

Disentangling Temporal Patterns in Elasticities: A Functional Coefficient Panel Analysis of Electricity Demand*

Yoosoon Chang[†], Yongok Choi[‡], Chang Sik Kim[§], J. Isaac Miller[¶], Joon Y. Park^{||}

Abstract

We introduce a panel model with a nonparametric functional coefficient of multiple arguments. The coefficient is a function both of time, allowing temporal changes in an otherwise linear model, and of the regressor itself, allowing non-linearity. In contrast to a time series model, the effects of the two arguments can be identified using a panel model. We apply the model to the relationship between real GDP and electricity consumption. Our results suggest that the corresponding elasticities have decreased over time in developed countries, but that this decrease cannot be entirely explained by changes in GDP itself or by sectoral shifts.

This Version: November 7, 2013

JEL Classification: C33, C51, C53, Q41

Key words and phrases: semiparametric panel regression, partially linear functional coefficient model, elasticity of electricity demand

*The authors are grateful for useful comments from seminar participants at the University of Missouri, Indiana University and the University of Kansas. We also thank Zongwu Cai for helpful discussions. We gratefully acknowledge Kyung-Mook Lim and Doosan Corporation for generous financial support.

[†]Department of Economics, Indiana University

[‡]Department of Economics, Indiana University

[§]Department of Economics, Sungkyunkwan University

[¶]Corresponding author. Address correspondence to J. Isaac Miller, Department of Economics, University of Missouri, 118 Professional Building, Columbia, MO 65211, or to millerjisaac@missouri.edu.

^{||}Department of Economics, Indiana University and Sungkyunkwan University

1 Introduction

A diverse literature addresses methods for handling structural change in the coefficients of econometric models, usually by allowing the coefficients to vary over time. Many of these approaches neglect one or both of two important aspects of structural change. First, such models do not typically allow changes in the specification of the functional form itself. Such misspecification may invalidate the economic interpretations of the coefficients when these interpretations are derived from partial derivatives, as is the case with elasticities. Second, few models of coefficient change aim to identify the underlying drivers of that change.

A functional coefficient with multiple arguments, consisting of the regressor itself and each potential driver of parameter change, remedies both of these deficiencies. Specifying the coefficient as an unknown function of the regressor explicitly allows for nonlinearity in the conditional mean of the regressand. The additional arguments further elucidate the underlying causes of the coefficient changes. However, a functional coefficient with more than one argument cannot be effectively estimated – especially when the arguments are highly correlated or share trends.

In order to operationalize such a model, we propose coupling a panel data approach to a nonparametric functional coefficient model. This approach provides several advantages in this context. First, the addition of the cross-sectional dimension of the data allows effective estimation of an unknown function of up to two variables. We consolidate the additional arguments into a single time trend to represent structural change. Thus, once the model is estimated, we may fix time and examine nonlinearity in the conditional mean. For a fixed regressor or a constant function of the regressor, on the other hand, the model reduces to a more standard model of temporal coefficient change. Such flexibility is not possible using only a time series or cross-sectional model. Second, the allowance of the coefficient to vary over *both* the regressor and time enabled by a panel allows dynamic misspecification, because the nonlinear function itself may evolve over time. Third, a panel provides a much larger number of observations to counter the well-known drawback of the slower rate of convergence of nonparametric estimators.

Once a temporal pattern in the coefficient is established and any nonlinearity is identified, further analysis may unlock distinct components of that pattern. These components could be of interest in their own right. For example, a policy maker considering a stimulus package to a particular economic sector might be interested in the effect on overall electricity consumption – especially in a country with a limited power grid that cannot import electricity, such as South Korea or Taiwan, or in which the constituents have substantial concerns about increasing pollution from fossil fuel consumption.

We apply our econometric approach to a panel of observations on electricity consumption across countries with disparate GDP levels. A stylized fact of developed economies is change over time of energy intensity, measured as the ratio of energy consumption to real GDP. Many of these countries have seen a decrease in energy intensity, often referred to as an autonomous energy efficiency increase (AEEI). Such changes have occurred not only with respect to overall energy consumption, but also with respect to consumption of individual energy sources, such as electricity. For example, over the period 1995-2010, our data suggest that this ratio (*electricity* intensity) has decreased by 14-17% for the US, UK, and Denmark, and decreased by 1-4% for Japan, Germany, and Belgium, but increased by 46% for Korea.

A common specification for modeling the relationship between electricity and GDP is a fixed coefficient regression of the log of electricity consumption per capita on the log of real GDP per capita and covariates.^{1,2} Holding the covariates constant, this specification assumes that the relationship is linear and stable. Galli (1998), Judson *et al.* (1999), and Medlock and Soligo (2001) document an inverted U shape in the relationship between log GDP and log energy consumption, which they attribute to changing patterns in electricity consumption as countries develop and especially to shifts in the compositions of national economies from more energy-intensive to less energy-intensive sectors. In other words, the relationship is both nonlinear and changing over time.

Even for a single economic sector, the relationship between log electricity consumption and log income, proxied by log GDP in the household sector, or log electricity consumption and log output, proxied by log GDP in the firm sector, may be linear but with unstable coefficients or may not be linear at all. The often assumed log-linearity of the aggregate household demand function or the aggregate firm conditional factor demand function results from multiplicative indirect utility or production functions.³ However, changes in utility, technology, policy, or other factors may shift or change these functions over time, inducing time-varying coefficients and even functional misspecification.

While an ideal specification might allow the coefficient on log GDP to be a function of GDP itself, utility, technology, energy policy, sectoral shares, and perhaps other factors, a model that can effectively identify all of these arguments is not feasible. Instead, we use

¹Initiated by Kraft and Kraft (1978), Granger causality is a major focus of the literature on this relationship. However, Granger causality is not a focus of the present analysis.

²For brevity, all further references to electricity consumption or GDP should be interpreted to mean electricity consumption per capita or real GDP per capita, unless otherwise specified.

³A large number of studies on household demand, including Halvorsen (1975), Maddala *et al.* (1997), and Silk and Joutz (1997), *inter alia*, have estimated a fixed coefficient on log income (income elasticity of demand), for which log GDP may be considered a proxy. In the firm sector, GDP may be a proxy for either output or income. Halvorsen (1978) included measures of both output and income in his model of commercial demand, while Berndt and Wood (1975) and Halvorsen (1978) used measures of output in industrial demand.

the panel nonparametric approach to create counterfactuals at fixed levels of development and time periods.

Our model generates a very clear empirical result and one that is expected from the discussion of AEEI's: income elasticities have been declining over time for developed countries.⁴ Our counterfactual analysis with time fixed and varying GDP suggests that economic development does not fully explain the declining elasticities in developed countries. The right-hand tail of the inverted U shape is almost flat and has become flatter over time – i.e., there is a threshold beyond which GDP (per capita and relative to other countries) barely affects the elasticity, and both the threshold and the decreasing effect have decreased over time.

Similarly, we construct counterfactuals in which GDP is fixed and time varies and find that the decreasing temporal pattern remains. Reliable sectoral data on electricity consumption for a subset of our panel that includes developed countries over a relatively recent time period allows further analysis of this decrease. We isolate the component of the time-varying elasticities for this subsample that cannot be explained by sectoral reallocations over time, and we find that the decreasing trend in elasticities of these developed countries still remains.

Having ruled out GDP changes and accounting for sectoral shifts as possible explanations for the evident decreasing pattern in elasticities, we must conclude that the salient residual decrease has been driven by one or more residual influences: utility, technology, policy, or something else proxied by time. It would indeed be difficult to further isolate the effects from these possible drivers, given the inherent difficulty in measuring these influences, and we leave this task to future research.

The remainder of the paper is organized as follows. Section 2 provides a short and general motivation of the panel nonparametric approach to modeling economic elasticities. In Section 3, we detail the construction and sources of our electricity panel. We present a basic benchmark model of electricity demand, discuss possible sources of coefficient instability in such a model, and introduce a functional coefficient panel model to better identify these sources. Our econometric methodology is described and empirical results are collected in Section 4, and we conclude with Section 5. Appendix A lists countries used to obtain our empirical results, and Appendix B discusses some additional technical details of the methodology and some ancillary empirical results.

⁴We will refer to the partial derivative as the *income elasticity*, or simply the *elasticity*, even though it reflects both an income elasticity and an output elasticity in sectorally aggregated data.

2 An Overview of Our Approach

Before introducing a specific model, we illustrate the basic idea of using panel data with a nonparametric approach by considering a single term given by $\beta(r, x)x$, perhaps of an otherwise linear model, where both the regressor x and the regressand $y(r, x)$ are in logs and r represents time. Because x is an argument of the coefficient $\beta(r, x)$, the coefficient is not equal to the partial derivative with respect to the regressor and therefore cannot be interpreted as an elasticity. The elasticity $\epsilon(r, x) = (\partial/\partial x)y(r, x)$, which is given by

$$\epsilon(r, x) = \beta(r, x) + x\beta_x(r, x), \quad (1)$$

where $\beta_x(r, x)$ refers to the partial derivative of $\beta(r, x)$ with respect to its second argument, clearly contains an additional term $x\beta_x(r, x)$. This term is zero only if the coefficient does not vary with the regressor – i.e., if the model is in fact linear in x – or, trivially, if the coefficient does not vary at all. We may estimate both $\beta_x(r, x)$ and $\beta(r, x)$ using our panel approach and compute the income elasticity $\epsilon(r, x)$ at each time r and x from (1).

In contrast, estimation approaches using individual countries cannot adequately identify the effects of both time and the regressor. With a single-country nonparametric approach, we have only one observation of x at any given time, and therefore, it is impossible to identify the effect of varying x while holding time fixed. In fact, if x is log GDP, as in our electricity consumption model below, there is an almost one-to-one relationship between time and log GDP, since log GDP has a dominant increasing linear time trend. Consequently, we may not be able to identify the time effect with the regressor fixed, either. Our approach encompasses the single-country approach, so we may relate elasticities estimated using our approach to those estimated in studies relying on time series data for a single country.

A time series (single-country) approach that allows the coefficient to vary only over time, such as that of Park and Hahn (1999), actually sets x as a function of time – i.e., $x = x(r)$ – and looks at the coefficient $\beta(r, x(r))$ as a univariate function of time r . In that case,

$$\frac{d}{dr}\beta(r, x(r)) = \beta_r(r, x(r)) + \dot{x}(r)\beta_x(r, x(r)), \quad (2)$$

where $\beta_r(r, x)$ refers to the partial derivative of $\beta(r, x)$ with respect to its first argument, and $\dot{x}(r) = (d/dr)x(r)$ denotes the instantaneous change in the regressor. As is evident from (2), the slope of the coefficient $(d/dr)\beta(r, x(r))$ in such a study does not truly represent the rate of time change $\beta_r(r, x(r))$ of the coefficient, which is identified in our approach separately from the secondary temporal effect due to the growth of x . Unless there is trivially no change in the regressor, so that $\dot{x}(r) = 0$, or unless the coefficient does not

depend explicitly on x and we have $\beta_x(r, x(r)) = 0$, the temporal effects are not identical. In our model below, in which x is log GDP, we expect the growth rate $\dot{x}(r)$ to be positive. Thus, when $\beta_x(r, x(r))$ is positive (negative), we expect the slope of the coefficient in a single-country study to be larger (smaller) than the rate of time change of our coefficient.

On the other hand, a single-country approach that allows the coefficient to vary only over the regressor is equivalent to setting time r as a function $r = r(x)$, say, of log GDP x . In this case, we have

$$\frac{d}{dx}\beta(r(x), x) = \dot{r}(x)\beta_r(r(x), x) + \beta_x(r(x), x), \quad (3)$$

similarly to (2), where $\dot{r}(x) = (d/dx)r(x)$ is the reciprocal of the regressor's growth rate. Such an approach allows coefficient changes *only* through regressor changes – i.e., non-linearity – but omits any temporal changes in the coefficient unrelated to the regressor itself. However, drivers of coefficient instability, such as technology and utility in the case of demand functions, do not need to relate specifically to the regressor (log GDP, e.g.), and therefore cannot be identified by such approaches. If $\beta_r(r, x)$ is positive for all r and x , as we find in our empirical analysis of electricity consumption, and if $\dot{r}(x)$ is positive for all x , we would expect the slope of the coefficient $(d/dx)\beta(r(x), x)$ estimated using a single-country approach to exceed the partial effect $\beta_x(r(x), x)$ estimated using our approach.

Neither of these single-country approaches can separately identify the two arguments of the coefficient. Further, we expect them to systematically underestimate or overestimate the rates of change in the coefficient with respect to its arguments. This bias may have a devastating effect on inferences or predictions from the model. In contrast, a panel approach allows us to create counterfactuals by fixing the regressor and allowing time to vary, identifying $\beta_r(r, x)$, or by fixing time and allowing the regressor to vary, identifying $\beta_x(r, x)$, at each and every combination of time r and regressor x . As a result, the change in elasticity is more accurately analyzed, the individual drivers of this change are better identified, and we may gain a much deeper understanding of the dynamics underlying the structural change.

3 Electricity Data, Models, and Sources of Instability

3.1 Data Sources and Construction

A model of electricity intensity requires at least two series: electricity consumption and GDP. Electricity consumption is measured implicitly by adding net exports to production and subtracting losses from transmission. Production data are available over a longer and

wider span than consumption,⁵ so we use production as a proxy for consumption. Because electricity must be transmitted by wire and large amounts are not transmitted under water, net exports are a very small percentage of production for most countries. According to data from the World Factbook,⁶ 20 of the 89 countries in our sample import or export more than 5% of production and only 7 import or export more than 20%. The Republic of Congo imports 84% of its electricity, but both imports and production are quite small. Paraguay exports 80% of its production, apparently to neighbors Brazil and Argentina. In fact, the absolute value of Paraguay's net trade is nearly the largest of any country's, but this is clearly an outlier.

We collect annual electricity production and GDP over the period 1971-2010 for 184 countries, although not all of these data are ultimately used. Electricity production in gigawatt hours (GWh) is used as a proxy for electricity demand, and almost all of these data originate from Enerdata,⁷ except those of a few countries with missing data for 1971 for which we use data from the World Bank. We calculate electricity production per capita using population in thousands. We then create an electricity production index with 2000 as the base year in order to eliminate some of the heterogeneity in per capita production across countries. GDP per capita is constructed from real GDP in 2005 million US dollars at constant purchasing power parities and using population in thousands. GDP per capita is thus expressed in thousands of 2005 US dollars.

We omit 62 countries with some missing electricity production or GDP data. Many of these are former Soviet bloc countries, for which data were missing or unreliable during the beginning of the sample. We then omit 33 countries that appear to have nonsensical (negative or statistically insignificant) cointegrating relationships between electricity production and GDP using a conventional fixed coefficient single-country model similar to the benchmark model (4) below. The countries in this group tend to be poorer countries with less developed markets, and the lack of a cointegrating relationship may reflect the lack of a long-run market equilibrium. After omitting these groups, we are left with 89 countries listed in Appendix A.

In order to ameliorate cross-country heterogeneity and generate more stable parameter estimates, we group countries according to GDP per capita in each year. The first group (Gr1) includes the US and any country with a higher GDP per capita. To form the remaining 9 income groups (Gr2-Gr10), we first compute the empirical distribution of the remaining

⁵Sectoral consumption data are available from the UN Energy Statistics and Enerdata. However, missing subsectoral data are treated as zeros in aggregating to the sectoral level, which generates substantial measurement error, especially as a series becomes missing or ceases to be missing over time.

⁶<https://www.cia.gov/library/publications/the-world-factbook/index.html>

⁷<http://yearbook.enerdata.net/>

Group	Percentiles of GDP per Capita
Gr2	[88,100]
Gr3	[77,88)
Gr4	[66,77)
Gr5	[55,66)
Gr6	[44,55)
Gr7	[33,44)
Gr8	[22,33)
Gr9	[11,22)
Gr10	(0,11)

Table 1: **Income Grouping Rules.** For each year, Gr1 is defined to be the US and countries with a higher GDP per capita than the US. Countries in Gr2-Gr10 are determined by the percentiles of the annual distribution of GDP per capita of the remaining countries.

countries. These countries are then assigned to 9 income groups according to the grouping rule given by the percentiles in Table 1. For each group in each year, we weight each country's GDP per capita by its GDP level as a ratio of the group's total GDP level, thus deriving an aggregate measure of GDP per capita for each of the 10 groups in each of the 40 years.

We use a similar procedure to weight the electricity production data, using exactly the same GDP-based weights. We thus have 400 total observations (40 years of 10 country groups) of electricity consumption (proxied by production) and GDP. In this way, the income represented by each group is relatively stable, even though the group members may change each year. The exception is that Gr1 contains the US by definition and is in fact dominated by the US, since the other members tend to be smaller wealthy countries, such as Bermuda and Singapore. Figure 1 shows the sample paths of the logs of the electricity consumption and GDP series.

Since both demand and supply of any good are functions of its price, energy prices are often used in models of energy consumption. Annual household electricity prices per kilowatt hour measured in 2005 US cents at constant purchasing power parities are available for a range of countries from Enerdata. We create a price index with 2000 as the base year. Using price necessitates a slightly different modeling strategy, because price data are available for only 25 of the 89 countries and no earlier than 1978 for these countries.

Table 4 (in Appendix A) shows the number of countries in each group and each year for which we have price data. We can see from the table that Gr1-Gr4 contain at least 2 of the 25 countries in each year. We therefore use the *same* groups defined above with the same group members for GDP and electricity production, but just omit Gr5-Gr10. We construct price

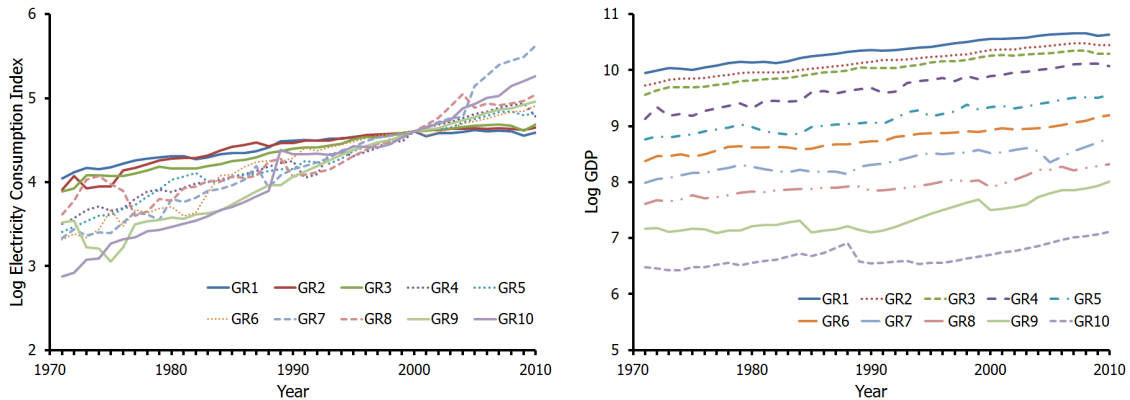


Figure 1: **Time Series Plots.** Time series plots of log electricity consumption indices with 2000 as the base year and log real GDP per capita for ten groups Gr1-Gr10.

data similarly to GDP and electricity production data for Gr1-Gr4, except that, because we have fewer group members with price, we re-weight using each country's GDP level as a ratio of the group total excluding those countries with no price data. We therefore have 132 total observations (33 years of 4 country groups) of electricity consumption (production), GDP, and price. The right panel of Figure 2 in Section 3.5 below shows the sample paths of the log price series for these groups.

3.2 A Benchmark Model of Total Electricity Consumption

A general model of the ratio of electricity consumption to GDP is of course linear when the data are expressed in logs. Assuming fixed coefficients, such a model for a single country that also includes price may be written as

$$y_t = \alpha + \beta x_t + \gamma p_t + \varepsilon_t, \quad (4)$$

for $t = 1, \dots, T$ years, where y_t , x_t , and p_t represent logs of the electricity consumption, GDP, and electricity price series discussed above. This fixed coefficient model may be interpreted as a model of electricity demand, in which GDP is a proxy for income, and β

may thus be interpreted as the income elasticity of electricity demand.^{8,9}

However, we note several obstacles to this interpretation of β . On the left-hand side, we use production data as a proxy for a country's demand. Recall that actual consumption is measured implicitly by adding net exports and subtracting transmission losses. As long as the omitted variables, net exports and transmission losses, do not have unit roots or deterministic trends, the regression is still cointegrating so that β is estimated consistently.

Further obstacles to associating β with an income elasticity may arise on the right-hand side from using GDP (output) as a proxy for income. By definition, GDP contains electricity sales as a final good and any net exports of electricity. We do not expect the percentage of GDP comprised of such sales to be very large for most countries.¹⁰

Countries with historically unreliable data may present a more subtle problem. Forecasters of GDP of autocratic countries, for example, have traditionally found electricity consumption to be a good predictor of GDP. Our estimates could be biased for countries in which electricity consumption is used – other than as a final good in the definition of GDP – to construct the publicly available measure of GDP. However, many of these countries have been excluded from our sample due to missing data.

Finally, we note again that β itself is not strictly an *income* elasticity, because electricity is consumed by both households and firms. As a proxy for income, GDP is an input into an aggregate household utility function, making the coefficient an income elasticity in the household sector. However, GDP is an output of an aggregate production function, of which energy is a factor of production, making the coefficient an output elasticity in the firm sector. Even within the firm sector, different technologies across firms induce different coefficients, and studies conducted at the sectoral level have uncovered substantial differences in elasticities across economic sectors within a given country. The coefficient β

⁸Such a model usually includes additional covariates that act as demand shifters. Going back to Halvorsen's (1978) monograph on energy demand and earlier, demand shifters have been used with varying degrees of success. Perhaps the most obvious ones that we omit are a measure of seasonality to account for electricity use in heating and cooling and the price of a close substitute. Since we examine annual data, such a seasonal measure is unnecessary. In a fairly homogeneous market, such as the US, natural gas or fuel oil may be the predominant substitute for electricity in heating applications. However, these data may not be as informative globally. With a wide range of countries and climates, a substantial amount of heating is unnecessary in some countries. Technological differences across time and across countries may further weaken the impact of the price of a heating substitute. Moreover, data availability in this panel precludes the use of such prices.

⁹The joint determination of electricity price and consumption means that γ cannot be interpreted as a price elasticity of demand, and that estimates of γ will likely be inconsistent. We do not expect such a problem for estimates of β , since there appears to be no co-movement between price and GDP in our data, as shown in Figure 2.

¹⁰The ratio is about 2% for the US, e.g., using data from the Energy Information Administration (EIA): <http://www.eia.gov/electricity/>; and the Federal Reserve Bank of St. Louis: <http://research.stlouisfed.org/fred2/>.

in a sectorally aggregated single-country model thus reflects all of these elasticities.

3.3 Coefficient Variation from Sectoral Shifts and GDP

A cost of using sectorally aggregated data is that, strictly speaking, the coefficient β is not a single elasticity as just mentioned. Aggregating data across sectors implies some type of aggregation of the coefficients for those sectors. Even if the coefficients for each sector are time-invariant, the aggregation weights may change as the focus of real economic activity shifts from one sector to another over time.

Medlock and Soligo (2001) explicitly tie sectoral shifts to economic development and argue for an econometric specification flexible enough to allow for nonlinearity in the relationship between energy demand and GDP. Denoting log GDP by x , those authors let $\beta = \beta(x) = \beta_0 + \beta_1 x$ so that $\beta(x)x$ is quadratic in x . The left-hand tail of the inverted U shape of the quadratic allows an increase in energy intensity at low income levels as countries build their industrial bases. In line with earlier authors, such as Brookes (1972), those authors emphasize the process of dematerialization, which drives sectoral dominance from energy-intensive heavy industry to light industry, and then eventually to the less energy-intensive commercial sector. Dematerialization is tied to high income levels, and the right-hand tail of the inverted U shape allows a decline in energy intensity as countries prosper. The nonlinearity in β for countries that have developed rapidly may be quite pronounced, as Galli (1998) notes for Korea and Taiwan.

Although GDP may be a good proxy for sectoral composition in some cases, temporal sectoral shifts do not have to be tied to economic development. For example, consider the sectoral reallocation of Eastern European countries in the 1990's following the demise of Soviet-style communism. From the starting point of high growth and an intensive industrial focus with little emphasis on energy efficiency under that system, the process of dematerialization was rapid in those countries that tried to adapt quickly, while the pace of economic development lagged behind.

More generally, increased globalization may drive countries to specialize in certain industries in order to compete. Sectoral shifts resulting from such specialization may indeed be correlated with economic development, such as in the case of specialization in manufacturing exports by Japan, Korea, Taiwan, and China. Alternatively, economic growth due to specialization might not be followed by dematerialization if a country does not diversify its main economic driver, as is often the case with exporters of basic resources, such as oil or minerals.

3.4 Coefficient Variation from Other Sources

Shifts in utility, technology, and energy policy may also drive coefficient change. In this case, a nonlinear function of GDP may not adequately capture such instability. Consider a residential consumer's choice between a cheaper durable that uses more electricity (lower fixed cost and higher total variable cost) and a more expensive durable that uses less electricity (higher fixed cost and lower total variable cost). Suppose that the expected total cost of each durable is equal, so that household income is irrelevant to the choice between them. For a given increase in household income after which such a choice becomes possible, the choice between which durable to consume – and thus the income elasticity of electricity demand – may be influenced by both technology and utility. Technology determines whether in fact such a trade-off is possible, while the diffusion of knowledge about the negative externalities associated with using exhaustible resources may influence utility.

In industry, consider the example of the rise of electric arc furnaces in steelmaking. While electric arc furnaces do not require the large amounts of coal needed for basic oxygen furnaces and decrease overall energy usage, they increase the need for electricity to obtain the same amount of output.¹¹ A further example is provided by increased computerization across all firm sectors, increasing productivity while most likely increasing reliance on electricity for power even while enabling more efficient usage of it.

Utility and technology are difficult to identify, because these terms refer to the functional forms and unknown parameters of the respective functions. In addition, energy policies have certainly affected how firms and households use electricity, thus influencing the elasticity. Policy is difficult to quantify beyond simple binary policy changes, and policies are not necessarily comparable across countries.

Because these additional drivers are difficult to measure, researchers often allow a function of time to capture residual influences on coefficient instability. For example, Webster *et al.* (2008) use a linear time trend to capture AEEL, although Kaufmann (2004) argues against such a specification. A more flexible alternative is offered by Chang and Martinez-Chombo (2003) and Chang *et al.* (2013), who model the coefficient β as a flexible function of time, such that $\beta = \beta(r)$ with r denoting time.

¹¹“Energy Trends in Selected Manufacturing Sectors: Opportunities and Challenges for Environmentally Preferable Energy Outcomes,” U.S. Environmental Protection Agency (March 2007): <http://www.epa.gov/sectors/pdf/energy/report.pdf>.

3.5 Econometric Model Specification

In order to segregate the influence of GDP from other time-varying factors, we consider a model in which $\beta = \beta(r, x)$ is a function of both GDP and time, as above. We refer to $\beta(r, x)$ as the *income coefficient of electricity demand*, or *income coefficient* for short. Specifically, we extend the time-series model in (4) to a panel model given by

$$y_{it} = \alpha' c_{it} + \beta(t, x_{it})x_{it} + \gamma p_{it} + u_{it}, \quad (5)$$

for $i = 1, \dots, N$ groups of countries, where c_{it} represents an $N \times 1$ vector of group binaries to capture cross-sectional heterogeneity, and α is an $N \times 1$ vector of group fixed effects. Because electricity price data pose a major constraint on data collection, it will also be useful to rewrite the model in (5) as

$$y_{it} = \alpha' c_{it} + \beta(t, x_{it})x_{it} + v_{it}, \quad (6)$$

where $v_{it} = \gamma p_{it} + u_{it}$.

Detailed empirical comparisons of these specifications are presented in Section 4. Note that the regressions introduced in (5) and (6) both involve variables having deterministic and stochastic trends and are interpreted as describing long-run relationships. In fact, tests for the presence of deterministic and stochastic trends, shown in Tables 5 and 6 in Appendix B, strongly and unambiguously show that (x_{it}) has both a stochastic trend and a linear time trend, while (p_{it}) has only a stochastic trend, for a majority of $i = 1, \dots, N$. Time series plots of (p_{it}) and linearly detrended (x_{it}) shown in Figure 2 provide additional evidence.

On the other hand, as shown in Figure 5 below, it appears that the fitted residuals from both regressions (5) and (6) have no time trends, and in particular, the series of fitted residuals from regression (5) is stationary. Therefore, we interpret regressions (5) and (6) as representing semiparametric long-run relationships with stationary and integrated errors respectively.¹² Clearly, regression (5) may be regarded as a semiparametric cointegrating regression if (u_{it}) is stationary. Regression (6) is also meaningful, since (x_{it}) has a linear trend, providing a stronger signal than the noise generated by an integrated error (v_{it}) . Although it is misspecified, we may estimate $\beta(r, x)$ consistently from regression (6).¹³

Unless $\gamma = 0$, we expect that regression (5) to provide a better estimate of $\beta(r, x)$ than

¹²Unfortunately, no formal test exists for the stationarity and integratedness of the error terms in regressions like (5) and (6).

¹³An integrated time series is of order \sqrt{T} , and it is therefore asymptotically negligible compared with a linear time trend growing at the rate of T . However, the parameter α cannot be estimated consistently in regression (6) unless $\gamma = 0$.

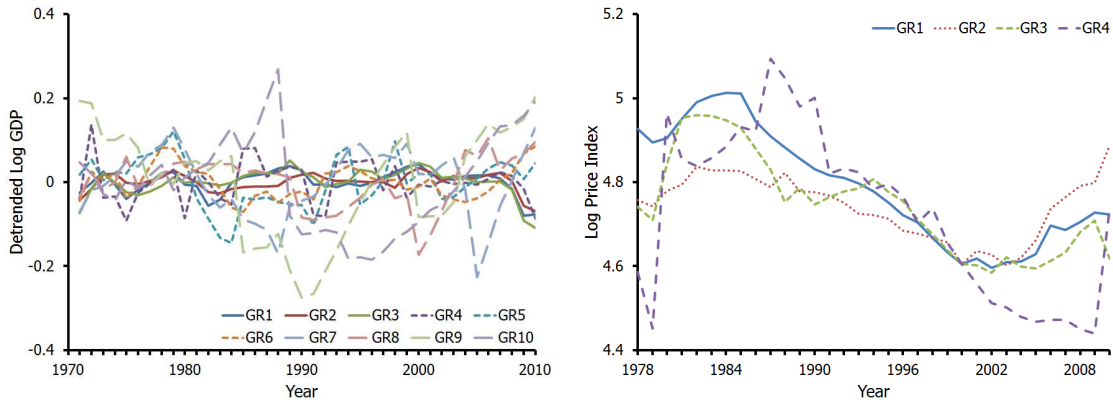


Figure 2: **Time Series Plots.** Time series plots of detrended log real GDP per capita for ten groups Gr1-Gr10 and log real electricity prices for four groups Gr1-Gr4, respectively, in the left and right panels.

regression (6). In our case, this is not necessarily true, since the observations we may use to fit regression (5) are only a fraction of the entire sample, due to the severely restricted availability of price data. Using regression (6) to estimate $\beta(r, x)$ incurs some bias. However, at the same time, we may drastically reduce the sample variance of the estimator of $\beta(r, x)$ by using observations on more years and many more countries and estimating $\beta(r, x)$ from regression (6).

Overall, regression (6) may yield an estimator of $\beta(r, x)$ with a smaller mean squared error than regression (5), if the reduction in sample variance due to the utilization of observations on more countries over a longer time span exceeds the magnitude of the squared bias resulting from the omitted variable problem in regression (6). Indeed, we show in Section 4.3 that regression (6) provides an improved estimator of $\beta(r, x)$ in terms of the bootstrap mean squared error.

4 Estimation Method and Empirical Results

4.1 Estimation Procedure

We estimate the nonlinear panel data models in (5) and (6) semiparametrically, allowing the most flexible form for the functional coefficient $\beta(r, x)$ and allowing us to glean substantial information about the dynamics of the functional coefficient using our aggregated but cross-sectionally diverse data. The regression model introduced in (5) has both a linear component $\alpha'c_{it} + \gamma p_{it}$ and a nonparametrically specified nonlinear component $\beta(t, x_{it})x_{it}$

with an income coefficient $\beta(t, x_{it})$ that is an unknown function of both time and GDP. The models and econometric methodologies used in the paper are mostly closely related to those developed by Cai (2007) and Cai *et al.* (2009). The essential methodology and intuition was developed by Robinson (1988) for the linear component and Stock (1989) for the nonparametric component. Interested readers are referred to Härdle (1992), Pagan and Ullah (1999), or Li and Racine (2007) for some general introductions and more detailed discussions.

To explain our estimation procedure, we introduce the weight function

$$W_{it}(r, x) = \frac{1}{h_r h_x} K\left(\frac{t-r}{h_r}\right) K\left(\frac{x_{it}-x}{h_x}\right), \quad (7)$$

where K is a kernel function and h_r and h_x are bandwidth parameters such that $h_r, h_x \rightarrow 0$ as $NT \rightarrow \infty$. The function $W_{it}(r, x)$ defines weights for time t and observations (x_{it}) of log GDP at each local point (r, x) , and the value of $\beta(r, x)$ is estimated from the regressions localized around each local point (r, x) of the argument of $\beta(r, x)$.

Bandwidth selection is very important in estimating any models involving nonparametric components such as ours. Unfortunately, there is no result available in the literature on optimal bandwidth selection for our nonstandard nonstationary model. In our estimation procedure, we therefore use the standard rates used for stationary models. More explicitly, we set $h_r = c_r(NT)^{-1/6}$ and $h_x = c_x(NT)^{-1/6}$ for some constants $c_r, c_x > 0$, and select the constants c_r and c_x by augmenting the cross validation method and the bias-corrected AIC approach implemented by Cai and Tiwari (2000) and Cai (2002). Details of our bandwidth selection procedure are provided in Appendix B.

Suppose for now that the values of the parameters α and γ are known. We could subtract the linear component from each side to rewrite model (5) as

$$y_{it} - \alpha' c_{it} - \gamma p_{it} = \beta(t, x_{it}) x_{it} + u_{it}, \quad (8)$$

and then estimate the income coefficient by

$$\tilde{\beta}(r, x) = \left(\sum_{i,t} x_{it}^2 W_{it}(r, x) \right)^{-1} \sum_{i,t} x_{it} (y_{it} - \alpha' c_{it} - \gamma p_{it}) W_{it}(r, x)$$

using the weight function $W_{it}(r, x)$ introduced in (7). Note that

$$\tilde{\beta}(r, x) = \hat{\beta}_y(r, x) - \alpha' \hat{\beta}_c(r, x) - \gamma \hat{\beta}_p(r, x), \quad (9)$$

where the components

$$\begin{aligned}\hat{\beta}_y(r, x) &= \left(\sum_{i,t} x_{it}^2 W_{it}(r, x) \right)^{-1} \sum_{i,t} x_{it} y_{it} W_{it}(r, x), \\ \hat{\beta}_c(r, x) &= \left(\sum_{i,t} x_{it}^2 W_{it}(r, x) \right)^{-1} \sum_{i,t} x_{it} c_{it} W_{it}(r, x), \text{ and} \\ \hat{\beta}_p(r, x) &= \left(\sum_{i,t} x_{it}^2 W_{it}(r, x) \right)^{-1} \sum_{i,t} x_{it} p_{it} W_{it}(r, x)\end{aligned}$$

are simply the functional coefficient estimates from three separate nonparametric regressions of (y_{it}) , (c_{it}) , and (p_{it}) on (x_{it}) alone. Note that while $\tilde{\beta}(r, x)$ is infeasible, because the values of parameters α and γ are indeed unknown, $\hat{\beta}_y(r, x)$, $\hat{\beta}_c(r, x)$, and $\hat{\beta}_p(r, x)$ are feasible estimators of the corresponding functional coefficients $\beta_y(r, x)$, $\beta_c(r, x)$, and $\beta_p(r, x)$ in these three separate nonparametric regressions.

We replace the functional coefficient $\beta(t, x_{it})$ in (8) with $\tilde{\beta}(t, x_{it})$ in (9) at $(r, x) = (t, x_{it})$ to obtain

$$y_{it} - \hat{\beta}_y(t, x_{it})x_{it} = \alpha'(c_{it} - \hat{\beta}_c(t, x_{it})x_{it}) + \gamma(p_{it} - \hat{\beta}_p(t, x_{it})x_{it}) + \tilde{u}_{it}, \quad (10)$$

where $\tilde{u}_{it} = u_{it} - (\tilde{\beta}(t, x_{it}) - \beta(t, x_{it}))x_{it}$. Under general regularity conditions, we may expect that the error incurred by replacing the true functional coefficient $\beta(t, x_{it})$ with its estimate $\tilde{\beta}(t, x_{it})$ is asymptotically negligible, since all functional coefficient estimators $\hat{\beta}_y(r, x)$, $\hat{\beta}_c(r, x)$ and $\hat{\beta}_p(r, x)$ consistently estimate the corresponding functional coefficients. As a result, we may consistently estimate the parameters α and γ by their least squares estimators $\hat{\alpha}$ and $\hat{\gamma}$ in regression (10), similarly to the time-series case considered by Cai *et al.* (2009).

With $\hat{\alpha}$ and $\hat{\gamma}$, we may consistently estimate the income coefficient by

$$\hat{\beta}(r, x) = \hat{\beta}_y(r, x) - \hat{\alpha}'\hat{\beta}_c(r, x) - \hat{\gamma}\hat{\beta}_p(r, x), \quad (11)$$

which is a feasible version of the infeasible estimator of the income coefficient in (9). It follows that

$$\begin{aligned}\hat{\beta}(r, x) &= \left(\sum_{i,t} x_{it}^2 W_{it}(r, x) \right)^{-1} \sum_{i,t} x_{it} (y_{it} - \hat{\alpha}'c_{it} - \hat{\gamma}p_{it}) W_{it}(r, x) \\ &= \beta(r, x) + \left(\sum_{i,t} x_{it}^2 W_{it}(r, x) \right)^{-1} \sum_{i,t} x_{it} (u_{it} - (\hat{\alpha} - \alpha)'c_{it} - (\hat{\gamma} - \gamma)p_{it}) W_{it}(r, x),\end{aligned}$$

and we may expect $\hat{\beta}(r, x)$ to estimate $\beta(r, x)$ consistently under suitable regularity condi-

tions.¹⁴

The estimator is intuitively similar to the familiar method of partitioned regression. In essence, we first “condition out” the nonparametric component involving GDP, treating it as a nuisance term in estimating α and γ , as Robinson (1988) did. With consistent estimators of the latter in hand, the nonparametric estimation of the nonlinear component is then easily accomplished in the second step, as Stock (1989) did.

The estimation procedure for model (6) without price is identical in every respect, except (p_{it}) is an omitted variable. The procedure is modified so that

$$y_{it} - \hat{\beta}_y(t, x_{it})x_{it} = \alpha'(c_{it} - \hat{\beta}_c(t, x_{it})x_{it}) + \tilde{v}_{it} \quad (12)$$

is estimated in the first step, where $\tilde{v}_{it} = v_{it} - (\tilde{\beta}(t, x_{it}) - \beta(t, x_{it}))x_{it}$ with $v_{it} = \gamma p_{it} + u_{it}$, as defined above, and then

$$\hat{\beta}(r, x) = \hat{\beta}_y(r, x) - \hat{\alpha}'\hat{\beta}_c(r, x). \quad (13)$$

Similarly to the procedure for model (5), we may write

$$\hat{\beta}(r, x) = \beta(r, x) + \left(\sum_{i,t} x_{it}^2 W_{it}(r, x) \right)^{-1} \sum_{i,t} x_{it} (v_{it} - (\hat{\alpha} - \alpha)'c_{it}) W_{it}(r, x),$$

and the second step will consistently estimate the income coefficient.¹⁵

With a consistent estimator of the income coefficient in hand, estimating the elasticity $\epsilon(r, x)$ requires a nonparametric derivative estimator of $\beta_x(r, x)$ in (1). We may estimate this derivative using

$$\hat{\beta}_x(r, x) = \frac{1}{2\delta} (\hat{\beta}(r, x + \delta) - \hat{\beta}(r, x - \delta))$$

with the choice of some small δ .¹⁶ The elasticity $\epsilon(r, x)$ may then be estimated by

$$\hat{\epsilon}(r, x) = \hat{\beta}(r, x) + x\hat{\beta}_x(r, x) \quad (14)$$

for any combination of r and x .

¹⁴Note that our models are nonstandard in the sense that they include nonstationary variables and represent long-run relationships. As a consequence, existing theories regarding functional coefficient models, such as those in Fan and Huang (2005) and Section 9.3.4 of Li and Racine (2007), are not applicable to our models. We believe that consistency can be formally established following Cai *et al.* (2009). However, a detailed proof is beyond the scope of this paper.

¹⁵As discussed above, our estimation procedure is expected to yield a consistent estimator $\hat{\beta}(r, x)$ of $\beta(r, x)$ in regression (6), as well as in regression (5), even though we have an omitted variable problem in regression (6). The omitted variable only affects the consistency of $\hat{\alpha}$.

¹⁶Specifically, we set $\delta = 0.01$.

	With Price		Without Price	
	estimate	s.e.	estimate	s.e.
γ	-0.114	0.055		
α_1	-2.232	1.479	1.217	1.745
α_2	-2.092	1.460	1.276	1.714
α_3	-2.063	1.446	1.284	1.692
α_4	-1.781	1.451	1.338	1.652
α_5			1.563	1.623
α_6			1.669	1.580
α_7			1.836	1.557
α_8			1.972	1.446
α_9			1.961	1.385
α_{10}			2.079	1.324

Table 2: **First-step Results.** Least squares parameter estimates from regressions given by (10), with prices included but Gr5-Gr10 excluded, and by (12), with prices excluded but Gr5-Gr10 included. Standard errors calculated using the bootstrap method described in Appendix B.

4.2 Estimated Income Coefficients

Recall that the first step of the partially linear procedure entails estimating α and γ . Table 2 shows the estimation results for the parametric parts of the models with and without price, by running least squares on (10) (with price) for model (5) and on (12) (without price) for model (6). Note that our models are nonstationary and that standard results for stationary models are not applicable. We therefore calculate asymptotic variances using a bootstrap method described in detail in Appendix B. All parameters in the linear part of model (5) appear to be estimated with reasonably good precision. The estimated price elasticity $\hat{\gamma}$ of electricity demand is negative and significant, and the estimates of the group effects are in general not negligible. In regression (6), the standard errors for the group effects have mostly larger standard errors and are generally insignificant, which is expected from the inconsistency of the estimator.

The top two panels of Figure 3 show the income coefficient estimates from model (5) with price. The top left panel shows income coefficient estimates for the exact data points observed (Gr1-Gr4 over 33 years, $t = 1978, \dots, 2010$) while the top right panel shows estimates evaluated over a grid of log real GDP per capita ranging from 9.3 (\$10,938) to 10.7 (\$44,356). The income coefficient estimates using model (6) without price are plotted in the bottom two panels of the figure, using all 10 country groups and $t = 1971, \dots, 2010$. In this case, the bottom right panel shows estimates evaluated over a grid of log real GDP per capita ranging from 6 (\$403) to 11 (\$59,874).

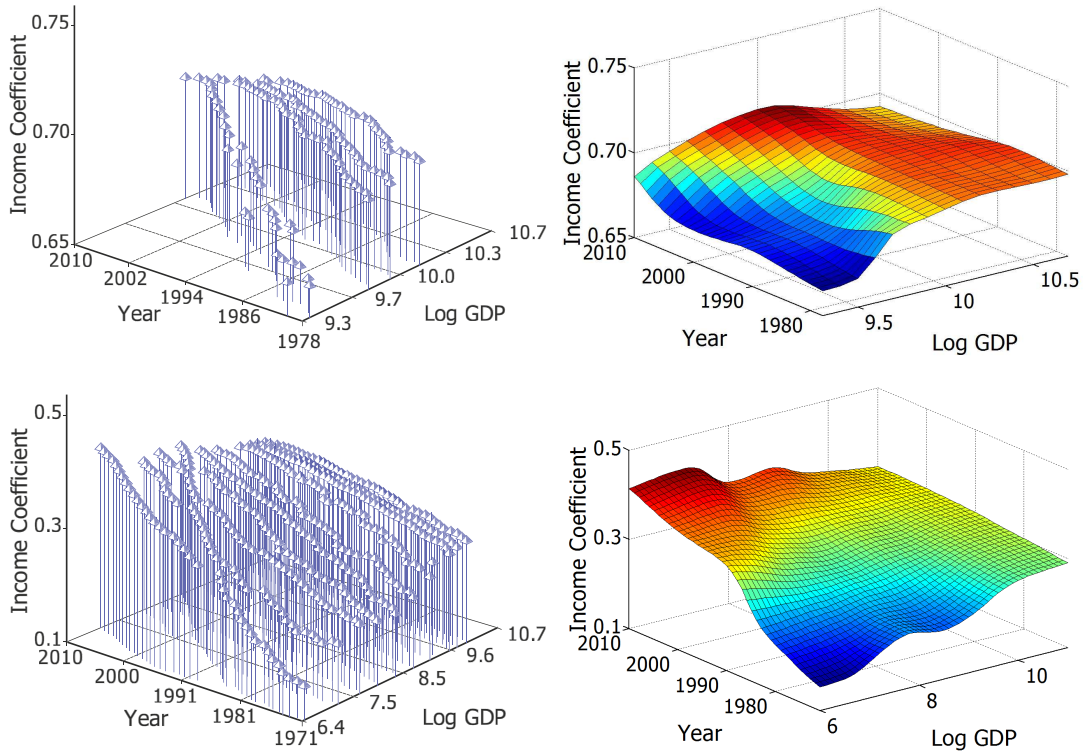


Figure 3: **Surface of Income Coefficient Estimates.** Surface plots of estimated income coefficient using (11) with price for model (5) (top panels) and (13) without price for model (6) (bottom panels), plotted at actual data (left panels) and interpolated over a grid (right panels). Actual data are Gr1-Gr4 over 1978-2010 (top left) and Gr1-Gr10 over 1971-2010 (bottom left). Grid points range from 9.3 (\$10, 938) to 10.7 (\$44, 356) using data from Gr1-Gr4 (top right), and from 6 (\$403) to 11 (\$59, 874) using data from Gr1-Gr10 (bottom right).

The most salient feature of the income coefficient estimates from either model (5) or (6) is that they are not constant functions of time and GDP, as a traditional fixed coefficient model assumes. The income coefficient estimates in the models with and without price appear to be flatter for richer countries than for poorer countries, emphasizing the importance of a sample with diverse GDP levels to evaluate the coefficient dynamics over time and GDP. Comparing the bottom left panel with the top left panel makes clear the increase in sample size from dropping price. Comparing the smoothness of the bottom right panel with the top right panel gives some idea of the efficiency gain from the larger sample.

Though not exactly the same, the topologies of the income coefficient estimates from regressions (5) and (6) are qualitatively very similar. On the other hand, the estimates from regression (6) based on a larger data set identify the income coefficient over a wider range

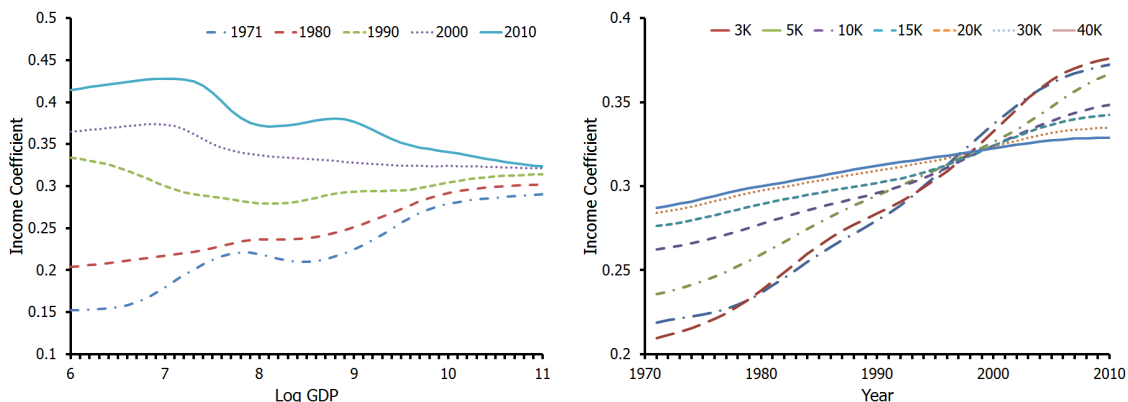


Figure 4: **Selected Income Coefficient Estimates.** Cross sections of income coefficient estimates at selected years and time series of income coefficient estimates at selected income levels, estimated using (13) without price.

of time and GDP. Furthermore, the estimates we obtain from regression (6) show a clearer pattern of variation of the income coefficient over the time and GDP, compared to the estimates from regression (5). Although a comparison of the vertical axes suggests that a scalar bias may exist from estimating the model without price, our bootstrap simulations in the next section conclusively demonstrate, to the contrary, that the severe restriction on the sample imposed by using price data generates a much larger variance. The evident vertical difference thus results more from variance than from bias, and the small bias incurred by using model (6) is therefore more acceptable.¹⁷

In order to examine the temporal patterns in the coefficients more closely, Figure 4 shows a two-dimensional representation of the bottom right panel of Figure 3 for fixed years and levels of GDP. Specifically, the left panel of Figure 4 illustrates $\hat{\beta}(r, x)$ holding time r fixed, while the right panel illustrates $\hat{\beta}(r, x)$ holding log GDP x fixed.

Holding time r constant, $\hat{\beta}(r, \cdot)$ appears to be mostly increasing in GDP during the 1970's and 1980's ($\hat{\beta}_x(r, \cdot) > 0$), but mostly decreasing in GDP during the 1990's and 2000's ($\hat{\beta}_x(r, \cdot) < 0$). In light of the role played by this term in equations (1) and (2), the elasticities and temporal change in elasticities could be quite a bit larger or smaller than those previously estimated using more restrictive models.

The coefficients appear to be increasing over time with GDP held constant: $\hat{\beta}_r(r, x) > 0$ for each (r, x) . This increase may reflect global diffusion of electronic technology in all

¹⁷The level of the income coefficient does not play any important role in our discussions in the paper, since our objective is to analyze the *changing patterns* in income elasticities of electricity demand, not their absolute levels.

sectors over the last four decades. In every country, both households' and firms' reliance on machinery and appliances that require electricity has increased. An increase in the income of a household in 1971 might have induced a large purchase of a durable good that did not require much electricity, resulting in a relatively small coefficient. On the other hand, an increase in the income of a household in 2010 might have induced a large electronics purchase, making that household's demand for electricity more sensitive to income changes and resulting in a relatively larger coefficient.

The increase over time appears to be much steeper for poorer countries than for richer countries. As relatively cheap electronic technology diffuses to poorer countries, their households demand more electricity to power these devices and their firms require more electricity to be globally competitive. Especially in markets in which the electricity price is heavily regulated and electronics are imported from abroad, we may expect the ratio of electricity usage relative to GDP per capita of these countries to grow disproportionately faster than that of richer countries.

4.3 Evaluation of Model Specification

Before making further inferences from our model, we analyze the fitted residuals from regressions (5) and (6) in order to investigate the adequacy of our model specifications and validity of these inferences. We interpret regression (5) as representing a cointegrating relationship, assuming that it has a stationary error term (u_{it}). Given that the omitted covariate (p_{it}) follows a unit root process, we do not have such a cointegrating relationship in regression (6).

As noted above, however, we expect that regression (6) also provides a consistent estimate of the income coefficient, since the covariate (x_{it}) has a dominant increasing trend and is of a bigger asymptotic order than that of the omitted (p_{it}). Of course, if (x_{it}) does not have a time trend, regression (6) becomes spurious and yields nonsensical results, as is well known. Unfortunately, there is no formal statistical procedure in the literature that we may use to test our claims more rigorously. We rely only on some qualitative regression diagnostics.

Figure 5 presents the fitted residuals from regressions (5) and (6). The fitted residuals (\hat{u}_{it}) from regression (5) (with price) generally show some persistence, though their sample paths are not what we typically expect for unit root processes.¹⁸ However, the magnitudes of the residuals (\hat{u}_{it}) are by far smaller than those of (x_{it}) and (p_{it}) (see Figures 1 and 2,

¹⁸Testing for a unit root using (\hat{u}_{it}) does not work, since the limit distribution of the unit root test in this case is affected by the limits of covariates, among other things. Also, there is no obvious way to implement any resampling method to compute the critical value of the test.

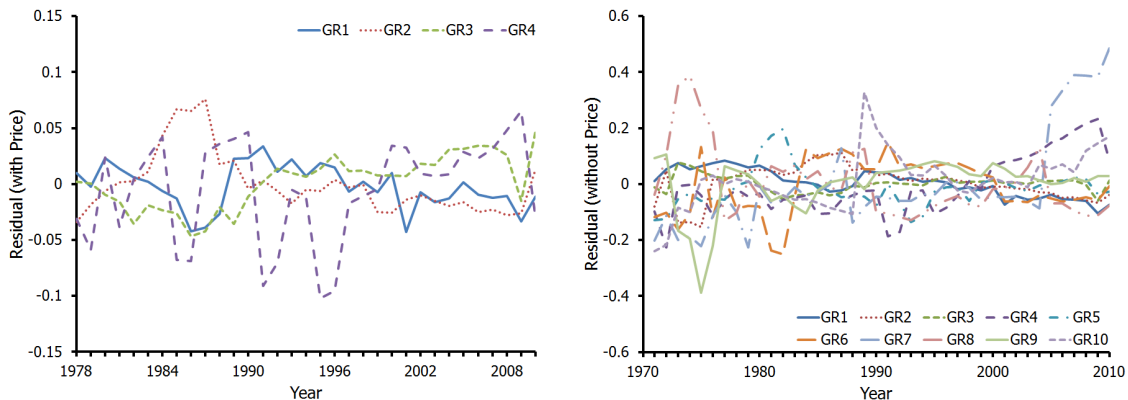


Figure 5: **Fitted Residuals.** Fitted residuals (\hat{u}_{it}) and (\hat{v}_{it}) from regressions (5) and (6), respectively, in the left and right panels.

respectively). Also, the persistence of (\hat{u}_{it}) in Figure 5 appears to be much weaker than that of (p_{it}) in Figures 2. These observations support, at least qualitatively, our specification of regression (5) as a cointegrating regression. The magnitudes of the fitted residuals (\hat{v}_{it}) from regression (6) are generally much bigger than those of (\hat{u}_{it}) from regression (5). Furthermore, for some groups, their time series paths show some strong persistence, as expected. Nevertheless, the range of (\hat{v}_{it}) is substantially smaller than that of (x_{it}) for most groups (see Figure 1). This suggests that (x_{it}) is indeed an order of magnitude bigger than (v_{it}) , which in turn implies that regression (6) is not spurious.

We evaluate the gains and losses in estimating the income coefficient $\beta(r, x)$ from regression (6), using a larger data set that includes observations for more groups spanning a longer period of time, in place of regression (5). For the estimation of the income coefficient $\beta(r, x)$, we expect the estimator from model (6) to yield a bigger bias due to the omitted variable but a smaller variance due to many more observations. In what follows, we compute and compare the bootstrap mean squared errors of two estimators, which can be decomposed into the bootstrap squared bias and bootstrap variance given respectively by

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left(\mathbb{E}^* \hat{\beta}^*(r, x) - \hat{\beta}(r, x) \right)^2 w(r, x) dr dx \quad (15)$$

and

$$\mathbb{E}^* \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left(\hat{\beta}^*(r, x) - \mathbb{E}^* \hat{\beta}^*(r, x) \right)^2 w(r, x) dr dx, \quad (16)$$

where $\hat{\beta}(r, x)$ and $\hat{\beta}^*(r, x)$ signify the sample and bootstrap estimates for the income coefficient $\beta(r, x)$, \mathbb{E}^* signifies the bootstrap expectation, and $w(r, x)$ signifies a weight function.

To evaluate the relative performance of the two regressions in terms of biases for their income coefficient estimates, we (i) define $\hat{\beta}(r, x)$ to be the sample estimate of $\beta(r, x)$ from regression (5) obtained using the smaller sample over the restricted time span, (ii) generate bootstrap samples from the fitted regression of (5), and (iii) compare the bootstrap squared biases of the two bootstrap estimates $\hat{\beta}^*(r, x)$ from regressions (5) and (6).¹⁹ We use the estimated joint density of (r, x) as our weight function.

Obviously, we expect the bootstrap estimate $\hat{\beta}^*(r, x)$ from regression (5) to have a smaller squared bias than that from regression (6). Indeed, the bootstrap squared bias from regression (6) is 5.1×10^{-3} , while that from regression (5) is 1.5×10^{-5} . Though the omission of (p_{it}) yields some bias, its magnitude appears to be small relative to the coefficient estimates in Figure 3. Moreover, the resulting bias is mainly caused by a vertical shift in the estimate of income coefficient, which does not affect the validity of our subsequent discussions on changing patterns in income elasticities. Therefore, we claim that the bias incurred by omitting (p_{it}) is not critical in our study.

The comparison of bootstrap variances using the smaller and larger samples is based on (16). For this comparison, we (i) generate two sets of bootstrap samples based on the fitted regression (6), one with the size of the smaller sample and the other with the size of the larger sample, and (ii) compare the bootstrap variances for the estimates from these two sets of bootstrap samples of different sizes. For each of the smaller and larger samples, the computed bootstrap variances are given respectively by 0.0763 and 0.0317. The reduction in sample variance by using the larger sample is clearly evident and substantial.

We conclude that the restricted regression in (6) with the larger sample is strongly preferred to the unrestricted regression in (5) with the smaller sample. With this comparison in mind, we henceforth restrict our attention to model (6) (without price).

4.4 Identifying the Effects of GDP and Time

Using the formula in (14) to calculate elasticities, the left panel of Figure 6 shows the time paths of the elasticities of the four groups with the highest GDP's per capita. Even though we aggregated individual countries into groups in order to avoid some heterogeneity in estimation, we can evaluate the elasticity of an individual country with non-missing GDP data. The right panel of Figure 6 shows the time paths of the elasticities of four countries: China, Korea, Japan, and the US.

The most obvious pattern that emerges suggests that elasticities of developed countries have been declining over time. Indeed, this finding reflects that of Brookes (1972), *inter*

¹⁹We use the two-stage bootstrap procedure described in Appendix B with the correction of the initial-ization bias computed in the first stage for both regressions.

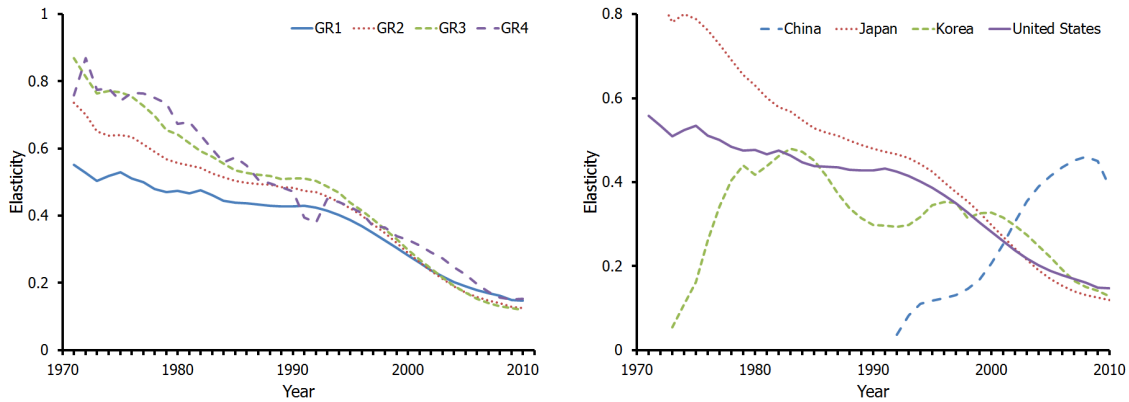


Figure 6: **Elasticities of Four Groups and Four Countries.** Estimated elasticities of four richest income groups: Gr1-Gr4, and four selected countries: China, Japan, Korea, and the US. Elasticities estimated using (14), with $\beta(r, x)$ and $\beta_x(r, x)$ estimated using (13) without price.

alia. Looking at individual countries, Japanese and US elasticities exhibit declines similar to those observed for the richest four groups. Indeed, these two countries are members of these groups throughout the sample.

Korea and China both clearly exhibit over *time* the inverted U shape over *GDP* discussed above. Korea has only been a member of the top five groups since 1983, after which its time path appears similar to that of the US and Japan. At about the same time, there is a clear switch from an increase to a decrease, suggesting dematerialization of its economy.

A similar pattern appears for China. Because it was so poor prior to 1991, we do not estimate an elasticity in order to avoid an empty bin problem for the nonparametric estimator. Since then, China's elasticity appears to have increased, surpassing those of the other three countries after 2002, and then appears to decline slightly by the end of the sample. The Chinese pattern appears similar to the Korean pattern, only shifted by about three decades. Although China's immense population makes it still quite poor on a per capita basis, a subset of its population is clearly not: Hong Kong moved into Gr1 by becoming richer per capita than the US in 2010.

The panel analysis allows us to construct meaningful counterfactuals, because for a fixed time period we estimate a range of coefficients that vary with GDP, and for a fixed country or GDP level we estimate a range of coefficients that vary over time. Figure 7 shows the elasticities over income groups but for fixed years (left panel) and over time but for fixed incomes (right panel).

Looking at the left panel, we clearly observe prior to the 1990's the inverted U shape noted by previous authors, with its peak shifting over time from Gr3, down to Gr4 and

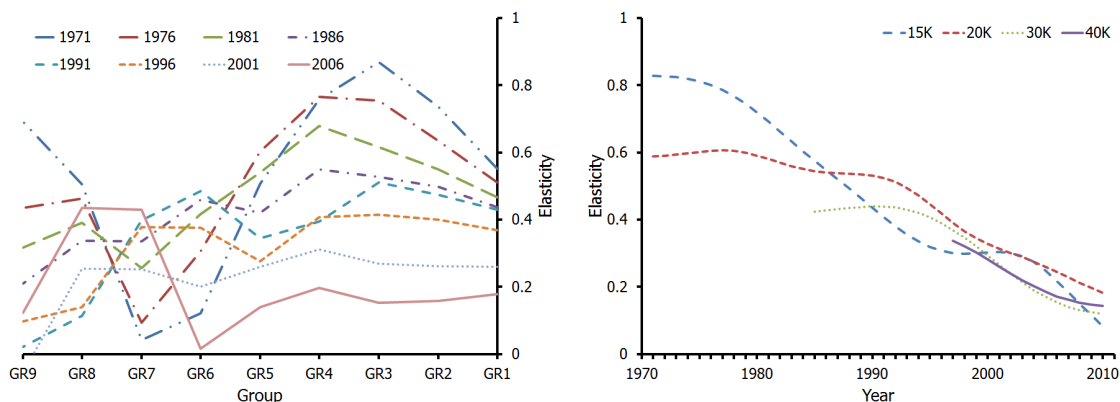


Figure 7: **Selected Elasticities.** Cross sections of elasticity estimates at selected years and time series of elasticity estimates at selected income levels. Elasticities estimated using (14), with $\beta(r, x)$ and $\beta_x(r, x)$ estimated using (13) without price.

below. The inverted U shape appears to break down by about 1991 and even inverts to an *uninverted* U shape by 2006, when the elasticities appear to *decrease* from Gr8 until about Gr6. This inversion may reflect the diffusion of electronic technology made possible both by the proliferation of such technology and by the internationalization of trade patterns during this period.

The elasticities are nearly flat across the four richest groups Gr1-Gr4 since about 1996. Such a pattern suggests that once a country attains a threshold level of development, GDP no longer seems to play an important role in determining the elasticity. Temporal patterns unrelated to GDP seem to matter much more for countries in these groups.

Median countries (primarily those in North Africa, Central America, and poorer countries in Asia and Europe) have less wealthy households, and firms use a larger share of electricity. As these countries industrialize, more sensitive industrial demand plays a larger role until household demand catches up, which seems to happen at about the income level of Gr4, which includes at different times Ireland, Israel, Korea, Mexico, and Portugal, among others.

Having effectively ruled out GDP as a major driver of electricity intensity in developed countries, we now examine the right panel of Figure 7. We do not plot years for which we have either no data or only data for Bermuda in excess of \$30,000 and \$40,000 in order to avoid biasing estimates by overweighting a small and atypical country that relies on only a few economic sectors. Specifically, the beginnings of these plots, 1985 and 1997 respectively, correspond to the real GDP per capita of Norway surpassing these thresholds. (That of the US surpasses \$30,000 in 1988 and \$40,000 in 2004.) With GDP fixed, decreasing temporal

patterns clearly remain. The general decline suggests the importance of alternative drivers of the coefficient instability.

It is interesting to compare the elasticities against the backdrop of the Kyoto Protocol, which was signed in 1997 and went into effect in 2005, marking a major milestone in awareness of the negative externalities associated with consuming fossil fuels. We expect such awareness to reduce the income elasticity, as richer countries may choose to pollute less, following the downward slope of a hypothetical environmental Kuznets curve. According to the left panel of Figure 7, such a decrease appears *prior* to 1996, when the downward slope in the elasticity with respect to income declined over time due to a decrease in the peak elasticity across income groups. In this light, we may perhaps interpret the Kyoto Protocol not so much as a binding agreement to shape future environmental sensitivity, but rather as the culmination of shifting attitudes up to that time.

Both panels of Figure 6 and the right panel of 7 suggest a different interpretation, however. The decline in many of the elasticities presented in these figure steepens during the 1990's. The fact that the decrease remains when holding income fixed (right panel of Figure 7) suggests that there is more to the story than an environmental Kuznets curve. Clearly, other temporal factors matter, and the timing of the evident break around the time of the signing of the Kyoto Protocol is very suggestive that sensitivity of electricity consumption to increases in GDP could be related to that agreement. Note that Korea reversed an increase in its elasticity in about 1997 and its subsequent decrease steepened in about 2005, suggesting that – at least in the case of Korea – the signing and enforcement of the Kyoto Protocol may have indeed affected future (as well as reflecting past) environmental awareness.

The right panel of Figure 6 contains elasticities of two countries for which the constraints of the Kyoto Protocol are not binding: China is exempt and the US did not ratify the agreement. To the extent that we can attribute the decline in elasticities to changing attitudes about emissions, there is a huge difference between these two countries. As a developing country, China continues to increase electricity intensity of its economy. Similarly to Korea, Japan, and Gr2-Gr4, the US shows a substantial change at about the same time, following a decline almost as steep as Korea and Japan. The decline in elasticities for Korea, Japan, and Gr2-Gr4 were such that the elasticities fell below that of the US by 2008, when emissions targets became binding. Interestingly, the slope of Japan's elasticity appears to flatten out shortly before the end of the sample, leading up to Japan's announcement shortly thereafter (in 2011) that it would not engage in additional second-round emissions cuts.

The proliferation of electronic devices (computers, cell phones, etc.) suggests an alternative explanation for the steepening of the decline in the elasticities of developed countries in

Figures 6 and 7 that started in the 1990's. Use of such devices has skyrocketed over time. The increased usage has perhaps rendered demand for electricity to power these devices less sensitive to income and has perhaps sped the process of dematerialization in developed countries by decreasing reliance on the industrial sector for economic growth. Of course, these two explanations, the Kyoto Protocol and proliferation of electronics, need not be mutually exclusive. Additional research would be required to further disentangle them.

4.5 Identifying the Effects of Sectoral Shifts

Previous authors tied sectoral shifts to GDP by way of sectoral dominance at different stages of economic development. However, sectoral shifts alone may drive changes in the aggregate elasticity, since different economic sectors do not generally have the same elasticities. For countries with reliable sectoral data, we may test the null that sectoral shifts explain the apparent decrease in the estimated elasticities by regressing them onto the sectoral shares and a simple time trend. The null corresponds to a zero coefficient on the time trend, while the one-sided alternative on that coefficient is negative. Although the test may have low power against more complicated temporal patterns, we expect that the test will distinguish trends as striking as the negative trends depicted in Figures 6 and 7.

The regression we run to test this hypothesis is given by

$$\hat{\epsilon}_t = \pi_o(t/T) + \pi_r s_t^r + \pi_c s_t^c + \pi_n s_t^n + \varepsilon_t$$

where $\hat{\epsilon}_t = \hat{\epsilon}(t, x_t)$ denotes the fitted values of the elasticity $\epsilon(t, x_t)$ from model (6) for a particular country,²⁰ and s_t^r , s_t^c , and s_t^n denote shares of electricity consumption by the residential, commercial, and industrial sectors of that country. Letting $\pi = (\pi_o, \pi_r, \pi_c, \pi_n)'$, its least squares estimator $\hat{\pi}$ has a limiting distribution given by

$$\sqrt{T}(\hat{\pi} - \pi) \rightarrow_d \mathbb{N}(0, (\text{plim} S_T)^{-1} \Omega (\text{plim} S_T)^{-1})$$

under general conditions, where $S_T = T^{-1} \sum_{t=1}^T s_t s_t'$, $s_t = (t/T, s_t^r, s_t^c, s_t^n)'$, and Ω is the asymptotic variance of $T^{-1/2} \sum_{t=1}^T s_t \varepsilon_t$ - i.e., $T^{-1/2} \sum_{t=1}^T s_t \varepsilon_t \rightarrow_d \mathbb{N}(0, \Omega)$. A consistent estimator of Ω can be easily obtained from a consistent estimator of the long-run variance of $(s_t \varepsilon_t)$.^{21, 22}

²⁰Here and subsequently, we suppress the subscript i for notational simplicity, since this part of our analysis is based on a single country.

²¹We use a standard long-run variance estimator with a triangular window and a bandwidth chosen by Andrews' (1991) automatic bandwidth selection procedure with a maximum of four lags.

²²A consistent estimator of Ω can also be obtained from a consistent estimator of the long-run variance Ω° of $(s_t^\circ \varepsilon_t)$, $s_t^\circ \varepsilon_t = (\varepsilon_t, s_t^r \varepsilon_t, s_t^c \varepsilon_t, s_t^n \varepsilon_t)'$. Indeed, if we let $\Omega = (\omega_{ij})$ and $\Omega^\circ = (\omega_{ij}^\circ)$, then we have $\omega_{11} = \omega_{11}^\circ/3$,

Belgium 1990-2010		Denmark 1990-2010		France 1990-2010		Germany 1990-2010	
estimate	s.e.	estimate	s.e.	estimate	s.e.	estimate	s.e.
-0.519***	0.022	-0.394***	0.067	-0.466***	0.032	-0.508***	0.040
0.838***	0.219	1.547***	0.599	1.127**	0.484	0.712**	0.320
1.066***	0.096	0.024	0.806	0.532	0.487	0.914**	0.376
0.253	0.162	-0.022	0.291	0.235**	0.105	0.154	0.175
Japan 1971-2010		Korea 1979-2010		United Kingdom 1990-2010		USA 1973-2010	
estimates	s.e.	estimates	s.e.	estimates	s.e.	estimates	s.e.
-0.794***	0.115	-0.912***	0.102	-0.574***	0.013	-0.230*	0.163
0.078	0.534	0.554	1.088	-0.011	0.244	-1.878***	0.534
1.804***	0.644	2.924***	0.637	2.821***	0.385	1.795*	0.951
0.977***	0.141	0.044	0.184	-0.622***	0.132	1.766***	0.567

Table 3: **Share Regressions: Selected Countries.** Regression of estimated elasticities on time trend and sectoral shares. Statistical significance shown for one-sided test of $H_0 : \pi_o = 0$ and two-sided tests of $H_0 : \pi_j = 0$ for $j = r, c, i$, denoted by *** for 1% significance, ** for 5% significance, and * for 10% significance.

Table 3 shows the results of the regressions for eight countries with reliable sectoral consumption data over sub-periods of our main sample.²³ All eight countries show $\hat{\pi}_o < 0$, and significantly so. Holding sectors constant, there still appears to be a downward trend in elasticities, and we reject the null of no trend against a downward trend for all eight of these countries.²⁴

We note that $|\pi_o|$ for the US is smaller but with a larger standard error than for the other countries. Figure 6 provides a possible explanation. Electricity intensity declined more slowly in the US than in Japan and in Gr2-Gr4. The 1990's marked a substantial change in the decrease in all of these elasticities, but the break seems more pronounced for

$\omega_{1j} = \omega_{j1} = \omega_{1j}^o/2$ for all $j \neq 1$, and $\omega_{ij} = \omega_{ij}^o$ for all $i, j = 2, 3, 4$. However, such an indirect estimator of Ω is not necessarily positive definite, and appears to work significantly worse in finite samples.

²³EuroStat (European Union): <http://epp.eurostat.ec.europa.eu/>; KEPCO (Korea): <http://cyber.kepco.co.kr/kepco/EN/main.do>; FEPC (Japan): <http://www5.fepc.or.jp/tokei-eng/>; EIA (US): <http://www.eia.gov/electricity/>

²⁴Some additional intuition about sectoral shifts may be gleaned from other coefficient estimates. We expect that income elasticities of residential demand are lower than the output elasticities in the firm sectors (see Chang *et al.*, 2013). As the residential share increases, we should expect the elasticity to decrease. However, the coefficient $\hat{\pi}_r$ does not have to be negative for this to hold. First, if the increase in the elasticity is linear, it is already accounted for by $\hat{\pi}_o$, which is negative for all eight countries. Second, an increase in the residential share must be accompanied by a decrease in one or both of the other shares.

the US. A smaller slope to begin with and a larger break at that time may underlie the smaller point estimate with the larger standard error for the US.

Overall, it is clear that although sectoral shares account for some of the temporal decrease in elasticities in some of the countries, they do not entirely account for this pattern in any of these countries. While this finding does not necessarily contradict the empirical findings of previous authors, it suggests that those findings are incomplete. Additional drivers, such as utility, technology, and policy, must account for the decline in elasticities over time.

5 Some Conclusions

As we wrote in the introduction, a general contribution of the paper is to demonstrate how we can use a panel nonparametric approach to identify nonlinearity and specific drivers of coefficient instability in an econometric model. We now summarize our conclusions drawn from our application of this approach to a model of electricity consumption. We have identified a clear, decreasing pattern in aggregate income/output elasticities of electricity demand, which is consistent with some of the previous empirical findings. The main empirical contribution of this research is to analyze this pattern by identifying some of its sources, and the nonparametric panel gives us enough flexibility to do so.

Among richer countries, GDP itself does not affect the elasticity very much – i.e., a linear specification may be reasonable for these countries. We have detected the expected inverted U shape in GDP, but only until the 1990’s, when this pattern appears to break down. We attribute this breakdown to increased environmental awareness and technology diffusion.

Having ruled out GDP as a primary driver of coefficient instability for richer countries, our ancillary regression results have shown that sectoral shifts alone are insufficient to explain the downward movement in elasticities for a selected number of these countries. We are left with residual explanations, such as consumer utility, consumer and producer technology, and policy. Further research would be needed for the potentially more difficult task of identifying the effects of these remaining drivers.

Appendices

A Included and Excluded Countries

A.1 Countries Excluded Due to Missing Data (62)

Missing electricity production: Armenia, Azerbaijan, Belarus, Benin, Bosnia-Herzegovina, Cambodia, Croatia, Eritrea, Estonia, Georgia, Kazakhstan, Kyrgyz Republic, Latvia, Lesotho, Lithuania, Macedonia, Moldova, Namibia, Netherlands Antilles and Aruba, Russia, Serbia, Slovenia, Tajikistan, Turkmenistan, Ukraine, Uzbekistan.

Missing GDP: Bhutan, Brunei, Cape Verde, Comoros, Cuba, Cyprus, Czech Republic, Djibouti, Dominica, Equatorial Guinea, Ethiopia, Grenada, Guinea, Iraq, Jordan, Lao, Lebanon, Macau, Maldives, Mongolia, Mozambique, North Korea, Poland, Romania, Sao Tome and Principe, Slovakia, Somalia, St. Lucia, Tanzania, Tonga, Uganda, United Arab Emirates, Vanuatu, Vietnam, Western Samoa, Yemen.

A.2 Countries Excluded Due to Spurious Regression (33)

Afghanistan, Albania, Belize, Botswana, Burundi, Central African Republic, Cote d'Ivoire, Gabon, Ghana, Guinea-Bissau, Haiti, Iran, Kiribati, Kuwait, Libya, Luxembourg, Madagascar, Malawi, Mauritania, Nicaragua, Niger, Nigeria, Papua New Guinea, Qatar, Rwanda, Saudi Arabia, Senegal, Solomon Islands, South Africa, Swaziland, Togo, Venezuela, Zimbabwe.

A.3 Countries Retained in the Sample without Price Data (89)

Algeria, Angola, Argentina, Australia, Austria, Bahamas, Bahrain, Bangladesh, Barbados, Belgium, Bermuda, Bolivia, Brazil, Bulgaria, Burkina Faso, Cameroon, Canada, Chad, Chile, China, Colombia, Congo, Congo Democratic Republic, Costa Rica, Denmark, Dominican Republic, Ecuador, Egypt, El Salvador, Fiji Islands, Finland, France, Gambia, Germany, Greece, Guatemala, Guyana, Honduras, Hong Kong, Hungary, Iceland, India, Indonesia, Ireland, Israel, Italy, Jamaica, Japan, Kenya, Liberia, Malaysia, Mali, Malta, Mauritius, Mexico, Morocco, Myanmar (formerly Burma), Nepal, Netherlands, New Zealand, Norway, Oman, Pakistan, Panama, Paraguay, Peru, Philippines, Portugal, Seychelles, Sierra Leone, Singapore, South Korea, Spain, Sri Lanka, St. Vincent and Grenadines, Sudan, Suriname, Sweden, Switzerland, Syria, Taiwan, Thailand, Trinidad and Tobago, Tunisia, Turkey, United Kingdom, United States, Uruguay, Zambia.

A.4 Countries Retained in the Sample with Price Data (25)

Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Japan, Mexico, Netherlands, New Zealand, Norway, Portugal, South Korea, Spain, Sweden, Switzerland, Thailand, Turkey, United Kingdom, United States.

Year	Gr1	Gr2	Gr3	Gr4	Gr5	Gr6	Gr7	Gr8	Gr9	Gr10
1978	2	9	8	2	1	2	0	1	0	0
1979	2	8	9	2	1	2	0	1	0	0
1980	3	8	8	3	0	2	0	1	0	0
1981	3	9	6	4	0	2	0	1	0	0
1982	3	9	6	4	0	2	0	1	0	0
1983	3	9	6	4	1	1	0	1	0	0
1984	3	9	6	4	2	0	0	1	0	0
1985	3	9	6	3	3	0	0	1	0	0
1986	3	9	6	3	3	0	0	1	0	0
1987	3	9	6	3	3	0	0	1	0	0
1988	3	10	5	3	3	0	1	0	0	0
1989	3	9	6	3	3	0	1	0	0	0
1990	3	10	5	3	3	0	1	0	0	0
1991	3	8	6	5	2	0	1	0	0	0
1992	3	9	5	5	2	0	1	0	0	0
1993	2	9	6	4	3	0	1	0	0	0
1994	2	9	6	4	3	0	1	0	0	0
1995	2	9	6	5	2	1	0	0	0	0
1996	2	9	6	5	2	1	0	0	0	0
1997	2	9	7	4	2	1	0	0	0	0
1998	2	9	7	3	3	0	1	0	0	0
1999	2	9	7	4	2	0	1	0	0	0
2000	2	9	7	4	2	0	1	0	0	0
2001	2	9	7	4	1	1	1	0	0	0
2002	2	10	6	4	2	0	1	0	0	0
2003	2	8	8	4	2	0	1	0	0	0
2004	2	9	8	3	2	1	0	0	0	0
2005	2	9	8	3	2	1	0	0	0	0
2006	2	9	8	3	2	1	0	0	0	0
2007	2	9	8	3	2	1	0	0	0	0
2008	2	9	8	3	2	1	0	0	0	0
2009	2	8	9	3	1	1	1	0	0	0
2010	2	9	8	3	1	2	0	0	0	0

Table 4: **Price Observation Counts.** Number of countries for which price data are available for each year and income group. **Bold** columns denote a positive count in every year.

B Additional Technical Notes and Empirical Results

B.1 Testing for Time Series Properties

In order to validate inferences from our models (5) and (6), we study the properties of the time series used in our empirical analysis. In particular, the presence or absence of a unit root and/or a linear time trend plays a critical role in interpreting our models and statistically evaluating our empirical results. As discussed in Section 3.5, we interpret our models as representing long-run relationships assuming that variables included in the models have deterministic and/or stochastic trends. Note from Figure 1 that (y_{it}) and (x_{it}) have conspicuous time trends for all i . To empirically support our interpretation, we formally test for the presence of a unit root and a linear time trend in (x_{it}) and (p_{it}) . The time series plots of (x_{it}) and (p_{it}) are given in Figure 2. To magnify the stochastic components of (x_{it}) , we remove their dominant linear trends by the usual detrending method.

We use the standard ADF (Augmented Dickey-Fuller) test with a maintained linear time trend for the unit root tests of (x_{it}) and (p_{it}) for each i . In Table 5, we report the test results based on the unit root regression augmented with one, two, and three lagged differences. The results for the unit root test are somewhat dependent upon the choice of the number of lagged differences in the unit root regression. All groups of both (x_{it}) and (p_{it}) fail to reject the presence of a unit root at the 5% significance level with three lagged differences. For both (x_{it}) and (p_{it}) , the unit root hypothesis is rejected for some groups if we include only one or two lagged differences. However, given that income growth and inflation – represented by the differences of (x_{it}) and (p_{it}) here – typically have slowly decaying autocorrelations, test results using only one or two augmented lagged differences do not seem very meaningful, and results using three lagged differences seem more so. We therefore interpret the evidence as decisively in favor of the presence of unit roots in (x_{it}) and (p_{it}) .

To test for the presence of a linear time trend in (x_{it}) and (p_{it}) , we test whether the means of their first differences are zero. Specifically, we use the regression $w_{it} = \mu_i + \varepsilon_{it}$ with $w_{it} = \Delta x_{it}$ or Δp_{it} , and we test whether $\mu_i = 0$ for each i using a consistent estimate for the long-run variance of (ε_{it}) .²⁵ It is quite clear from Figure 1 that (x_{it}) has an increasing trend for each i , with the possible exception of the last group. Our test results for linear time trends, summarized in Table 6, support this notion. For all groups but Gr10, the evidence for a linear time trend is very strongly supported. On average, they all have an increase of approximately 2% per year. In sharp contrast, the test results for (p_{it}) are unanimously insignificant. None of the four groups show any evidence of a linear time trend in price.

²⁵See Footnote 21 for the details of our window and bandwidth selection.

GDP				Price			
group	lag	<i>t</i> -stat	<i>p</i> -value	group	lag	<i>t</i> -stat	<i>p</i> -value
Gr1	1	-3.155	0.094	Gr1	1	-1.859	0.676
	2	-2.374	0.394		2	-2.113	0.539
	3	-2.067	0.564		3	-2.062	0.567
Gr2	1	-3.530	0.036	Gr2	1	0.359	0.997
	2	-2.016	0.593		2	0.087	0.995
	3	-1.571	0.804		3	-0.04	0.994
Gr3	1	-1.271	0.895	Gr3	1	-4.892	0.000
	2	0.026	0.995		2	-3.515	0.038
	3	0.641	0.997		3	-2.959	0.144
Gr4	1	-4.644	0.001	Gr4	1	-4.612	0.001
	2	-4.102	0.006		2	-1.335	0.879
	3	-3.034	0.123		3	-1.277	0.894
Gr5	1	-3.043	0.121				
	2	-2.619	0.271				
	3	-2.064	0.566				
Gr6	1	-2.036	0.582				
	2	-2.356	0.403				
	3	-1.901	0.654				
Gr7	1	-2.565	0.296				
	2	-2.403	0.378				
	3	-2.417	0.371				
Gr8	1	-1.814	0.698				
	2	-2.253	0.460				
	3	-2.037	0.581				
Gr9	1	-1.711	0.746				
	2	-1.423	0.854				
	3	-1.365	0.871				
Gr10	1	-1.467	0.840				
	2	-1.109	0.928				
	3	-0.959	0.949				

Table 5: **Unit Root Test Results.** ADF test results for (x_{it}) and (p_{it}) with a maintained time trend and augmented with one, two, and three lagged differences. The *p*-values are obtained using the approximation provided by MacKinnon (1994).

GDP				Price		
group	estimate	<i>t</i> -stat	<i>p</i> -value	group	<i>t</i> -stat	<i>p</i> -value
Gr1	0.018	5.486	0.000	Gr1	-0.834	0.404
Gr2	0.018	6.654	0.000	Gr2	0.516	0.606
Gr3	0.019	5.240	0.000	Gr3	-0.381	0.703
Gr4	0.024	4.058	0.000	Gr4	0.264	0.792
Gr5	0.020	2.486	0.013			
Gr6	0.021	4.292	0.000			
Gr7	0.021	2.151	0.031			
Gr8	0.018	2.678	0.007			
Gr9	0.022	1.920	0.055			
Gr10	0.016	1.494	0.135			

Table 6: **Linear Trend Test Results.** Test results for the existence of a linear trend in (x_{it}) and (p_{it}) based on the assumption that they have a unit root in their stochastic components.

B.2 Bandwidth Selection

Existing theory for optimal selection of bandwidths of the type we employ in estimation is derived under stationarity and is not valid for our nonstationary models. However, we apply standard methodologies to obtain some baseline values of bandwidths, which we then modify. We consider widely used procedures based on cross validation and biased-corrected AIC. We do not claim that the bandwidths that we obtain from either method are truly optimal in any sense for our models. In fact, in the presence of nonstationarity, we are generally advised to choose a bigger bandwidth than the optimal value given under the assumption of stationarity, because nonstationary data are more sparsely and unevenly distributed than stationary data. See, for instance, Phillips and Park (1998).

Our baseline bandwidth selection is done in two steps. In the first step, we start with a reasonable bandwidth and estimate the parameters in the linear part of our models (5) and (6). The estimates of parameters in the linear part are known to be insensitive to the selection of the bandwidth used in estimating the nonparametric part, unless it is too small. See, e.g., Bickel and Kwon (2001). In the second step, we fix our parameter estimates for the linear part of our regressions obtained in the first step and perform the bandwidth selection procedure for the nonparametric part of our regressions.

It is possible to independently choose the bandwidth constants c_r for time and c_x for log GDP. However, the independent selection of two bandwidth constants is unstable in some cases. Therefore, we also impose the restrictions $c_x = c_r/10$ and $c_x = c_r/2.6$ for the

estimation of models (5) and (6) respectively. Normalized by its maximum, (x_{it}) has a sample standard deviation of roughly 0.029 for model (5) and 0.113 for model (6) – about 1/10 and 1/2.6, respectively – times the standard deviation $1/\sqrt{12} \approx 0.29$ of the uniform distribution over the unit interval.

In our analysis, the standard cross validation method generally yields bandwidth values that are smaller than those using the bias-corrected AIC procedure. For the estimation of models (5) and (6), cross validation yields the values 0.026 and 0.015 respectively for the bandwidth constant c_r . However, these values are clearly too small and produce overly noisy estimates of the income coefficient. The bias-corrected AIC procedure appears to work better than the cross validation method and provides larger bandwidth constants for our models. In particular, the bias-corrected AIC procedure chooses the bandwidth constant c_r to be 0.085 and 0.076 for models (5) and (6) respectively.

In all of our results, we use three times the baseline bandwidths obtained from the application of the bias-corrected AIC procedure. The magnification takes into account that our models are nonstationary and our observations are somewhat locally clustered and not evenly spread out. Although we do not report the details to save space, we carefully investigate the effects of varying bandwidths as a robustness check. All of our discussions and conclusions remain valid – at least qualitatively – for a wide range of bandwidth parameter values, including those we obtain from the application of the standard cross validation and bias-corrected AIC procedures. As we decrease the size of the bandwidths, our results simply become noisier.

B.3 Bootstrap Procedure

We use a sieve bootstrap based on a fitted AR model to compute the standard errors of our estimates for regressions (5) and (6) reported in Table 2 and to compare the two specifications in (5) and (6) discussed in Section 4.3. For a unit root process (z_{it}) , $z_{it} = x_{it}$ or p_{it} , we impose the unit root and fit an AR model for (w_{it}) , $w_{it} = \Delta z_{it}$, and then we generate bootstrap samples (z_{it}^*) for (z_{it}) from bootstrap samples (w_{it}^*) of (w_{it}) with the starting value given by the initial observation z_{i0} of the original sample.

It is not clear whether or not we should impose the unit root in bootstrapping the fitted residuals (\hat{v}_{it}) from regression (6), since the error term (v_{it}) , $v_{it} = \gamma p_{it} + u_{it}$, in regression (6) has the stationary component (u_{it}) , as well as the unit root component (p_{it}) . On one hand, it seems more appropriate to impose the unit root, since it is an integrated process. On the other hand, however, we cannot fit (Δv_{it}) as a finite AR process, since it has an MA unit root. Consequently, we do not impose the unit root and simply fit (\hat{v}_{it}) as a finite

AR process in our bootstrap.²⁶ We explain our bootstrap procedure in detail for regression (5) below. The bootstrap procedure we employ for regression (6) is entirely analogous.

To bootstrap regression (5), we first let the series of fitted residuals (\hat{u}_{it}) be an AR(2) process given by

$$\hat{u}_{it} = \delta_{i1}\hat{u}_{i,t-1} + \delta_{i2}\hat{u}_{i,t-2} + \eta_{it}, \quad (\text{B.1})$$

we define $w_{it} = \Delta x_{it} - T^{-1} \sum_{t=1}^T \Delta x_{it}$ or $w_{it} = \Delta p_{it}$, and we fit (w_{it}) as an AR(1) process

$$w_{it} = \delta_i w_{i,t-1} + \varepsilon_{it} \quad (\text{B.2})$$

for $i = 1, \dots, N$. To ensure that the fitted residuals from regressions (B.1) and (B.2) have zero sample mean, which is essential for the validity of the bootstrap, we include a constant term in the regressions. We obtain the least squares estimates $\hat{\delta}_{i1}$, $\hat{\delta}_{i2}$ and $\hat{\delta}_i$ of the parameters and the fitted residuals $(\hat{\eta}_{it})$ and $(\hat{\varepsilon}_{it})$ from regressions (B.1) and (B.2) augmented with a constant term.

Second, we resample $(\hat{\eta}_{it})$ and $(\hat{\varepsilon}_{it})$ to obtain bootstrap samples $(\hat{\eta}_{it}^*)$ and $(\hat{\varepsilon}_{it}^*)$, from which we obtain bootstrap samples (u_{it}^*) and (w_{it}^*) of (u_{it}) and (w_{it}) using

$$u_{it}^* = \hat{\delta}_{i1}u_{i,t-1}^* + \hat{\delta}_{i2}u_{i,t-2}^* + \eta_{it}^* \quad (\text{B.3})$$

and

$$w_{it}^* = \hat{\delta}_i w_{i,t-1}^* + \varepsilon_{it}^* \quad (\text{B.4})$$

for $i = 1, \dots, N$. Bootstrap samples (z_{it}^*) of (z_{it}) , $z_{it} = x_{it}$ or p_{it} , are then constructed from (w_{it}^*) in (B.4) using

$$z_{it}^* = z_{i0} + \sum_{j=1}^t w_{ij}^*$$

with $z_{i0} = x_{i0}$ or p_{i0} . To preserve the correlation between (η_{it}) and (ε_{it}) in the sample, we bootstrap them jointly within each group i , for $i = 1, \dots, N$.

Our bootstrap samples of (x_{it}) and (p_{it}) are all initialized at the starting values of the original sample. The parameters in our models are not invariant with respect to the initializations of these covariates, and we find that this initialization yields some bias in our bootstrap estimates of the parameters in the models, and especially so for the estimate of α in regressions (5) and (6). To avoid this bias, we use a two-stage bootstrap procedure for the estimation of α . In the first stage we compute the bias in the estimate of α from

²⁶In an unreported simulation, we also bootstrap (\hat{v}_{it}) with the unit root restriction. The unit root restriction appears to have little effect, as long as we start the bootstrap samples from the origin. If we initialize them at the starting value of the original sample, they have a nonzero mean that significantly biases the estimator of α .

the bootstrap, i.e., the difference between the mean of estimates from the bootstrap and the corresponding sample estimate of α . Then we use the computed bias to correct for the bias in the bootstrap estimates of α at the second stage. Once the bias-corrected estimate of α is obtained, we use it to re-estimate the other parameters $\beta(r, x)$ and γ in the models. We employ the two-stage bootstrap not to correct for any bias in the coefficient estimates themselves, but only to adjust for the bias in the bootstrap estimates caused by our initialization of the unit root bootstrap samples, which are used to calculate the standard errors of the coefficient estimates.

References

- Andrews, D.W.K. (1991). "Heteroskedasticity and autocorrelation consistent covariance matrix estimation," *Econometrica* 59, 817-858.
- Berndt, E.R. and D.O. Wood (1975). "Technology, prices, and derived demand for energy," *The Review of Economics and Statistics* 57, 259-268.
- Bickel, P. J. and J. Kwon (2001). "Inference for semiparametric models: some current questions and an answer (with discussion)," *Statistica Sinica* 11, 863-960
- Brookes, L.G. (1972). "More on the output elasticity of energy consumption," *Journal of Industrial Economics* 21, 83-94.
- Cai, Z. (2002). "A two-stage approach to additive time series models," *Statistica Neerlandica* 56, 415-433.
- Cai, Z. (2007). "Trending time-varying coefficient time series models with serially correlated errors," *Journal of Econometrics* 136, 163-188.
- Cai, Z., Q. Li and J.Y. Park (2009). "Functional-coefficient models for nonstationary time series data," *Journal of Econometrics* 148, 101-113.
- Cai, Z. and R.C. Tiwari (2000). "Application of a local linear autoregressive model to BOD time series," *Environmetrics* 11, 341-350.
- Chang, Y., C. Kim, J.I. Miller, S. Park, and J.Y. Park (2013). "Time-varying long-run income and output elasticities of electricity demand," mimeo, Department of Economics, University of Missouri.
- 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.
- Fan, J. and T. Huang (2005). "Profile likelihood inferences on semiparametric varying-coefficient partially linear models," *Bernoulli* 11, 1031-1057.
- Galli, R. (1998). "The relationship between energy intensity and income levels: Forecasting long term energy demand in Asian emerging countries," *The Energy Journal* 19, 85-105.
- Halvorsen, R. (1975). "Residential demand for electric energy," *The Review of Economics and Statistics* 57, 12-18.

- Halvorsen, R. (1978). *Econometric Models of U.S. Energy Demand*. New York: Lexington, MA: D.C. Heath and Company.
- Härdle, W. (1992). *Applied Nonparametric Regression*. Cambridge: Cambridge University Press.
- Judson, R.A., R. Schmalensee, and T.M. Stoker (1999). "Economic development and the structure of the demand for commercial energy," *The Energy Journal* 20, 29-57.
- Kaufmann, R.K. (2004). "The mechanisms for autonomous energy efficiency increases: A cointegration analysis of the US energy/GDP ratio," *The Energy Journal* 25, 63-86.
- Kraft, J. and A. Kraft (1978). "On the relationship between energy and GDP," *The Journal of Energy and Development* 3, 401-403.
- Li, Q. and J.S. Racine (2007). *Nonparametric Econometrics Theory and Practice*. Princeton: Princeton University Press.
- MacKinnon, J. G. (1994). "Approximate asymptotic distribution functions for unit-root and cointegration tests," *Journal of Business and Economic Statistics* 12, 167-176.
- Maddala, G.S., R.P. Trost, H. Li, and F. Joutz (1997). "Estimation of short-run and long-run elasticities of energy demand from panel data using shrinkage estimators," *Journal of Business and Economic Statistics* 15, 90-100.
- Medlock, K.B. and R. Soligo (2001). "Economic development and end-use energy demand," *The Energy Journal* 22, 77-105.
- Pagan, A. and A. Ullah (1999). *Nonparametric Econometrics*. Cambridge: Cambridge University Press.
- Park, J.Y. and S.B. Hahn (1999). "Cointegrating Regressions With Time Varying Coefficients," *Econometric Theory* 15, 664-703.
- Phillips, P.C.B. and J.Y. Park (1998). "Nonstationary Density Estimation and Kernel Autoregression," Cowles Foundation Discussion Paper #1181.
- Robinson, P.M. (1988). "Root-N consistent semiparametric regression," *Econometrica* 56, 931-954.
- 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.

Stock, J.H. (1989). "Nonparametric policy analysis," *Journal of the American Statistical Association* 84, 567-575.

Webster, M., S. Paltsev, and J. Reilly (2008). "Autonomous efficiency improvement or income elasticity of energy demand: Does it matter?" *Energy Economics* 30, 2785–2798.