Roger Fouquet

Consumer surplus from energy transitions

Article (Accepted version)
(Refereed)

Original citation:
DOI: 10.5547/01956574.39.3.rfou

Reuse of this item is permitted through licensing under the Creative Commons:

© 2018 IAEE
CC BY

This version available at: http://eprints.lse.ac.uk/86502/
Available in LSE Research Online: October 2018

LSE has developed LSE Research Online so that users may access research output of the School. Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. You may freely distribute the URL (http://eprints.lse.ac.uk) of the LSE Research Online website.
Consumer Surplus from Energy Transitions

Roger Fouquet*

ABSTRACT

Energy transitions have led to major advances in human wellbeing. However, little evidence exists about the scale of the net benefits. By developing a new method for identifying the demand curve, this paper estimates the consumer surplus associated with heating, transport and lighting over more than two hundred years and identifies the gains from key energy transitions. For certain energy transitions, the increase was dramatic, reflecting the transformations in society and lifestyles that mobility and illumination provided in the nineteenth and twentieth centuries. Yet, the net benefits related to heating technologies only rose modestly. Finally, due to saturation effects of the demand for energy services, future technological developments and energy transitions may benefit consumers (though not necessarily society as a whole) less than those in the past.

Keywords: Energy demand, Energy transition, Technological change, Consumer surplus

https://doi.org/10.5547/01956574.39.3.rfou

1. INTRODUCTION

The recent focus on the impacts of climate change has drawn attention to the negative aspects of energy use, overshadowing the positive benefits to consumers (Nordhaus 2014, Weitzman 2014). As a result, expectations about the advantages of a potential low carbon transition tend to focus on the declines in external costs associated with the energy system. Nevertheless, past energy transitions generated large benefits to consumers from switching from one energy source (and technology) to another (Fouquet and Pearson 2006, Gordon 2016). Similarly, the continually upward trend in energy consumption, particularly since the Industrial Revolution in developed economies and more recently in industrializing countries, reflect value to users, as well as the role energy plays in economic growth (Stern and Kander 2012). However, very little evidence exists about the scale of the net benefits to consumers from rising fuel use, related technological improvements and energy transitions, and how these benefits vary in the long run. A better understanding may offer insights about how future technological developments and energy transitions affect consumer welfare, and about ways of reducing carbon dioxide emissions while minimizing the burden to consumers.

A major problem with comparing old and new goods is that they have different characteristics and levels of quality (Breshnahan and Gordon 1997, Friedman 2016). Nordhaus (1997) offered a solution to this problem by comparing the common energy service that different technologies provide. With this approach, he was able to compare the price of lighting from candles, gas lamps, and electric bulbs. By taking account of the different efficiencies with which these technologies provided the service, he showed how using the price of fuels underestimated the fall in the price of lighting by

* Grantham Research Institute on Climate Change and the Environment. London School of Economics and Political Science (LSE), Houghton Street, London WC2A 2AE, United Kingdom. E-mail: r.fouquet@lse.ac.uk.

The Energy Journal, Vol. 39, No. 3. All rights reserved.
between 25- and 750-fold, depending on the method used. Building on his insight, this paper uses this “energy service” approach to compare different technologies over more than two hundred years.

With this in mind, the overall purpose of the paper is to estimate how consumer surplus for heating, transport, and lighting, and their associated energy technologies and sources, changed over the last two hundred years. Using a unique historical data set and developing a new method to identify the demand curve, the results indicate that the consumer surpluses (as a share of GDP) for energy services have risen from the early nineteenth century until the second-half of the twentieth century and, then, have started to decline. This suggests that, as an economy develops, consumers gain greatly from rising energy use. However, it also highlights that, at higher levels of economic development and material comfort, rising consumption brings more limited benefits.

The changes over time also reflect the peculiarities of specific energy technologies and the services they provide. For instance, the estimates suggest that transport and lighting technologies provided large and rising consumer surplus, until the 1960s and 1950s, respectively. For heating technologies, the rise in consumer surplus was more modest. The results indicate that not all technological innovations are equal when it comes to their potential to increase consumer surplus. Some innovations, such as buses and gas heating, merely act as substitutes, and bring little additional consumer surplus. By contrast, major potential improvements in welfare occur when goods and services, such as key transport and lighting technologies, go beyond the simple services they provide, and offer new opportunities to transform lives.

The value of the present study is first to provide information about the magnitude of the benefits to society from particular energy transitions experienced in the nineteenth and twentieth centuries. It informs us about the stream of net benefits resulting from new energy sources and technologies. For instance, the declining (relative) utility from energy consumption at higher levels of economic development signals that there may be less profits to be made by companies from new energy technologies and transitions. Finally, this study also offers a method for anticipating consumer benefits from future technological development and energy transitions, and insights about carbon mitigation strategies that minimize consumer losses.

The paper is organized as follows: Section Two outlines briefly the framework for analyzing the demand for and supply of household-produced energy services against the backdrop of technological change. Section Three introduces a simple method for constructing a demand curve using historical evidence, and suggests a procedure for calculating the associated consumer surplus. Section Four outlines the data used in the study. Section Five presents estimates of the demand curve for heating, transport, and lighting, shows how demand shifted between 1830 and 2010, offers estimates of how the related consumer surplus varied over roughly two hundred years, and indicates the impact particular technological developments and energy transitions have had on well-being initially and in the long run. The final section draws conclusions and highlights the limitations of the study—it is important to stress that this paper aims to provoke thought by presenting trends in consumer surplus (and the reader is encouraged to focus on them rather than on values for individual years) and to stimulate further research and debate.

2. THE DEMAND FOR AND SUPPLY OF HOUSEHOLD PRODUCED SERVICES

To estimate the consumer surplus, locating the full demand curve is crucial, in order to quantify the area underneath it and above the price line (Marshall 1890 p.78, Willig 1976). This section briefly outlines a simple model of the demand for household-produced services, and the framework for jointly analyzing old and new technologies. This framework emphasizes that, to meet the
demand, the household produces services by combining technologies with inputs. For more detail, Quigley and Rubinfeld (1989) and Nordhaus (1997) outline the supply side; Hunt and Ryan (2015) model the demand side; and Fouquet (2016) connects the two sides, as is done here.

The objective of a consumer or, here, a household, is to maximize utility by combining service consumption \(s_t\) and all other goods and services, illustrated by the composite variable \(x_t\):

\[
Max U_t = u(s_t, x_t)
\]

subject to the budget constraint

\[
y_t = p_s s_t + p_x x_t
\]

where \(p_s\) and \(p_x\) refer to the prices of the services and prices of the composite goods, and \(y_t\) is the household’s budget or income. Utility depends indirectly on prices and income, and the indirect utility function can be represented as

\[
v_t = V(p_s, p_x, y_t).
\]

This representation of the indirect utility function implies (from Roy’s Identity) that the demand function for services can be investigated within this utility maximization framework:

\[
s_t = f(p_s, p_x, y_t).
\]

Observing how service consumption varies as constraints change offers an opportunity to identify the relationships between consumption and constraints. The effects of these changing constraints can be summarized in the form of the own price elasticity of demand (in any particular year \(t\)):

\[
\varepsilon_{ps} = \frac{\partial s_t}{\partial p_s} \frac{s_t}{p_s}
\]

and the income elasticity of demand (in any particular year \(t\)):

\[
\varepsilon_{sy} = \frac{\partial s_t}{\partial y_t} \frac{s_t}{y_t}
\]

This standard model of the household demand for services links with the supply of services which is rooted in the theory of household production and consumption technology developed by Becker (1965) and Lancaster (1966). In this theory, households produce their own services by combining factors, which might include labor, capital, and/or energy. The latter input is particularly important in the production of heating, transport, and lighting, which are the focus of this paper.

To simplify, many models of the household production related to services tend to include only capital, \(k_t\), and energy used, \(e_t\) (Quigley and Rubinfeld 1989). Over the last five hundred years, technological change related to the capital has led to a substitution away from labor inputs for many household services, and toward inputs of physical capital and energy, such that many energy-related services, such as heating and lighting, are now provided with virtually no labor needs (Mokyr 2000, Fouquet 2008). Driving and cycling are the only modes of transport in which labor provides a substantial input to produce the service today—taking a bus involves time waiting, though not active labor. Inevitably ignoring the role played by labor in the provision of these services affects the analysis. In particular, the consumption of energy services would be in competition with other activ-
ities in a person’s daily life, such as working, eating, hygiene, socializing and sleeping—certainly, the introduction of driverless or autonomous vehicles would help reduce the competition between travel and some of these other activities, with the implication that travel would be likely to increase considerably. Nevertheless, in this study, the modeling convention will be followed and it will be assumed that the role of labor inputs is zero.

The relationship between inputs and outputs depends on the efficiency of the technology for each service ($\phi_{st}$)—that is, the amount of services generated by a specified quantity of energy. Although a choice can be made about the technology used, an important feature is that, once chosen, the relationship tends to be fixed at any point in time by the technology (Hunt and Ryan 2015, p.274). Thus, the provision of services can be determined by the energy consumption ($e_t$) multiplied by the efficiency of the appliance:

$$s_t = \phi_{st} e_t.$$  

(7)

Improvements in energy efficiency may be associated with higher capital costs (Frondel et al. 2008); thus, the capital costs of generating services ought to be taken into account. However, a simplifying assumption is that the price of services is determined by the marginal cost of production, which is often measured as the price of energy ($p_o$) divided by the technical efficiency of the appliance being used (see Nordhaus 1997)—although this does ignore depreciation associated with appliances:

$$p_o = p_a / \phi_{st}.$$  

(8)

Feeding equations (7) and (8) into equations (5) and (6) enable own price and income elasticity of demand for certain household-produced services, such as heating, transport, and lighting, to be estimated.

3. EMPIRICAL STRATEGY TO IDENTIFY DEMAND AND ESTIMATE CONSUMER SURPLUS

The central objective of this paper is to estimate the long run trend in the Marshallian consumer surplus associated with household-produced services, and how they varied with technological development. Hausman (1997) offers a straightforward way of thinking about consumer surplus. Hausman’s method estimates the consumer surplus of the service ($CS_{st}$) as a function of the share of consumer expenditure on the service and the price elasticity of demand ($e_{pst}$) in equation (5)):

$$CS_{st} = \frac{1}{2} \frac{(p_o s_t)/y_t}{e_{pst}}.$$  

(9)

Although helpful for thinking about consumer surplus, the Hausman equation (9) is limited by the linear assumption about the relationship between price and quantity consumed. Hausman (1997, 1999) presented this model for the estimation of the consumer surplus related to new goods, such as IT and flavored Cheerios, where consumption levels were low, the distance from the y-axis was not great, and the linear relationship was probably an acceptable approximation.

However, for many goods and services which have been consumed for a long time, including the services considered here, the demand curve may display more curvature, and the linear

1. Bockstael and McConnell (1983) show that welfare measurements in the context of household production generate results identical to those for a consumer with a utility function defined exclusively by consumption decisions.

2. Greenwood and Kopecky (2013) develop an alternative non-linear way to estimate consumer surplus, however, it imposes a structure to the utility function and requires careful calibration to avoid spurious results.
approximation may be less appropriate. Indeed, most models of consumer theory assume a convex relationship between price and quantity. Ignoring the high willingness-to-pay values for low levels of marginal consumption implies that the Hausman (1997, 1999) model underestimates the consumer surplus\(^3\).

To improve on this linear approximation, here, an attempt will be made to trace out the full demand curve for each year studied (i.e., from the nineteenth to the twenty-first century). The novel method used here is a benefit transfer performed through time (i.e. a temporal benefit transfer)—rather than spatially, which has been developed for non-marketed goods and environmental resources (Smith et al. 2002).

A benefit transfer involves the use of existing information (often associated with willingness-to-pay values) designed for one specific context to address questions in another context. Given the cost of performing new stated-preference studies, and the lack of circumstances where preferences are revealed about non-marketed goods, transferring benefits from one source to another is a useful tool to estimate willingness-to-pay values. As a result, benefit transfer methodology forms the basis of many economic analyses and policy assessments (Loomis 1992, Johnston et al. 2015).

There are various approaches to benefits transfers. The first method is the ‘direct unit’ value transfer, taking a value from one study and applying it without adjustment to the new study, thus, taking no account for differences between the original study and the site of interest. The second method is the ‘adjusted unit’ transfer, which takes a value from one study (or a group of studies), and adjusts it to take account of differences in important factors. The most common approach involves adjusting for differences in income (e.g. \(WTP_a = WTP_b \frac{y_a}{y_b}\) where \(y\) is income, \(a\) is the site of interest, and \(b\) is the original study). A third, more rigorous method is the “value function” transfer, in which a value function in one study is estimated and applied to the site of interest. However, this depends on having gathered data on relevant variables, which may include income, age and education, and estimated the coefficients indicating the relationship between these variables and willingness-to-pay, such as income, age or education elasticities (Johnston and Rosenberger 2010).

As will be discussed later, it will be possible to use a value function transfer, because Broadberry et al. (2015) offers annual data on per capita GDP and Fouquet (2014) provides annual estimates of the income elasticities of demand for different services. Now, income elasticity of demand for a good or service is defined as the percentage change in the quantity demanded for a given increase in income at constant prices—not as the change in price or willingness-to-pay for a given quantity. Randall and Stoll (1980) explain that the income elasticities of willingness-to-pay (also known as the price flexibility of income, \(\epsilon_{yWTP}\)) indicates how changes in income affect the amount the consumer is willing to spend to consume a specific quantity of a good, service or amenity and is an analogous concept to the income elasticity of demand for goods and services which are purchased facing a particular price. Hanemann (1991) shows that this income elasticity of willingness-to-pay (or price flexibility of income) is analytically equivalent to the ratio of the income elasticity of demand for the good to the elasticity of substitution between the good or service of interest and the composite good. Flores and Carson (1997 p.293) highlight that the two income elasticities can indeed be different, but in conclusion their analysis “suggests that the ... income elasticity for most values of \(q\) are reasonably close in magnitude to the income elasticity of willingness-to-pay”. Thus, the crucial issue is the size of the elasticity of substitution. Here, the assumption is that the composite good is a decent but not perfect substitute for the energy services, that the elasticity of substitution between \(s\) and \(x\) is close to one and,

3. Evidence about the convexity of the demand curve is discussed in Section Five.

All rights reserved.
therefore, the income elasticities of demand, $\varepsilon_{yst}$, and of willingness-to-pay, $\varepsilon_{yWTPst}$, are close in magnitude and can act as proxies for the purpose of this exercise.

Information about the relationship with age and education are not available, but, there is no reason for expecting these to affect the willingness-to-pay values. As Bateman et al. (2011, p.384) emphasize, functions should be ‘constructed from general economic theoretic principles to contain only those variables about which we have clear, prior expectations.’ In traditional (i.e. spatial) benefit transfers, institutional and cultural factors differ greatly, yet, transferred willingness-to-pay values are considered broadly acceptable when care is taken about transferability issues (Nelson and Kennedy 2009). Here, the benefit transfer is made for the same consumers valuing the same service, one year later. Thus, despite limits associated with using historical information, this temporal benefit transfer should be more suitable than traditional spatial benefit transfers, and this method offers a broadly acceptable way of eliciting willingness-to-pay values.

This “temporal benefits transfer” method can be considered more formally, building on the discussion in Section 3. The underlying indirect utility function, $V(.)$, is represented by equation (3), and, from Roy’s Identity, it is possible to determine the Marshallian demand curve and, therefore, potentially each optimal level of consumption at different prices for a particular income level. Assuming the market outcome occurs where demand and supply meet, this implies that the market and equilibrium price, $p_{st}$, reflects the monetary indicator of the household’s utility generated, its reservation price and, therefore, the willingness-to-pay, $WTP_{smt}$, from consuming different (m) marginal levels of the service, $s_{mt}$. That is,

$$WTP_{smt} = p_{st} \, \, (10)$$

Now, assuming the indirect utility function, $V(.)$, remains constant from year $t$ to year $t+1$, the willingness-to-pay value in year $t$ can provide information about the following year’s willingness-to-pay for the same level of consumption, $s_{mt+1}$. For example, in the extreme case that all variables remain unchanged, then it can be inferred that the willingness-to-pay next year at that level of consumption, $WTP_{smt+1}$, will be the same as this year’s willingness-to-pay, $WTP_{smt}$. Alternatively, assuming income is the only variable to change in year $t+1$, then with knowledge of the change in income, $\Delta y_{t+1}$ and the income elasticity of willingness-to-pay, $\varepsilon_{yWTPst}$, the new willingness-to-pay at the same marginal level of consumption can be inferred.

More generally, the willingness-to-pay value in year $t=1$ can provide information about the willingness-to-pay value in year $t=2$, year $t=3$, etc... using the value function. Thus, with knowledge of the equilibrium quantity in year 1 and, hence, the first marginal level of consumption ($s_{m1}$) and the equilibrium price ($p_{m1} = WTP_1(s_{m1})$), as well as percentage changes in income and the income elasticity of willingness-to-pay, the new willingness-to-pay in year $r$ for the first level of marginal consumption, $WTP_r(s_{m1})$, can be estimated for each year

$$WTP_r(s_{m1}) = WTP_1(s_{m1}) \left[ \sum_{t=1}^{r} \varepsilon_{yWTPt} \frac{y_t - y_{t-1}}{y_{t-1}} + 1 \right] \, \, (11)$$

As the equilibrium level of consumption changes over time, different willingness-to-pay values are revealed for particular marginal levels of consumption in particular years—based on equation (10). So, more generally, for each marginal level of consumption, $s_{mt}$, for which knowledge of the equilibrium price exists (as well as about the changes in income and the income elasticities of willingness-to-pay), it would be possible (making the same assumptions) to quantify the willingness-to-pay in subsequent years and, therefore, perform a temporal benefit transfer:

All rights reserved.
For each year, a series of points indicates the willingness-to-pay, $WTP_r(s_m)$, at different particular marginal quantities of the service, $s_m$. In other words, this series “identifies” (or approximates) a portion of the demand curve for marginal quantities where market information was available.

As the range of different marginal quantities (for which this temporal benefits transfer is performed) increases, a greater portion of the demand curve is “identified”. Thus, because of the dramatically lower equilibrium levels of consumption in the past, historical data reveals scarce information about the willingness-to-pay values for low marginal levels of consumption and a rare opportunity to identify a large portion of the demand curve.

Having produced these household demand curves, the next step is to calculate the area under them and above the price to reveal the consumer surplus. Here, rather than using integration as is normally done when a functional form is known or assumed (see, for example, Sleznick 1998 or Just et al. 2004), a piecewise linear approach will be used, and the consumer surplus, $CS_{st}$, will be the sum of series of measurable areas, each representing marginal increases in consumption:

$$CS_{st} = \frac{1}{2} \frac{(WTP_r(s_1)s_1)}{y_2} + \sum_{m=1}^{m-1} \frac{(WTP_r(s_m) - p_r)\Delta s_m}{y_t} + \frac{1}{2} \frac{(WTP_r(s_{m-1}) - WTP_r(s_m))\Delta s_m}{y_t}$$

Once each household’s consumer surplus has been quantified, they are summed across households to estimate the aggregate consumer surplus.

To summarize, for each year, the method (i) estimates the new (and generally increased) willingness-to-pay for a series of marginal levels of consumption; (ii) connects these willingness-to-pay values in each year to trace out the annual demand curve, then (iii) estimates the areas under small portions of the demand curve (based on the marginal levels of consumption) and above the price, (iv) sums the small portioned areas to estimate the consumer surplus, (v) aggregates for all households and (vi) divides the monetary value of the consumer surplus by GDP to offer a comparison as incomes change. The first two of these steps locate the demand curve, the middle two steps estimate the consumer surplus, and the latter two make the results interpretable. All of these steps will be discussed in Section Five—for more details on the method, please see the appendix, and its limitations, please see the conclusion. Before outlining these steps, the data will be presented.

4. DATA

Identifying trends in the consumer surplus related to services requires extensive information on consumption, prices, and efficiencies related to agricultural and energy commodities and technologies. Schools, colleges, hospitals, and government departments around the United Kingdom offer remarkable records of the history of the country’s commodity prices going back almost one thousand years (Rogers 1886, Beveridge 1894, Mitchell 1988). The volumes of the History of the British Coal Industry (Flinn 1984, Church 1987, Hatcher 1993) pull together most statistics on coal prices and consumption over the last 500 years. The Statistical Abstracts of the British Parliamentary Papers and then of the Ministry of Fuel and Power provide data beginning in the
mid-nineteenth century, and were forerunners of the current Digest of United Kingdom Energy Statistics, which provides annual data on all energy sources. This data have now been combined and are available in two annual data sets for the United Kingdom on service prices between 1300 and 2010 (Fouquet 2011a), and service consumption between 1700 and 2010 (Fouquet 2014). The data sources and methodologies used are extensively explained in Fouquet (2008), and the following discussion summarizes their construction.

As shown in equation (7) and (8), consumption and prices of service can be calculated by combining fuel consumption and prices with technical efficiency estimates. Heating and lighting data were not directly available, and were quantified in this way. For example, in the early nineteenth century, a ton of coal could be burned in a traditional fireplace, generating around 10 percent of a ton of coal’s useful heat (Fouquet 2008). Using this information and the price of one ton of coal (£115 in 2000 values), the price of one ton of coal equivalent of useful heat (or of ‘effective heating’) was estimated to be about £(2000)1,150^4.

A similar process was used for lighting, which was measured in lumen-hours—one lumen-hour is equivalent to the illumination generated from a candle burning at a distance of one meter for one hour, and one million lumen-hours is equivalent to leaving on a 100-watt incandescent bulb for 30 days. For example, a candle produced 13 lumen-hours per kWh; a typical town gas lamp from the late 1820s generated 130 lumen-hours per kWh; by 1916, the ‘Welsbach Mantle’ gas lamp produced more than six times more light, 870 lumen-hours per kWh (Nordhaus 1997). With this information, and data on gas prices, the price of lighting can be estimated. The price of gas lighting in 1830 was £2,700 (in 2000 money) for one million lumen-hours, and, in 1920, it was £40. Today, with LED lighting generating 66,000 lumen-hours per kWh, the same amount of illumination costs under £1. In a similar way, prices and consumption of gas and electricity (or other fuels) can be combined with the efficiency of the technology to estimate lighting or heating use, or the consumption of other services.

The time series for average lighting and heating efficiency is then assembled using the efficiency estimates in Nordhaus (1997) for lighting, and Billington (1982) for heating, and simple technological diffusion models (Fouquet 2008). This provides the values for \( \phi_s \), the efficiency of the technology for each service. These values are fed into equations (7) and (8) to estimate the consumption and price of the service.

Data on transport use were directly available for most of the technologies. In particular, direct data were available for stagecoaches (Chartres and Turnbull 1983) and railways (Mitchell 1988). Estimates of car and bus use before 1952 were created by combining annual statistics on vehicle ownership (Mitchell 1988) with a model of average distance traveled per vehicle to produce an estimate of the billions of passenger kilometers (bpk) back to 1904. DoT (2002) and DfT (2010) presented data on passenger travel between 1952 and 2010.

As shown in the examples above, a key advantage of focusing on energy services, rather than on fuels, is that the demand for services remains comparable with the introduction of new goods and technologies. For instance, as Nordhaus (1997) showed, because of improvements in lighting efficiency between the nineteenth and twenty-first centuries, it is difficult to properly compare long-run behavior without focusing on the service. Similarly, changes in vehicle efficiency pre- and post-1973 imply that the utility to a car user from consuming one liter of gasoline has greatly changed over the last forty years, making it difficult to analyze long-run demand using direct

---

4. Throughout, prices are quoted in real terms (i.e., in £ in year 2000 values). The retail price index is from the data used in Allen (2007), and then updated using Office of National Statistics (ONS 2012) data. This means that the costs of producing services are broadly comparable across time.
analysis of car-user fuel consumption. Instead, focusing on the services provided (e.g., the passenger-kilometers or lumen-hours) helps to identify very long-run patterns in consumption that would be hidden by focusing only on the changing uses of different commodities.

Elasticity estimates were generated using annual time series data on energy service consumption, prices, and GDP per capita from 1700 to 2010 (Fouquet 2014, Broadberry et al. 2015). An important assumption (confirmed through statistical tests) was that the causality ran from income and prices to consumption, and not from consumption to income and prices—implying that any changes in income and prices were exogenous. Indeed, the likelihood of, say, lighting (which is less than 1 percent of total final-user energy consumption, which in itself is only 10 percent of aggregate consumer expenditure) altering GDP, GDP per capita, and consumers’ budgets is low. Similarly, these technologies certainly boosted business and GDP; however, this is an independent line of causality and should not alter the relationship between GDP and residential energy service consumption. Thus, there is unlikely to be a problem of endogeneity when looking at residential service consumption at a disaggregated level.

The key observation is that, as the economy developed over the last two hundred years, trends in income elasticities of demand followed an inverse U-shaped curve (see Figure 1, top-half). For instance, they reached a peak (about 2.3, 3.0 and 4.0 for income elasticities of demand for heating, transport, and lighting, respectively) in the nineteenth century—at levels of GDP per capita below £(2000)3,000 (or $(2010)6,000). After the peaks, initially rapid declines occurred, followed by more gradual declines. Income elasticities of demand took almost 100 years to reach unity, in the mid-twentieth century, at between £(2000)4,500–6,000 (or $(2010)9,000–12,000) per capita. Similarly, price elasticities also peaked (at values of about −1.5) at levels of per capita income of between £(2000)2,000–2,500 (see Figure 1, bottom-half).

These results offer the beginnings of a stylized fact about the relationship between elasticities of demand and economic development. That is, at very low levels of economic development, consumers focused on meeting basic needs, particularly food and cooking. As income grew, so did

![Figure 1: Trends in Income and Price Elasticity of Demand for Heating, Transport and Lighting, 1800–2010](image-url)
the focus on making shelter and heating homes. As income rose further, these demands started to grow less proportionately than income (e.g., income and price elasticities for heating fell). In turn, other demands were met, such as mobility and illumination (implying rising income elasticities for transport and lighting demand). The income elasticities for transport and lighting may have been particularly high because greater consumption created opportunities to alter lifestyles (e.g., suburbanization, or working and entertaining into the night). As income increased further, these income and price elasticities started to fall (below 1, in absolute terms, from the 1950s onwards) as saturation kicked-in.

5. CONSUMER SURPLUS OF ENERGY SERVICES AND OF ENERGY TRANSITIONS

Having presented the data sets available on prices and consumption of services, as well as estimates of the income and price elasticities of demand for these services, the first task is to locate the full demand curves for heating, passenger transport, and lighting in each year, broadly between the early nineteenth and early twenty-first century. The length of the series depends on the availability of data on income and price elasticity to estimate individual willingness-to-pay values for a series of marginal quantities consumed, using the value function transfer described in equation (12). Thus, an average person’s willingness-to-pay and demand curve for each year are located based on changes in income and the income elasticity data.

Figure 2 shows these individual demand curves for residential heating, passenger transport, and lighting, and how they have shifted from 1830 to 2010. It is important to note that, given the large increase in values over time, the axes are in log form. As expected from Figure 1, the slopes of the demand curves are not linear, with high values for very small levels of the services.

Also, as expected, demand curves have shifted upwards and to the right over the last two hundred years. The size of the shift was determined by the income elasticities, which were based on the evidence presented in Figure 1. As a result, the largest shifts in demand curves were in the second-half of the 1800s. Another observation from Figure 2 is that the demand curve may have become less linear and more convex over the decades, implying that the linear Hausman (1997, 1999) method becomes a less reliable approximation over time.6

The final step is to estimate the consumer surplus by calculating the area below the demand curve (i.e. the benefits), presented in Figure 2, and above the price line (i.e. the costs). Figure 3 presents the consumer expenditure related to these services. Consumer expenditure on transport rose in the early nineteenth century with the expansion of stagecoach networks, peaking in the 1820s as stagecoach companies raised journey prices to recuperate their stranded investments following the threat from superior railways entering the market. Consumer expenditure on transportation was between 7 percent and 11 percent of GDP during most of the nineteenth century, with the advent of railways; and between 5 percent and 7 percent during the twentieth century, associated with the use

5. Annual data on income elasticities of transport demand were only available back to 1850. In order to push back the estimates of consumer surplus for transport, and to evaluate the role of the railway from its introduction, assumptions were made about the income elasticities between 1830 and 1850. In particular, the elasticities in the first half of the nineteenth century appear to have been lower than in the second half (Fouquet 2014). So, from a value of 2.6 in 1850, it was assumed to have risen from 2.5 in 1840 and 2.4 in 1830. These assumed values also mirror the apparent trend in income elasticities for transport demand (and for other services, more generally) of a peak in the 1860s, preceded by lower values, as shown in Figure 1. Using the trend to make assumptions about the income elasticities also provides a method for anticipating the consumer surplus associated with future technologies. This use of the method will be discussed in the conclusion.

6. Cohen et al. (2016) offers another rare attempt to identify the full demand curve associated with transport services.
of the internal combustion engine—with another peak in the 1920s. For heating, consumer expenditure was around 4 percent of GDP in the eighteenth century, then peaked at 6 percent in the 1820s with the introduction of more-efficient fireplaces, and declined to 2 percent as efficiencies improved further. The introduction of gas heating led to a new peak in the 1930s, but in a declining trend in expenditure since the mid-twentieth century. Expenditure on lighting was around 2 percent of GDP in the nineteenth century, and it fell below 1 percent from the 1950s. Overall, rises in consumer expenditures have been associated with new technological developments, but also with increases in prices, and do not directly reveal the utility generated from consumption.

Using equation (13), the information on prices, income per capita, and income and price elasticities of demand generated estimates of long run trends in consumer surplus for heating, passenger transport and lighting (see Figure 4). A general observation is that relative consumer surpluses increased with economic development, peaked, then stabilized, and then showed signs of decline. Specifically, the net benefits to consumers resulting from heating, transport, and lighting rose until the 1990s, 1960s, and 1950s, respectively. In other words, the peaks were relatively late in terms of economic development—so, it might be too early to be confident that these are the peaks. Furthermore, despite the apparent relationship, the trends are quite different. For heating, the consumer surplus increased very gradually, while for transport it increased very rapidly, and for lighting it initially increased quite quickly, and then declined considerably. Given this variation, examining the individual experiences in more detail is worthwhile.
The consumer surplus resulting from transport services appears to be on a rising trend from under 5 percent in the 1830s, towards 20 percent in the late nineteenth century, on to nearly 30 percent in the mid-twentieth century, and then declining a little from the 1960s. The long-term increase mostly reflects the more-than-unit income elasticity of demand until the 1950s (see the top of Figure 1). This trend contrasts with the broadly declining expenditure on transport since the mid-1800s (see Figure 3). Consumer surplus for lighting also rose—rapidly to about 10 percent just after 1900 and then more gradually to 12 percent by the end of the 1930s. It declined, reflecting the fall in income elasticities. Interestingly, consumer surplus for domestic heating remained relatively stable throughout the nineteenth and twentieth centuries, offering between 6 percent and 8 percent of GDP.

The different trends appear to represent how these developments affected lives. Increased mobility radically altered society in the nineteenth century. From the 1840s, many upper- and upper-middle-class households chose to move away from the crime, sewage, and smoke of the cities. The expansion of urban railway networks made it possible for them to move to the suburbs. Income was the key to escaping the disamenities of the city, and as more people reached higher income levels, they moved to the suburbs and traveled more. The Cheap Trains Act of 1883, and a rapid expansion of British suburban housing in the 1890s offered an opportunity for lower-middle-class families to live in the suburbs, and to commute to the city (Jackson 2003). Thus, transport transformed lives, and improved people’s well-being.

Figure 5(b) shows the role of railways in increasing consumer surplus, while horse-drawn carriages continued to add value to society. However, the introduction of buses in the early twentieth century only appears to have offered a substitute to railways. Then, the growing use of cars, especially from the 1950s, enabled personalized and flexible transport, which added greatly to people’s lives. Also, it is reassuring to note that this paper’s estimate of the consumer surplus for cars in 1923 (1.6 percent of GDP) is almost identical to the estimate for cars in 1923 (1.8 percent of GDP) produced by Leunig and Voth (2011 p.13).
Similarly, gas lighting revolutionized lives (see Figure 5(c)). Cheaper gas was an ‘enabler’ of or complement to other goods and services (Fouquet and Pearson 2006). Work, education and social activities all became much easier to undertake (or cheaper to ‘produce’) at night. The expansion of street lighting also promoted urbanization, and reduced urban crime. Interestingly, lighting improvements also changed sleeping behavior. Prior to the growth in lighting use, the long nights were often broken up into two periods of sleep; however, as the day was ‘lengthened,’ work, education, and social activities replaced sleep, which became more concentrated (Koslofsky 2011). However, by the 1950s, those changes had occurred, and better or cheaper lighting had more modest effects on welfare.

Cheaper and more heating, on the other hand, did not lead to major transformations in people’s lives (see Figure 5(a)). The advent of new heating methods certainly made people more comfortable. The introduction of central heating in the 1980s did reverse the decline in consumer surplus. But these improvements did not engender new experiences and lifestyles. Thus, for the consumer, gas and electric heating mostly substituted for coal heating.

So, to summarize, the shape of the inverse-U relationship that appears with economic development varies greatly between services. For lighting and transport, the increase in the consumer surplus was dramatic. For heating, the increase was modest. The differences may also explain why the sizes of the consumer surplus estimates (relative to income) were so different—reaching 30 percent of GDP for transport, 12 percent for lighting and 9 percent for heating.

Putting the results in perspective, in 2010, the average consumer received the equivalent of £(2000)1,500, £(2000)5,800 and £(2000)1,300 in net utility from residential heating, transport, and lighting, respectively. That is, the average consumer would be willing to pay these amounts in addition to his consumer expenditure to keep his consumption of these services.

The source of this net benefit is enlightening, and relates to the difference between the value of basic and average levels of consumption. For transport, 80 percent of the consumer surplus is a result of the first 10 percent of average consumption. That is, the average consumer hugely val-
ues the basic service of transport—particularly the first 1,200 km per year—and, at current prices, this generates a great deal of net utility or consumer surplus (more than £(2000)4,600). However, the next 10,800 km per year generate only one-fifth of the net utility to the average consumer. Similar ratios occur for basic to average levels of consumption for heating and lighting. This highlights that it is crucial to not lose basic levels of service provision. It also shows that consumers will not necessarily greatly suffer if they reduce consumption by, say, 10 percent. Thus, in a market where there were marginal external costs associated with consumption (Fouquet 2011b), a reduction in consumption may have been socially optimal.

6. CONCLUSION

This paper provides estimates of the consumer surplus associated with energy services and key energy transitions occurring in the last two hundred years, and how these surpluses changed over time, and, arguably, at different phases of economic development. Inevitably, attempts to estimate the consumer surplus from major technological changes face a number of challenges and limitations. In particular, the disruptive nature of new technologies is the first challenge that needed to be addressed. Building on Nordhaus’ (1997) insight on lighting prices, this paper uses detailed data on the price and consumption of heating, transport, and lighting over the last two hundred years as the starting point.
The second challenge was the lack of information about the shape of the demand curves, and how they shift over decades. The paper offers a novel method, using historical values of the willingness-to-pay for marginal changes in the consumption of these services, combined with benefits transfers, to locate the full shape of the demand curve for heating, transport, and lighting. The method uses a behavioral model of consumer theory, rather than being based on first principles—this awaits further research.

The shifts in the demand curve use estimates of the income and price elasticities of demand for these services to produce long-run trends in consumer surplus. It is reassuring to note that the estimate of the consumer surplus for cars in 1923 (1.6 percent of GDP) provided here is almost identical to the estimate for cars in 1923 (1.8 percent of GDP) produced by Leunig and Voth (2011 p.13), which is the only study and year for which a comparison is possible. This comparison and others related to the willingness-to-pay values indicate that the method appears broadly valid within the limits of our current knowledge about how to quantify these net benefits.

The paper estimates dramatic increases in consumer surplus due to the transitions in transport services from stagecoaches (4 percent of GDP in 1830) to railways (nearly 20 percent of GDP in 1900) to cars (close to 30 percent of GDP in 2000); and in lighting services from candles (1 percent of GDP in 1800) to gaslight (10 percent of GDP in 1900) to electric lighting (13 percent of GDP in 1950). These increases reflected the transformations in societies and lifestyles that mobility and illumination provided.

The evidence also shows that not all innovations increased consumer surplus (relative to income); some innovations acted only as substitutes rather than as substitutes that provided a springboard to generate transformative forces. That is, these technologies did not significantly increase the consumer surplus derived from services provided.

Nevertheless, the results indicate the substantial benefits to society from consuming energy and producing energy services. Crucially though, as shown in Figure 4, the evidence indicates that the consumer surplus related to key energy services follows an inverse-U shape with economic development.

A first implication of the evidence is that consumers in developing economies are likely to gain greatly from growth in energy service (and also energy) consumption, although it may take a number of years for the net benefits to be observed. Another implication is that future technological development and energy transitions in industrialized countries may benefit consumers less than they did in the past.

Given the limited understanding of how the net benefits from energy transitions and R&D investment change over time, this paper also offers a practical method for modeling the long-run net benefits of new goods, technologies and services. As discussed in footnote 4, for a particular period in which the data were unavailable, it was possible to produce estimates of the consumer surplus by inputting assumed values of the income elasticities. Since the values of future income elasticities are unknown, an understanding of the trend in income elasticities can help develop plausible assumptions; thus, the method offers an opportunity to forecast the long-run net benefits of new energy technologies and transitions. This understanding and method may help to provide guidance to policymakers seeking to perform a cost-benefit analysis of future energy transitions or major technological developments, such as autonomous vehicles, or climate mitigation policies.

Now, it is worth highlighting the limitations of the results. First, estimating the net benefits to consumers is fraught with difficulties. In particular, the use of consumer surplus as a measure of utility should always be used cautiously (Silberberg 1972, Willig 1976). Second, the data have been collected from numerous sources, and the early data are subject to error. While efforts have been made to select data from reliable sources, and to create consistency, the outcome is inevitably less
than perfect. More detail on the data sources and methodology used to create the series can be found in Fouquet (2008).

Also, GDP may not be the ideal denominator. Aggregate consumer expenditure is frequently seen as a more suitable denominator as it can take account of variations in consumers saving and borrowing (Sleznick 1998). Unfortunately, aggregate consumption is not available. Alternatively, the estimates could be presented without a denominator, as a monetary value, but this limits comparison over the long run.

Another issue relates to the method used. The temporal benefit transfer offers a new way to identify the demand curve. It depends on strong assumptions, including the assumption that indirect utility remains constant over time and the assumption that in the indirect utility function, in equation (3), only the price of the service and income change. If these assumptions hold, the estimates should provide reliable information about the willingness-to-pay at different marginal levels of consumption. However, over time, they are less likely to hold for a number of reasons. First, as discussed earlier, the income elasticity of demand is not always equal to the income elasticity of willingness-to-pay (Flores and Carson 1997). Here, the assumption was that the estimates in Fouquet (2014) could be used as proxies for the increases in willingness-to-pay as income rises. If there is very limited substitutability between the energy services and the composite good, then the income elasticity under-estimate the actual values, and, if they are very substitutable, then the estimates are over-estimated.

Second, the price of the composite good, $p_{nt}$, may vary also. The price of the composite good reflects the average of many goods and services, so, it is likely to be relatively stable over a decade or two. However, in the long run (i.e., over the 50, 100 or 150 years covered in this paper), prices are likely to have varied. In fact, the average cost of living has fallen substantially over the last two hundred years. Through the income effect, this would have fed through into higher willingness-to-pay values. Now, the impact of any variation price of the composite good on the optimal level of consumption of the service and on the willingness-to-pay values depends on the cross-price elasticity. Given the previous discussion on cross price elasticities for services such as heating, transport and lighting, there may well be some impact on consumption in the long run. Hence, it is important to be aware that the results may well have under-estimated the actual willingness-to-pay and consumer surplus.

Finally, this paper discusses the welfare gains to the consumer and these should be compared with the producer surplus, government revenue and the external costs. Due to length, this paper did not include other welfare indicators. In future studies, it would be worth taking account of the non-negligible external costs associated with these technologies and energy sources. For the services discussed here, the air pollution-related costs in the United Kingdom were estimated to be close to 10 percent of GDP in the late nineteenth century, though much lower today (Fouquet 2011b). Thus, in general, these external costs appear to be smaller than the dramatic gains to society from these technological developments and energy transitions.

Nevertheless, the growing external costs associated with climate change highlight the tension between the victims and the consumers of energy use. So far, the debate has emphasized the negative aspects. To push forward the debate, it might be valuable to better understand the incentives driving consumers to continue to pollute. An ambition of this study is to help move forward the debate by providing a way to identify the “low-hanging fruit” in terms of reductions in energy use and carbon dioxide that impose minimal burden on consumption.
APPENDIX

This appendix explains the method for estimating the location of the demand curve and the consumer surplus in more detail. The central concept underlying the method is the “temporal benefit transfer” function, which assumes that willingness-to-pay values can be transferred from one time period to another. For example, in year 2, it will be assumed that the willingness-to-pay for one marginal unit of the good or service at consumption level \( s_1 \) will increase, relative to year 1, by an amount determined by the percentage increase in income, \( \frac{y_2 - y_1}{y_1} \), multiplied by the income elasticity of willingness-to-pay, \( \varepsilon_{yWTP} \), such that:

\[
WTP_2(s_1) = WTP_1(s_1)\left[\varepsilon_{yWTP} \cdot \frac{y_2 - y_1}{y_1} + 1\right]
\]

This increase in the willingness-to-pay is presented in Figure A1(a), as a shift from point \( WTP_1(s_1) \) to point \( WTP_2(s_1) \). In year 2, there is an additional point associated with the equilibrium quantity and price, \( WTP_2(s_2) \)—creating a demand curve, \( Demand_2 \). Figure A1(b) shows another shift due to the increase in income in year 3 and the income elasticity of willingness-to-pay in year 3, creating the curve \( Demand_3 \).

As shown in Figure A1(b), with knowledge of the equilibrium quantity\(^7\) in year 1 and, thus, the first marginal level of consumption \( s_1 \) and price \( p_1 = WTP_1(s_1) \), as well as percentage changes in income and the income elasticity of willingness-to-pay in year 3, new willingness-to-pay in year \( r \) for the first level of marginal consumption, \( WTP_r(s_1) \), can be identified for each year

\[
WTP_r(s_1) = WTP_1(s_1)\left[\sum_{r=1}^{r} \varepsilon_{yWTP} \cdot \frac{y_r - y_{r-1}}{y_{r-1}} + 1\right]
\]

Similarly, for each marginal level of consumption, \( s_m \), for which knowledge of the equilibrium price exists, it would be possible to calculate the willingness-to-pay in different years, provided the percentage changes in income and the income elasticities of willingness-to-pay are known, and, therefore, a temporal benefit transfer can be performed.

\[
WTP_r(s_m) = WTP_m(s_m)\left[\sum_{r=m}^{r} \varepsilon_{yWTP} \cdot \frac{y_r - y_{r-1}}{y_{r-1}} + 1\right]
\]

In other words, for each year, a series of points indicates the willingness-to-pay for a particular marginal quantity of the service. This approximates the demand curve, using the temporal benefits transfer method. It is worth noting that the demand curve and shifts in it are not affected by changes in prices of the service, since the demand curve reflects the quantity consumed at different prices. The shifts are influenced by changes in income, and other factors altering the relationship between quantity consumed and price, such as changes in the price of substitutes and complements, here represented by the composite good, \( p_x \), and in tastes. Here, as explained in the conclusion, the assumption is that only income changes affect the demand curve and the price of the composite good and tastes remain unchanged or do not affect the demand curve. This is naturally a strong assumption and is a limitation of the exercise when considering changes over decades.

---

\(^7\) As a reminder, it has been assumed that the income elasticity of willingness-to-pay (or price flexibility of income (Hanemann 1991)), \( \varepsilon_{yWTP} \), is equal to the income elasticity of demand, \( \varepsilon_{yD} \).

\(^8\) It is assumed that the price, \( p_r \), represents the point of equilibrium of demand and supply, and, thus, the consumer’s willingness-to-pay at the level of consumption, \( s_r \).
Using these demand curves, the next step is to calculate the area under them and above the price to reveal the consumer surplus. Here, rather than using integration as is normally done when a functional form is known or assumed (see, for example, Sleznick 1998), a piecewise linear approach will be used, and the consumer surplus will be broken down into small, measurable areas. For example, as shown in Figure A1(a), the area under the demand curve in year 2 (Demand$_2$) down to WTP$_2$(s$_1$) for the portion of consumption between 0 and s$_1$ can be estimated, using the Hausman (1997, 1999) method$^9$,$^{10}$.

To quantify the full consumer surplus in year 2, two additional areas (Area B and Area C) need to be calculated (see Figure A2(a)). Indeed, for the portion of consumption between 0 and s$_1$, it is also necessary to calculate the rectangular area under this triangle (Area A) and above the price line—that is, connecting WTP$_3$(s$_1$) and p$_2$, and 0 and s$_1$ (Area B). In addition, Area C needs to be calculated as the triangle between the horizontal distance s$_2$ and s$_1$ and vertical distance p$_2$ and WTP$_2$(s$_1$). Thus, the consumer surplus can be estimated as:

9. As mentioned before, this linear approximation may well be acceptable for small levels of consumption, and will be used only for the very first marginal level of consumption.

10. It will be assumed that the appropriate price elasticity to use for this first marginal level of consumption is the original one in year 1, $\varepsilon_{p1}$, not $\varepsilon_{p2}$ for year 2, or for any subsequent year. This assumption is based on the argument that the price elasticity of demand for this historical basic level of consumption is a more appropriate indicator of consumer response to price changes for basic levels of consumption today than the price elasticity at the (contemporary but much higher) equilibrium level of consumption. This assumption is intended to present an opportunity to raise an interesting empirical question, rather than to offer the definitive answer.
Figure A2: (a) Estimating the Consumer Surplus after Shifting the Demand Curve; (b) Estimating the Consumer Surplus after Shifting Multiple Demand Curves

\[ CS_{s_2} = \text{AreaA} + \text{AreaB} + \text{AreaC} \] (A4)

\[ \text{AreaA} = \frac{1}{2} \frac{(WTP_1(s_1) - s_1)}{\epsilon_{p1}} y_2 \] (A5)

\[ \text{AreaB} = \frac{(WTP_2(s_1) - p_2)(s_1)}{y_2} \] (A6)

\[ \text{AreaC} = \frac{1}{2} \frac{((WTP_2(s_1) - p_2)(s_2 - s_1)}{y_2} \] (A7)

Figure A2(b) shows the multiple triangular and rectangular areas to be calculated to estimate the consumer surplus as the demand curve shifts outwards from the original demand curve. Area A is a unique triangle connecting the first marginal level of consumption with the intercept based on Hausman’s approach—and, in this study, will reflect very low levels of marginal consumption. Area D is a similar rectangular area under a triangle as Area B. Area E is a similar triangular area as Area C.

This can be generalized, based on equation (12), which is the same as equation (A2), which estimates a whole set of willingness-to-pay values for particular marginal levels of consumption, \( WTP(s_m) \), to estimate the consumer surplus values \( CS(s_m) \), based on equation (13) or, in a simplified form, (A8)–(A11), for all demand curves. The consumer surplus will be the sum of a single initial triangle (Area A; see equation (A9)), a series of rectangular areas (B type Areas; see equation (A10))
under the Area A triangle and the C types Area triangles, and then a series of triangular C type Areas (see equation (A11)):

\[
CS_{st} = AreaA + \sum_{r=m}^{r-1} BTypeAreas + \sum_{r=m}^{r-1} CTypeAreas
\]  
(A8)

\[
AreaA = \frac{1}{2} \frac{(WTP_{s}(s_{r})s_{r})/y_{r}}{e_{ps}}
\]  
(A9)

\[
BTypeAreas = \sum_{m=1}^{r-1} \frac{(WTP_{m}(s_{m}) - p_{r})(s_{m} - s_{m-1})}{y_{r}}
\]  
(A10)

\[
CTypeAreas = \sum_{m=2}^{r-1} \frac{(WTP_{m}(s_{m-1}) - WTP_{m}(s_{m}))(s_{m} - s_{m-1})}{y_{r}}
\]  
(A11)

ACKNOWLEDGMENTS

I would like to thank Elizabeth Baldwin, Eric Beinhocker, Alex Bowen, Rafael Calel, Mona Chitnis, Antoine Dechezlepretre, Simon Dietz, Dave Donaldson, Doyne Farmer, Eli Fenichel, Meredith Fowlie, Tim Foxon, Ken Gillingham, Michael Greenstone, Jim Hammitt, David Hendry, Cameron Hepburn, Jonathan Hersh, Daniel Heyen, Matt Kotchen, Ralf Martin, Kyle Meng, Tanya O’Garra, Peter Pearson, Carlota Perez, Robert Pindyck, Neil Rickman, Lutz Sager, Gregor Semieniuk, Steve Sorrell, Richard Tol, Frederick van der Ploeg, Dirk Jan van de Ven, Catherine Wolfram and Dimitri Zenghelis for their comments and discussions. Naturally, the usual disclaimer applies. Support for this research from the ESRC and the Grantham Foundation is gratefully acknowledged.

REFERENCES


All rights reserved.


