DEPARTMENT OF FOOD AND RESOURCE ECONOMICS
UNIVERSITY OF COPENHAGEN

IFRO Working Paper

Can pecuniary and environmental incentives via SMS messaging make households adjust their intra-day electricity demand to a fluctuating production?

Niels Framroze Møller
Laura Mørch Andersen
Lars Gårn Hansen
Carsten Lynge Jensen

2018 / 06

IFRO Working Paper 2018 / 06

Can pecuniary and environmental incentives via SMS messaging make households adjust their intra-day electricity demand to a fluctuating production?

Authors: Niels Framroze Møller, Laura Mørch Andersen, Lars Gårn Hansen, Carsten Lynge Jensen

JEL-classification: C2, C22, C5, C55, Q4

Published: May 2018

See the full series IFRO Working Paper here:
www.ifro.ku.dk/english/publications/ifro_series/working_papers/

Department of Food and Resource Economics (IFRO)
University of Copenhagen
Rolighedsvej 25
DK 1958 Frederiksberg DENMARK
www.ifro.ku.dk/english/

Can pecuniary and environmental incentives via SMS messaging make households adjust their intra-day electricity demand to a fluctuating production?

Niels Framroze Møller*, Laura Mørch Andersen†,

Lars Gårn Hansen†, Carsten Lynge Jensen†

* Management Engineering, Technical University of Denmark

†Department of Food and Resource Economics, University of Copenhagen

Abstract

The increasing deployment of renewables introduces substantial variability into the production of electricity, requiring demand to be more movable across time. We analyze data from a large Danish field experiment (2015-2016) to investigate whether households can be prompted, via SMS messages, to move electricity consumption, and if so, whether these are motivated by pecuniary or environmental motives. To take heterogeneity fully into account we first use *general-to-specific-based* automatic model selection which allows for a different time-series regression for each of the 1488 households studied. From this we obtain a cross-section of estimated SMS effects which we then regress on the motive type. Since households can opt out there is a risk of self-selection. We therefore control for the size, income and average consumption of the household, and the age, educational- and labor market status of the SMS recipient. The results suggest that SMS messages can to some extent motivate households to move consumption. A stronger financial motive seems more effective, whereas a purely environmental motive actually reduces the displaced amount. However, mixing financial and environmental motives seems the most effective. Finally, women and elderly people are more inclined to move consumption.

1 Introduction

The integration of fluctuating renewable energy sources, like wind, solar and wave, in the production of electricity, means that supply will vary much more than previously, when production came from controllable oil- and coal-fired power plants. Depending on the extent to which electricity can be stored in an economic sustainable way, this implies that demand has to become more flexible, i.e. movable across time, in order to balance supply at a given point in time.[1]

In this paper we consider the electricity demand of households. We analyze data from a large Danish field experiment conducted in the period 2015 (June) - 2016 (June), in order to investigate whether SMS messages can induce households to shift their consumption along the time axis.[2] In particular, within the same day. We also identify which particular motives, i.e. purely financial, purely environmental, or a mix of these, are the most effective. The data set for each participating household contains time series of electricity consumption in 3-hour periods and information about SMS messages, suggesting to move consumption into or away from these periods. We will focus on the time periods 10-13 (noon) and 18-21 (evening).

The econometric analysis has two stages: First, for each household we estimate a dynamic time-series regression model that includes the SMS indicator variables. A key point to note is that the models are allowed to differ across households (i.e. in terms of dynamics and deterministic components). In the second stage, we then regress the estimated SMS effects on motives and a set of socioeconomic and demographic variables. The latter are included to reduce the self-selection bias on the estimated effect of motives, but also to throw light on which types of households are more responsive than others, which can be important for the targeting of SMS messages.

The *potentially different* modeling of the large number of individual time series is practically feasible if based on automated general-to-specific (GETS) model selection al-

[1]Flexibility on the demand side, often referred to as, *demand response*, is an extensively studied field. See e.g. Caves, Christensen, and Herriges (1984), Faruqui and George (2002), King and Chatterjee (2003), Faruqui and George (2005), and for a more recent survey, Heshmati (2014).

[2]The experiment is described below and in Andersen, Hansen, Jensen, and Wolak (2017) in detail.

gorithms (see e.g. Hoover and Perez 1999, Hendry and Krolzig 2005, Doornik 2009, Hendry and Doornik 2014). A main advantage of this approach is that it avoids imposing homogeneity restrictions across the models for the individual households, as is standard in panel analyses. Here, we use the GETS algorithms for dynamic regression models as recently implemented in the software, R (see Pretis, Reade, and Sucarrat 2016). Starting from a common broadly specified regression model which has a good chance of nesting the model of any household, these algorithms can be used to remove insignificant regressors, as well as to detect and account for both outliers and temporary and permanent level shifts in the individual time series, based on indicator saturation, see Hendry, Johansen, and Santos (2008) and Johansen and Nielsen (2009).[3] The latter is particularly useful in the present context in that such level shifts vary across households with respect to number, location and duration.

The data from the experiment has been analyzed recently in Andersen et al. (2017). There, results from the experiment are reported focusing on the difference in consumption shifts into target hours relative to consumption shifts away from target hours. That paper also finds that the incentives to shift consumption into a set of target hours yields reduced consumption in the hours of the day that surround these target hours. The contribution of the present paper relative to Andersen et al. (2017) is twofold. First, here we use automatic model selection algorithms to take into account the individual time-series characteristics, by allowing for dynamic correlations, breaks and outliers in the individual series. Second, the regression on motives, socioeconomic and demographic explanatory variables in the second step of the analysis, also allows us to assess what these variables imply for the willingness to move consumption.

There is a large literature studying the impact of economic incentives on household-level energy consumption. This includes analyses of the effect of time-varying electricity prices and feedback information via in-house displays, mail, home pages, SMS and e-mail messages (see e.g. Faruqui and Malko 1983, Fischer 2008, Faruqui and Sergici 2010, Newsham and Bowker 2010, Faruqui, Sergici, and Sharif 2010, Delmas, Fischlein, and

[3]See Castle, Doornik, Hendry, and Pretis (2015) for an application.

Asensio 2013 and Vine, Buys, and Morris 2013, which together survey a vast amount of experiments and studies).[4] However, to the best of our knowledge the present research is the first to take fully into account the heterogeneity (across households) in consumption patterns despite the large number of time series, by use of automatic model selection algorithms. In addition, although not specifically designed for this type of analysis the data from the experiment provide a unique opportunity to gain insights into the potential for demand shifting via different incentives sent by SMS messages. In particular, this experiment is the first to consider how SMS messages can induce consumption shifts with only a few hours warning.

From an applied methodological perspective, we believe that the present research demonstrates a large potential of GETS automatic modeling as a means of analyzing large collections of time series for which modeling the individual time series *differently* is of interest. This is becoming increasingly relevant with the prevalence of Big Data. In fact, extracting the value of Big (time series) Data often lies in the detection of the differences between the individual time series. Examples of this relate to consumer profiling and segmentation in general. Moreover, related to household-level energy consumption in particular, this could be segmentation based on smart meter data.[5] Within that field, automatic modelling used carefully may greatly assist utility companies in grouping customers according to differences in consumption patterns as estimated by a dynamic regression model, say. This can then be used to tailor time-of-use tariffs and dynamic pricing prompting consumption in off-peak hours which is becoming increasingly important in economies with fluctuating renewable-based power production.

The remainder of the paper is organized as follows. In the next section we briefly describe the field experiment and the data that we use from this. Section 3 contains the first part of the empirical analysis which concerns the automatic modeling of the time series corresponding to the individual households. In Section 4 we then regress the estimated SMS effects from the first step on the motive type and the control variables for

[4] A study in this literature analyzing Danish data is Gleerup, Larsen, Leth-Petersen, and Togeby (2010), who study the impact on total household electricity consumption of feedback by SMS and e-mail messages.

[5] A recent survey of the literature on smart meter segmentation is found in Tureczek and Nielsen (2017)

demographic and socioeconomic factors. Finally, we sum up and conclude in Section 5.

2 The field experiment and the data

The field experiment is described in detail in Andersen et al. (2017) which also contains an account of the exact contents of the e-mail texts. Here, we just provide a brief description of this and the associated time-series data that we analyze. In the experiment there are two artificial "peak periods" both of which are three-hour periods. A "high peak" for which consumers are encouraged by an SMS message to move their consumption away from, and a "low peak" period in which consumers are encouraged to place their consumption. Households receive the SMS messages some hours in advance and this includes a short text with a given economic and/or environmental motive. The set of households we analyze can be divided into eleven groups according to the particular contents of the SMS text they receive. Of the eleven groups, three groups contain the households that receive purely financial motives, with different rates of discounts across these three groups. Two of the groups concern purely environmental motives, while six groups contain the households who have recieved some variant of a mixed financial and environmental motive (see below). Together, the eleven groups comprise 1488 individual households. For a given individual household the SMS text remains the same throughout the experimental period.

To ensure comparability across different households, initially it was necessary to exclude a minor share of these: First, households with an installed solar power system were not included, as their power meter does not show their total amount of consumed kWhs. Secondly, households with either extremely low or extremely high average consumption levels were also excluded as these most likely do not correspond to an actual household. In particular, a very low level most likely corresponds to a single appliance whereas an unusually high level probably reflects some sort of commercial or professional activity. Third, a negligible share of the participants moved during the experiment period while others changed their mobile number or their E-mail address and both of these groups were discarded. Fourth, a smaller part of the households accepted to receive an electronic device in the beginning of the experiment, which enables remote control of their consumption. Finally, only few time-series observations were missing for some of the households and hence these were therefore simply interpolated. Table 1 contains a description of the eleven groups in terms of the particular motive and the number of households, after exclusion according to the above-mentioned criteria. Each group is identified by a combination of letters and numbers, describing the motive for the households belonging to this group.

Table 1: A description of the 11 different motive groups of the participating households in the Danish field experiment.

Motive	**Financial**			**Mixed**						**Environmental**	
Group	f5	f20	f50	fe5	fe20	fe50	ef5	ef20	ef50	eA	eB
#	276	154	100	167	87	51	159	83	50	246	115

Notes: f5 denotes a financial (f) motive of 5% discount etc. Mixed financial and environmental (e) motives are denoted ef5 and fe5 etc., where the order of e and f indicates which of the motives is mentioned first in the text message. Purely environmental motive groups are denoted eA and eB. The number of households (#) in the groups are those that result after excluding households with solar power systems, extreme consumption levels, those that moved during the experiment, and finally those that received the automatic device.

In particular, f5 denotes a financial (f) motive of 5 percent discount etc. Mixed financial and environmental (e) motives are denoted ef5 and fe5 etc., where the order of e and f indicates which of the motives is mentioned first in the text message. Finally, the two groups eA and eB receive purely environmental motives.[6]

3 Individual time-series models for 1488 households

Since the number of time-series observations for each household is sufficiently high, we fit a dynamic regression equation for *each* of these, allowing us to take heterogeneity explicitly into account.[7] Given the many different households we study, this approach may at first sight seem infeasible or at least excessively cumbersome. That is, for different households it is most likely that we need different models with respect to regressor set, e.g. SMS indicators and lagged consumption. Moreover, obtaining a well-specified

[6] The difference between the latter two groups is minor and relates only to the graphical layout of the invitation sent to these households.

[7] All households studied have at least 100 observations. Note also, that by "sufficiently high" we mean (loosely), high enough in order for the asymptotic approximative inference to work satisfactory.

statistical model also requires taking account of level shifts and outliers, which also vary across households, with respect to number and location. As a result, in practice it is far too time consuming to manually identify a statistically valid model specification for each household. This is of course exacerbated as there will often be several regressors which are highly correlated (particularly in dynamic models), implying that the order in which insignificant regressors are removed matters. However, modeling of large number of individual time series can be facilitated by using automated general-to-specific GETS model selection algorithms (Doornik 2009, Hendry and Doornik 2014). Such algorithms allow us to start with many candidate regressors in order to *search* for a parsimonious encompassing representation of the local data generation process. Studies have shown that the cost of searching is often relatively low, meaning that we can expect to retain only few irrelevant regressors (see e.g. Hendry 2009). Moreover, as these algorithms can handle more regressors than observations, indicator saturation may be used to detect outliers and level shifts (see Hendry, Johansen, and Santos 2008 and Johansen and Nielsen 2009). In fact, the automatic detection of level shifts is indispensable since such shifts are typically caused by holiday spells which vary between households with respect to number, location and duration. Here, we use the GETS modeling tools for dynamic regression models as recently implemented in R (see Pretis, Reade, and Sucarrat 2016). These tools allow us to use impulse- and step-indicator saturation in order to detect and account for outliers and level shifts in the individual time series, respectively (see Castle, Doornik, Hendry, and Pretis 2015).

Our approach is to loop over all households and in each iteration, i.e. for each household, there are two steps: First, we start from a sufficiently general dynamic regression model which is saturated with a full set of indicators for level shifts and impulses. This model serves as a common point of departure for all households and has a good chance of nesting the regression model corresponding to each of these. We run the automatic algorithm on the saturated model, i.e. the *isat-algorithm* to detect and insert the household-specific indicators. In this way we end up with a General Unrestricted Model (GUM) for each household which we require to pass the specified model diagnostics.[8] In the second step, we apply the *getsm-algorithm* to simplify the model further (see Pretis, Reade, and Sucarrat 2016). Reflecting heterogeneity, the final estimated model will then vary across consumers with respect to the regressor set, i.e. lags of consumption, the various SMS indicators as well as the number and location of impulse- and step-indicators.

Below we first describe the form of the GUM that results in the first step. We then describe and motivate the particular choices of regressors for the GUMs used in the estimation for the noon (10-13) and evening (18-21) consumption.

3.1 The model for household i

For household i, the GUM that results in the first step is typically household-specific (depends on i), in that it includes the set of detected impulse- and step-indicators for the particular household. These are needed for almost all households to maintain the error term assumptions. The GUM has the form,

$$y_{i,t} = \delta_{i,t} + \alpha_i(L)y_{i,t} + \boldsymbol{\beta}'_i(L)\mathbf{z}_{i,t} + \nu_{i,t}, \tag{1}$$

for $t = 1, ..., T_i$, and where $y_{i,t}$ is the logarithm of electricity consumption for household i in a given 3-hour period of day t. Below we consider the two periods 10-13, referred to as *noon*, and 18-21, referred to as *evening*. The intercept, $\delta_{i,t}$, is given by, $\delta_{i,t} \equiv \boldsymbol{\phi}'_i\mathbf{d}_{i,t}$ where $\mathbf{d}_{i,t}$ $(h_i \times 1)$ vector of deterministic terms and $\boldsymbol{\phi}_i$ $(h_i \times 1)$ the corresponding coefficients. The deterministic terms comprise a constant term, a sine-cosine term, which accounts for the annual cyclic influences, e.g. due to lighting and heating, and indicators for distinguishing between weekends and work days. In addition, $\delta_{i,t}$ includes the detected impulse- and step-indicators. Note that, just as the indicator for weekends, which obviously is common across household, one could likewise construct an indicator for holiday periods, in particular, for the summer holidays. However, this is left to the automatic step-indicator saturation to detect, which seems reasonable as different consumers spend

[8] We mainly focus on the assumption of uncorrelated and non-ARCH error terms. As shown below, for around 90% of the households studied, we are able to obtain a GUM that passes the diagnostics.

their holidays in different periods.

The auto-regressive polynomial is, $\alpha_i(L) \equiv \alpha_{1,i}L + \alpha_{2,i}L^2 + ... + \alpha_{P_i,i}L^{P_i}$, with L being the lag- or backshift operator, i.e. $Ly_{i,t} \equiv y_{i,t-1}$ etc. In the initial GUM, $P_i = 7$, for all i, but in the estimated regression for household i that we end up with in the second step, P_i generally varies across i. Note that, differences in lag length P_i reflect individual differences in consumption patterns. For example, some households display a more pronounced weekly pattern, which may be described by $\alpha_{7,i} \neq 0$ but with most of the remaining $\alpha_{\cdot,i} = 0$, whereas the consumption behavior of other households is more unsystematic, corresponding to all $\alpha_{\cdot,i} = 0$ (no dynamic dependence).

Exogenous regressors are included in the $K_i \times 1$ vector, $\mathbf{z}_{i,t}$. In the initial GUM, the regressor set is the same for all i, but again will vary across i in the model we end up with. The set comprises indicators for the receipt of an SMS message and electricity consumption from the preceding hours from the same day and from previous days. Table 2 contains an overview of the particular regressors included in $\mathbf{z}_{i,t}$, for the regressions corresponding to the noon and evening consumption, respectively.

Table 2: The regressors included in the vector, $\mathbf{z}_{i,t}$, in the regressions corresponding to consumption in the noon (10-13) and evening (18-21), respectively.

Response variable:	
Noon consumption (10-13)	Evening consumption (18-21)
Regressors in $\mathbf{z}_{i,t}$:	
SMS indicators (current period) IN10, AW10 SMS indicators (later periods) IN15, AW15, IN18, AW18, IN21, AW21	SMS indicators (current period) IN18, AW18 SMS indicators (later periods) IN21, AW21
Consumption from: - all previous hours from current day - all hours outside 10-13 in the last 7 days	Consumption from: - all previous hours from current day - all hours outside 18-21 in the last 7 days

Notes: A regressor, e.g. IN10 (AW10) is an indicator, being 1 if the household has received a text message suggesting to move consumption into (away from) the time interval beginning at 10 o'clock, and so forth. For indicators which refer to later treatment periods than the one that the regression corresponds to, we include only those for which the text message has been received early enough.

The regressor INXX (AWXX) denotes an indicator, being 1 if the household has received a text message suggesting to move consumption into (away) the time interval beginning at XX o'clock. Generally speaking, one may initially include all potentially relevant available variables known to the respondent in the beginning of the particular time interval. However, we have kept the specifications relatively simple. Nevertheless, besides the indicators for SMS messages corresponding to the time interval studied, we have included SMS indicators corresponding to the subsequent/later time intervals, provided that the consumer has been informed about this, before the beginning of the particular time interval under study. For example, in the analysis of consumption in the period 18-21, we not only include AW18 and IN18. In particular, if the consumer has received a text message before 18 o'clock, suggesting to change consumption in the period 21-24 the same day, this will also enter. Including treatments for later periods of the same day allows us to gain insight into whether or not consumers shift consumption across time periods without changing the overall consumption. This distinction is important if the goal is to induce consumers to reduce consumption in peak periods. On the one hand the consumer may reduce consumption in some peak period in response to a message, but not compensate by increasing consumption correspondingly in another time period the same day. On the other, if consumption is increased (replaced) by the same amount in the interval, the peak problem may in principle have been pushed to another peak period and effectively not resolved.

For the regressor set we also considered the following: First, as a number of errors occurred during the experiment in connection with sending/receiving the SMS messages, we considered indicators for this. However, as these errors are transitory in nature and few in number they seem of minor importance. In any case, if they do in fact have a role to play this will be picked up by the indicator saturation in the first step. Second, some of the text messages were also followed by a reminder while others were not. To assess whether reminding had any impact we chose initially to distinguish between SMS messages that were reminded and not reminded. In particular, initially we performed a statistical test of the hypothesis that the effect of an SMS which is not reminded is the same as that which is reminded. Since this could not be rejected, we chose to abandon this distinction. Finally, yet another regressor we considered is the time difference between peak time, and the time when the consumer receives the text message. Intuitively, the longer this time period is the more time to respond. However, we have not attempted this since, for many consumers there is little variation in this variable over the sample period. As a result it is often close to being proportional to the respective treatment dummy, creating a high degree of collinearity.

3.2 The estimations for consumption in the noon and evening

To get an overview of the estimation results corresponding to the many households, the tables in this section provide a rough summary characterization of the distributions of estimates of the INTO and AWAY SMS indicators for each group/motive. Tables 3 and 4 contain the results from the regressions of consumption in the noon period (10-13). Tables 5 and 6, which have a similar structure, contain the corresponding results for consumption in the evening period (18-21).

Starting from the left in Table 3 the first column shows the particular group. The notation for the groups was described in Table 1. Corresponding to the two sections of the table, IN10 and AW10, consider the former as the latter has a similar structure: The column "pos" shows the percentage share of households that have a coefficient estimate on IN10 which is positive. The columns SpXX (SnXX in the case of AW18) show the percentage share of households that have a significantly positive (negative) coefficient at the XX% level. Finally, for each group, the numbers, n_I and n_F, in the last column, denote the number of households available for *initial* estimation (i.e. remaining after cleaning the data as described in Section 2). Of these, it was possible to reach a *final* model in n_F cases, which is based on a GUM, that passes the chosen model diagnostics. Note that, for all groups the automatic algorithm was capable of identifying well-specified GUMs for the vast majority of households.

Table 3: A summary description of the distributions (in terms of the percentage share of households) of the estimated coefficients on IN10 and AW10 in the regression of the consumption in the noon period (10-13).

Group	**IN10**				**AW10**				
	pos	$Sp25$	$Sp10$	$Sp5$	*neg*	$Sn25$	$Sn10$	$Sn5$	n_I/n_F
f5	65.9	45.1	31.8	27.8	56.9	34.9	14.1	9.0	276/255
f20	67.6	55.2	35.9	26.2	59.3	33.8	19.3	13.1	154/145
f50	71.7	53.3	39.1	32.6	62.0	35.9	9.8	7.6	100/92
fe5	66.7	45.8	29.4	27.5	64.7	37.3	20.9	16.3	167/153
fe20	83.1	58.4	45.5	39.0	61.0	32.5	19.5	7.8	87/77
fe50	72.9	60.4	43.8	39.6	70.8	50.0	25.0	14.6	51/48
ef5	63.1	50.4	36.2	28.4	57.4	34.0	14.2	9.2	159/141
ef20	70.9	51.9	34.2	30.4	57.0	29.1	8.9	6.3	83/79
ef50	72.3	55.3	53.2	40.4	70.2	53.2	29.8	14.9	50/47
eA	64.9	47.3	32.4	27.0	56.3	34.7	13.1	5.9	246/222
eB	60.7	39.3	23.4	16.8	55.1	29.0	15.9	9.4	115/107

Notes: In the section IN10 (AW10 has a similar structure), the column "pos" shows the percentage share of households that have a *positive* coefficient estimate on IN10. The columns SpXX (SnXX in the case of AW10) show the percentage share of households with a *significantly positive* (negative) coefficient at the $XX\%$ level. The notation for the groups was described in Table 1. For each group, the number n_I denotes # households, *initially* available for estimation (i.e. the data after having been cleaned cf. the criteria in Section 2). n_F is # households for which it was possible to obtain a GUM passing the chosen diagnostics.

Consider the results in Table 3. For IN10 we expect a positive impact. As appears from the column "pos", we see that for a majority of the households, ranging from 60.7% to 83.1%, depending on the different groups, this is the case. However, one has to remember that under the null hypothesis of no effect there would supposedly be around 50% positively signed on average. In terms of statistical significance, consider the column Sp5 which shows the share of households with a positively significant effect at the 5% level. The range across groups is 16.8%-40.4%. Again, these numbers have to be viewed in light of around 5% under a null of no effects. For some households the amount of electricity consumption that is flexible, i.e. can be shifted across time, may be small in percentage terms. This implies that, a significance level of 5% may not always be entirely reasonable, as the test of the null hypothesis can be expected to suffer from low power against positive alternatives. The table therefore also contains the columns Sp10 and Sp25, which accordingly show the shares with significant positive coefficients at the 10%

and 25% levels, respectively.

When going from one row to another the motive and/or the rate of discount changes. Although an overall impression from the table may be that a stronger financial motive is slightly more effective, we emphasize that we should not make such comparisons until we have conditioned on other relevant variables that may change when going from one motive group to another. This is a main purpose of the next section. The problem is essentially that of self-selection which can be expected to be inevitable given that the participation is voluntary.

Consider AW10 in which case we expect a negative sign. As we see, the shares of households with a negative sign lie in the range of 55.1%-70.8%. Hence, the support to the expected sign is less pronounced compared with the positive effect of IN10. Hence, at a glance, the evidence seems to suggest, albeit vaguely, that households are more willing to move their consumption into a given time interval as opposed to moving consumption away from this (see also Andersen et al. 2017). However, an alternative plausible explanation of this is that, in the former case, the potential amount of consumption to be shifted is expected to be larger than in the latter case. For example, in the former case, a household can move all planned consumption from the whole day after the interval 10-13 and into this time interval. In contrast, if consumption is to be moved away from this time interval, the potential is less, due to the simple fact that the interval is shorter. However, note that the conclusion eventually depends on the shape of the consumption profile over the day.

As explained in Section 3.1, in the regression of the consumption in the noon period we also included the SMS indicator variables corresponding to the time intervals during the remainder of the day. The summary description of the distribution of the t-values is given in Table 4. To save space only the pos/neg column and the 5% column is reported. The table is otherwise similar in structure as Table 3.

Table 4: A summary description of the distributions (in terms of the percentage share of households) of the estimated coefficients on AW15, IN15, AW18, IN18, AW21 and IN21 in the regression of the consumption in the noon period (10-13).

group	**AW15**		**IN15**		**AW18**		**IN18**		**AW21**		**IN21**	
	pos	*Sp*5	*neg*	*Sn*5	*pos*	*Sp*5	*neg*	*Sn*5	*pos*	*Sp*5	*neg*	*Sn*5
f5	49.8	10.2	59.2	11.4	46.3	9.4	58.0	9.4	36.1	5.5	55.3	7.8
f20	46.2	9.7	62.1	7.6	46.2	10.3	66.9	8.3	45.5	8.3	53.8	6.9
f50	56.5	10.9	59.8	8.7	51.1	10.9	55.4	17.4	31.5	7.6	52.2	4.4
fe5	53.6	10.5	54.9	7.2	37.9	8.5	50.3	2.6	37.3	7.2	49.0	7.2
fe20	50.6	3.9	50.6	7.8	37.7	5.2	53.2	10.4	54.5	11.7	50.6	3.9
fe50	47.9	14.6	62.5	12.5	50.0	12.5	50.0	10.4	25.0	0.0	68.8	2.1
ef5	47.5	9.2	53.9	7.1	49.6	6.4	58.2	6.4	34.0	8.5	58.2	7.8
ef20	55.7	8.9	57.0	11.4	46.8	7.6	55.7	6.3	48.1	7.6	60.8	11.4
ef50	68.1	10.6	68.1	10.6	31.9	10.6	55.3	4.3	40.4	10.6	48.9	8.5
eA	57.2	11.7	56.3	7.2	44.6	8.1	55.9	7.7	34.7	6.8	55.9	5.9
eB	51.4	12.2	60.7	7.5	49.5	10.3	54.2	4.7	40.2	8.4	45.8	8.4

Notes: For notational details see Tables 1 and 3. The SMS indicators corresponding to these time intervals which come after 10-13 have only been included provided that the corresponding SMS has been received before 10 o'clock.

The coefficients on the AWAY indicators, corresponding to the later intervals, 15-18, 18-21 and 21-24, are expected to be non-negative, whereas those on the INTO indicators are expected to reduce consumption in the interval 10-13 and hence to be non-positive. Recalling that we include only these indicators if the corresponding SMS was sent to the respondent before 10 o'clock, it appears for example that 59,8% of the households in group f50 have a negative coefficient estimate on IN15 indicator in the regression of consumption in 10-13. Although the numbers in the table are not so pronounced as in Table 3, there seems to be a slight indication that households are more willing to reduce their consumption in 10-13 when an INTO SMS for the later periods is received, whereas they are less inclined increase it as a result of receiving an AWAY SMS corresponding later periods of the day. This suggests that households are on average a bit more willing to postpone consumption relative to the opposite of moving planned consumption closer to the present.

Consider now the estimation for the evening period 18-21.

Table 5: A summary description of the distributions (in terms of the percentage share of households) of the estimated coefficients on IN18 and AW18 in the regression of the consumption in the evening period, 18-21.

Group	**IN18**				**AW18**				
	pos	*Sp*25	*Sp*10	*Sp*5	*neg*	*Sn*25	*Sn*10	*Sn*5	n_I/n_F
f5	67.1	44.2	27.9	22.5	60.5	36.8	21.7	13.6	276/258
f20	68.8	50.4	33.3	29.1	63.1	39.0	19.9	14.9	154/141
f50	71.0	50.5	39.8	33.3	60.2	30.1	12.9	7.5	100/93
fe5	66.7	39.2	28.1	20.9	60.8	38.6	19.6	15.0	167/153
fe20	70.0	48.8	30.0	21.2	60.0	35.0	18.8	11.3	87/80
fe50	83.0	70.2	53.2	48.9	61.7	44.7	21.3	4.3	51/47
ef5	65.0	47.6	29.4	23.8	57.3	32.2	18.9	11.9	159/143
ef20	64.0	52.0	40.0	34.7	50.7	30.7	10.7	8.0	83/75
ef50	68.9	53.3	42.2	35.6	66.7	46.7	20.0	13.3	50/45
eA	62.6	42.0	29.2	21.0	53.0	32.9	17.8	10.5	246/219
eB	64.4	45.2	26.9	16.3	63.5	37.5	20.2	11.5	115/104

Notes: The table has a similar structure as that in Table 3.

For IN18 we expect a positive impact. As appears from the column "pos" in the table, we see that for a majority ranging from 62.6% to 83.0% of the households depending on the different groups, this is the case. In terms of statistical significance, consider the column Sp5 which shows the share of households with a positive coefficient which is significant at the 5% level. The range across groups is 16.3%-48.9% (again 5% under a null of no effects).

Consider AW18 in which case we expect a negative sign. As we see, the shares of households with a negative sign lie in the range of 50.7%-66.7%. Hence, the support to the expected sign is a little less pronounced compared with the positive effect of IN18, resembling the findings for the noon period consumption.

Turning to the estimates of the coefficients on IN21 and AW21 (Table 6), there can be arguments both in favor of a positive sign as well as a negative sign. Starting with IN21, it would perhaps be the most natural to expect a negative sign. In particular, it seems reasonable to expect that, as long as households are informed before 18 o'clock, they can move their consumption into the time interval 21-24, by reducing their consumption in the time interval 18-21. On the other hand, one could also imagine that households would shift some of their consumption during the day time to a time interval which is not

21-24 exactly, but rather overlapping with the time interval 18-21 also. In such a case, consumption would increase in the latter time interval.

Table 6: A summary description of the distributions (in terms of the percentage share of households) of the estimated coefficients on AW21 and IN21 in the regression of the consumption in the evening period, 18-21.

Group	**AW21**				**IN21**			
	pos	$Sp25$	$Sp10$	$Sp5$	*neg*	$Sn25$	$Sn10$	$Sn5$
f5	51.2	28.7	15.1	9.3	53.5	31.0	14.7	10.1
f20	48.9	27.7	12.8	9.2	56.7	32.6	14.9	7.1
f50	48.4	34.4	16.1	12.9	49.5	23.7	11.8	4.3
fe5	56.9	29.4	11.8	5.9	58.8	32.7	13.1	6.5
fe20	40.0	18.8	12.5	7.5	61.2	42.5	26.2	10.0
fe50	46.8	25.5	14.9	8.5	63.8	36.2	12.8	6.4
ef5	45.5	21.0	11.9	7.7	55.2	32.9	14.7	9.1
ef20	57.3	29.3	16.0	9.3	56.0	26.7	13.3	9.3
ef50	42.2	22.2	13.3	11.1	48.9	35.6	24.4	11.1
eA	41.1	21.5	10.5	8.7	56.6	36.5	16.9	8.7
eB	43.3	20.2	9.6	5.8	57.7	31.7	17.3	6.7

Notes: For details see Table 4.

As Table 6 (column "neg") shows, there is a mild indication that on average, IN21 will decrease consumption in the time interval 18-21. Concerning AW21, there is no indication that an AW21 SMS will increase consumption in 18-21. Hence, these results also mimic those corresponding to the noon consumption.

4 The regression of estimates on motives and socioeconomic and demographic controls

In this section we regress the estimated SMS coefficients against motives and socioeconomic and demographic explanatory variables. We will focus on the regression of the IN10 and AW10 for the noon interval and IN18 and AW18 for the evening interval, implying four different regressions. We have two main purposes of this section. First, we want to investigate whether differences in motives have an effect on the propensity to move consumption across time periods. Here, we want to investigate which of the motive types, i.e. environmental, financial, or a mix of these, is the most effective. Since participation in the experiment is voluntary, there is a risk of a self-selection bias. In particular, the participating households know the type of the particular motive that they will receive already at the invitation, and they always have the possibility to opt out. For example, it could be the case that people with higher education/incomes or a different political standpoint are overrepresented in certain groups as opposed to others. Hence, the motive can potentially be correlated with such variables and we should therefore attempt to reduce the bias by including the ones that are observable. Second, the influences from these explanatory variables are also of interest per se: For example, we may want to assess the influence on the propensity to move consumption of the age, gender, labor market status, education, income of the respondent, for the purpose of better targeting SMS messages in the future.

As explanatory variables, we include first of all, the indicator variables for the different motive groups. DXX denotes such indicator when the observation corresponds to a household in group XX. The benchmark category is chosen to be group f5, i.e. the purely financial motive with the lowest discount (5%). Secondly, we include the gender (equals 1 if female), the age and the labor market status of the recipient of the text message. The indicator for labor market status is equal to 1 for retirees, long-term unemployed (i.e. more than 6 months), early retirees, cash beneficiaries, national pensioners, persons receiving sickness benefits, education allowances etc. Accordingly, this indicator therefore equals zero if the recipient is employed more than half-time, including self-employment. It is included to take into account whether or not the recipient has a higher chance of being at home when receiving the SMS message, and can thus manually change consumption. We also include the educational level of the recipient, as measured by the highest completed education. In particular, we graded education into 6 levels or categories corresponding to the variables edu1 through edu6, (see Table 7), where edu1 is the lowest and edu6 is the highest level. We include edu2 through edu6 so that edu1, corresponding to primary school, is the benchmark category. This variable is included as a rough proxy for unobserved variables such as knowledge and understanding of environmental problems, intelligence, political standpoint etc. which may have an influence on the willingness to move electricity consumption across time periods. We also included household income as well as the total number of persons in the household. Finally, since some of the included variables, e.g. household size/income and labor market status, but also the motive, may correlate strongly with the general level of electricity consumption for the household, we included the average level of consumption for the time interval in question.

Table 7 contains the estimates and associated t-values for the regression of the estimated coefficient on IN10 on the explanatory variables.

Table 7: OLS regression of the IN10 coefficient on explanatory variables for motive group, household size, household income and average electricity consumption level as well as and gender, age and labor market status of the recipient.

Regressor	**Estimate**	**t-value**	**Regressor**	**Estimate**	**t-value**
Constant	-0.05	-1.02	Gender	0.03	2.60
Df20	0.02	0.75	Age	0.00	4.43
Df50	0.05	2.09	Labor market status	0.01	0.78
Dfe5	0.00	0.05	Avg. consumption	-0.02	-2.94
Dfe20	0.08	3.09	Household income	0.00	-0.86
Dfe50	0.10	2.96	Household size	0.01	0.78
Def5	0.01	0.55	edu2	-0.02	-1.02
Def20	0.01	0.35	edu3	0.04	1.31
Def50	0.06	1.90	edu4	-0.04	-1.34
DeA	-0.02	-0.86	edu5	-0.03	-1.42
DeB	-0.04	-1.80	edu6	-0.01	-0.25

Notes: 1356 observations after 10 outliers have been removed according to a Bonferroni limit. Influential diagnostics were checked but showed no evidence of particularly influential observations.

There are 1356 after the removal of 10 outliers. Diagnostics suggested homoschedasticity whereas normality was not supported. For all regressions in this section we also assessed the influence of single observations based on a set of diagnostics. However, there was no evidence of any individual highly influential observations. Table 8 shows estimation results after the insignificant variables have been removed (joint exclusion was accepted with a p-value of 0.48).

Based on Table 8 the following conclusions about the influence of motives emerge: First, the significance of Df50 indicates that a stronger financial motive is more effective, recalling that the benchmark motive group is f5. On the other hand, the negative estimate on De2 suggests that a purely environmental motive, reduces the amount of consumption moved, relative to the benchmark.

Table 8: As in Table 7 with insignificant regressors removed (p=0.48 for the joint test).

Regressor	**Estimate**	**t-value**
Constant	-0.06	-1.85
Df50	0.05	2.19
Dfe20	0.08	3.22
Dfe50	0.09	2.94
Def50	0.06	2.01
DeB	-0.05	-2.13
Gender	0.03	2.77
Age	0.002	6.39
Avg. consumption	-0.02	-3.08

However, from the positive estimates on Dfe20, Dfe50 and Def50, it seems that mixing financial and environmental motives is the most effective. In other words, these results together support the hypothesis that respondents are willing move their consumption for the sake of the environment, but only if they are compensated financially. Alternatively, the interpretation could also be that since the savings on the electricity bill in absolute terms are rather limited other non-pecuniary motivates are necessary. Turning to the remaining explanatory variables the positive estimate on the gender indicator suggests that women are more inclined to move consumption than men. In addition, the older the respondent is the more is moved (the positive estimate on the age coefficient). Note in particular that, the age effect is obtained despite that we have conditioned on labor market status (see Table 7). Finally, a higher average level of consumption seems to reduce the consumption change. As mentioned, since the change is in percentage terms this could suggest that there is little variation in the absolute changes in consumption across households with different average consumption levels. This, in turn, is consistent with moving "flexible consumption" such as laundry and dishwashing, which implies the same kwh consumption (per run) across households. Note that although a large household has more laundry and dish washing to do per unit of time, compared to a single retiree, say, this has nothing to do with whether consumption is moved or not when receiving and SMS.

For the corresponding regression of the AW10 coefficients on the same explanatory variables it turned out that all of these could be excluded jointly, and hence this is not reported.

Consider now the regression corresponding to consumption movement in the evening period, 18-21.

Table 9: OLS regression of the estimated IN18 coefficient on explanatory variables for motive group, household size, household income and average electricity consumption level, and the gender, age and labor market status of the recipient.

Regressor	**Estimate**	**t-value**	**Regressor**	**Estimate**	**t-value**
Constant	-0.01	-0.24	Gender	0.04	3.63
Df20	0.01	0.60	Age	0.00	3.27
Df50	0.04	1.76	Labor market status	0.00	0.32
Dfe5	-0.03	-1.45	Avg. consumption	-0.03	-4.09
Dfe20	0.00	-0.04	Household income	0.00	-0.13
Dfe50	0.13	4.21	Household size	0.01	0.88
Def5	0.00	0.08	edu2	0.01	0.90
Def20	0.03	1.00	edu3	-0.06	-1.88
Def50	0.04	1.33	edu4	0.00	0.11
DeA	-0.03	-1.71	edu5	0.00	0.11
DeB	-0.02	-0.78	edu6	0.00	-0.05

Notes: 1350 observations after 8 outliers have been removed according to a Bonferroni limit. Influential diagnostics were checked but showed no evidence of particularly influential observations.

Table 10: As in Table 9 with insignificant regressors removed (p=0.93 for the joint test).

Regressor	**Estimate**	**t-value**
Constant	0.02	0.65
Df50	0.04	1.69
Dfe5	-0.04	-2.02
Dfe50	0.13	4.40
DeA	-0.04	-2.34
Gender	0.04	3.59
Age	0.00	4.12
Avg. consumption	-0.03	-4.24
edu3	-0.07	-2.20

The results in Table 10 seem by and large to support the same conclusions as for the noon period in Table 8. However, the differences are that the conclusion that mixed motives is not as clear cut and that there is an indication, albeit vague, that education could have an influence. Concerning the role of mixing motives, we see that the positive estimate on Dfe50 is still obtained whereas now Dfe5 is significant and negative. However, comparing both the magnitudes and the t-values of these coefficients the conclusion for the former seems more convincing. Concerning the educational level, it appears that edu3, which indicates the “lower to intermediate” level, has a negative coefficient estimate. However, since the educational variable is first of all a rather composite variable, supposedly covering several influences, it may to some extent be viewed more as a control variable in the regression on motive and not have a clear interpretation in itself.

5 Conclusions

To an increasing extent, power generation comes from renewables, like solar, wind and wave. This introduces significant intra-day variability into the supply of electricity, which implies that demand has to be more flexible or movable across time. In the present research, we have considered the electricity consumption of households and investigated whether SMS messages (containing various motives) can induce these to change their consumption in different periods during the day. We have also investigated the role of motives in this context. In particular, which type of motive, i.e. purely pecuniary (financial), purely environmental, or a mix of these, is the most effective.

Taking the heterogeneity across individual households explicitly into account, we used automatic model selection (see Doornik 2009 and Hendry and Doornik 2014) to model each of 1488 household-level time series from a large Danish field experiment conducted in the period June, 2015 through–June 2016 (Andersen et al. 2017). From this, we obtained a cross-section of estimated SMS effects, which we then regressed on indicators for the different motives. Since participation in the experiment is voluntary, households have the possibility to opt out at any time. This may potentially create a self-selection bias on the estimated coefficient on the motive-indicators in this cross-sectional regression. In an attempt to reduce this, we therefore included socioeconomic and demographic explanatory variables. In particular, we controlled for the size, income and average consumption of the household, as well as the age, educational- and labor market status of the SMS recipient.

The results suggest that SMS messages can motivate households to move consumption within the day. There is some support to the claim that a stronger financial motive is more effective. Interestingly, a purely environmental motive seems to reduce the displaced amount. However, when mixing environmental motives with private financial gains, this seems to be the most effective. Hence, it seems that respondents are willing to be flexible and move their consumption for the sake of the environment, *but* only if they are compensated financially. Alternatively, the interpretation could also be that since the savings on the electricity bill in absolute terms are rather limited other non-pecuniary motivates are necessary. Finally, the inclusion of socioeconomic and demographic explanatory variables suggested that, women and elderly people are more inclined to move consumption.

To some extent, we view these results as tentative since the self-selection problem may go beyond the above-mentioned and we therefore suggest that future research could dig more systematically into this. Nevertheless, we believe that the insights provided may still give a good idea of which "types" of households could be targeted with text messages in the future and with what motives.

Acknowledgement

We would like to thank in particular Morten Nyland Christensen for excellent research assistance and Geraldine Henningsen and Frits Møller Andersen for useful comments. We are grateful to Innovation fund Denmark for funding this research.

References

Andersen, L. M., L. G. Hansen, C. L. Jensen, and F. Wolak (2017). Using Real-Time Pricing and Information Provision to Shift Intra-Day Electricity Consumption: Evidence from Denmark. Working paper, Stanford University.

Castle, J. L., J. A. Doornik, D. F. Hendry, and F. Pretis (2015). Detecting location shifts during model selection by step-indicator saturation. *Econometrics 3*(2), 240–264.

Caves, D. W., L. R. Christensen, and J. A. Herriges (1984). Consistency of residential customer response in time-of-use electricity pricing experiments. *Journal of Econometrics, Elsevier 26*(1-2), 179–203.

Delmas, M. A., M. Fischlein, and O. I. Asensio (2013). Information strategies and energy conservation behavior: A meta-analysis of experimental studies from 1975 to 2012. *Energy Policy 61*, 729 – 739.

Doornik, J. (2009). *Autometrics*, pp. 88–121. Oxford University Press.

Faruqui, A. and S. George (2005). Quantifying customer response to dynamic pricing. *The Electricity Journal 18*(4), 53 – 63.

Faruqui, A. and S. S. George (2002). The value of dynamic pricing in mass markets. *The Electricity Journal 15*(6), 45 – 55.

Faruqui, A. and J. Malko (1983). The residential demand for electricity by time-of-use: A survey of twelve experiments with peak load pricing. *Energy 8*(10), 781–795.

Faruqui, A. and S. Sergici (2010, Oct). Household response to dynamic pricing of electricity: a survey of 15 experiments. *Journal of Regulatory Economics 38*(2), 193–225.

Faruqui, A., S. Sergici, and A. Sharif (2010). The impact of informational feedback on energy consumptionŮa survey of the experimental evidence. *Energy 35*(4), 1598–1608.

Fischer, C. (2008). Feedback on household electricity consumption: a tool for saving energy? *Energy Efficiency 1*(1), 79–104.

Gleerup, M., A. Larsen, S. Leth-Petersen, and M. Togeby (2010). The effect of feedback by sms-text messages and email on household electricity consumption: Experimental evidence. *Energy Journal 31*(3), 113–132.

Hendry, D., S. Johansen, and C. Santos (2008). Automatic selection of indicators in a fully saturated regression. *Computational Statistics 23*(2), 317–335.

Hendry, D. and H.-M. Krolzig (2005). The properties of automatic "gets" modelling.

Economic Journal 115(502), C32–C61.

Hendry, D. F. (2009). *The Methodology of Empirical Econometric Modeling: Applied Econometrics Through the Looking-Glass*, pp. 3–67. London: Palgrave Macmillan UK.

Hendry, D. F. and J. A. Doornik (2014). *Empirical Model Discovery and Theory Evaluation: Automatic Selection Methods in Econometrics*. MIT Press.

Heshmati, A. (2014). Demand, customer base-line and demand response in the electricity market: A survey. *Journal of Economic Surveys 28*(5), 862–888.

Hoover, K. D. and S. J. Perez (1999). Data mining reconsidered: encompassing and the general-to-specific approach to specification search. *Econometrics Journal 2*(2), 167–191.

Johansen, S. and B. Nielsen (2009). *An analysis of the indicator saturation estimator as a robust regression estimator*, pp. 1–36. Oxford University Press.

King, C. S. and S. Chatterjee (2003, July). Predicting California Demand Response. *Public Utilities Fortnightly*, 27–32.

Newsham, G. R. and B. G. Bowker (2010). The effect of utility time-varying pricing and load control strategies on residential summer peak electricity use: A review. *Energy Policy 38*(7), 3289 – 3296.

Pretis, F., J. Reade, and G. Sucarrat (2016, April). General-to-Specific (GETS) Modelling And Indicator Saturation With The R Package Gets. Economics Series Working Papers 794, University of Oxford, Department of Economics.

Tureczek, A. and P. Nielsen (2017). Structured literature review of electricity consumption classification using smart meter data. *Energies 10*(5).

Vine, D., L. Buys, and P. Morris (2013). The effectiveness of energy feedback for conservation and peak demand : a literature review. *Open Journal of Energy Efficiency 2*(1), 7–15.