Economic Issues, Vol. 3, Part 2, September 1998

Estimation of the Demand for Cigarettes: A Review of the Literature

Sam Cameron[1]

Abstract

This paper reviews work on the demand for cigarettes showing considerable agreement on a statistically significant inelastic price response which is much greater in the long-run. It also appears that, on balance, there are significant negative effects of health scares. The evidence on the effects of advertising is inconclusive. Restrictions on smoking in public places appear to decrease demand. At present, the rational addiction model looks set to dominate the field. Unfortunately it has, in common with the earlier work, received an uncritical response.The results gained are highly questionable; the time-series studies were carried out without reference to the literature on cointegration. Further, the rational addiction models have encountered acute problems with implausible discount rates and insignificant price terms.

1. Introduction

Economists have been estimating demand functions for cigarettes since (Schoenberg) 1933. The number of studies has mushroomed since (Bass) 1969, spurred by an interest in detecting the influence of advertising expenditure, the health warnings issued since 1953, restrictions on the right to smoke and controls on advertising. The cigarette industry has been worried by these studies and has sponsored work attacking some of the literature (Sinnott *et al*, 1979; Johnston, 1980) and apparently refused advertising data to McGuinness and Cowling (1980). To date there has not been a comprehensive survey of the literature. The most recent efforts have been a selective review by Godfrey (1989), a table of price and income elasticities, updated from a PhD by Russo, in Viscussi (1992), and a review of the effect of advertising by Smee (1992). This paper focuses on the many recent studies, especially using rational addiction theory, which are obviously omitted in the above three papers. The focus is on price elasticities, advertising, health scares and regulatory measures.[2]

It is important to consider the theory which underlies the estimates in the literature. Unfortunately this is somewhat difficult as many studies fail to give an explicit theoretical background. The apparent theory in the earlier studies (which continues to be used in some cases) is static neo-classical partial equilibrium single equation demand models derived from utility functions with exogenously given tastes. These studies have eschewed an explicit treatment of the physiological and psychologically addictive properties of tobacco with the exception of the overlooked work of Simonich (1991) who includes nicotine in his demand equation but fails to obtain a significant coefficient. As economists do not try to measure biological or social psychological determinants of addiction the only way for them to model addiction is

as a habit where consumption depends on the consumption of other periods. Harking back to the idea of herd behaviour, an interesting point (neglected in the literature, see Deaton and Muellbauer, 1980) is that dependence of consumption on that of reference groups (which is well known in the non-economic literature on smoking) can lead to the widely adopted form of an infinite distributed lag model as in (5) below.

The following sections review, respectively, the theoretical background, problems with data and estimation, price elasticities, the effect of health scares, and the influence of advertising and government regulation.

2. Theoretical background

Theoretical work, emphasizing habits in consumption, has been around since the early 1950's but the pioneering empirical work was by Houthakker and Taylor (1966) who estimated single equation dynamic demand models for 83 commodities including tobacco products. They retain the static intertemporally separable utility function but graft an adjustment function based on a 'psychological' stock of habits on to the derived demand function which is thus, in linear form:

$$Q_t = b_o + b1P_t + b2X_t + \gamma S_{t-1} + U_t \qquad (1)$$

where Q = consumption, P = relative price, S is the stock of habits, U is the classical disturbance term in a regression equation, the b's are parameters, γ is a parameter representing the influence of the stock of habits on consumption, t is the time subscript. S is formed according to the stock adjustment equation:

$$S_t = (1-\delta)S_{t-1} + Q_t \qquad (2)$$

where δ is the rate at which habits decay.

Solving (1) and (2) gives a reduced form for estimation (a form of the derivation is given in Thomas (1987)):

$$Q_t = b_o\delta + b1P_t - (1-\delta)b1P_{t-1} + b2X_t - (1-\delta)b2X_{t-1} + (1+\gamma-\delta)Q_{t-1} + U_t - (1-\delta)U_{t-1} \qquad (3)$$

the b's give the short-run effects while long-run impacts are given by:

$$b/\gamma-\delta \qquad (4)$$

Houthakker and Taylor (1966) estimated equation (3) but excluded prices because of implausible results; they also had problems of obtaining a negative depreciation rate (δ) and did not include any shift variables for health warnings. Their equation has not been much used in the literature. Only three articles in the whole pre- 'rational addiction' literature give explicit consideration to the choice of dynamics for smoking demand. Equations of the type of (3) are reported by Hamilton (1972) and Leu (1984) whilst Zanias (1987) reports having estimated it and rejected it and various other specifications in favour of the alternative given by (5). Assuming a 100 per cent rate of decay (δ = 1) then equation (3) becomes:

$$Q_t = b_o + b1P_t + b2X_t + \gamma Q_{t-1} + U_t \qquad (5)$$

It is formal equivalents of this which provide the results, apart from those using the rational addict model (equation 7), reported in Table 2. This is simply a model with a fixed propensity to habituation or addiction which is carried over from period to period. It is thus open to Becker's fundamental objection that a rational individual will observe that they have a habit or addiction and will seek to modify it by adjusting their intertemporal consumption trajectory if this can lead to utility gains.

It is rare to find instances in the literature of equation (5) being presented as estimates of a restricted Houthakker-Taylor model. For example, Goel and Morey (1995) fail to give any theoretical justification for the lagged dependent variable whilst McGuinness and Cowling (1975) and Baltagi and Levin (1986) interpret it as the speed of adjustment in Nerlove's partial adjustment model which is problematic for as Johnston (1980, p.121) points out:

> The partial adjustment model .. is commonly employed when the costs of adjustment are fairly high, whether in terms of financial, legal or institutional considerations. It is difficult to find ... reasons why the adjustment of cigarette consumption to the level indicated by current price, income and tastes should be a lag process.

He goes on to point out that McGuiness and Cowling's empirical work supports this as the estimated coefficient on the lagged dependent variable is insignificant. Adopting the habits interpretation of the lagged dependent variable estimating equation seems much more plausible and would lead us to expect a high value of γ. It would also lead us to expect it to decline over time if smoking is becoming less socially acceptable and less addictive. This is borne out in the varying parameter study of US data by Tegene (1991) in which it decreases steadily from 0.66 in 1956 to 0.15 in 1985. The coefficient is generally about 0.5 - 0.7, and highly significant, although in Seldon and Boyd (1991), Goel and Morey (1995) and Conniffe (1995) it is zero (insignificant *t*-ratio).

The theory of rational addiction was formally developed in Becker and Murphy (1988) and applied to cigarette smoking by Chaloupka (1991, 1992), Keeler *et al* (1993), Becker, Grossman and Murphy (1994), Sung *et al* (1994), Conniffe (1995) and Cameron (1997). The model in Becker *et al* (1994) assumes that smoking has no influence on future earnings, perfect foresight, discounting of the future at the market rate of interest and a stable lifetime utility function specified as:

$$U = U(C_t, C_{t-1}, Y_t, e_t) \qquad (6)$$

where C is consumption of cigarettes, Y is other goods and e is a stochastic term for unmeasured life cycle influences. The e term is an intrinsic part of the model as distinct from the u term in the earlier models which is grafted on afterwards to measure all omitted factors. Note that last period consumption enters directly into the utility function rather than indirectly through a stock of habits equation such as (3) which is equivalent to assuming a 100 per cent depreciation rate on the stock of habits ($\delta = 1$).

(6) is assumed to be quadratic and discounted at the rate of interest over an infinite life time. Maximising this subject to constraints gives a linear difference equation:

$$Q_t = \theta Q_{t-1} + \beta\theta Q_{t+1} + \theta 1 P_t + \theta 2 e_t + \theta 3 e_{t-1} \qquad (7)$$

which is the basis of the estimating equation. (5) is nested within this so that a *t*-test of the coefficient on Q_{t+1} should in principle decide between the models. Becker interprets (7) restricted to (5) as involving 'myopic' behaviour as the discount rate equals zero (infinite rate of time preference).

In the rational addict model, plausibility can also be tested by calculating the discount rate. Becker *et al* (1994) find that their model performs quite badly on this count as all the discount rates, except in the OLS estimation,

imply highly implausible interest rates. Problems with implausible discount rates are also encountered in Chaloupka (1991), Chaloupka (1992), Keeler *et al* (1993), Labeaga (1993), Sung *et al* (1994) and Cameron (1996). A notable result in Chaloupka (1991) is that estimating the model for those who currently smoke, on individual data which is surely the best test of the model, gives a discount rate of 182.5 per cent. Implausible discount rates seem to cause little concern in this new wave of smoking research. In a study which features negative discount rates, Sung *et al* (1994, p.91) proclaim in their abstract that 'These results support the theory of addiction'. Using highly suitable data (a six times observed panel for individual households) Labeaga (1993) also finds negative discount rates and seems unconcerned by this.

Conniffe (1995) finds the rational addict model to be rejected, in aggregate time-series data, on grounds of the insignificance of the lead consumption term. Chaloupka (1991) also finds this for some groups in his individual survey data.

Outside of the above developments, Young (1983), Godfrey (1986, reported in 1989) and Conniffe (1995) are the only authors to use an asymmetric specification allowing for different magnitudes of response to price fall and price rise as a reflection of addiction.

It is quite evident that the rational addiction model implies a negative price elasticity. As regards the other variables surveyed here: advertising, health information and smoking restrictions, these should generally follow the expectations developed in the earlier static and myopic studies. Advertising which does not convey information on health risks should lead to increased consumption whilst health risk information and restrictions would be expected to lead to a fall in consumption. In the neo-classical model, especially in its Austrian and Chicagoan brands, advertising serves a positive role of providing information and might thereby be expected to increase consumption for individual brands which are 'new and improved' in some way. Competition in the industry generally takes place through product differentiation rather than price. More realistically, it might be felt that advertising promotes an image that smoking is an exciting activity intimately linked with sexuality; this is particularly so in adverts aimed at youth (see Kinman and Vinten, 1995). This being the case, advertising might raise aggregate consumption by increasing the participation rates of young smokers who then become addicted. Doroodian and Seldon (1991) argue that US advertising also could have had a negative effect on consumption because of the 1965 Cigarette Labeling and Advertising Act which required a printed health warning on all advertisements. The same could apply in the United Kingdom which has similar requirements. It is possible that advertising may lead to a rise or a fall in consumption as the promotion of 'safer' brands may encourage people to break their addiction in the long run.

To add to the vagueness of the possible impact of advertising, we may note that in the non-economics literature, and in the opinion of the industry (see Kinman & Vinten, 1995), advertising redistributes demand amongst brands rather than adding to total demand thus having no aggregate impact. This opinion is also frequently stated in the econometric literature when an insignificant coefficient on an advertising variable is unearthed.

In such models advertising and health campaigns perform the role of providing information or of changing tastes. Regulations such as banning smoking in public places are always assumed, in the literature (see e.g. Sung *et al*, 1994), to have a negative effect

without the mechanism being spelt out. It is not axiomatic that such regulations will induce a reduced equilibrium smoking rate per smoker. For one thing the smoker may compensate by smoking more away from the workplace, restaurant or aeroplane. In England, phalanxes of smokers huddled outside offices for their cigarette break are becoming an increasingly common sight. In the absence of total control, the smoker may also select in favour of non-regulated environments which will generally involve a loss of welfare but it need not lead to reduced demand. Anti-smoking regulations may induce total cessation through a wish to conform to herd behaviour.

The majority of studies use aggregate data in which the non-smokers are lumped in with smokers so that it is impossible to distinguish variations in quitting and starting from variations in rates of consumption. The only aggregate time series study which separates these decisions is Jones (1989a). Chaloupka (1991) and Lewitt and Coate (1982) study the separate decisions in survey data. Jones (1989b) uses individual data to model the tripartite sequence of starting/-consumption/quitting. Douglas and Hariharan (1994) are the first to model duration of smoking behaviour. Some literature from marketing and advertising journals has been excluded as, although it uses econometric methods, the specifications or methods would be considered unacceptable by economists, for example, the exclusion of an income variable. A summary of these papers can be found in Simonich (1991, Appendix A). There are variations in the choice of dependent variable although the usual choice is packs *per capita* of adult age population, variously cut off at 14-18. The regressions are invariably run in levels; only Peto (1974) and McCleod (1986) use changes in the dependent variable and run all their regressions in changes. The overwhelming majority of studies use packets of cigarettes in the dependent variable without adjustment for the strength or tar content. Schneider, Murphy and Klein (1981) and Porter (1986) include the market shares of filter tip and low tar cigarettes as controls along with tobacco per cigarette whilst Simonich (1991) uses the amount of nicotine per cigarette. These are the only studies which take any account of cigarette heterogeneity. Experimentation with tobacco *per capita* rather than cigarettes *per capita* is rare but can be found in Schneider *et al* (1981), Skegg and Atkinson (1973), McCleod (1986), and Porter (1986). As Porter points out, tobacco seems a more appropriate dependent variable so it is somewhat unfortunate that the focus has been on cigarettes especially as the two series can diverge quite radically.

3. Econometric models

3.1 Multicollinearity

Time-series studies have almost all suffered problems of multicollinearity; in many there is a severe problem of collinearity between price and income, e.g. in Sinott *et al* (1979) $r = 0.9$. This leads some authors to use extraneous estimates for the income variable, e.g. Schneider *et al* (1981), Zanias (1987). Hamilton (1972) even uses extraneous estimates for price and income. Schneider, Klein and Murphy claim that there is no problem of multicollinearity in their time series of American data for 1930-1978 but no evidence is given in support of this claim.

Attempts to overcome multicollinearity using ridge regression are made by Fujii (1980) and Young (1983), which is a replication of Fujii using an asymmetric model. This is not a generally advisable 'solution' to the multicollinearity problem as results can vary drastically depending on the

value of k selected and as Greene (1990, p.283) points out '... it is difficult to attach much meaning to hypothesis tests which are biased in an unknown direction'.

3.2. Simultaneity

Most studies use single equation models, thus ignoring simultaneity issues. Three stage least squares has been used by Bishop and Yoo (1985) to estimate a conventional supply and demand system. Price is also treated as endogenous by Kao and Tremblay (1988), Porter (1986) and Goel and Morey (1995) who use instrumental variables rather than specifying a supply equation. There are good grounds for ignoring the supply function/ endogeneity of prices as the industry is not perfectly competitive. In support of this, Doroodian and Seldon (1991) perform a test that shows prices to be exogenous although the opposite is found by Goel and Morey (1995). In many of the European countries which have recently been studied the cigarette supplier is a state monopoly thereby eliminating any need for supply modelling. The most complete systems are those of Simonich (1991) who has three equations: a demand equation, an advertising expenditure equation and a nicotine content per cigarette equation and Porter (1986) who uses demand, advertising and price equations. 2SLS is used as the advertising and nicotine content variables are treated as endogenous variables in all of these equations. None of the other published studies since Schmalensee (1972) treat advertising as endogenous despite the fact that it is a strategic reaction to falls in consumption induced by health scares and government regulation. Goel and Morey (1995) jointly estimate the demand for cigarettes with the demand for liquor on the presumption that the goods are substitutes, which is borne out by their empirical analysis. Jones(1989a), Lewitt *et al* (1981) and Lewitt and Coate (1982) use system methods to jointly estimate frequency of smoking and participation in smoking equations.

3.3 Mis-specification

Many of the time series studies suffer from apparent mis-specification in the form of low Durbin-Watson statistics (those including lagged dependent variables seldom report an h or other appropriate test) but do not go on to correct for serial correlation with the exception of Simonich (1991), Ippolito *et al* (1979), Bishop and Yoo (1985), Kao and Tremblay (1988), Keeler *et al* (1993), Porter (1986), Doroodian and Seldon (1991), Sung *et al* (1994). This may be just as well given the inefficiency (relative to OLS) of methods which 'correct' for autocorrelation with trended data, using small samples, under classical linear regression assumptions. There is also a more general problem of the bias and inconsistency of AR(1) GLS estimators (Mizon, 1995) even with stationary data series when the 'correct' model is a dynamic one. As yet no one has examined the stationarity of the individual series or the presence of cointegration in the model with the exception of the Smee (1992) report which does not actually report the unit root or cointegration tests. Schneider *et al* (1981) have persistently low D-W statistics and report that the error process (p.599) is MA(1) rather than AR(1). They do not explain how they come to this conclusion and do nothing to deal with the problem which is surely indicative of omitted dynamic processes. One study only (Yucelt and Kaynak (1984) has negative serial correlation (D-W = 2.98) which persists in all the specifications they report although they make no attempt to deal with it. Their model excludes the effects of health scares and regulation which are found by Seldon and Boyd (1991) to be crucial to the stability of the demand equation over time.

Table 1: Cigarette demand estimates - static studies

Author/date	Data	Period	Price elasticity	Health scare	Adv*	Anti-adv**	Media ban
Stone/1945	UK A	1920-38	0.5				
Prest/1949	UK A	1870-38	0.22				
Koutsoyannis/ 1963	US Q	1950-59	0.94				
Sumner/1971	UK A	1951-67	0.24	1962-			
Skegg and Atkinson/1973	UK A and Q	1951-70	0.00 men 0.35 women	1962 1962	n s n s		- n s
Peto/1974	UK A	1952-70	0.37	1962-			-
Johnson/1988	Australia A	1961-86	n c	-	n s		0
McLeod/1986	Australia A	1954-83	0.3	-			-
Witt and Pass/1981	UK A	1955-75	0.3	1962-	+		
Witt and Pass/1983	UK A	1955-75	0.21	1962-	+		-
Townsend/1987	UK A	1961-75	0.00	1962- (socio-econ group 2 only) 1964- (socio-econ group 4 only)			
Cox and Smith/1984	15 OECD A	1962-80	0-0.5 ave. 0.21	intervention works (different method used)			
Hu *et al*/1995	California Q	1980-93	not given	anti-smoking adv. significant by state			
Leu/1984	Sweden A	1954-81	0 (real) 1 (nom.)	1964- 1979-	1968-		
Jones/1989a	UK A	1954-86	0.37 (per smoker) 0.19 (ptcptn)	62-71-77 62-71-77	 n s		
Bishop and Yoo/1985	US A	1954-80	n s	1964	n s	+	
Lewitt *et al*, 12-17 yr old/1981	US Q	1966-70	1.44 (per smoker) 1.2 (ptcptn)		n s	n s	
Schneider *et al*/1981	US A	1930-78	1.22	1964-	n s	n s	

table continued....

Table 1 cont... Cigarette demand estimates - static studies

Author/date	Data	Period	Price elasticity	Health scare	Adv*	Anti-adv**	Media ban
Lewitt and Coate, 20-70 yr old/1982	US	1976	0.42 (all)				
Ippolito, Murphy and Sant/1979	US A	1925-75	0.81	1953-	1964-	n s	n s
Ippolito and Ippolito/1984	US A	1934-80	0.48 (in 1980)	1964-variable for reports of harm since 1953			
Porter/1986	US A	1947-82	0.28	1953- 1964-	+	-	
Simonich/1991	US Q	1959-83	0.37	1964- 1966 n s 1972 n s			
Reekie/1994	South Africa A	1970-89	0.88				
National health interview study	UK	'76, 80, 83, 85, 89	0.23 0.28	1985 1989			
Wasserman *et al*/ 1991	USA	1970	n s	1970-83			

Notes: Blank spaces mean variables not included in study; n s means included but not significant; * advertising; ** generally the 'fairness doctrine'.

4. Price effects

There is an overwhelmingly strong negative relationship between prices and demand. Virtually all the studies cited here obtain a significant negative coefficient on the relevant parameter.

What makes cigarettes interesting to economists is their habitual nature. None of the models really deal with addiction as it is normally understood. The static models have no explicit treatment of the effect of addiction on equilibrium, preferring to use standard neoclassical demand theory; addiction is referred to in discussing the elasticities with a low value being regarded as a sign of addictiveness. Table 1 shows the range of short-run price elasticities. The typical value, in aggregate studies, seems to be about 0.2 with only one exceeding 0.5 (Schneider *et al*, 1981, who obtain a value of 1.22). The composition of the smoking population is ignored with the exception of Townsend (1987) who finds the price elasticity to be zero for all social classes implying that the demand curve is vertical as in the heroin literature and Skegg and Atkinson (1973) who find a vertical demand curve for men and an elasticity of 0.35 for women. Coate and Lewitt find an inelastic demand curve for women, men over 25 and men who smoke. Lewitt *et al* (1981) look at teenage smoking and find an elasticity of over 1. Using a different selection of teenagers (17 years +) Wasserman *et al* (1991) find an insignificant

Table 2: Cigarette demand estimates - dynamic studies

Author/date	Data	Period	Price elasticity	Health scare	Adv*	Anti-adv**	Media ban
Sackrin/1962	US A		0.29/0.45				
Houthakker and Taylor/1966	US A	1929-61	0.13/0.54				
Sutton/1974	US A		0.45/1.27				
Fujii/1980	US A	1929-73	0.45/0.67	1953 + 1964 n s	n s	n s	
Young/1983	US A	1929-73	s/r 0.475	1953 + 1964 n s	sig		
Kao and Tremblay/1988	US A	1954-80	0.5/0.77	1964-			
Stavrinos/1987	Greece A	1961-82	0.1/0.14	n s			
Zanias/1987	Greece A	1954-81	0.14/0.5	n s			
Doroodian and Seldon/1991	US A	1952-84	0.1/0.32		+ to 1963 0 after	-	-
Chaloupka/1991. Survey of high school educ. 17-24 and 65-73	US n=28,000	1976-80	0.2/0.45				
Chaloupka/1992 (same sample)	S	1976-80	0.27/0.36 men 0 women				
Keeler *et al*/1993	Calif. M	1980-90	0.36/0.58				
Hamilton/1972	US A	1926-70	uses prior estimate	1953- 1964-	+		
Baltagi and Levin/1986. Panel of state data.	US A	1963-80	0.35-0.48 /1.18-1.7				
Warner/1977	US A	1947-70	not given	1953- 1964-			
Becker *et al*/1994. Panel of state data.	US A	1955-85	0.4/0.75				
Goel and Morey/1995	US A	1959-82	0.28/0.34				
Sung *et al*/1994. Panel of 11 western states	US A	1967-90	0.4/0.48				n s

table continued....

Table 2 cont... Cigarette demand estimates - dynamic studies

Author/date	Data	Period	Price elasticity	Health scare	Adv*	Anti-adv.**	Media ban
McGuinness and Cowling/1980	UK Q	1957-68	0.99/0.99	1962-	+		
Radfar/1985	UK Q	1965-80	0.2	1971- 1977-	+		
Sinnott *et al*/1979	UK Q	1958-78	n s	n s	n s	n s	
Seldon and Boyd /1991	USA A	1953-84	0.22/0.37	1962-	+	-	-
Tegene/1991	USA A	1929-86	0.66/1.23(1956) 0.42/0.7(1966) 0.28/0.36(1971) 0.15/0.17(1985)		n s		
Yucalt and Kaynak/1984	USA A	1955-80	0.18/0.96		n s		
Tansel/1993	Turkey A	1960-88	0.21/0.37	-			
Valdes/1993	Spain A	1964-88	0.6/0.69				-
Conniffe/1995	Rep of Ireland A	1960-90	0.32-0.38	-			
Labeaga/1993. 413 households observed 6 times.	Spain	1978-83	Price coefficients generally **positive** and insignificant				

Notes: Blank spaces mean variables not included in study; n s means included but not significant; * advertising; ** generally the 'fairness doctrine'.

price effect[3] which is echoed in Douglas and Hariharan (1994).The usual policy conclusion drawn from the static studies is that taxes on cigarettes are a good means of raising revenue and correspondingly that taxes are not a very effective means of improving health. In the myopic and rational addiction models (see table 2 the lagged dependent variable represents the influence of habit as pioneered by Houthakker and Taylor (1966). The short-run elasticity is similar to those found in the static studies with none exceeding 0.5 with the exception of McGuiness and Cowling (1975) who obtain a value of 0.99. Chaloupka (1992) finds price insignificant for women and a long-run elasticity of around 0.3 for men. Chaloupka (1991) finds price insignificant for the high-school educated, those aged 17-24 and those aged 65-73. The long-run elasticities naturally exceed the short-run values as the habit term has a positive coefficient. The values are again inelastic with only the early study by Sutton (1974) reporting figures in excess of 1, and the recent paper by Baltagi and Levin (1992) which does not discuss the surprisingly high[4] elasticities (all over 1) from recent US data. Indeed they do not even calculate the elasticities; the figures in table 2 have been estimated from their results. Sinott *et al.* (1979) and Labeaga (1993) are the only authors to obtain a value of zero (insignificant

coefficient). The policy conclusions are still that the tobacco trade is a good tax base. Only two studies (Tegene, 1991; Wasserman *et al*, 1991) address the question of what is happening to the price elasticity over time. The evidence conflicts; in the individual-level data of Wasserman *et al*, it is zero from 1970 until 1985 whilst in the aggregate data, varying-parameter, model of Tegene it declines steadily from 1956 to 1985.

5. Health scares

Health scares arise because of the publication of prominent official reports linking smoking to diseases or a cumulation of media reports on the health effects of smoking. In Spain and South Africa there have not been any official health scares, leading Valdes (1993) and Reekie (1994) to omit this from their models. Although this appears sensible at first glance, it is surely questionable that Spanish and South African smokers made no adjustment to rising perceptions of the fear of death and illness from their habit. The health fears have been spread through general reports in the media. The initial approach to modelling health scares was to use dummies and trends with modifications for spread and return effects, as people go back to smoking once they accomodate the health information. These measures have generally proved significant and substantial with exceptions such as the Greek study by Stavrinos (1987) (which contradicts the other Greek study, using a roughly similar time frame, by Zanias (1987)) and the UK paper by Townsend (1987) who finds that the 1962 health scare dummy decreases smoking significantly for socio-economic group 2 and the scare in 1971 for socio-economic group 4 but there is no effect for other socio-economic groups. Skegg and Atkinson (1973) find that health scares are not significant for men but are for women. In most cases there is more of a tendency to find the early 1960's scares significant but not the later ones which is what we would expect. An exception to this is the work of Fujii (1980) and the replication by Young (1983) both of which find a positive and significant coefficient on the 1953 warning dummy. Given our earlier remarks on ridge regression, this should probably be disregarded. There is a problem of *ad hoc* specification searches when spread and return trend effects are allowed so some authors have striven to produce a better measure of health scare effects by using measures based on the switch to low tar brands and filter tip cigarettes (Schneider *et al*, 1981). These have a significant negative effect. However these two series are monotonically trended and are probably standing in for the trend terms so scorned by Schneider *et al* in the work of others. It is likely that they reflect corporate adaptation to a secular decline in sales of cigarettes rather than an exogenous influence on smoking *per se*. Porter (1986) uses basically the same data and model as Schneider *et al* and finds significant shift dummy effects for the health scares even when the low tar and filter tip variables are included. McGuiness and Cowling (1975), Radfar (1985), Doroodian and Seldon (1991), Seldon and Boyd (1991) and Valdes (1993) interact health scares with advertising so that they decrease the effectiveness of advertising rather than having a shift effect. In general, the health scares seem to have had an effect but it is very limited in duration. I have been investigating the data of Simonich, which is rare in being quarterly, and have been unable to find much of an impact of any kind of health scares and health information except for a one quarter only effect of the 1964 Surgeon General's health report.

6. Regulation

There have been a number of government measures to restrict smoking.

(i) Antismoking advertisements.

Hu *et al* (1995) find that anti-smoking advertising campaigns in California have a significant negative effect whilst Sinnott *et al* (1979) found no effect in Britain. From July 1 1967 to January 2 1971, the US government operated the 'Fairness Doctrine' under which antismoking messages were broadcast free at a ratio of one per three cigarette commercials broadcast. The measure was then replaced by a total advertising ban. The dominant impression from the literature is of no effect although some find a negative effect (e.g. Doroodian and Seldon, 1991, and Seldon and Boyd, 1991) although Goel and Morey (1995) find a significant positive effect. As indicated above, a positive effect might occur if smokers switch to more cigarettes but less tobacco/nicotine/tar in the shape of filter and low tar brands. If this is the case then the models have been mis-specified by not modelling such switching (see next paragraph).

(ii) Tv and radio broadcasting bans.

These have been introduced in the UK, USA, Australia and Greece. Bans have been found to have significant and negative effects by all those who have studied them except Witt and Pass (1981, 1983), Schneider, Klein and Murphy (1981), and Baltagi and Levin (1986). Schneider *et al* follow earlier studies by adding a ban dummy to the model which (*ibid* p.580) makes the advertising variable significantly negative. This leads them to point out the weakness of the ban dummy which '..does not measure changes in the average productivity of advertising expenditures which is the primary mechanism by which the ban is likely to directly affect consumption' (ibid.p.589). Accordingly they specify a model in which the productivity of advertising shifts after the ban but there is no dummy variable. Finding that advertising is not significant they go on to conclude that the advertising ban did not have any direct impact.

Recent papers by Seldon and Boyd (1991), Doroodian and Seldon (1991) and Goel and Morey (1995) ignore the paper by Schneider, Murphy and Klein and continue to use the advertising plus broadcasting ban dummy specification used by earlier writers. There is a severe deficiency in the United States advertising data, ignored in all the econometric work, which makes it very difficult to assess the impact of the advertising ban. In 1975 the Federal Trade Commission greatly expanded its coverage of advertising expenditures with the consequence that the series, which slumped after the ban, climbs steadily after 1978.

(iii) Warning labels on packs of cigarettes.

The US government introduced warning labels in January 1965, subseqently modified the wording and in January 1971 required tar and nicotine listings. In the only study of these policy interventions, Simonich (1991) finds no effect. This is undoubtedly due to severe collinearity with other constructed variables for health scares and policy information.

(iv) Anti-smoking ordinances.

Increasingly, restrictions on smoking have been introduced in restaurants and workplaces. Six studies, of the United States where such measures began in 1974, have taken account of this: Simonich (1991), Wasserman *et al* (1991), Chaloupka (1992), Keeler *et al* (1993), Sung *et al* (1994) and Leigh (1995). All obtain some indication of a significant negative effect. Chaloupka finds that the effect exists for men but not for women.

Using the same measure, Leigh finds a significant effect for women as well using the 1986 wave of the Panel Study of Income Dynamics although he reports only an equation for probability of smoking rather than rate of consumption. As indicated earlier it is not clear why regulations reduce aggregate smoking and these studies do not shed any light on the mechanism.

Given their endorsement of the widespread belief that smoking causes life threatening disease, economists have welcomed the general finding that policy intervention works. Unlike most other researchers, Schneider *et al* (1981), following Hamilton (1972), conclude that the advertising ban in the USA has been a bad thing. It is worth quoting them in full:

> Given the trivial effect of advertising on aggregate cigarette consumption, governmental prohibition of broadcast cigarette advertising could not have had any significant effect in terms of reducing demand. Instead we find that the policy actually increased demand producing detrimental effects on both social health and consumer welfare. The decrease in the quantity and effectiveness of advertising expenditures produced by the ban is likely to have lowered the quality (utility) of cigarettes consumed by making it more costly for consumers to buy the 'advertising services' they demand. In addition, the decrease in advertising expenditures appears to have led to a lower (non-quality-adjusted) cigarette price and hence to increased consumption. Further, the ban increased consumption via the elimination of the antismoking commercials. Finally, the ban substantially increased the cost to firms of introducing new low-tar brands and the cost to consumers of obtaining information about these newer brands, thus slowing down the movement to these lower tar cigarettes. (Schneider *et al*, 1981, p.610).

It is not at all clear that these inferences can be drawn from the data. It is not conclusively established that advertising does not affect aggregate consumption (see below). The other factors mentioned have not been included in the empirical work reported by Schneider *et al.* and are thus pure speculation.

7. Advertising

Results on advertising have been very mixed although there is a general feeling (Doroodian and Seldon, 1991 p.363, Johnson, 1988) that the aggregate effects are zero. Reekie (1994) dismisses the idea of even including advertising in his model on the grounds that all previous studies have found it to be insignificant. Inspection of Tables 1 and 2 clearly shows that this is nonsense. There have been four methods used to measure advertising:

(i) The current level of real advertising expenditures defined as nominal advertising expenditure divided by the retail price index (Witt and Pass, 1981, 1983) with a significant positive effect being found.

(ii) The use of distributed lags on real advertising expenditure e.g. Hamilton (1972) who finds a small positive effect; Doroodian and Seldon (1991) who find that advertising increased aggregate demand up to 1963, in the US, after which time the effects have been nullified by health scares and anti-smoking so that further advertising expenditures are simply redistributing market share within the industry. Baltagi and Levin fail to find a significant effect. Brown (1978) shows that

advertising does provide extra market share for individual brands.

(iii) Some estimate a 'stock of advertising' variable which declines at a rate fixed by the researcher (Schneider *et al,* 1981; Porter, 1986; Simonich, 1991). This is highly questionable for studies using annual data, and even for those using quarterly data, as the general finding in the literature for non-durables is that any effect of advertising is a matter of months rather than years (Ashley *et al*, 1980).

(iv) Telser (1962), McGuiness and Cowling (1975), Radfar (1985) and Seldon and Boyd (1991) are the only authors to use a measure of the stock of advertising 'messages', defined as nominal advertising expenditure divided by an advertising price index, which they find to be positive and significant.

Of the more recent US pooled cross-section/time-series studies, Goel and Morey (1995) find a significant positive effect of advertising in contrast to a broadly similar study by Baltagi and Levin (1986); the main difference being that Goel and Morey use a one-period lag on advertising. It is also notable that, unlike all of the other studies, Goel and Morey include sports promotion expenditures as well as media advertising. It is also worth re-emphasising the point made above that none of the studies cited in this paragraph make any allowance for the purely artificial jump in the advertising expenditure series after 1975. It should also be noted that advertising data are only available at national level.

Of the other studies which do find a positive impact of advertising, those by Fujii (1980) and Young (1983) may well depend crucially on the use of ridge regression whilst that of Witt and Pass (1983) depends on the way in which the shifts due to health scares are modelled. In Witt and Pass a number of different specifications of limited duration health scare effects are reported. Most of these have unsatisfactorily low Durbin-Watson statistics and insignificant advertising coefficients. I focus on the one which has a satisfactory Durbin-Watson value which turns out to have a significant positive coefficient on advertising.

None of the advertising stock or distributed lag of advertising expenditure studies have found a significant effect with the exception of Porter (1986). Simonich (1991) argues that concentrating on advertising *per se* ignores the strategic reaction of firms to the threat of lost revenue from health information by introducing new lower risk cigarettes with filters and/or less tar and nicotine. His is the only model to include the number of new brands as a determinant of consumption, with it being found to be a significant positive influence. The significance of the new brands variable is a little puzzling as the majority of the new brands have failed totally to capture any of the market (Porter, 1986, p.450). Valdes (1993) reports that the 1978 partial ban on TV and radio advertising in Spain lead to the retirement of strong cigarette brands and their replacement by filter cigarettes. In 1982 the government showed recognition of this when their latest control measure expressed a willingness not to damage this shift; this would have pleased Schneider, Murphy and Klein.

In conclusion, we may note that the most forceful evidence in favour of the advertising-causes-sales hypothesis is that the best measure of advertising (advertising messages) is found to have a significant positive coefficient. Nevertheless the simultaneity problem which has not really been explored adequately since Schmalensee (1972) hangs over this. Also there are problems of