(I_t = 1|F, G, t) = \Phi(\delta_4 + \delta_5 F_t + \delta_6 G + \delta_7 t) \quad (5)$$

$$Pr(Y_t = 1|X, T, t) = \arg \max_{m=1}^3 g \left[ \alpha + \sum_{i=1}^5 \beta_i X_i t + \gamma_1 T_t + \gamma_2 t + \gamma_3 \hat{\epsilon} + \gamma_4 T \hat{\epsilon} + \gamma_5 \lambda \right]^c \quad (6)$$

where Equation 4 is the reduced form model of telecare use, Equation 5 is the sample selection model and Equation 6 is the second stage model;  $T$  is

an indicator for telecare;  $F, G$  and  $t$  respectively denote a vector of exogenous variables, a vector of instruments and a time trend variable (unit of time = 1 week);  $I_t$  is a variable for inclusion into the study sample at time  $t$ ;  $g$  follows a multivariate normal distribution such that  $g \sim N(0, \sigma^2)$  and  $i = 1, 2, 3$ ;  $c = 1, 2, 3$  are the three categories of  $Y_t$  as shown in Equation 3;  $X_1$  is age in years;  $X_2$  is a measure of comorbidity;  $X_3$  is a nominal variable for client group;  $X_4$  is a binary variable for sex;  $X_5$  is binary variable for rural residence;  $\hat{\epsilon}$  are the residuals obtained from the reduced form model in Equation 4, and  $\lambda$  is the Amemiya-Heckman correction for sample selection. We shall later provide more elaborate variable definitions in Table 2.

Note that Equations (4), (5) and (6) are specified as probit models since in all cases  $Pr(.) = \Phi(.)$  or  $Pr(.) = g(.)$ , where  $\Phi$  is the cdf of the standard normal distribution and  $g$  is multivariate normal. Given that probit models are typically estimated via Maximum Likelihood Estimation (MLE) and that MLE for such models imposes specific values for the means and standard deviations of the underlying processes generating  $T, I_t$  and  $Y_t$  i.e. *variance* = 1 and *mean*=0, we formulate our empirical models of interest as latent variable models. The intuition behind this approach is that the actual data generating process is unobservable but can be linked to  $T, I_t$  and  $Y_t$  via the following measurement equations (Greene, 2012; Long and Freese, 2006):

$$T = \begin{cases} 1 & \text{if } \zeta_T > 0 \\ 0 & \text{otherwise} \end{cases} \quad (7)$$

$$I_t = \begin{cases} 1 & \text{if } \zeta_s > 0 \\ 0 & \text{otherwise} \end{cases} \quad (8)$$

$$Y_t = \begin{cases} 1 & \text{if } \zeta_t^1 > \zeta_t^2, \zeta_t^3, \\ 2 & \text{if } \zeta_t^2 > \zeta_t^1, \zeta_t^3, \\ 3 & \text{if } \zeta_t^3 > \zeta_t^1, \zeta_t^2. \end{cases} \quad (9)$$

Throughout, the notation  $\zeta$  refers to the linear predictor functions for the actual data generating processes and the superscripts in these functions denote the three categories of  $Y_t$  unless otherwise specified. In Equation 9, therefore,  $\zeta_t^1$ ,  $\zeta_t^2$  and  $\zeta_t^3$  refer to the linear predictor functions for day case charges, elective admissions and emergency admissions at time  $t$ .

The linear predictors are, in turn, expressed as linear combinations of the independent variables in Equations (4), (5) and (6) as follows:

$$\zeta_T = \delta_0 + \delta_1 F + \delta_2 G + \delta_3 t + \epsilon_1 \quad (10)$$

$$\zeta_s = \delta_4 + \delta_5 F_t + \delta_6 G + \delta_7 t + \epsilon_2 \quad (11)$$

$$\zeta_t^c = \left[ \alpha + \sum_{i=1}^5 \beta_i X_i t + \gamma_1 T_t + \gamma_2 t + \gamma_3 \hat{\epsilon} + \gamma_4 T \hat{\epsilon} + \gamma_5 \lambda + \epsilon_3 \right]^c \quad (12)$$

where  $\epsilon_1$ ,  $\epsilon_2$  and  $\epsilon_3$  are random error terms.

Substituting Equations (7), (8) and (9) into their corresponding linear predictor functions we obtain the predicted probabilities of observing  $T$ ,  $I_t$  and  $Y_t$ .

$$Pr(T = 1|F, G) = Pr(\zeta_T > 0|F, G) \quad (13)$$

$$Pr(I_t = 1|F, G) = Pr(\zeta_s > 0|F, G) \quad (14)$$

$$Pr(Y_t = 1|X, T) = Pr(\zeta_t^1 > \zeta_t^2, \zeta_t^3|X, T) \quad (15)$$

$$Pr(Y_t = 2|X, T) = Pr(\zeta_t^2 > \zeta_t^1, \zeta_t^3|X, T) \quad (16)$$

$$Pr(Y_t = 3|X, T) = Pr(\zeta_t^3 > \zeta_t^1, \zeta_t^2 | X, T) \quad (17)$$

Notice from the above equations, specifically Equation 17 that  $Pr(Y_t = 3|X, T)$  shows the predicted probabilities for unplanned hospitalization. If we let  $Pr(Y_t = 3|X, T)$  be denoted by  $\theta$ , then the change in  $\theta$  that is attributable to variable  $T$  is the marginal effect of telecare and one could thus interpret the treatment effect in this way. In this paper, however, we directly interpret  $\gamma_1$  in Equation 6 so as to show the effect of telecare on the likelihood of admission to hospital as an emergency case with reference to hospital admission as an elective inpatient. Here, a negative sign for  $\gamma_1$  implies a negative effect, whereas a positive sign shows a direct relationship with the outcome variable. It is important to note that if we were to take the partial derivative of Equation 6 with respect to  $T$  and instead interpret our substantive model using marginal effects, then the empirical results would include the predicted probabilities of the three categories of  $Y_t$  as shown in Equations (15), (16) and (17). In such a case, therefore, there would be no base category.

## 4 Results

### 4.1 Descriptive statistics

As pointed out in Section 2, the data for this study is drawn from four different information sources: the Scottish Morbidity Records, the Homecare Census data for 2011, prescribing data and Self-Directed Support data. The descriptive statistics are presented in Table 1. The table contains the means (or medians and proportions where applicable) and the number of observations for the study variables. The variable definitions are presented in Table 2. In

this paper, we consider two treatment variables: one for the use of community alarms (henceforth referred to as community alarm use) and another for the use of the other telecare devices besides community alarms.

The results in Table 1 show that while the study population has 49,025 observations, the study sample has only 21,580. This implies that a selectivity issue could arise since the sample covers less than half of the study population. We can also observe from the table that the data has missing values since there are some variables with different number of observations. The distribution of the observations for the homecare clients in the study sample by client group shows that about 6% of the observations were for the individuals with dementia or other mental health conditions, 17% for the physically disabled homecare clients, 50% for the frail elderly and 27% for those with learning disabilities.

Other highlights from the table include that there were more females (65%) than males (35%); approximately 9% of the individuals included in the sample were rural residents; 67% of the homecare clients in the study sample used a community alarm, and a substantial proportion of the individuals in the sample were admitted in the emergency department.

#### *4.2 Empirical results*

The estimated coefficients for the empirical model defined in Equation 6 are presented in Tables (3) and (4). In each table, we report the effect of the treatment variables on emergency admission with reference to admission to hospital as an elective inpatient. We also present four variants of the empirical model of interest. These models are labeled as (1), (2), (3) and (4). The

Table 1  
Descriptive statistics

	Study population		Study sample	
	N	M	N	M
Age	48,789	75.60	21,580	74.97
Comorbidity status	49,025	0.40	21,580	0.43
Dementia and Mental Health	49,025	0.10	21,580	0.06
Learning disability	49,025	0.30	21,580	0.27
Physical disability	49,025	0.20	21,580	0.17
Frail elderly	49,025	0.50	21,580	0.50
Male	48,789	0.40	21,580	0.35
Female	48,789	0.70	21,580	0.65
Rural residence	48,584	0.10	21,517	0.09
Telecare use	49,025	0.10	21,580	0.10
Community alarm use	49,025	0.70	21,580	0.67
SIMD-Decile	48,644	5.00	21,580	5.00
Day case charges	21,580	0.21	21,580	0.21
Elective admissions	21,580	0.10	21,580	0.10
Emergency admissions	21,580	0.70	21,580	0.69
Proportion of telecare users	49,025	0.12	21,553	0.11
Trend	49,025	40.0	21,580	28.0
Number of homecare clients		25,982		10,590

**Notes:** The actual size of the study population is 27,800 as stated in Section 2 but only 25,982 clients could be identified in the dataset.  $M$  denotes arithmetic mean, median or proportion. We report the arithmetic mean for the age variable and the variable for the proportion of telecare users; the medians for SIMD-Decile and the variable for trend since they are ordinal and count variables respectively, and proportions for the other variables since they are categorical.

first model shows the basic model without  $\lambda$ ,  $\epsilon$  and  $T'\hat{\epsilon}$ . The second model controls for potential endogeneity of the treatment variable using the approach discussed in Section 3.2. The third model controls for potential endogeneity of the treatment variable and corrects for potential sample selection bias using the Amemiya-Heckman method. The fourth column extends the third model by accounting for potential unobserved heterogeneity.

Table 2  
Variable definitions

<i>Variable</i>	<i>Definition</i>
Age, $X_1$	Age at admission to hospital in years at time $t$ .
Comorbidity status, $X_2$	1 if an individual has three or more comorbid conditions, 0 otherwise.
Client group, $X_3$	1 if client group is ‘Dementia and Mental Health’, 2 if client group is ‘Learning disability’, 3 if client group is ‘Physical disability’, 4 if client group is ‘Frail elderly’.
Female, $X_4$	1 if female, 0 otherwise.
Rural residence, $X_5$	1 if area of residence is rural, 0 otherwise.
Telecare use, $T_1$	1 if for the use of other telecare devices besides the commun
Community alarm use, $T_2$	1 for community alarm use only, 0 otherwise.
SIMD-Decile, $G_2$	Ten categories of the Scottish Index of Multiple Deprivation.
Admission to hospital, $Y_t$	1 for day case charges, 2 for elective inpatient admissions and 3 for emergency admissions.
Status of admission	A nominal variable indicating the status of admission i.e. whether formal or informal.
Inclusion into the sample, $I_t$	1 for day case charges, elective inpatient admissions and emergency admissions at time $t$ , 0 otherwise.
Proportion of telecare users, $G_1$	A variable indicating the proportion of telecare users in each local council area.
Project ID	A unique reference number for each individual in the dataset.
Residuals, $\hat{\epsilon}$	The residuals of the reduced form models of telecare/community alarm use.
The potentially endogenous variables interacted with their residuals, $T'\hat{\epsilon}$	The potentially endogenous variables interacted with their residuals.
The Amemiya-Heckman selection term, $\lambda$	A control for potential sample selection bias.
Trend, $t$	A time trend variable where the unit of time is 1 week.

Table 3

Estimated coefficients for the probability of admission to hospital in the emergency department,  $Z$  statistics in parentheses

	<i>Base category: elective inpatient admission</i>			
	(1)	(2)	(3)	(4)
Age, $X_1$	-0.040 (4.19)	-0.052 (5.09)	-0.040 (3.52)	-0.040 (3.58)
Square of age, $X_1^2$	0.0004 (6.48)	0.001 (6.82)	0.0004 (4.51)	0.01 (4.58)
Comorbidity status, $X_2$	1.060 (24.15)	1.042 (23.33)	1.050 (23.49)	1.056 (23.49)
Dementia and Mental Health, $X_{3,1}$	-0.040 (0.68)	0.012 (0.23)	-0.100 (1.41)	-0.099 (1.32)
Learning disability, $X_{3,2}$	0.430 (3.99)	0.322 (2.93)	0.200 (1.60)	0.183 (1.56)
Physical disability, $X_{3,3}$	-0.170 (2.53)	-0.119 (1.70)	-0.134 (1.93)	-0.132 (1.85)
Female, $X_4$	-0.004 (0.07)	0.057 (1.04)	0.071 (1.17)	0.067 (1.19)
Rural residence, $X_5$	0.120 (1.55)	0.173 (2.09)	-0.173 (2.06)	0.170 (2.04)
Telecare use, $T_1$	0.190 (2.60)	0.190 (2.62)	0.193 (2.63)	0.520 (1.68)
Trend, $t$	-0.005 (4.25)	-0.005 (4.19)	0.030 (2.29)	0.028 (2.27)
Telecare residuals, $\hat{\epsilon}$		3.999 (2.83)	3.842 (2.72)	4.240 (2.86)
Correction for sample selection, $\lambda$			1.650 (2.71)	1.640 (2.68)
Telecare interacted with its residuals, $T_1 \hat{\epsilon}$				-2.690 (1.08)
Number of observations	21,517	21,488	21,488	21,488
Number of homecare clients	10,590	10,578	10,578	10,578
<i>Wald test for weak instruments; <math>\chi^2(p - value)</math></i>			28.10(0.00)	
<i>Wald Chi-square test; <math>\chi^2(p - value)</math></i>	1621(0.00)	1632(0.00)	1654(0.00)	1656(0.00)

The  $Z$  statistics are clustered by Project ID.  $Z$  statistics greater than or equal to 1.96 imply statistical significance at 5% level of significance, whereas the  $Z$  statistics greater than or equal to 1.64 imply statistical significance at 10% level of significance. The reference group for client group is 'Frail elderly'.



Table 4

Estimated coefficients for the probability of admission to hospital in the emergency department,  $Z$  statistics in parentheses

	<i>Base category: elective inpatient admission</i>			
	(1)	(2)	(3)	(4)
Age, $X_1$	-0.032 (3.85)	-0.088 (4.78)	-0.073 (3.73)	-0.075 (3.81)
Square of age, $X_1^2$	0.0004 (6.22)	0.001 (6.42)	0.001 (4.52)	0.01 (4.56)
Comorbidity status, $X_2$	1.062 (24.09)	1.076 (23.95)	1.081 (24.06)	1.082 (24.05)
Dementia and Mental Health, $X_{3,1}$	-0.020 (0.34)	-0.651 (3.42)	-0.700 (3.52)	-0.660 (3.42)
Learning disability, $X_{3,2}$	0.411 (3.85)	0.805 (4.98)	0.680 (3.83)	0.697 (3.97)
Physical disability, $X_{3,3}$	-0.162 (2.42)	-0.529 (3.87)	-0.506 (3.67)	-0.502 (3.66)
Female, $X_4$	-0.0004 (0.01)	-0.182 (2.78)	-0.160 (2.36)	-0.160 (2.39)
Rural residence, $X_5$	0.125 (1.55)	0.020 (2.09)	0.030 (2.06)	0.029 (2.04)
Community alarm use, $T_2$	-0.098 (2.60)	-0.113 (2.62)	-0.116 (2.63)	0.061 (1.68)
Trend, $t$	-0.005 (4.22)	-0.006 (4.16)	0.018 (1.37)	0.017 (1.37)
Community alarm residuals, $\hat{\epsilon}$		2.610 (3.34)	2.376 (2.97)	2.550 (2.31)
Correction for sample selection, $\lambda$			1.098 (1.77)	1.092 (1.76)
Telecare interacted with its residuals, $T_2\hat{\epsilon}$				-0.271 (0.70)
Number of observations	21,517	21,488	21,488	21,488
Number of homecare clients	10,590	10,578	10,578	10,578
<i>Wald test for weak instruments; <math>\chi^2(p - value)</math></i>		52.21(0.00)		
<i>Wald Chi-square test; <math>\chi^2(p - value)</math></i>	1583(0.00)	1601(0.00)	1620(0.00)	1651(0.00)

The  $Z$  statistics are clustered by Project ID.  $Z$  statistics greater than or equal to 1.96 imply statistical significance at 5% level of significance, whereas the  $Z$  statistics greater than or equal to 1.64 imply statistical significance at 10% level of significance. The reference group for client group is 'Frail elderly'.

From the results presented in Tables (3) and (4) we draw the following conclusions:

- (1) The use of telecare devices is an endogenous determinant of admission to hospital in the emergency department since the coefficients of  $\hat{\epsilon}$  in Columns (2), (3) and (4) are statistically significant at 5% level of significance.
- (2) The empirical model in Table 3 suffers from sample selection bias since the coefficient of  $\lambda$  is statistically significant at 5% level of significance in both Columns (3) and (4).
- (3) There is sufficient evidence at 10% level of significance to suspect the existence of sample selection bias in Table 4.
- (4) The unobservable factors that bring about endogeneity do not lead to unobserved heterogeneity. This is because the coefficients of  $T'\hat{\epsilon}$  in Column 4 of Tables (3) and (4) are not statistically significant at 5% level of significance.
- (5) The observed probability of admission to hospital as an emergency case was in part due to unexplained trend variations since the coefficient of  $t$  is statistically significant at 5% level of significance across all variants of the model. We can, however, also observe from Table 3 that the sign of the coefficient changes from negative to positive once we correct for sample selection bias.
- (6) Community alarm users have a comparatively low probability of admission to hospital in the emergency department, all things equal; whereas the users of the other telecare devices are more likely to be hospitalized as emergency cases than the non-users, controlling for other factors.
- (7) The preferred model in Table 3 is the one presented in Column 3 since it simultaneously controls for the endogeneity of telecare use and sample

selection bias.

- (8) The preferred model in Table 4 is the one that controls for the endogeneity of using a community alarm.

A close look at the results of our preferred models suggests that comorbidity status and age are significant predictors of admission to hospital in the emergency department. Specifically, we can observe that, other factors held constant, the older a particular individual is, the lower the likelihood that the individual was admitted to hospital as an emergency case and the individuals with three or more comorbid conditions are more likely than their counterparts to be emergency cases, controlling for the other independent variables.

An even closer look at the results indicates that the coefficient of the variable for rural residence has a negative sign in Table 3 but a positive sign in Table 4; whereas that of female gender is statistically significant in Table 4 but not in Table 3. Moreover, the results also show that the coefficients of the three categories of the variable for client group are statistically significant in Table 4 but not in Table 3.

We conduct the Wald Chi-square test to determine whether the estimated coefficients are simultaneously equal to zero. The results of this test show that the models in Tables (3) and (4) have sufficient explanatory power. We also test for instrument relevance as explained in Section 3.3. The results of this test ( $\chi^2 = 28.10$  and  $p - value = 0.00$  in Column 3 of Table 3, and  $\chi^2 = 52.21$  and  $p - value = 0.00$  in Column 2 of Table 4) show that our chosen instrumental variables are relevant.

## 5 Discussion

This paper investigates the effect of using telecare devices on unplanned hospitalization in Scotland. Unlike the previous studies in the literature that are mainly randomized controlled trials and quasi-experiments with small samples and high attrition rates, we use large scale non-experimental data to address our study objectives. We also adopt an estimation strategy that simultaneously controls for various methodological challenges that might preclude the estimation of a consistent treatment effect. Because our results contain the treatment effects for both the use of community alarms only and the use of the other telecare devices besides community alarms, we also extend the studies in the literature that do not focus on the impacts of different telecare devices.

The empirical results show that age, client group, rural residence, the use of telecare devices including community alarms and comorbidity status are significant predictors of emergency admission. An interesting point to note, however, is that the effects of some independent variables differ depending on the empirical model being estimated. For example, we find the coefficient of the sex variable to be statistically insignificant at 5% level of significance in the model that controls for rural residence, comorbidity status, client group age, the use of other telecare devices besides the community alarm, sample selection bias and the endogeneity of the treatment variable (see Table 3) but statistically significant when we control for the same covariates and replace the treatment variable with the variable for community alarm use (see Table 4).

The results also show that the treatment effect varies for different telecare

devices. Notice, for example, that the coefficient of the treatment variable in Table 3 is positive (implying that the use of other telecare devices besides the community alarm is associated with an increased likelihood of emergency admission), whereas that of the treatment variable in Table 4 is negative (implying that the use of a community alarm only is associated with a lower likelihood of being admitted to hospital in the emergency department). Based on these results, therefore, policy makers should design policies around the devices that do actually reduce unplanned hospitalization.

### **Acknowledgements**

I would like to acknowledge NHS National Services Scotland for providing the linked administrative health and social care data for this study, and eDRIS (electronic-Data Research and Innovation Service) for providing technical support while conducting the data analysis in the National Safe Haven.

### **Endnotes**

- (1) We test for parameter stability by interacting the covariates of the substantive models with the variable for time trend. The results of this test show that the estimated parameters are stable over time since the coefficients of the interaction terms are not statistically significant at 5% level of significance. The results also suggest that there are no estimation biases that emanate from using the Scottish Morbidity Records to generate time series data since the signs and statistical significance of the other estimated coefficients remain unchanged.
- (2) Throughout the regression analyses in this paper, we include the quadratic

term of age as an additional independent variable in order to control for the potential non-linear effects of age. This follows Grossman's theory of the demand for health capital, where it is hypothesized that an individual's stock of health capital depreciates over time, *ceteris paribus*. Accordingly, we expect the propensity of using health care services so as to improve health status (which in our case is indicated by unplanned hospitalization) to be progressively higher with every unit increase in age, other factors held constant.

- (3) The residuals included in our empirical models are constructed in such a manner as to eliminate the data points that do not contribute to the log-likelihood function. To illustrate this, notice from the first two columns of Tables (3) and (4) that the sample size decreases from 21,517 to 21,488; implying that 29 observations are dropped from the regression analyses. Some of these observations include emergency admissions due to road traffic accidents and admissions to hospital as a result of assault or accidental poisoning at work etc. that may not be explained by our study covariates.
- (4) Recall from Section 3.2 that we control for potential endogeneity of telecare use by solving two equations simultaneously i.e. a reduced form model of the use of telecare devices and a model relating the treatment variable to unplanned hospitalization. One advantage of obviating the bias in this way is that we are able to derive a consistent treatment effect in the presence of simultaneity, which arises from the fact that individuals could use telecare devices in a bid to reduce unplanned admissions to hospital and a high likelihood of hospitalization could also lead to greater use of telecare devices.

## References

- [1] Akematsu Y., Nitta S., Morita K., Tsuji M. (2013). Empirical analysis of the long-term effects of telecare use in Nishi-aizu town, Fukushima, Prefecture, Japan. *Technology & Health Care*, 21(2), 173-182.
- [2] Akematsu, Y. and Tsuji, M. (2012a). Does telecare reduce the number of treatment days? *Studies in Health Technology & Informatics*, 180, 507-511.
- [3] Akematsu Y. and Tsuji, M. (2012b). Measuring the effect of telecare on medical expenditures without bias using the propensity score matching method. *Telemedicine Journal & E-Health*, 18(10), 743-747.
- [4] Amemiya, T. (1980). Selection of regressors. *International Economic Review*, 21, 331-354.
- [5] Andersen, R. (1995). Revisiting the Behavioral Model and Access to Medical care: Does it Matter? *Journal of Health and Social Behavior*, 36(1), 1-10.
- [6] Barnett, K., Mercer, S., Norbury, M., Watt, G., Wyke, S. and Guthrie, B. (2012). Epidemiology of multi-morbidity and implications for health care, research and medical education: a cross sectional study. *The Lancet*, 380(9836), 37-43.
- [7] Billings, J., Dixon, J., Mijanovich, T., Wennberg, D. (2006). Case finding for patients at risk of readmission to hospital: development of an algorithm to identify high risk patients. *British Medical Journal*, 333(7563), 327-333.
- [8] Brownsell, S., Blackburn, S., Hawley, M. (2008). Evaluating the impact of 2nd and 3rd generation telecare services in older people's housing. *Journal of Telemedicine & Telecare*, 14(1), 8-12.
- [9] Cameron, A. and Trivedi, P. (2005). *Microeconometrics: Methods and applications*. Cambridge University Press.

- [10] Cannolly, J., Lowe, D., Anstey, K. (2006). Admissions to hospital with exacerbations of chronic obstructive pulmonary disease: effect of age related factors and service organisation. *Thorax*, 61(10), 843-848.
- [11] Coburn, A., Keith, R., Bolda, E. (2002). The impact of rural residence on multiple hospitalizations in nursing facility residents. *The Gerontologist*, 42(5), 661-666.
- [12] Greene, W. (2002). *Econometric analysis*. Fifth edition. Prentice Hall Publishers.
- [13] Greene, W. (2012). *Econometric analysis*. Seventh edition. Pearson Education, pp. 810-822.
- [14] Grossman, M. (1972). On the concept of health capital and demand for health. *Journal of Political Economy*, 80(2), 223-255.
- [15] Henderson, C., Knapp, M., Fernandez, J., Beecham, J., Hirani, P., Beynon, M., Cartwright, M., Rixon, L., Doll, H., Bower, P., Steventon, A., Rogers, A., Fitzpatrick, R., Barlow, J., Bardsley, M., Newman, S. (2014). Cost-effectiveness of telecare for people with social care needs: The whole systems demonstrator cluster randomized trial. *Age & Ageing*, 43(6), 794-800.
- [16] Hirani SP, Beynon M, Cartwright M, Rixon L, Doll H, Henderson, C., Bower, P., Steventon, A., Rogers, A., Fitzpatrick, R., Barlow, J., Bardsley, M., Newman, S. (2013). The effect of telecare on the quality of life and psychological well-being of elderly recipients of social care over a 12-month period: The whole systems demonstrator cluster randomized trial. *Age & Ageing*, 43(3), 334-341.
- [17] Long, J.S. and Freese, J. (2006). *Regression models for categorical dependent variables using Stata*. Second edition. College Station, TX: Stata press.
- [18] Lyon, D., Lancaster, G., Taylor, S., Dowrick, C., Chellaswamy, H., (2007). Predicting the likelihood of emergency admission to hospital of older people:



development and validation of the Emergency Admission Risk Likelihood Index (EARLI). *Family Practice*, 24(2), 158-167.

- [19] McCullagh, P., Nelder, J. (1999). *Generalized Linear Models: Monographs on statistics and applied probability*. Second edition. Chapman and Hall.
- [20] McDonald, J. (2014). *Handbook of Biological Statistics*. Third edition. Sparky House Publishing.
- [21] Murray, M. (2006). Avoiding invalid instruments and coping with weak instruments. *Journal of Economic Perspectives*, 20(4), 111-132.
- [22] National Records of Scotland. (2015). *Age-standardised Death rates calculated using the European standard population*. Available at <http://www.nrscotland.gov.uk> [Accessed 29th February 2016].
- [23] Pokropek, A. (2016). Introduction to instrumental variables and their application to large scale assessment data. *Large-Scale Assessments in Education*, 4(4), 1-20.
- [24] Santolino, M., Bolance, C., Alcaniz, M. (2011). Factors affecting hospital admission and recovery stay duration of inpatient motor victims in Spain. *Research Institute of Applied Economics*, 19, 1-27.
- [25] Shelton, P., Sager, M., Schrader, C. (2000). The community assessment risk screen (CARS): identifying elderly persons at risk for hospitalization or emergency department visit. *American Journal of Managed Care*, 6(8), 925-933.
- [26] Steventon, A., Bardsley, M., Billings, J., Dixon, J., Doll, H., Cartwright, M., Bower, P., Rogers, A., Fitzpatrick, R., Barlow, J., Bardsley, M., Newman, S. (2013). Effects of telecare on use of health and social services; findings from the whole systems demonstrator cluster randomized trial. *Age & Ageing*, 42(4), 501-508.

- [27] Terza, J., Basu, A. and Rathouz, P. (2008). Two-stage residual inclusion estimation: addressing endogeneity in Health Econometric Modeling. *Journal of Health Economics*, 27(3), 531-543.
- [28] Verdaguer, A., Peiro, S., Libero, J. (2003). Variations in the use of hospital resources in treating patients with chronic obstructive pulmonary disease. *Arch Bronconeumol*, 39(10), 442-448.
- [29] Winner, L. (2004). Introduction to biostatistics. *Lecture Notes, University of Florida*, 1-204.
- [30] Wolff, J., Starfield, B. and Anderson, G. (2002). Prevalence, expenditures and complications of multiple chronic conditions in the elderly. *Arch Intern Med*, 162(20), 2269-2276.
- [31] Wooldridge, J. (2002). *Econometric analysis of cross-section and panel data*. MIT Press.