9th Postal Economics Conference on E-commerce, Digital Economy and Delivery Services

University of Toulouse, Mar 31 - Apr 1, 2016

The Impact of E-substitution on the Demand for Mail: Some Results from the UK *

Frank Rodriguez (Oxera), Soterios Soteri (Royal Mail Group) and Stefan Tobias (CER)

1. Introduction

In recent years there has been a significant reduction in the volume of addressed letter mail in most developed countries including the UK (PwC, 2013). Much of this decline has been due to the substitution of letter mail by electronic modes of communication, which we refer to as e-substitution. Econometric estimates of this effect in the UK using data from Royal Mail and methods outlined in Veruete-Mckay et al. (2011) and Jarosik et al. (2013) indicate that in the past few years this process has been advancing rapidly although other factors such as increases in GDP have mitigated some of the negative impact on letter mail volumes from e-substitution.

It is clear that the prospects for the addressed letter mail market in the UK and elsewhere will depend fundamentally on the future course of e-substitution. In order to consider how this process might develop in future years it is important to gain a better understanding of the pattern and extent of e-substitution to date. The impact of e-substitution on the demand for mail potentially varies across different types of mail; for example, its impact on social mail or on advertising mail may differ from that on business (or transactional) mail both in scale and process (PwC, 2013; USPS, 2010). In this paper, which is empirical in nature, we focus on addressed business to consumer (B2C) business mail, which constitutes a little under a half of all addressed inland mail in the UK, and consider the development over time of e-substitution for this type of mail.

We use survey data from Royal Mail to disaggregate in a number of ways the overall trend in e-substitution as estimated using an updated version of the econometric model reported in Veruete-McKay et al. (2011). In particular we provide indicative estimates of the extent to which e-substitution has impacted B2C business mail by content type (for example, bills, financial statements, legal documents and other business letters), sender group (for example, banks, government, retailers and other industry sectors) and age group (in particular, age groups from 16-44 year olds up to 65 and over).

One theme we emphasise in the paper is more apparent when considering segmentations of mail sent and received: whether addressed mail is sent often depends not only on decisions of senders but also on the ability and willingness of recipients to accept communications electronically instead of by letter mail (Nikali, 2008). In the current paper's context of business mail, this is an area considered in theoretical terms by De Donder et al. (2015) who show that the demand for letters

* The views expressed in this paper are those of the authors and do not necessarily reflect those of their affiliated organisations.

versus electronic communications, which is driven by a requirement to complete business transactions, may lead to substantial variations in the extent to which letter volumes will decline in different segments. We indeed find considerable differences in the impact of e-substitution on business mail segmented in the three ways outlined above (content type, sender group and age group). In these results we add at a more disaggregated level to those of Nikali who examined broad sender-recipient groups (e.g. total B2C versus total B2B groups) and found substantial variations in the extent of e-substitution at that higher level of aggregation.

We proceed as follows. Section 2 reports historical estimates derived from econometric analysis of the overall impact of e-substitution on business mail. Section 3 uses survey based data to disaggregate these estimates for B2C business mail by content type, sender group and age group and discusses reasons for the significant differences in the extent of e-substitution for each of these segmentations of traffic. Section 4 summarises briefly and draws together our conclusions.

2. Overall Trend in the E-substitution of Business Mail in the UK

The process of the substitution of business mail by electronic alternatives has been at work at least since the 1980s (via, for example, the use at that time of facsimile machines and introduction of direct debit payments) and possibly earlier. At the same time, technology has also encouraged some growth in mail volumes as well as contraction and the overall net effect of technology on mail volumes began to have a discernible negative impact only later. The technologies (both hardware and software) underlying this process have developed and changed over time moving through telephone, fax, electronic banking and dial-up Internet to Internet access by broadband and smart phones. The diffusion path of each of these technologies through time can be represented as being S-shaped in form and the progression of these technologies can be captured as a sequence of time dependent diffusion curves which form a "large corrugated S-curve" (Nikali, 2008).

Two implications follow from this representation of the process of e-substitution for the econometric modelling of the demand for mail over time. First, it is not possible, at least in some cases, to identify directly the replacement of a specific communication which previously had occurred through paper mail but which now takes place through an electronic medium. In turn, that implies that it is not possible to observe and so measure directly an unambiguous aggregate of "e-substituted mail". Second, given that the process leading to the e-substitution of mail is multi-dimensional, it is unlikely that in the time series modelling of the demand for mail a single variable will be able to adequately proxy the effects of e-substitution on mail volumes over time.[1] Given these points, instead we adopt an approach of estimating an econometric model which includes not only variables that are standard in demand for mail models such as GDP, prices and demography but also linear "unexplained" time trends that can be added to capture structural breaks in the time series. Interpretation is required of such breaks but we draw on other sources of information that indicate these are likely to be due to the effects of e-substitution.

[1] Examples of studies which have had some success in modelling the effects of e-substitution when using a single variable (at least for a time before diffusion curves begin to flatten) include: Trinkner and Grossman (2006) whose main variable to proxy the impact of e-substitution was the use of e-banking; Nikali (2008) (indicators of e-mail use by sender- recipient type); Soteri et al. (2009) (internet advertising as a proportion of total advertising); and Boldron et al. (2010) (proportion of households with broadband subscription).

Using annual data on letter traffic from Royal Mail, Veruete-McKay et al. (2011) estimated a model of this type of the demand for mail in the UK which segmented overall mail volumes into three main categories: commercial mail, social mail and advertising mail. The first of these is predominantly business (or transactional) mail of which, in turn, about three quarters is B2C.[2] Results for the commercial mail model are reported in Table 1. Their equation for commercial mail included estimates of elasticities for GDP, household numbers, mail prices, telecomm prices and quality of service with the elasticities for prices and quality of service (as well as the changes in these variables during the sample period) being relatively small compared with those for GDP and household numbers. The model also included an "unexplained" time trend term which indicated a structural break in the time series in 2002 and was attributed to the effects of e-substitution. In the early 2000s, for example, there was an acceleration in access to the Internet and the development of broadband services in the UK. The percentage of households with access to the Internet in the UK rose from 13% in 1999 to 25% in 2000, 36% in 2001 and 42% in 2002 (Office for National Statistics, 2015a). More particularly, broadband access rose from virtually 0% of households in 2001 to 11% by the end of 2003 and reached 50% by the start of 2007 (Ofcom, 2005, 2014).

The results for an updated version of that model are also reported in Table 1 where the data set has been extended by five years to 2012. The sample period now includes data points for the years of the great recession during which GDP in the UK fell by close to 5% over the two years 2008 and 2009. This compares with average GDP growth during the preceding five years of a little under 3% per annum and so represented a swing during 2008 and 2009 of about 10% in the level of GDP from a trend which would have implied growth of rather more than 5% over the two years.[3] Despite the scale of the shock from the dramatic downturn in the UK economy, the elasticities for GDP, prices and quality of service and the coefficient on the trend term beginning in 2002 are all reasonably stable when estimated over the extended sample period. However, there was a major change in the model as part of its re-estimation. Letter volume data and business information on individual customer communication strategies suggested an increase in the rate of e-substitution during this period. This hypothesis was tested in the econometric model using a trend break term in 2010 and could not be rejected, so adding a second and large trend term to the model. It seems likely that the great recession led firms to place even greater emphasis on lowering cost levels and increased their use of electronic means of communication as part of that process. In the next section we report estimates of the extent to which this acceleration in the rate of e-substitution varied across various disaggregations of B2C business mail.

The estimated coefficients of these time trend terms can be converted into an e-substitution index in year t, E_t. The index is defined as E_t = (1 - the proportionate loss of mail to e-substitution) where $(0 < E_t \leq 1)$ such that E_t = 1 represents a year when there had been no overall net negative impact from e-substitution while E_t = 0 would be a year where all mail would have been lost to e-substitution. From the results reported in Table 1, the last year for commercial mail for which E_t = 1 is estimated to have been 2001 and since then E_t has decreased as calculated from the time trends in the model. For the purpose of understanding the impacts of e-substitution on B2C business mail

[2] In addition to business mail, commercial mail also included relatively small volumes of publishing material and lightweight parcels.

[3] Calculated from time series for UK GDP at market prices , variable "ABMI", data release 23 Dec 2015 (ONS, 2015b)

Table 1: Model of Demand for Commercial Mail Per Household: Econometric Results

	Veruete-McKay et al. 2011[1]	Updated model
Estimation period	1980/81 to 2007/08	1980/81 to 2012/13
Estimated coefficients (t-statistics in brackets)		
Economic activity[2]	0.97 (7.5)	0.90 (8.3)
Mail price index[3,4]	-0.19 (-1.2)	-0.13 (-0.9)
Telecomm price index[3]	0.10 (2.0)	0.18 (3.1)
Quality of service	0.34 (5.4)	0.19 (2.4)
Time trends, estimates p.a.[5]	2002 onwards -2.9% (-9.4)	2002 onwards -3.4% (-12.8) 2010 onwards -5.5% (-5.3)
Total net impact of all time trends from 2010 onwards[6]	-2.9%	-8.9%
Diagnostic tests and goodness of fit		
R^2 adjusted	0.99	0.99
Reg SE	0.014	0.018
Durbin Watson	1.52	1.75
Serial correlation (p-value)	0.42	0.68
Heteroscedastity (p-value)	0.29	0.76

Notes

1. Veruete-McKay et al. combined business data and information from two surveys to derive time series data for letter volumes by content type. Two different methodologies (M1 and M2) were used to combine information from the surveys and these yielded similar results. The estimates reported above refer to coefficients estimated by method M1.. Cazals et al. (2011) report results from a model using additional statistical techniques to combine methods M1 and M2, again, this yielded similar results.
2. As measured by Gross Domestic Product (GDP)
3. Deflated by the all items Retail Prices Index.
4. The use of survey information to create time series data by content type adds a greater degree of noise to the data set. In order to allow for this in the econometric analysis, the critical values for the price variable t-ratios were relaxed.
5. Time trends for 2002 and 2010 onwards refer to UK financial years 2002/03 and 2010/11 respectively. The 2002 time trend variable is therefore equal to 1 in 2002/03; 2 in 2003/04; 3 in 2004/05..... Similarly, for the time trend variable for 2010 onwards.
6. The total net time trend impact for each content type is equal to the sum of the individual content time trend effects.

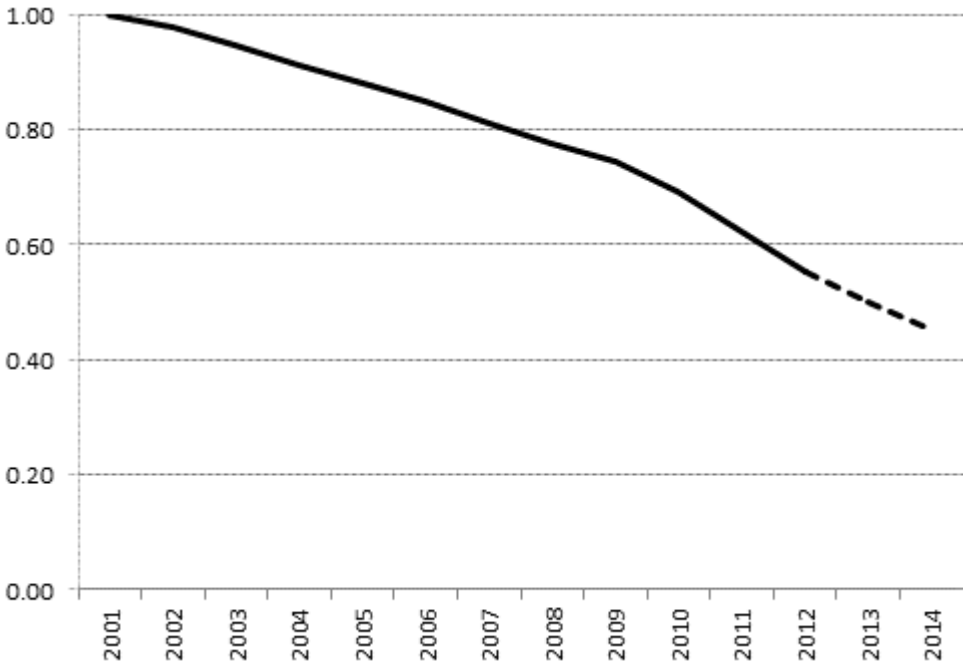
and given the very high proportion of the volume of commercial mail accounted for by B2C business mail, we use the series estimated from the data set for commercial mail to proxy that for B2C business mail and also, for simplicity, incorporate within the index the relatively small impact from prices on volumes estimated by the model. Define T_1 as the econometric estimate of the unexplained annual time trend from 2002 (expressed as a proportion, with this term including also the relatively small annual impact from prices in the updated econometric equation in Table 1) and T_2 as the estimate of the additional annual time trend from 2010 where $T_1 < 0$ and $T_2 \leq 0$ over the period since 2001. Then:

$$E_t = (1 + T_1)^{n1t} . (1 + T_2)^{n2t} \quad (1)$$

where $n1t$ is the number of years in year *t* since 2001 after which the negative time trend is first estimated to have come into effect (for example, in 2007 $n1t$ equals 6) and $n2t$ is the number of years in year *t* since $T_2 \neq 0$ (that is, 2009) and *n2t* is equal to zero previous to 2010 (so that, for example, in 2007 $n2t$ equals 0 but in 2011 $n2t$ equals 2). The e-substitution index so calculated has declined since 2001, as shown in Figure 1. From the results in Table 1, after an initially relatively small negative impact e-substitution of business mail has accelerated since the late 2000s.

In 2009, the value of the E-index was estimated to have been 0.74, that is, business mail was estimated to be only 74% of the level it would have been if there had been no electronic substitution of such mail (based on the coefficients in the econometric model including the negative time trend

Figure 1: Estimates of E-substitution Index, E_t , for Business Mail to 2012 (2001=1)*



Source: Royal Mail

* E_t equals (1 - proportionate loss of mail to e-substitution) where E_t = 1 implies no e-substitution (last such year estimated as 2001) and E_t = 0 implies complete loss of all mail. Estimates for E_t based on econometric equation for commercial mail, see Table 1.

from 2002 implying an average decline due to e-substitution (including price effects) of rather less than 4% per annum between 2001 and 2009).[4] By 2012, the index for B2C business mail in total was estimated to have been 0.55 such that the rate of decline due to e-substitution accelerated to about 9% per annum between 2010 and 2012 (based on the impact of the second time trend in the econometric equation from 2010). Note that while the time paths of the explanatory variables used in models to try to capture the effects of e-substitution (such as the proportion of households with

[4] As noted, these price impacts were relatively small and nearly all of the effect on E_t was from the estimated time trend term in the econometric equation.

access to the Internet) are generally S-shaped and so point to diminishing e-substitution as diffusion proceeds and matures, in fact e-substitution has continued to proceed at a rapid rate. This suggests strongly that multiple effects are likely to be at work in line with the construct of a corrugated S-shaped curve as a representation of the overall development of e-substitution. More recently, business mail in the UK has continued to decline broadly in line with its post-2010 historical trend which suggests also that the negative impact of e-substitution on mail volumes, while advancing at a rapid rate, has not been increasing further. This is shown in Figure 1 by the dashed line covering the period after 2012.

These estimates of overall substitution are used as an input to the analysis in Section 3. There they are disaggregated to give estimates of e-substitution of B2C business mail for specific content types, sender groups and age groups.

3. Disaggregated Estimates of the E-substitution of B2C Business Mail

i) Data and methodology. There have been marked differences in the time paths of e-substitution when disaggregated across different segmentations of B2C business mail traffic. Three are considered in this paper. First, differences by the content type of B2C business mail, *i* (*i* = 1, ... 6: Bills, invoices; Business Letters; Insurance/legal/financial documents; Other financial correspondence; Financial statements; and All other content types); second, by sender group, *j* (*j* = 1, ... 6: Banks; Government (including Health and Education); Insurance; Retail; Utilities; and All other sender groups); and, third, by age group, *k* (*k* = 1, ... 3: 16-44; 45-64; and 65 and over).

Data to estimate the time path of e-substitution for each of these segmentations were derived from a survey of mail sent and received by households.[5] The last full calendar year of data available at the start of this study was the data set for 2012. Respondents in the survey completed a detailed diary of mail sent and received each day with the questionnaire used in the survey containing questions not only about the number of items of mail sent and received but also the overall content of the mail (content type, *i*) and its origin (sender group, *j*). The questionnaire also recorded information on the individuals completing the questionnaire, including the age of the recipient of mail (age group, *k*). Given that the data for producing these estimates of e-substitution were collected through a continuing survey of a sample of individuals, the estimates are best viewed as indicative of trends over time and subject to some element of error and noise.[6]

The starting point for the disaggregated estimates of e-substitution by the three segmentations was the time series for the overall e-substitution of business mail, E_t , reported in Figure 1 where, as noted in Section 2, this measure was taken as a proxy for the time series of E_t for B2C business mail. This was used to derive time series of e-substitution indices for each of the segmentations. In order to derive estimates that were consistent across the three segmentations we considered, first, traffic by content type (*i* = 6) by sender group (*j* = 6) jointly or 36 (*i* X *j*) sub-groupings of traffic in total. The overall method is outlined here for the disaggregation by one content type by sender

[5] The survey has been in place since the 1990s and is managed and operated by an outside market research agency for Royal Mail.

[6] Various challenges were faced in constructing the time series from the survey data. These included periodic changes to the questions asked in the survey, changes in the definitions of groupings of mail and sampling error. The estimates of e-substitution reported tried to account as fully as possible for these factors.

group pair *ij* (for example, Business Letters sent by Banks). The total volume of B2C business mail of type *ij* received by households in year *t* (> 2001), Q_{ijt} , can be expressed as follows:

$$Q_{ijt} = Q_{ij,t=0} \,.\, E_{ijt} \,.\, \left(1 + g_{ij}.G_t\right).(1 + p_{ij}.P_t) \quad (2)$$

where $Q_{ij,t=0}$ is an estimate of the total volume B2C business mail of type *ij* received by households in year t = 0; E_{ijt} is an index of e-substitution for B2C business mail of type *ij* relative to the base year (E_{ijt} = 1 in 2001 and $0 < E_{ijt} \leq 1$ for *t* > 2001) of the same form as E_t; g_{ij} and p_{ij} are elasticities of the volume of mail of type *ij* with respect to, respectively, GDP and population; and G_t and P_t are the cumulated proportionate changes in GDP and population by year *t* from *t* = 0 (2001). In (2), therefore the volume of a particular type *ij* of B2C business mail in year *t* is expressed as its volume in year 2001 multiplied by three factors impacting on that level in subsequent years: e-substitution (including price effects), GDP and population. Note that, consistent with Veruete-McKay et al. (2011), population is introduced separately to reflect approximately delivery point growth and its additional effect on demand for mail rather than the direct impact of population on total economic activity which is captured by the GDP term.[7]

The volume of B2C business mail of type *ij* that would have been sent if there had been no impacts from population or GDP but only from e-substitution, Q'_{ijt} , is given by the base year volume, $Q_{ij,t=0}$, multiplied by the e-substitution index, E_{ijt} , or:

$$Q'_{ijt} = Q_{ij,t=0} \,.\, E_{ijt} \quad (3)$$

Substituting into (2) and rearranging, an estimate of the volume of mail of type *ij* that would have been sent in year *t* if there had been no impact on volumes from GDP or population but just e-substitution is given by:

$$Q'_{ijt} = Q_{ijt} / \{\left(1 + g_{ij}.G_t\right).\left(1 + p_{ij}.P_t\right)\} \quad (4)$$

Although ideally there would be econometric estimates of the elasticities g_{ij} and p_{ij} for each traffic type *ij*, in practice these are not available due to lack of data and it was assumed that g_{ij} = g and p_{ij} = p for all *ij* pairs.

Define $a_{ij,t=0}$ as the share of total B2C business mail volume of type *ij* in year *t* = 0 where $Q_{t=0}$ is that total and a'_{ijt} as this share in year *t* after deducting the estimated impacts of GDP and population since year *t* = 0 where Q'_t is the corresponding total volume of B2C business mail across all content types. The assumption that GDP and population elasticities are equal across types *ij* of B2C business mail traffic implies that the volume share that allows for the effects of GDP and population, a_{ijt} , (which is the share observed empirically) is equal to that after deducting GDP and

[7] The demographic variable in Veruete-McKay et al. (2011) is the number of households rather than population but the latter is used here as a proxy.

population effects, a'_{ijt} , as the scalar transforming Q_{ijt} into Q'_{ijt} in (4) would be the same for all types *ij*. Then rearranging (3), substituting and simplifying leads to:

$$E_{ijt} = (a'_{ijt} \,.\, Q'_t)/(a_{ij,t=0} \,.\, Q_{t=0}) = (a_{ijt} \,.\, E_t \,.\, Q_{t=0})/(a_{ij,t=0} \,.\, E_0 \,.\, Q_{t=0}) = E_t(a_{ijt}/a_{ij,t=0}) \quad (5)$$

as E_0 = 1 by definition. From (5), the estimates of e-substitution by type *ij* in each year *t*, E_{ijt}, were derived by factoring the overall index of e-substitution shown in Figure 1, E_t , (interpreted here as a proxy for E_t for B2C business mail in total) by the ratio of the volume share of that type *ij* in year *t* to the equivalent volume share in *t* = 0.

The application of this method was more complicated in practice because of the nature of the data set given that it was sourced from a sample survey undertaken over time. This involved applying various techniques to allow the estimates of E-indices to satisfy a number of constraints by re-estimation and re-allocation of the volume shares used to calculate E-indices in (5). These constraints were based on overall model assumptions (notably that by definition $E_{ijt} \leq 1$[8] and further that the process of e-substitution as it has unfolded is essentially uni-directional and not reversible so that E-indices were constrained to be non-increasing over the sample period) as well as internal survey data from other sources. The time series for the e-substitution indices for content type, E_{it} , and sender group, E_{jt} , were calculated from aggregations of volume shares so estimated at the *ij* level. The E-indices by type of traffic, *ij*, were disaggregated further to estimate a time series by age group of recipients. Volume shares by age group for each *ij* pair were calculated but given the small sample size in some of these disaggregations and data issues generally, there was inevitably a degree of noise in these estimates. Volume shares were re-estimated and re-allocated after applying constraints similar to those for traffic types *ij* based on overall model assumptions and other internal survey data, here at a segmentation with content type, *ik* (for example, other survey results confirmed that the extent of e-substitution generally diminishes as the age of recipients of mail increases and this was applied by constraining estimates of E-indices not to decrease with age in a given year).

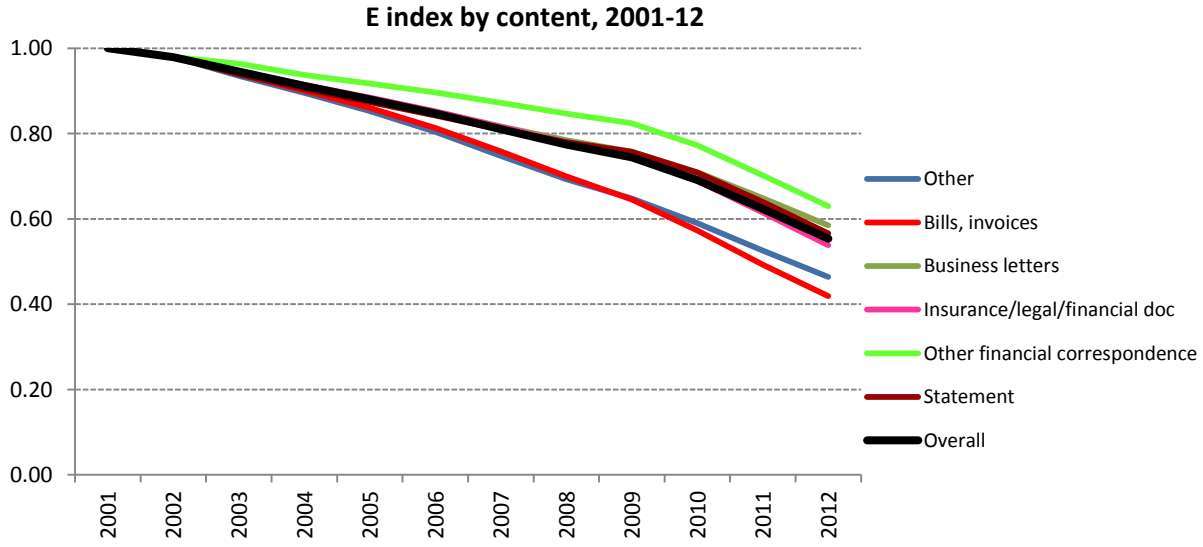
ii) <u>Estimates</u>. We begin by reporting our indicative estimates of e-substitution by content type, *i*. The content type of an item of mail reflects the purpose in sending mail and so is an important segmentation in understanding mail flows and the process of e-substitution. Figure 2 reports the estimated time series of the E-indices for content types *i* including also that inferred for B2C business mail as a whole as reported in Figure 1. One of these content types is the residual grouping "Other" which included categories such as mailings of acknowledgements, cheque books and payments. All of the time series show substantial declines over time, particularly after about 2009 in line with the overall average, indicating the significant impact of e-substitution across all content types. However, there are some clear differences between these. E-substitution is estimated to have advanced most in two: "Bills, invoices" and "Other". Its impact on volume appears to have affected least the content types "Other financial correspondence" and "Business letters".

[8] In particular, constraints were introduced to prevent E_{ijt} going above 1 in the early 2000s (for example in the Government sector).

Consider first the content types where e-substitution has advanced furthest. "Bills, invoices" represents mail that contains clear standardised instructions and requests for action. For example, senders, such as credit card providers or utilities, wish to be paid an outstanding sum of money by the recipient of the communication. Typically there is an expected and regular communication requesting a specific payment which can be automated by both sides (the sender via electronic bills and the recipient by setting up a direct debit or other electronic transfer mechanism) and such situations may be more likely with mailers who send large volumes of mail. This repeated interaction may build trust between the two parties such that the recipient agrees to provide the sender with an e-mail address and personal banking details allowing businesses to automatically deduct variable funds, or arrange for specific electronic payments to be made, for convenience or environmental reasons. It also may be the case that this is a condition for the sender to be willing to provide its service to the recipient. The overall effect of these factors is to increase e-substitution of "Bills, invoices" traffic. With regard to "Other", for many years banks and companies in sender groups such as "Utilities" have been actively discouraging the use of cheques and settlement of payments through mail. A number of banks no longer issue cheque books offering instead electronically based alternatives. These developments are likely to be the main reasons for the significant advance in e-substitution for this content type.

The two content types where e-substitution is estimated to have advanced least are "Business letters" and "Other financial correspondence" which together accounted for about a half of all B2C business mail in 2012. For these types of mail, the frequency of mailings to the same recipient is

Figure 2: Estimates of E-substitution Index by Content Type, E_{it} , to 2012 (2001=1)*



Source: Royal Mail

* E_t equals (1 - proportionate loss to e-substitution) relative to a base year, where E_t = 1 implies no e-substitution relative to that base year (here 2001) and E_t = 0 implies complete loss of all mail.

likely to be considerably less and cover a more diverse and *ad hoc* range of communication than, for example, bills and invoices. Examples would include communication to satisfy legal requirements or information requested by the recipient or required by the sender. The greater degree of non-standard communication for these categories and the lower degree of contact with recipients is

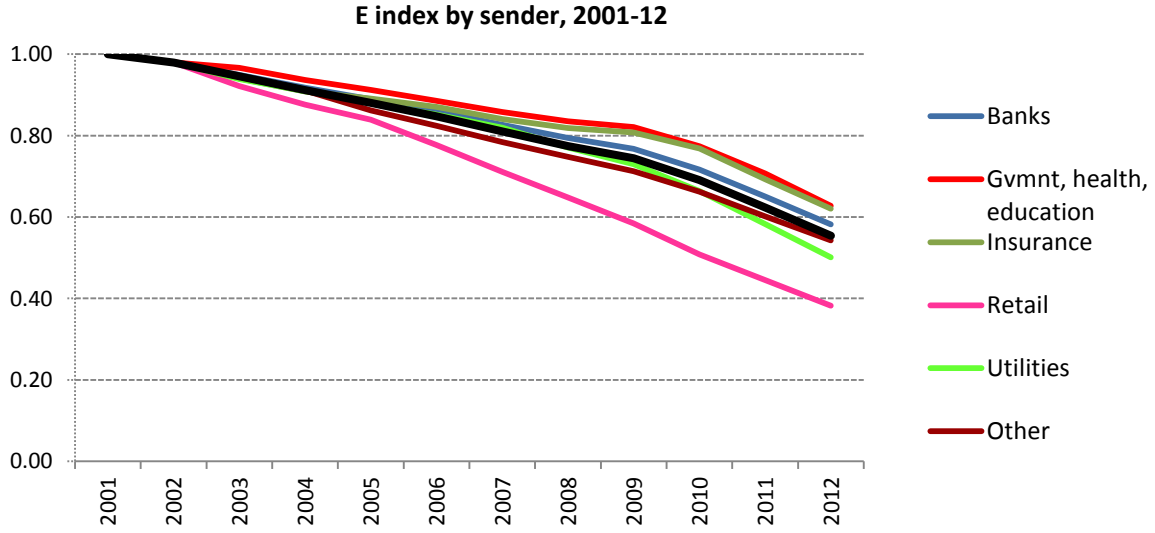
likely to mean that senders have less information on a potential recipient (for example, only a postal but not an email address is known) and the fact that contact is intermittent limits both the opportunity and perhaps also the incentive to communicate electronically.

Of the remaining content types, where the effect of e-substitution is estimated to have been roughly in line with the overall average for B2C business mail, "Statement" represented about a quarter of B2C business mail in 2012. In some ways this content type is similar to "Bills, invoices" in that it consists of communication which in most cases involves regular contact between the sender and recipient and can be, and in some cases is, automated. However, the reasons for sending such mail are more wide-ranging including legal requirements, good business practice, meeting customer needs and enhancing the sender's brand. To 2012, at least, e-substitution of statements may not have progressed further because it is perceived by some senders to be fulfilling business requirements and objectives.

We consider next the extent of e-substitution by sender group *j*, indicative estimates for which are shown in Figure 3. As in the case of disaggregation of traffic by content type, there are estimated to be marked differences in the impact of e-substitution by sender group. E-substitution appears to be most advanced among senders in the "Retail" sector and "Utilities" and least developed in "Government (including Health and Education)" and "Insurance".

The differences in the extent of these estimates of e-substitution by sender group recorded in Figure 3 may in part reflect variations in the content type of the mail sent by each group. As an example, the sender group "Government (including Health and Education)" sends disproportionately large

Figure 3: Estimates of E-substitution Index by Sender Group, E_{jt}, to 2012 (2001=1)*



Source: Royal Mail

* E_t equals (1 - proportionate loss to e-substitution) relative to a base year where E_t = 1 implies no e-substitution relative to that base year (here 2001) and E_t = 0 implies complete loss of all mail.

volumes of business letters which reflect the purposes for which the sector is sending mail. E-substitution is estimated to have advanced less rapidly for this type of mail due to factors outlined above and so it would be expected that B2C business mail volumes overall from this sender group would be affected less by e-substitution than the average. Conversely, the "Retail" sender group

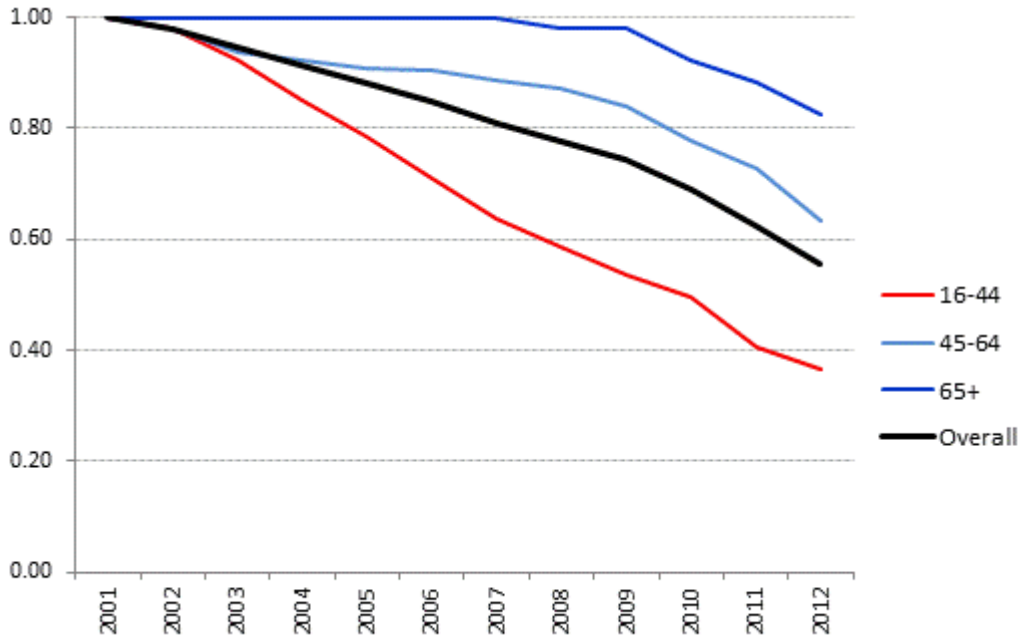
posts relatively low proportions of business letters and other financial correspondence that are content types estimated to have been less affected by e-substitution and more of those that have been affected to a greater extent, such as bills and invoices. As a result it would be expected that this sender group would be affected more by e-substitution than the average.

While the mix of traffic by content type may partly account for the differences in E-values estimated across sender groups, this is likely to be only one factor. Some sender groups such as "Government (including Health and Education)" and "Insurance" are estimated to have been less affected by e-substitution than their mix of traffic alone might have indicated while others such as "Utilities" and, particularly, "Retail" have been more affected by e-substitution than might have been expected from this single causal factor. Part of the reason for these differences may be due to the heterogeneity of traffic within content types which it is not possible to observe clearly or measure effectively. However, it suggests as well that factors specific to particular sender groups may also matter significantly; for example, the speed of adoption of new forms of communication. Indeed, causation may run in the other direction as well. For example, the content type "Bills, invoices", where e-substitution is estimated to have advanced furthest, is sent in particularly large amounts by the "Retail" and "Utilities" sectors. If it has been the case that "Retail" has been quicker to substitute out of letter mail than other sectors, say because of the degree of cost competition in the sector, then content types such as "Bills, invoices" will be those most affected by e-substitution.

Our third segmentation of B2C business mail is by age group of recipients where, our indicative estimates suggest there to be clear and substantial differences in the extent of e-substitution. As shown in Figure 4, e-substitution of B2C business mail is estimated to have been most marked among younger age groups (here including respondents up to the age of 44, who received about a quarter of all B2C business mail in 2012) and least advanced for recipients over the age of 64 (who also received about a quarter of all B2C business mail in 2012) and particularly those aged 75 and over. The estimated differences in the extent of e-substitution of mail received on average by younger and older individuals are quite striking. To some extent this is unsurprising for in part it reflects marked differences in access to the Internet and so the ability of individuals to receive communication electronically. This is lower for older age groups, as reported in Table 2, and particularly so in the case of individuals aged 75 and over.

But the differences in the extent of e-substitution by age group of recipient are estimated to be greater than those arising from differences in ability to receive electronic communications alone. In 2012, for example, about 85% of those in the UK aged 45-64 had used the Internet in the preceding three months compared with about 97% for those under the age of 45. But the E-indices for the two groups were estimated to be 0.63 and 0.37 respectively and the difference between these indices was over twice that in the rates of access to the Internet. In addition to the ability to receive electronic communications, a second important factor underlying the difference in the extent to which e-substitution is estimated to differ by age of recipient is likely to be the *willingness to receive* communication electronically even where older individuals have access the Internet. This may reflect factors such as older age groups on average being less comfortable and/or effective in the use of

Figure 4: Estimates of E-substitution Index by Age Group of Recipients, E_{kt}, to 2012 (2001=1)*



Source: Royal Mail

* E_t equals (1 - proportionate loss to e-substitution) relative to a base year where E_t = 1 implies no e-substitution relative to that base year (here 2001) and E_t = 0 implies complete loss of all mail.

electronic means of communication; less willing to change from previous methods of communication unless required or incentivised to; and having greater concerns about the use of electronic media on grounds of security or privacy.

Table 2: Estimates of Access by Individuals in the UK to the Internet by Age Group, %*

	Age Group							
	16-24	**25-34**	**35-44**	**45-54**	**55-64**	**65-74**	**75+**	**All**
2012 Q3	98	97	95	90	80	58	26	82
2015 Q1	99	99	97	94	87	71	33	86

* Percentage of individuals using the Internet by any device in the preceding 3 months.
Source: Office for National Statistics (2013, 2015c).

4. Conclusions

In this paper we have provided empirical evidence on the development of e-substitution in the UK in two main ways. First, we have reported estimates of the overall extent of e-substitution of business (or transactional) mail in the UK using results from an econometric model of a time series of this type of traffic. After beginning in the early 2000s, it is estimated that from about 2010 there has been a substantial increase in the displacement of mail to electronic substitutes, although the decline in business mail overall has been lessened considerably by the continuing positive impact on mail volumes from other factors, primarily economic and demographic growth.

Second, in the UK about three quarters of business mail is sent by businesses to consumers (B2C). Within the envelope of our overall estimate of e-substitution of business mail, we report indicative estimates of the impact of e-substitution on B2C business mail at a disaggregated level. This analysis uses a time series of survey data of consumers. Three segmentations are considered for the period up to 2012, the last year of data available at this disaggregated level at the start of the study. The segmentations are by content type of B2C business mail; by sender group; and by age of recipient. On content type we find evidence that e-substitution has advanced furthest for bills and invoices and least for business letters and elements of financial correspondence. The large category of statements has moved approximately in line with the trend for e-substitution of B2C business mail overall. Among sender groups, e-substitution appears most developed in the retail and utilities sectors and least for senders from government and insurance sectors. However, perhaps some of the most pronounced differences in the extent of e-substitution of B2C business mail are estimated to be by age of recipient. E-substitution appears to have advanced most among younger age groups (aged under 45) while for older groups (over 64) e-substitution is estimated to have commenced later and developed by much less.

This latter result is consistent with the perspective that whether business mail is replaced by electronic communication depends not only on the ability and willingness of senders to substitute an e-communication for letter mail but also on both the ability, particularly through access to the Internet, and willingness of recipients to accept an e-communication in place of letter mail. For e-substitution of B2C business mail to continue at a rapid rate it is likely to require that e-substitution continues to advance for mail received by older age groups where two processes can be thought of as being at work. For a given population cohort, born during a particular period of time, its ability to access electronic communication via the Internet and willingness to accept such communication in place of letter mail may both rise over time ("acceptance" effect). Additionally, a given cohort ages (today's 55-64 year olds in ten years will be 65-74) bringing higher levels of access to the internet and acceptance of electronic communication into that older age group in the future (an "ageing effect"). The dynamics of these processes suggest that the working through of these acceptance and ageing effects and so of e-substitution may be drawn out over a substantial period of time.

Our results also point to possibilities for further research and analysis. One area involves the linkages between e-substitution of business mail by content type and by sender group. Some content types of mail are estimated to date to have been more subject to e-substitution than others (for example, bills and invoices compared with individual items of financial correspondence) so that loss of letter mail has proved to be more rapid for those sender groups that send more of that type of mail. But sender groups may vary in their ability and, particularly, willingness to substitute out of letter mail and this would impact on the extent of e-substitution by content type, in part because the content type of mail varies substantially across sender groups. The directions of causation here are quite complex and may be worth examining in greater depth. More generally, the analysis we have developed focuses on business mail. However, there is likely to be potential also to consider trends and developments in other high level categorisations of mail such as social mail using similar techniques.

It is also clear that to better understand the time path of the e-substitution of business mail it is essential to monitor and analyse evidence both in aggregate and also by tracking and reviewing developments at a disaggregated level. At an aggregate level, it is important to update on a regular

basis the econometric model that underpins the e-substitution index reported in Section 2. Further, segmentation of trends in e-substitution by content type, sender group and age of recipient provides valuable evidence on the process itself and helps to inform the overall extent of the e-substitution of business mail and possible path, whether accelerating , stable or decelerating, over the long term. Here, for example, it would be valuable to try to assess further the degree to which elasticities vary across the segmentations of B2C mail (an important assumption used in the methodology for estimating the extent of e-substitution as reported in Section 3). Similarly, complementary evidence to augment indicative estimates of e-substitution by these segmentations could be sought through qualitative market research. Examples might include market research on reasons for differences by age of recipient in willingness to accept e-communications in place of letter mail or the factors that lead various sender groups to substitute out of letter mail by more or less than the overall average across all groups.

References

Boldron, F., C. Cazals, J-P Florens and S. Lecou (2010), 'Some dynamic models for mail demand: the French case', M. A. Crew and P. R. Kleindorfer (eds), *Heightening Competition in the Postal and Delivery Sector*, Edward Elgar

Cazals, C., J-P Florens, L. Veruete-McKay, F. Rodriguez and S. Soteri (2011), 'UK letter mail demand: a content-based time-series analysis using overlapping market survey statistical techniques', M. A. Crew and P. R. Kleindorfer (eds), *Reinventing the Postal Sector in an Electronic Age*, Edward Elgar

De Donder, P., H. Cremer, F. Rodriguez, S. Soteri and S. Tobias (2015), 'Analysing the prospects for transactional mail using a sender-recipient framework', M. A. Crew and T. J. Brennan (eds), *Postal and Delivery Innovation in the Digital Economy*, Springer

Jarosik, M., J. Nankervis, J. Pope, S. Soteri and L. Veruete-McKay (2013), 'Letter traffic demand in the UK: some new evidence and review of econometric analysis over the past decade', M. A. Crew and P. R. Kleindorfer (eds), *Reforming the Postal Sector in the face of Electronic Competition*, Edward Elgar

Nikali, H. (2008), 'Substitution of letter mail for different sender-receiver segments', M. A. Crew and P. R. Kleindorfer (eds), *Competition and Regulation in the Postal and Delivery Sector*, Edward Elgar

Ofcom (2005, 2014), *Communication Market Report 2005 (2014)*

Office for National Statistics (2013), *Internet Quarterly Update, Q1 2013*

Office for National Statistics (2015a), *Internet Access - Households and Individuals, 2015*

Office for National Statistics (2015b), *Quarterly National Accounts, 23 Dec 2015 Release*

Office for National Statistics (2015c), *Internet Users, 2015*

PwC (2013), The Outlook for UK Mail Volumes to 2023, available at http://www.royalmailgroup.com/sites/default/files/The20%Outlook20%for20%UK20%mail20%volumes20%to20%2023.pdf

Soteri, S., F. Feve, J-P Florens and F. Rodriguez (2009), 'Internet advertising and direct mail: trends and analysis for the UK', M. A. Crew and P. R. Kleindorfer (eds), *Progress in the Competitive Agenda in the Postal and Delivery Sector*, Edward Elgar

Trinkner, U. and M. Grossman (2006), 'Forecasting Swiss mail demand', M. A. Crew and P. R. Kleindorfer (eds), *Progress Toward Liberalization of the Postal and Delivery Sector*, Springer

USPS (2010), 'Projecting US mail volumes to 2020: compendium', available at http://www.prc.gov

Veruete-McKay, L., S. Soteri, J. Nankervis and F. Rodriguez (2011), Letter traffic demand in the UK: an analysis by product and envelope content type', *Review of Network Economics*, 10 (3)