A New Approach to Modeling the Effects of Temperature Fluctuations on Monthly Electricity Demand

Yoosoon Chang[*], Chang Sik Kim[†], J. Isaac Miller[‡]

Joon Y. Park[§], Sungkeun Park[¶]

Abstract

This paper proposes a novel approach to measure and analyze the effect of temperature on electricity demand. This temperature effect is specified as a function of the density of temperatures observed at a high frequency with a functional coefficient, which we call the temperature response function. This approach contrasts with the usual approach to model the temperature effect as a function of heating and cooling degree days. We further investigate how non-climate variables, which include the price of electricity relative to that of substitutable energy and latent variables such as preferences and technology that we proxy by a linear time trend, affect the demand response to temperature changes. Our approach and methodology are demonstrated using Korean electricity demand data for residential and commercial sectors.

[*]Department of Economics, Indiana University

[†]Corresponding author. Address correspondence to Chang Sik Kim, Department of Economics, Sungkyunkwan University, Seoul 110-745, Korea, or to skimcs@skku.edu

[‡]Department of Economics, University of Missouri

[§]Department of Economics, Indiana University and Sungkyunkwan University

[¶]Korea Institute for Industrial Economics and Trade

1   Introduction

In households and firms in modern economies, electricity is one of the most essential goods consumed. It is certainly no surprise that there is an extensive literature that seeks to explain the variability of electricity demand across markets or in a given market over time. There is a long tradition in this literature, going back at least to Engle *et al.* (1989), of modeling the long-run and short-run effects of economic covariates, such as price and income, using an error-correction model. See also Silk and Joutz (1997), Beenstock *et al.* (1999), *inter alia.*

Because of the obvious effects of temperature on the demand for electricity in heating and cooling, these studies typically employ some temperature-based metric to control for short-run temperature-induced fluctuations in demand, which occur at seasonal and higher frequencies. Controlling instead for long-run influences on electricity demand, we focus on modeling these short-run demand fluctuations, which we may think of as the high-frequency (HF) component of electricity demand. We may view the response of the HF component to temperature as a *temperature response function* (TRF).[1]

In modeling temperature effects, researchers have long recognized the inadequacy of temporally aggregated measures of temperature, such as a monthly average. A linear TRF based on a monthly average temperature suffers from at least two major well-known deficiencies: linearity fails to capture increased demand at both very high and very low temperatures, and the average over a month may not adequately reflect usage during periods of temperature extremes in a given month.

The standard method for handling these deficiencies has been to employ heating degree days (HDD) and cooling degree days (CDD), which measure the number of degrees that the daily average temperatures in a given period – say, a month – fall below (for HDD) or rise above (for CDD) a threshold value, usually 18°C or 65°F. (See, e.g., Gupta and Yamada, 1972; Al-Zayer and Al-Ibrahim, 1996; Sailor and Muñoz, 1997; Fan and Hyndman, 2011.) Using these metrics in an otherwise linear model replaces a linear TRF with a piecewise linear TRF with a break point at the threshold temperature, addressing the first deficiency, while indirectly employing intramonthly data (daily averages), addressing the second deficiency.

Of course, piecewise linearity of the TRF and an arbitrary specification of the threshold may still be inadequate, and there are a number of studies aimed at improving the functional form by way of more sophisticated nonlinear parametric methods or even nonparametric

---

[1]Our approach does not explicitly model a demand response from temperature fluctuations at periodicities longer than seasonal, because we do not differentiate between the distribution of temperatures in January of one year from that January in another year.

methods, including Engle *et al.* (1986), Filippini (1995), Pagá and Gürer (1996), Henley and Peirson (1998), Valor *et al.* (2001), Pardo *et al.* (2002), and Moral-Carcedo and Vincens-Otero (2005).

The second deficiency, using a temporal aggregate, seems to have received less attention. Perhaps the indirect use of daily data by way of the HDD and CDD (H/CDD) metrics is viewed as adequate to capture intramonthly fluctuations, and perhaps the lack of econometric methods to deal with data observed at different sampling frequencies has been an obstacle to using intramonthly temperature data. Nonetheless, the fact that temporal aggregation may have a deleterious effect on inference is well known.

Two examples illustrate the inadequacy of using H/CDD data. First, suppose that two months have the same number of CDDs (20), but that one has 20 days on which the average temperature is $19°C$ with the remaining days at or below $18°C$, but the other has one day on which the average temperature is $38°C$ but with the remaining days at or below $18°C$. A deviation from the threshold of a single degree would not likely increase electricity usage much if at all, while a deviation of $20°C$ would very likely induce a massive increase in cooling. Introducing piecewise linearity into the TRF by way of CDDs cannot adequately capture this difference, because the number of CDDs is the same in both months.

As a second example, suppose that temperature fluctuations within a day are substantial, as may be the case in continental climates, such as the Midwestern US. On a given day, the average may show $18°C$, while the fluctuation over the course of that day may be $\pm 8°C$.[2] Monthly measures of HDD and CDD would not count that day, even though automated thermostats may switch on the heat, the air conditioning, or even both during the course of that day.

There is a third – perhaps more subtle – deficiency of standard temperature response functions. A TRF based only on temperature does not take into account economic or other non-climate covariates, such as the price of electricity. The subtlety lies in the fact that demand models typically *do* include these covariates. However, controlling for high-frequency temperature fluctuations separately from these covariates means that the impact of cold weather, for example, must be the same regardless of the price of electricity. Since the price of electricity relative to an alternate heating source, such as city gas, may influence an economic agent's use of electricity at a given cold temperature, we should not expect the TRF to be stable as relevant economic covariates evolve.

Further, the effect of price in such models must be the same regardless of season. Never-

---

[2]According to the US National Weather Service, http://www.srh.noaa.gov/ama/?n=50ranges, accessed October 10, 2014, average fluctuations of $30°F$ ($16.68°C$, or roughly $\pm 8°C$) are common for some parts of the Midwest (High Plains region) in March.

theless, if the electricity price is less expensive relative to rival fuels, demand for electricity in heating may increase during the winter time, even though the effect of changes in price may be negligible during the spring and summer time when there is little demand for heating. Fan and Hyndman (2011) estimate price elasticities that are in fact different in the winter and than in the summer.

In related research (Chang *et al.*, 2014) focusing on time-varying coefficients in an error-correction model, we employ a semiparametric functional coefficient approach to the temperature response function that maps hourly and geographically disaggregated temperature observations onto a monthly measure of the seasonal component of electricity demand. This mixed sampling frequency functional coefficient approach easily addresses the first two deficiencies of the standard H/CDD approach mentioned above: the semiparametric specification allows for nonlinearity in the spirit of Engle *et al.* (1986), *inter alia*, while the functional coefficient explicitly utilizes intramonthly temperature data.

In this paper, we focus only on the HF component of demand, and our main aim is to address the third deficiency in addition to the first two. In place of a TRF, we introduce a new concept: the cross-temperature response function (CTRF). The CTRF employs economic covariates directly in the component temperature response functions, both allowing the seasonal demand component to respond to non-climate variables and allowing the effects of non-climate variables to affect the response of the HF component of demand to temperature.

We decompose the effect of temperature on the HF component of electricity demand into three different components: a pure temperature effect, a price-temperature effect, and a time-temperature effect. We investigate the effect of temperature conditional on price and other factors proxied by time, so that the pure temperature effect can be identified.

We apply our model to Korean residential and commercial electricity demand, finding that non-climate variables have particularly substantial effects on changes in the temperature response function of the commercial sector.

The rest of the paper is organized as follows. In the next section, we introduce the TRF and CTRF, novel measures of seasonality using the entire intramonthly temperature distribution for each month, and we show how they generalize extant measures of seasonality, average temperature and H/CDD data. We discuss data for our application to Korean electricity demand in Section 3 and our estimation results in Section 4. Section 5 concludes.

2 Measurement of the Temperature Effect

2.1 Temperature Response Function

The temperature response function was used by subsets of the present authors in previous work (Chang and Martinez-Chombo, 2003, and Chang *et al.*, 2014). Because this concept is critical in developing our analysis, we provide here all of the details for the reader's convenience and in fact a more extensive discussion that supersedes the discussions of the temperature response function in those papers.

Consider a hypothetical measure y of the HF component of electricity demand observed as often as temperature. Such a HF measure abstracts from demand changes directly due to slowly evolving economic covariates, such long-run income changes. We will refer to this component of demand simply as the *HF component*. Our main purpose is to estimate the mean of y conditional on temperature and economic covariates that may fluctuate frequently. Setting aside the possibility of economic covariates for now, we define the *temperature response function* (TRF) g to be a possibly nonlinear function that maps the temperature to a response of the HF component of demand.

More realistically, the measure of the HF component of electricity demand is available only monthly, and we denote it by y_t for $t = 1, \ldots, T$. Letting f_t denote the density of temperature observations in month t, an estimator for the conditional mean of y_t given f_t is given by

$$\tau_t = \int f_t(r)g(r)dr, \tag{1}$$

where r is a dummy of integration, but we may think of r as representing intramonthly temperature observations, and we are integrating over all temperatures in month t.

More formally, we may write

$$g(r_0) = \int \delta_{r_0}(r)g(r)dr,$$

where δ_{r_0} is the dirac delta function at r_0 – i.e., the function that has a spike at a point r_0 and integrates to 1. We may interpret the value of function g at r_0 as the temperature effect on the HF component of electricity demand when the temperature distribution is hypothetically concentrated at r_0 – i.e., when the temperature density is given by δ_{r_0}.

Note that τ_t captures both the inherent nonlinearity in the relationship by way of g and the available intramonthly data by way of the functional approach. For a given TRF g, the relationship between the density f and temperature effect τ is linear, i.e., if the temperature densities f_1 and f_2 yield temperature effects τ_1 and τ_2, then the temperature

effect of $c_1 f_1 + c_2 f_2$ becomes $c_1 \tau_1 + c_2 \tau_2$ for any constants c_1 and c_2. In this context, we may simply regard the TRF g as a functional coefficient of temperature density.

Suppose instead that we aggregate the temperature data into a single average temperature datum for month t, and then rely on a nonlinear function h to estimate the temperature effect. The average temperature in a given month is $\int r f_t(r) dr$, so that single-frequency parametric or nonparametric methods discussed in this literature could be used with $h\left(\int r f_t(r) dr \right)$ to estimate h. However, h is not the TRF – it does not estimate the response to temperature as g does, unless g and h are both (unrealistically) linear. Rather, h estimates the *aggregate* response to the *average* monthly temperature, and a temperature measurement of, say, 18°C means that demand must respond as if the temperature were constant at 18°C for the whole month.

Using H/CDD data in place of a monthly average improves the situation. These measures may be written as

$$ HDD_t = \int h_H \left(\int r f_t(r) dr \right) ds \quad \text{and} \quad CDD_t = \int h_C \left(\int r f_t(r) dr \right) ds $$

where h_H and h_C are functions defined as $h_H(z) = \max(18 - z, 0)$ and $h_C(z) = \max(z - 18, 0)$ with the commonly used threshold temperature of 18°C, and where the integral across r denotes a daily average of intradaily temperatures, while the integral across s denotes a monthly sum of daily h_H and h_C. H/CDD data are often used directly, or else $h(HDD_t, CDD_t)$ may be estimated. Because h_H and h_C are piecewise linear functions, it is possible to write τ_t as $c_1 HDD_t + c_2 CDD_t$ for constants c_1 and c_2 (linear h) for a piecewise linear V-shaped g. The coefficients c_1 and c_2 allow the desirable asymmetry of the V shape often discussed in the literature.

Both of the preceding examples, monthly average and H/CDD, are very special cases. The efforts to move away from linear functions h and/or g in favor of smooth functions – U-shaped instead of V-shaped – without a fixed threshold temperature clearly undermine the use of a monthly average and even undermine the use of a smooth nonlinear function of H/CDD data.

Using intramonthly temperature data allows us to estimate (1) directly, more precisely estimating the response g of demand to the actual temperatures observed within a given month. To estimate the TRF g, we set

$$ y_t = \tau_t + \varepsilon_t = \int f_t(r) g(r) dr + \varepsilon_t, \tag{2} $$

where ε_t is a mean-zero error term independent of f_t for $t = 1, \ldots, T$.

We approximate the TRF g by a flexible Fourier functional (FFF) form, which decomposes the function g as a linear combination of a polynomial and pairs of trigonometric functions.[3] For our subsequent analysis, we normalize the temperature so that the temperature densities (f_t) (and also the TRF g correspondingly) are defined on the unit interval $[0, 1]$. Though not absolutely necessary, the normalization will greatly simplify our presentation below. If the raw temperature r is observed in an interval given by $[a, b]$ for some constants a and b, the required normalization may be done by setting $s = (r - a)/(b - a)$ and make a change of variables from r to s. For our empirical analysis, we use $a = -20$ and $b = 40$ in degrees Celsius, because all of the temperatures in our data lie between between $-20°$C and $40°$C.

To be more explicit, we momentarily denote the densities for raw and normalized temperatures respectively by (f_t^R) and (f_t^N), and the corresponding TRFs respectively by g_R and g_N. If the raw temperature density f_t^R is given, then we may easily obtain the corresponding density for normalized temperature as $f_t^N(s) = (b - a)f_t^R(a + (b - a)s)$ by the change of variables formula for each $t = 1, \ldots, T$. On the other hand, once we obtain the TRF g_N corresponding to (f_t^N) from our subsequent analysis, we may easily find the TRF g_R corresponding to (f_t^R) by $g_R(r) = g_N\big((r - a)/(b - a)\big)$. Clearly, the temperature effects (τ_t) defined in (1) are not affected by our normalization here. In what follows, we will simply denote the normalized densities and the normalized TRF simply by $f_t(s)$ and $g(s)$ instead of $f_t(r)$ and $g(r)$ for $t = 1, \ldots, T$. This notational convention should cause no confusion.

Under our normalization, the TRF g is defined on the unit interval $[0, 1]$ and therefore it can be approximated as

$$g(s) = \sum_{i=0}^{p} c_i s^i + \sum_{j=1}^{q} \left[c_{1j} \cos(2\pi j s) + c_{2j} \sin(2\pi j s) \right], \tag{3}$$

where c_i, c_{1j} and c_{2j} are unknown coefficients and p and q are the orders of the polynomial and trigonometric terms in our approximation.[4]

---

[3]The FFF form (Gallant, 1981) is well known in semiparametric economic analysis, and has been used in the energy literature – e.g., by Serletis and Shahmoradi (2008) to model interfuel substitution in a full energy demand system for the US, by Park and Zhao (2010) to model gasoline demand, and more specifically by Chang *et al.* (2014) for a TRF for electricity demand.

[4]We may approximate the raw TRF g using the trigonometric pairs with frequencies $2\pi j/(b - a)$ for $j = 1, 2, \ldots$, in place of those with frequencies $2\pi j$ for $j = 1, 2, \ldots$ used to approximate the normalized TRF g in (3).

It follows that

$$\int f_t(s)g(s)ds = \sum_{i=0}^{p} c_i \int s^i f_t(s)ds$$
$$+ \sum_{j=1}^{q} \left[c_{1j} \int f_t(s)\cos(2\pi js)ds + c_{2j} \int f_t(s)\sin(2\pi js)ds \right], \qquad (4)$$

so that we may readily derive the regression model

$$y_t = \sum_{i=0}^{p} c_i x_{it} + \sum_{j=1}^{q} \left[c_{1j}x_{1jt} + c_{2j}x_{2jt} \right] + \varepsilon_t^{pq} \qquad (5)$$

from (2) and (4), where

$$x_{it} = \int s^i f_t(s)ds, \quad x_{1jt} = \int \cos(2\pi js)f_t(s)ds, \quad x_{2jt} = \int \sin(2\pi js)f_t(s)ds$$

and ε_t^{pq} differs from ε_t by an approximation error that vanishes as $p, q \to \infty$. Practical determination of p and q is discussed below. We refer to the regression model in (5) as the *TRF model*.

We may estimate the regression in (5) by the conventional least squares method. Of course, the regressors x_{it}, $i = 1, \ldots, p$, and x_{1jt} and x_{2jt}, $j = 1, \ldots, q$, are not directly observable. However, they can easily be computed numerically, once we obtain estimates $\hat{f}_t$ of the temperature densities f_t for $t = 1, \ldots, T$, which may be accomplished by the usual nonparametric kernel method (Li and Racine, 2007, e.g.) using intramonthly (e.g., hourly) temperature observations collected in each month t.

The TRF g can then be estimated from the least squares estimates $\hat{c}_i$, $\hat{c}_{1j}$ and $\hat{c}_{2j}$ of the regression coefficients c_i, c_{1j} and c_{2j} in (5) for $i = 0, \ldots, p$ and $j = 1, \ldots, q$ as

$$\hat{g}(s) = \sum_{i=0}^{p} \hat{c}_i s^i + \sum_{j=1}^{q} \left[\hat{c}_{1j}\cos(2\pi js) + \hat{c}_{2j}\sin(2\pi js) \right] \qquad (6)$$

using the approximation of g in (3).

2.2 Cross-Temperature Response Function

Naturally, we may expect that non-climate variables (economic covariates), such as energy price, preference, technology, and policy, affect not only energy demand but also the temperature effect on demand. These variables change over time.

We can modify the TRF accordingly by letting it vary over time, more generally modeling it as

$$g_t(s) = \sum_{k=0}^{m} w_t^k g^k(s), \tag{7}$$

where, by setting $w_t^0 \equiv 1$, g^0 signifies the time-invariant component of the TRF, and g^k denotes the TRF measuring the temperature-dependent effect of covariate w_t^k on electricity demand for $k = 1, \ldots, m$. We refer to g^0 as the *base TRF* and to g^k as the *TRF with respect to w_t^k*. More specifically, in our application using time and relative electricity prices, we refer to these as the *time TRF* and *price TRF* respectively. In general, we refer to $g_t(s)$ as the *cross-temperature response function* (CTRF).

With the CTRF in (7), the total temperature effect becomes

$$\int f_t(s)g_t(s)ds = \sum_{k=0}^{m} w_t^k \int f_t(s)g^k(s)ds. \tag{8}$$

In particular, if we set $f_t = \delta_{s_0}$, where as before δ_{s_0} denotes the dirac-delta function at s_0, then we have

$$\int \delta_{s_0}(s)g_t(s)ds = \sum_{k=0}^{m} w_t^k g^k(s_0),$$

which shows the effect of a spike at temperature s_0 on energy demand. The corresponding temperature effect is therefore given by a linear function of the covariates w_t^k with coefficients given by $g^k(s_0)$ and an intercept of $g^0(s_0)$. Note that these intercept and coefficients are functions of temperature.

By approximating the TRF g^k as

$$g^k(s) = \sum_{i=0}^{p_k} c_i^k s^i + \sum_{j=1}^{q_k} \left[c_{1j}^k \cos(2\pi j s) + c_{2j}^k \sin(2\pi j s) \right]$$

for $k = 0, \ldots, m$ as in (3), we may write

$$\int f_t(s)g_t(s)ds = \sum_{k=0}^{m} \sum_{i=0}^{p_k} c_i^k w_t^k \int s^i f_t(s)ds$$

$$+ \sum_{k=0}^{m} \sum_{j=1}^{q_k} \left[c_{1j}^k w_t^k \int f_t(s) \cos(2\pi j s)ds + c_{2j}^k w_t^k \int f_t(s) \sin(2\pi j s)ds \right]$$

similarly to (4). We may now construct a regression given by

$$y_t = \sum_{k=0}^{m} \sum_{i=0}^{p_k} c_i^k x_{it}^k + \sum_{k=0}^{m} \sum_{j=1}^{q_k} \left[c_{1j}^k x_{1jt}^k + c_{2j}^k x_{2jt}^k \right] + \varepsilon_t^{p_k q_k}, \qquad (9)$$

where

$$x_{it}^k = w_t^k \int s^i f_t(s) ds, \quad x_{1jt}^k = w_t^k \int f_t(s) \cos(2\pi j s) ds, \quad x_{2jt}^k = w_t^k \int f_t(s) \sin(2\pi j s) ds,$$

similarly to the model in (5). We refer to the regression model in (9) as the *CTRF model.*

As with the TRF model in (5), the CTRF model in (9) can be estimated by least squares, given orders p_k and q_k of the polynomial and trigonometric terms in the TRF with respect to the k-th covariate for $k = 0, \ldots, m$ and given estimates of the temperature densities f_t for $t = 1, \ldots, T$.

Once we fit the regression in (9), the TRFs with respect to each covariate is readily estimated. Specifically, if we denote the resulting least squares estimates by $\hat{c}_i^k, \hat{c}_{1j}^k$ and $\hat{c}_{2j}^k$ for $k = 0, \ldots, m$, $i = 0, \ldots, p_k$ and $j = 1, \ldots, q_k$, then we may use

$$\hat{g}^k(s) = \sum_{i=0}^{p_k} \hat{c}_i^k s^i + \sum_{j=1}^{q_k} \left[\hat{c}_{1j}^k \cos(2\pi j s) + \hat{c}_{2j}^k \sin(2\pi j s) \right] \qquad (10)$$

for $k = 0, \ldots, m$ to estimate the TRF g^k with respect to the k-th covariate in (6).

We may set the support of some TRF to be a proper subset of the unit interval $[0, 1]$. In fact, there is a good reason to restrict the support of the price TRF. The reason is that gas is used extensively in heating but not as much in cooling. Therefore, we do not expect the HF component of electricity demand to respond to the price of electricity relative to city gas at temperatures warmer than some threshold $\bar{r}$. This implies that the normalized TRF has support contained in $[0, \bar{s}]$ with $\bar{s} = (\bar{r} - a)/(b - a) < 1$. With this restriction in place, we may estimate the price TRF using the terms

$$\left(\frac{s}{\bar{s}} \right)^i 1\{0 \le s \le \bar{s}\} \quad \text{and} \quad \left(\cos\left(\frac{2\pi j}{\bar{s}} s \right) 1\{0 \le s \le \bar{s}\}, \sin\left(\frac{2\pi j}{\bar{s}} s \right) 1\{0 \le s \le \bar{s}\} \right),$$

instead of s^i and $\left(\cos(2\pi j s), \sin(2\pi j s) \right)$, where $1\{0 \le s \le \bar{s}\}$ denotes the indicator function taking value 1 if $0 \le s \le \bar{s}$ and 0 otherwise.

3 Data

Our temperature distribution and measure of the HF demand component are identical to those used in our previous work (Chang *et al.*, 2014). We use distributions of hourly temperatures sampled from 5 geographically distributed cities in Korea. Because demand data are available only in 21 overlapping billing cycles, rather than monthly, the monthly national temperature density is given by $f_t(s) = \sum_{a=1}^{5} \sum_{b=1}^{21} w_{at} w_b f_{ab_t}(s)$, where w_{at} and w_b are weights assigned to each city and each billing cycle, and $f_{ab_t}(s)$ is the density for each city a in each billing cycle b ending in month t.

There are consequently 105 densities of hourly temperature observations for each month in the sample. Issues relating to the use of billing cycle data were discussed by Train *et al.* (1984), and our geographic weighting of temperature data is similar to that of Moral-Carcedo and Vicéns-Otero (2005) for Spain. However our approach using temperature distributions is quite a bit different from these approaches.

We obtain Korean residential and commercial electricity sales in megawatt hours (MWh) from Korea Electric Power Corporation (KEPCO). The billing cycle issue naturally pertains to the construction of our measure of HF component of electricity demand, and the rather involved construction of this measure attempts to take into account calendar effects from HF cycles in a workday and throughout a week, but with different numbers of weeks and workdays in each billing cycle and in each month. The problems of different loads on such days in constructing demand measures have been addressed by Pardo *et al.* (2002) and Moral-Carcedo and Vicéns-Otero (2005), *inter alia*.

Once a monthly demand measure is constructed, we take natural logs and subtract out the 12-month moving average of the series in logs in order to eliminate any stochastic or deterministic trends and thus isolate the HF component. The interested reader is referred to Chang *et al.* (2014) for a more complete discussion of how these series were constructed.

Figure 1 shows the resulting HF components of electricity demand for the residential and commercial sectors. If the HF component was created in such a way to be uncorrelated with the low-frequency trends proxied by the 12-month moving average, we could interpret a unit change in the HF component as an approximation to a percentage change in monthly demand, because the demand measure is in logs. Instead, we interpret a unit change to be an approximation to a percentage change in the HF component.

In our analysis of Korean electricity demand, we set $w_t^1 \equiv t/T$, so that the first covariate is given by time. We include time as a proxy for changes in preferences, technology, government energy policy, among other latent variables, as many previous authors have done, including Watts and Quiggin (1984), Jones (1994), Hunt *et al.* (2003), and Halicioglu
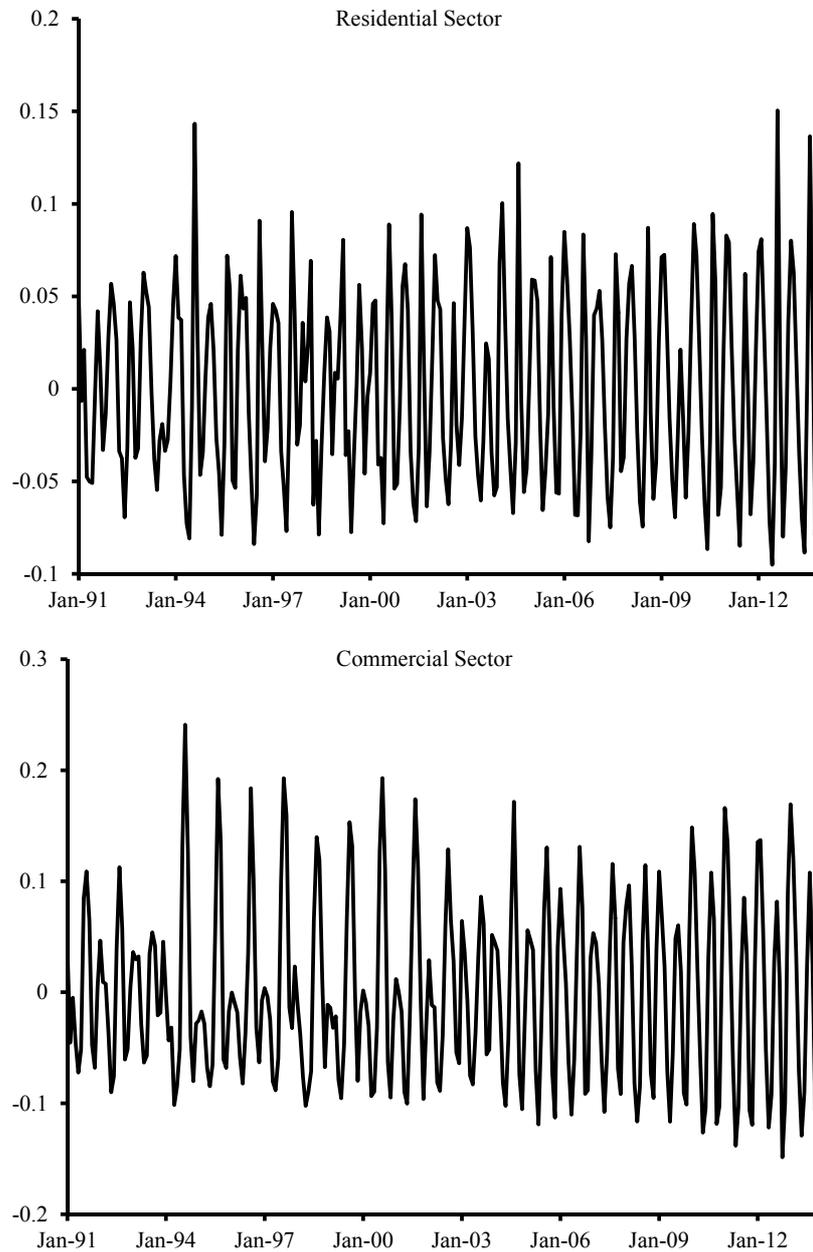
Figure 1: **HF Component of Electricity Demand in the Korean Residential and Commercial Sectors**. Data constructed as deviations from a twelve-month moving average of a measure of monthly national sectoral electricity demand (Chang *et al.*, 2014).

(2007).

For additional covariates in our application, we consider $w_t^2 = PR_t \ln RP_t$, where RP_t is the relative price of electricity to city gas, and PR_t is the penetration rate of city gas. We also estimated a less parsimonious model with $w_t^2 = \ln RP_t$ and $w_t^3 = PR_t$, we but do not report the results.

In Korea, city gas is the closest substitute for electricity, so these variables are expected to play important roles in determining electricity demand. The functional form implies that if the price of electricity relative to city gas increases by 1%, for instance, the effect on electricity demand is given by the fraction of 1% equal to the penetration rate. If penetration rate goes up, then there will be higher substitutability in gas consumption (instead of electricity), so the effect of the relative price of electricity on the HF component of electricity demand should increase. However, because of the substitutability, the effect of cold temperatures on the HF component should decrease with both penetration and electricity price.

We obtain electricity and city gas price indices from the Korean Statistical Information Service (KOSIS) and relative price is constructed as the electricity price index divided by the gas price index. The penetration rate of city gas is from the Korea City Gas Association. These series are displayed in Figure 2. City gas penetration relative to electricity has increased dramatically over the sample period, while the relative price of electricity has decreased dramatically.

Gas cooling equipment is less inefficient by 30-40% than electric cooling equipment, so gas cooling systems are currently used only for some public buildings to lower the summer peak of electricity demand in Korea. Therefore, we do not model any substitutional price effect in cooling demand.

Our final data set includes $T = 276$ monthly observations running from 1991:01 to 2013:12, since penetration rate data are available from 1991.

4 Estimation Results

4.1 Residential Temperature Response Function

We first analyze the temperature effect in residential electricity demand in Korea using the TRF model. To determine the orders p and q of the polynomial and trigonometric terms in our approximation of the TRF g in (3), we use the cross-validation criterion suggested by Burman *et al.* (1994) and choose p and q over the ranges of $p \in \{1, 2\}$ and $q \in \{0, 1\}$. The results suggest the choice of $p = 2$ and $q = 1$, i.e., the use of second order polynomial with one pair of trigonometric functions.
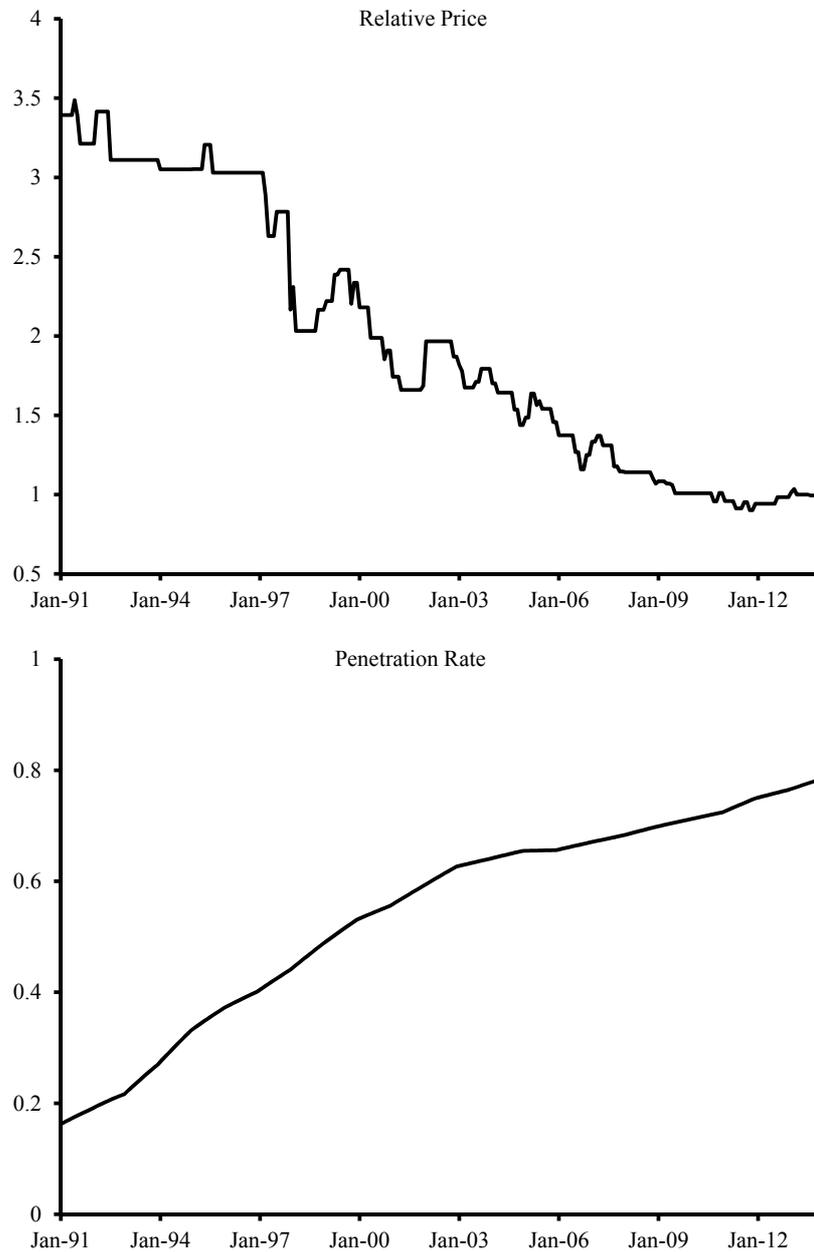
Figure 2: **Relative Price of Electricity (RP) and Penetration Rate of City Gas (PR).**
RP constructed as electricity price index divided by gas price index from the Korean Statistical
Information Service (KOSIS). PR from the Korea City Gas Association.

Coeff.	Est.	t-value
		$g(s)$
c_0	0.699	2.219
c_1	−4.695	−2.776
c_2	5.404	3.202
c_{11}	−0.272	−2.370
c_{21}	0.225	16.880
R^2	0.825	
$\bar{R}^2$	0.822	

Table 1: **Estimation Results for the Residential Sector TRF Model.** Using least squares with robust standard errors on the regression in (5) with temperature densities estimated using a normal kernel with plug-in bandwidth.

The least squares estimates for the regression coefficients are reported in Table 1, and the corresponding estimate of the TRF is presented in Figure 3. The estimated TRF has a shape that we normally expect. It is U-shaped taking values increasing as the temperature gets below and above a comfortable range. The temperature effects caused by heating and cooling needs appear to be asymmetric, the latter generating substantially more demand than the former.

The estimate of the TRF can be useful in many different contexts. First, the TRF itself provides some useful information on the intensities of the heating and cooling energy demands. If we look at 18°C, 23°C, and 28°C, the estimated values of the TRF are −0.09, −0.05, and 0.10. A 5°C increase in temperature from 18°C to 23°C increases (the HF component of) demand by 0.04, or 4%.

However, an increase in the same magnitude from 23°C to 28°C drives an increase in the HF component of demand of 15%. If the temperature instead drops from 13°C to 8°C, the increase is only 6%. These examples show both the asymmetry in the slopes and the nonlinearity both above and below the threshold temperature. Clearly, demand responses to otherwise equal temperature changes depend on the current temperature in a more complicated way than can be handled using H/CDD data.

Second, we may identify the temperature effect as in (1) using the estimated TRF and temperature densities. Analysis of the temperature effect in energy demand is very critical in forecasting peak load and deciding how to optimally employ a mix of power plants in electricity supply.

Third, we may perform some informative counterfactual analysis on temperature-related electricity demand. For instance, we may forecast the temperature effect assuming the temperature distribution will be the same as the average of temperature distributions in
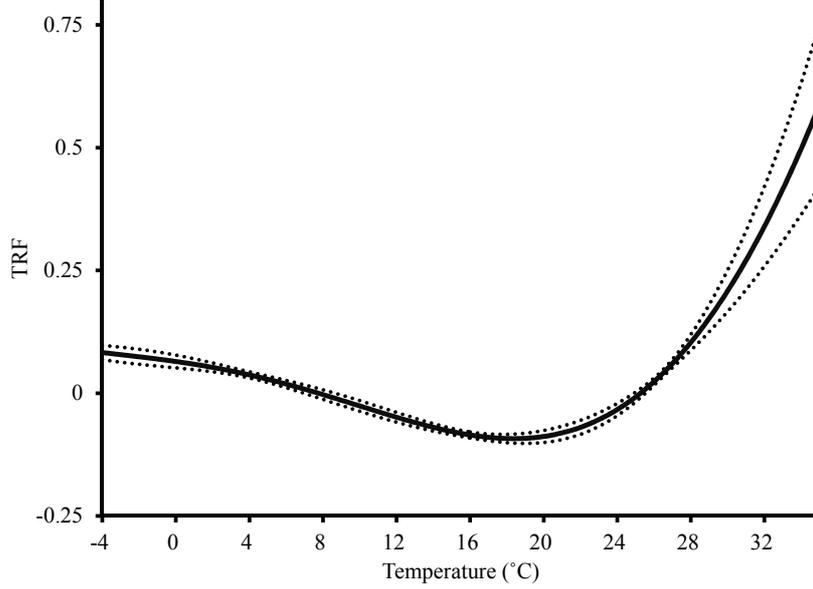
Figure 3: **Estimated Residential Sector TRF Model with 95% Confidence Bands.** Using the coefficient estimates in Table 1 and the TRF in equation (6) with rescaling. Confidence bands calculated according to Park (2010).

past years, or we may predict the effect of an increase in temperature. If the temperature distribution at time t is shifted to the right by u units of normalized temperature, we would have an increase in the temperature effect of $\int f_t(s - u)g(s)ds - \int f_t(s)g(s)ds$. Note that $f_t(\cdot - u)$ denotes the temperature distribution with mean temperature increased by u, compared to the temperature distribution represented by $f_t(\cdot)$, since $\int s f_t(s - u)ds = \int (s + u)f_t(s)ds = \int s f_t(s)ds + u$.

We also estimated a CTRF for the residential sector using the methodology described below for the commercial sector. We found that the confidence bands for the time TRF and price TRF contained a zero demand response for every temperature, suggesting that only the base TRF is useful in explaining the HF component of electricity demand. In light of the facts that residential electricity prices are kept artificially low and residential consumers are too small to warrant demand charges, the insignificance of the price TRF is not surprising. The residential time TRF exhibits a declining pattern similar to the commercial time TRF discussed below, but with much larger uncertainty.

4.2 Commercial Cross-Temperature Response Function

4.2.1 Estimation and Empirical Analysis of the CTRF Model

We first estimate the TRF model to find the threshold temperature $\bar{r}$ to use in estimating the price TRF. We determine p and q using cross-validation for the TRF, and then we set $\bar{r} = 14.2°$C, where the estimated TRF is minimized. Note that $\bar{s} = (14.2+20)/60 = 0.57$ for the price TRF. Next, we choose p_k and q_k for each TRF. In doing so, we consider $p_k \in \{1, 2\}$ and fix $q_k = 1$, and the cross-validation criterion selects $p_0 = 2$, $p_1 = 1$, and $p_2 = 1$.

To compare the results of the TRF and CTRF models, we include a time trend and $PR_t \ln RP_t$ as covariates alongside the TRF in the TRF model. In other words, to estimate the TRF, we are actually restricting the CTRF model by setting $p_1 = p_2 = 0$ and $q_1 = q_2 = 0$, with the convention that $q_k = 0$ means no trigonometric terms, but letting p_0 and q_0 (in the base TRF) exceed zero. Fixing $p_1 = p_2 = 0$ means that only a constant c_0^k is allowed in the TRFs with respect to time and price, and these constants become coefficients of these covariates in (9), since $\sum_{k=0}^r c_0^k x_0^k = \sum_{k=0}^r c_0^k w_t^k$. With the addition of the covariates, we refer to this as the *TRF+ model*.

The estimated results of TRF+ and CTRF models for commercial demand are summarized in Table 2, and the TRFs in the TRF+ and CTRF models are given in Figures 4 and 5 respectively. A Wald test allows a formal comparison of the two models. Using the values of R^2 for each model in the table, a Wald test may be constructed as $(0.920 - 0.771)/(1 - 0.920) \times (276 - 13) = 489.84$, easily beating the χ_7^2 critical value of 14.07 for a size-0.05 test. The TRF+ model is thus rejected in favor of the CTRF model.

The shapes of the estimated TRF in Figure 4 and analogous base TRF in Figure 5 are both U-shaped in the range of temperatures with the scale reflecting the fluctuations of the HF component of electricity demand. The only noticeable difference between the TRF and base TRF of the CTRF is that the latter appears to flatten out rather than continue to increase at the lowest temperatures.

As we can see in Table 2, the effect of time in the TRF+ model is estimated to be insignificant. Keeping in mind that the HF component of demand is detrended, this finding is not surprising. In the CTRF model, the effects of time are estimated to be *significant*. The time TRF and confidence intervals in Figure 5 better illustrate the effects.

The time TRF takes positive values in the range of 1°C or less, close to zero in the range of 1-24°C, most of the temperature spectrum, and negative values in the range exceeding 24°C. Consider for example the temperatures of $-4°$C and $34°$C, at which $\hat{g}^1$ is about 0.16 and -0.18. A change of ten years (a change in t/T of 120/276) increases the response of the HF component at $-4°$C by 7.0% but *decreases* the response at $34°$C by 7.8%. These

	TRF+ Model		CTRF Model	
	Est	t-value	Est.	t-value
Coeff.	$g(s)$		$g^0(s)$	
c_0	-0.339	-3.184	-2.549	-5.585
c_1	0.885	4.607	10.245	5.114
c_2			-7.966	-4.416
c_{11}	0.217	11.999	0.686	5.424
c_{21}	0.255	6.667	0.576	7.128
	t		$g^1(s)$	
c_0	-0.002	-0.131	0.662	2.912
c_1			-1.161	-2.757
c_{11}			0.123	3.115
c_{21}			-0.178	-1.942
	$PR_t \log RP_t$		$g^2(s)$	
c_0	0.027	0.962	0.004	6.689
c_1			-0.334	-3.775
c_{11}			0.300	6.795
c_{21}			-0.059	-0.580
R^2	0.771		0.920	
$\bar{R}^2$	0.767		0.917	

Table 2: **Estimation Results for the Commercial TRF and CTRF Models.** Using least squares with robust standard errors on the regression in (9) with temperature densities estimated using a normal kernel with plug-in bandwidth.

compare with base responses (from the base TRF) of 11.7% and 43.5% at $-4°$C and 34$°$C respectively.

These results suggest that, over a long span of time, the seasonality of commercial electricity demand in South Korea has increased in the winter time, but decreased in the summer. That is, the growth rate of heating demand has exceeded that of the average load, which is mainly due to a rapid increase in the supply of electric heating appliances in recent years so that consumers have switched their heating systems to electricity. However, the growth rate of the cooling demand is lower than that of the average load, which reflects the technical progress in electric cooling appliances so that consumers have replaced their cooling appliances by more energy efficient ones.

We can also see in Table 2 that the (short-run) price elasticity is estimated to be (insignificantly) positive in the TRF+ model – certainly the opposite sign of what we should expect. However, the CTRF model estimates a more sensible range of price elasticities. As shown in Figure 5, the price TRF is estimated to be significantly negative at temperatures under approximately 7.5$°$C – that is, 95% confidence interval does not include zero below
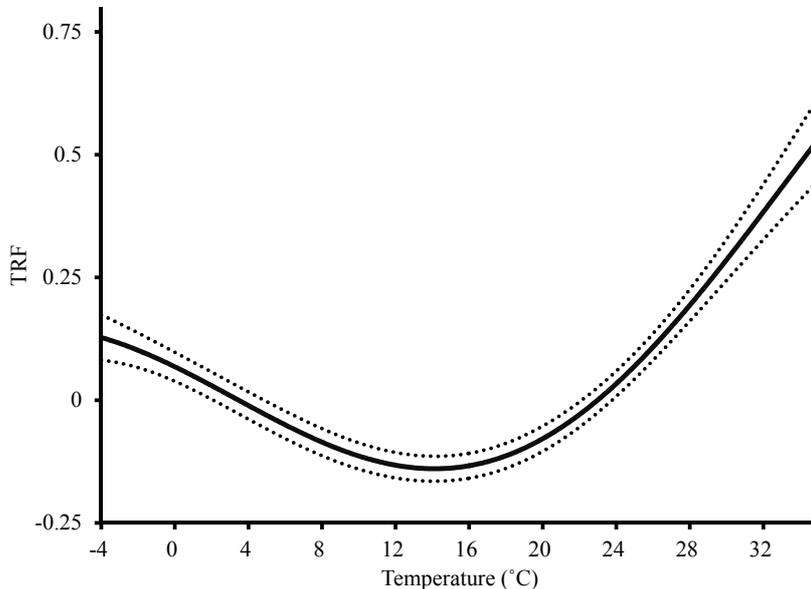
Figure 4: **Estimated Commercial Sector TRF with 95% Confidence Bands.** Using the coefficient estimates from the TRF+ model in Table 2 with the base TRF defined by $g^0(r)$ in equation (10) with rescaling. Confidence bands calculated according to Park (2010).

approximately $7.5°C$. Above this temperature, electricity price relative to that of city gas has no significant impact on commercial consumption of electricity, since city gas is not needed for heating.

A more interesting result is that the magnitude of the price TRF increases as temperature decreases below $7.5°C$, which means that the substitution effect in heating demand becomes clearer as the temperature becomes lower. Indeed, this result helps to explain the flattening of the base TRF discussed above: commercial consumers respond less to cold temperatures when accounting for the price of electricity relative to that of city gas.

In fact, electricity sales to the commercial sector of January 2012 increased by 39.0% compared with that of January 2006, whereas natural gas sales for the commercial sector grew by 0.7% during the same period. Meanwhile the electricity price index increased by 10.7% and gas price index grew by 61.5% between January 2006 and January 2012. The estimated price TRF clearly reflects this shift to electric heating from gas heating.

We illustrate in more detail how one can interpret the estimated price TRF in Figure 5. The relative price elasticity of the HF component of demand is given by $\hat{g}^2 PR_t$. For example, if $PR_t = 1$ and we look at $-4°C$ and $10°C$, the estimated values of $\hat{g}^2 PR_t$ are approximately -0.46 and -0.03 respectively. Hence, if temperature changes from $10°C$ to $-4°C$, then the relative price elasticity of the HF component will change from nearly
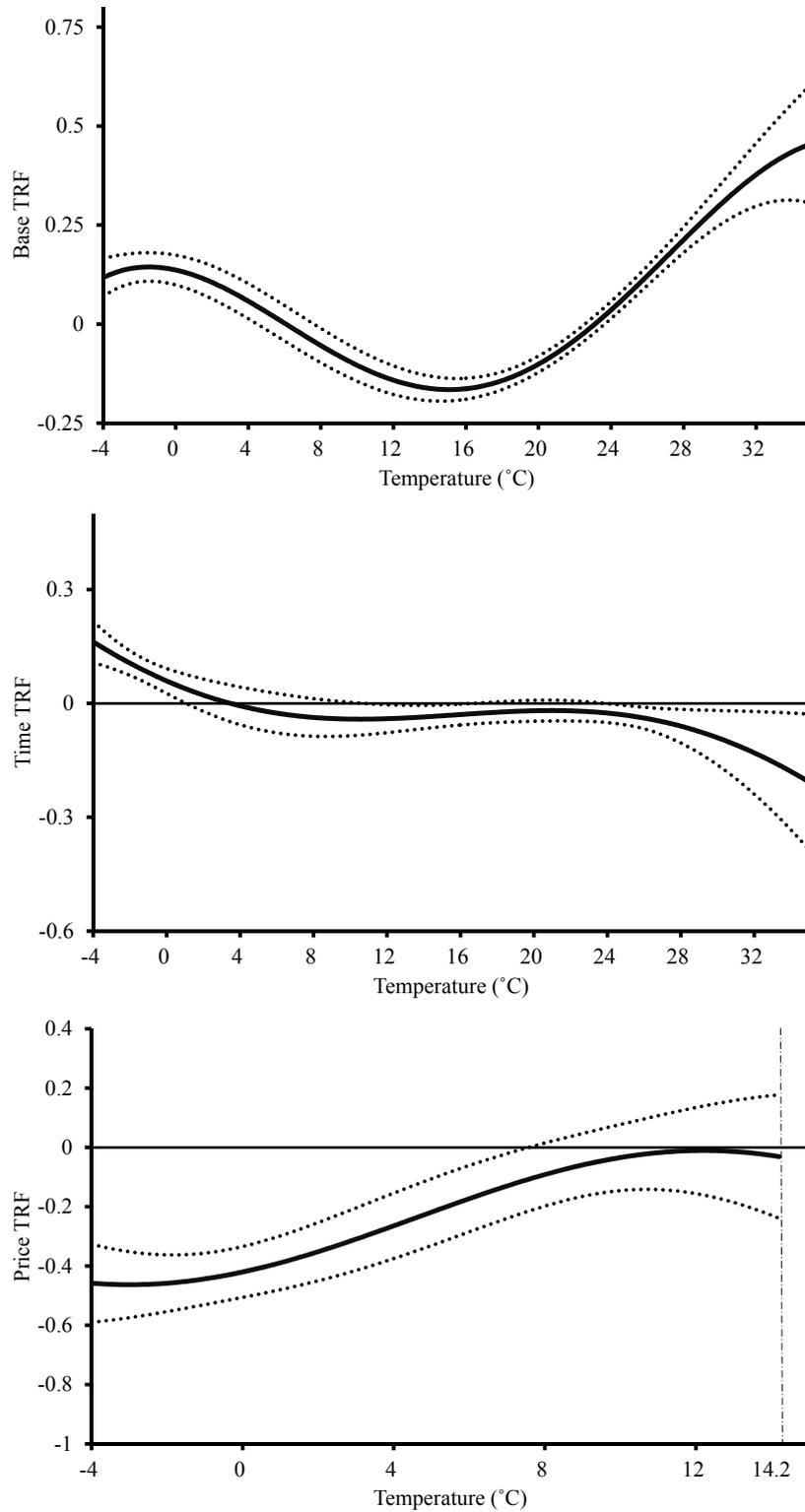
Figure 5: **Estimated Commercial Sector CTRF with 95% Confidence Bands.** Using the coefficient estimates from the CTRF model in Table 2 and TRFs defined by $g^k(r)$ in equation (10) with rescaling. Confidence bands calculated according to Park (2010).

zero (completely price inelastic) to -0.46. Moreover, if the electricity price index decreases by 10% and the natural gas index is unchanged (a change in $PR_t \ln RP_t$ of 0.10), the HF component of electricity demand for the commercial sector would increase by 4.6% at $-4°$C but only by 0.3% at $10°$C, which shows quite different substitution patterns at the different temperature levels. These compare with base responses (from the base TRF) of 11.7% and -10.4% at $-4°$C and $10°$C respectively.

Looking at the whole CTRF in the CTRF model as the sum of the individual TRFs, we can make another comparison with the TRF+ model. For example, at the counterfactual temperature of $-4°$C in January 2002 the sum of the base TRF and time-weighted time TRF is $11.7\% + 7.0\% \times (133/276) = 15.1\%$. At a penetration rate of $PR_t = 1$, a relative price decrease of 10% increases the response of the HF component by an additional 4.6%, so that the total response is 19.7% more than that of the response at an average temperature with no price change. In contrast, the TRF+ model suggests a response of $12.8\% - 0.2\% \times (133/276) = 12.7\%$ at $-4°$C in January 2002, but that a relative price decrease of 10% *decreases* the response of the HF component by 0.3% (but not significantly). The aggregate response is therefore predicted to be only 12.4% above an average temperature with no price change.

4.2.2   Seasonal and Temporal Analyses

Figure 6 shows the mean absolute error (MAE) of estimated residuals by months, and it shows that the CTRF model outperforms the TRF+ model in all months except April when MAE's for the two models are very close. The MAE's of January, February, March and August of CTRF model are 61%, 65%, 55% and 55% smaller than those of TRF+ model respectively, which shows rather clear price- and time-dependent temperature effects in the commercial sector in those months. We can deduce from our above results that time affects the temperature response in both winter and summer, while relative price also affects the temperature response in the winter.

The CTRF model enables us to decompose the monthly temperature effects into a price-dependent factor and a time trend-dependent factor, allowing us to better identify the aggregate changes in temperature effects due to time and relative price. The temperature effects in (8) may be written as

$$\int f_t(r)g_t(r)dr = \int f_t(r)g^0(r)dr + \frac{t}{T}\int f_t(r)g^1(r)dr + PR_t \ln RP_t \int f_t(r)g^2(r)dr$$

using our covariates $w_t^0 = 1$, $w_t^1 = t/T$ and $w_t^2 = PR_t \ln RP_t$.

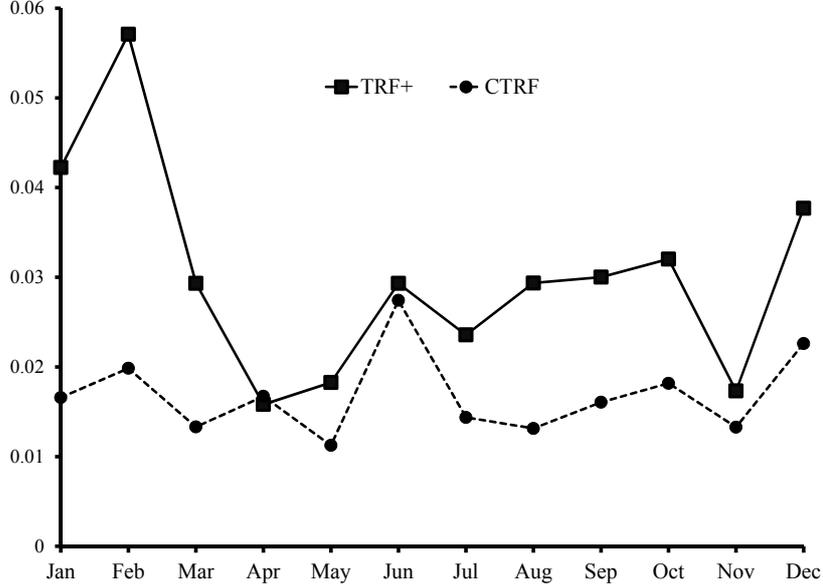Consider temperature effects for each month $M = 1, ..., 12$ constructed from this CTRF

Figure 6: **Mean Absolute Error by Months.**

using:

1. $\overline{TE}_{M0}$: the time index for month M in 1991 and the penetration rate and relative price for month M in 1991,

2. $\overline{TE}_{M1}$: the time index for month M in 1991 and the penetration rate and relative price for month M in 2013, and

3. $\overline{TE}_{M2}$: the time index for month M in 2013 and the penetration rate and relative price for month M in 2013.

The difference $\overline{TE}_{M2} - \overline{TE}_{M0}$ indicates the total change between 1991 and 2013. A component of the total change, $\overline{TE}_{M1} - \overline{TE}_{M0}$ is the change in the temperature effect due to the change in $PR_t \log RP_t$ between 1991 and 2013 while holding constant other temporal drivers proxied by a time trend. Similarly, $\overline{TE}_{M2} - \overline{TE}_{M1}$ is the change caused by these other drivers, while holding the price covariate constant.

To estimate these effects, we estimate g^k and f_t as described above, except we pool observations in month M across all 23 years in the sample to estimate f_M for $M = 1, ..., 12$. By using the same temperature density for the same month in all years, the changes that we identify over time given by $\overline{TE}_{M2} - \overline{TE}_{M1}$ can be attributed to temporal drivers other than possible long-run temperature changes.

Month	$\overline{TE}_{M0}$	$\overline{TE}_{M1}$	$\overline{TE}_{M2}$	$\overline{TE}_{M2} - \overline{TE}_{M0}$	$\overline{TE}_{M1} - \overline{TE}_{M0}$	$\overline{TE}_{M2} - \overline{TE}_{M1}$
January	−0.018	0.045	0.135	0.153	0.062	0.091
February	−0.013	0.044	0.124	0.137	0.058	0.079
March	−0.027	0.022	0.037	0.064	0.049	0.015
April	−0.085	−0.065	−0.088	−0.002	0.021	−0.023
May	−0.109	−0.105	−0.135	−0.026	0.004	−0.029
June	−0.046	−0.046	−0.075	−0.028	0.000	−0.028
July	0.044	0.044	0.010	−0.034	0.000	−0.034
August	0.128	0.128	0.080	−0.048	0.000	−0.048
September	0.047	0.047	0.012	−0.035	0.000	−0.035
October	−0.085	−0.083	−0.110	−0.025	0.001	−0.027
November	−0.102	−0.087	−0.112	−0.009	0.015	−0.024
December	−0.036	0.006	0.028	0.064	0.043	0.022

Table 3: **Decompositions of Monthly Temperature Effects.**

Table 3 shows the decompositions between 1991 and 2013 and their differences. The total change in temperature effect over the sample is positive in the winter months of December, January, February, and March, but negative in all other months.

The breakdown of the positive changes in the winter months are 66% price and 34% other factors proxied by time for December, 41% and 59% for January, 42% and 58% for February, and 76% and 24% for March. We may interpret this to mean that the overall increases in the HF component of demand in winter months may be attributed to both increases due to price and penetration changes over the sample and those due to changes in other temporally varying non-climate, non-price variables. The increases due to price during these months given by $\overline{TE}_{M1} - \overline{TE}_{M0}$ have roughly the same magnitudes (4.3%-6.2%), but the change from the other factors is less important in the warmer months of December and March (2.2% and 1.5%) than in January or February (9.1% and 7.9%).

Looking at the summer months of June, July, August and September, we see no effect on the HF component from price changes. This result is an artifact of our modeling strategy, since we set the price TRF to zero at temperatures above 14.2°C. Those four months are unlikely to have enough temperature observations below this threshold to make any substantial difference. We are essentially imposing that there will be no long-run effect of relative prices in summer months, except possibly through the time TRF if the price covariate has a time trend.

The remaining spring and fall months, April, May, October, and November show decreases in the temperature effect overall, but driven primarily by negative effects from non-climate, non-price factors and countervailed by *increases* due to price changes. In other

words, relative price changes have led to only small increases (0.1%-2.1%) in the HF component of electricity demand in those months – due to decreases in the relative price from increases in natural gas prices – while other factors have driven more substantial decreases (2.3%-2.9%).

5   Conclusions

In this paper, a general model is proposed in order to estimate and identify temperature effects in a short-run electricity demand function. We adopt a new approach using temperature densities to estimate a cross temperature response function, which exploits the fact that the non-climate variables have different effects on electricity demand at different temperatures.

We fit our proposed model (CTRF) and a benchmark model (TRF+) to the Korean commercial electricity demand data over 1991:01-2013:12. The effect of relative price is shown to be significant for heating demand. Moreover, technical progress in electric appliances and changes in consumption habits (proxied by the time trend) has lowered the growth rate of the cooling demand and has increased the growth rate of the heating demand.

6   References

Al-Zayer, J. and A.A. Al-Ibrahim (1996). "Modelling the impact of temperature on electricity consumption in the eastern province of Saudi Arabia," *Journal of Forecasting* 15, 97-106.

Beenstock, M., E. Goldin, and D. Nabot (1999). "The demand for electricity in Israel," *Energy Economics* 21, 168-183.

Burman, P., E. Chow, and D. Nolan (1994). "A cross-validatory method for dependent data," *Biometrika* 81, 351-358.

Chang, Y. and E. Martinez-Chombo (2003). "Electricity demand analysis using cointegration and error-correction models with time varying parameters: The Mexican case," mimeo, Department of Economics, Rice University.

Chang, Y., C.S. Kim, J.I. Miller, J.Y. Park and S. Park (2014). "Time-varying long-run income and output elasticities of electricity demand with an application to Korea," *Energy Economics* 46, 334-347.

Engle, R.F., C.W.J. Granger, J. Rice, and A. Weiss (1986). "Semiparametric estimates of the relation between weather and electricity sales," *Journal of the American Statistical Association* 81, 310-320.

Engle, R.F., C.W.J. Granger, and J.S. Hallman (1989). "Merging short and long-run forecasts: An application of seasonal cointegration to monthly electricity sales forecasting," *Journal of Econometrics* 43, 45-62.

Fan, S. and R.J. Hyndman (2011). "The price elasticity of electricity demand in South Australia," *Energy Policy* 39, 3709-3719.

Filippini, M. (1995). "Swiss residential demand for electricity by time-of-use," *Resource and Energy Economics* 17, 281–290.

Gallant, A.R. (1981). "On the basis in flexible functional forms and an essentially unbiased form: the Fourier flexible form," *Journal of Econometrics* 15, 211-245.

Gupta, P.C. and K. Yamada (1972). "Adaptive short-term forecasting of hourly loads using weather information," *IEEE Transactions on Power Apparatus and Systems* 5, 2085-2094.

Halicioglu, F. (2007). "Residential electricity demand dynamics in Turkey," *Energy Economics* 29, 199-210.

Henley, A. and J. Peirson (1998). "Residential energy demand and the interaction of price and temperature: British experimental evidence," *Energy Economics* 20, 157-171.

Hunt, L.C., G. Judge, and Y. Ninomiya (2003). "Underlying trends and seasonality in UK energy demand: A sectoral analysis," *Energy Economics* 25, 93-118.

Jones, C.T. (1994). "Accounting for technical progress in aggregate energy demand," *Energy Economics* 16, 245-252.

Li, Q. and J.S. Racine (2007). *Nonparametric Econometrics Theory and Practice*. Princeton: Princeton University Press.

Moral-Carcedo, J. and J. Vicéns-Otero (2005). "Modelling the non-linear response of Spanish electricity demand to temperature variations," *Energy Economics* 27, 477-494.

Pagá, E. and N. Gürer (1996). "Reassessing energy intensities: A quest for new realism," *OPEC Review* 20, 47-86.

Pardo, A., V. Meneu, and E. Valor (2002). "Temperature and seasonality influences on Spanish electricity load," *Energy Economics* 24, 55-70.

Park, C. (2010). "How does changing age distribution impact stock prices? A nonparametric approach," *Journal of Applied Econometrics* 25, 1155-1178.

Park, S.Y. and G. Zhao (2010). "An estimation of U.S. gasoline demand: A smooth time-varying cointegration approach," *Energy Economics* 32, 110-120.

Sailor, D.J. and J.R. Muñoz (1997). "Sensitivity of electricity and natural gas consumption to climate in the USA – methodology and results for eight states," *Energy* 22, 987–998.

Serletis, A. and A. Shahmoradi (2008). "Semi-nonparametric estimates of interfuel substitution in U.S. energy demand," *Energy Economics* 30, 2123-2133.

Silk, J.I. and F.L. Joutz (1997). "Short and long-run elasticities in US residential electricity demand: A co-integration approach," *Energy Economics* 19, 493-513.

Train, K., P. Ignelzi, R.F. Engle, C.W.J. Granger, and R. Ramanathan (1984). "The billing cycle and weather variables in a model of electricity sales," *Energy* 9, 1041-1047.

Valor, E., V. Meneu, and V. Caselles (2001). "Daily air temperature and electricity load in Spain," *Journal of Applied Meteorology* 40, 1413-1421.

Watts, G. and J. Quiggin (1984). "A note on the use of a logarithmic time trend," *Review of Marketing and Agricultural Economics* 52, 91-99.