

Quantitative analyses of VOD services

Introduction

1. In this appendix, we describe the quantitative analyses of VOD services that we have conducted that are relevant to our assessment of market definition and competitive effects (see paragraphs 4.75 to 4.84).
2. We gathered viewing data from four different sources and analysed each separately using the following model of consumer behaviour. If an increase in content available from provider X is associated with a fall in provider Y's demand, we interpreted this as being consistent with customers substituting between the two—from Y to X. On the other hand, if an increase in content available from provider X is associated with no change in provider Z's demand, we viewed this as being inconsistent with customers substituting between the two.
3. The quality of the data varied. The first dataset from the BBC and the parties was based of daily observations. This dataset was limited to the extent that it relied on observations of four newly-established VOD services during a period of market expansion. The other datasets were constructed from observations of viewing of different content providers on the same platform. That allowed us to isolate the effect of different providers from any possible platform effects. One of these datasets (iTunes data) consisted of daily data for an 11-month period. The other two ([X]) and Virgin Media) were constructed from less frequent observations, or from a shorter period of time.
4. We note that we must interpret the results with caution. First, in a nascent market it is likely that the launch of a new VOD retail offer (such as BBC iPlayer) causes demand to grow overall, making it difficult to isolate and identify consumer switching. Second, VOD retailers might not immediately identify the most appropriate model of supply in terms of content and price. They might also experience teething problems with their technology. These issues would tend to obscure consumer switching that we might observe in more established markets. Nevertheless, we can infer consumer substitution patterns to the extent that we are able to identify statistically significant effects that are not attributable to market expansion effects.

Section 1: Analysis of BBC iPlayer, ITV.com, 4OD and C4C.com viewing data

5. In this section, we present our estimates of the effect of changes in the VOD content supplied by alternative providers on the viewing of the content of a particular provider.
6. We first describe the methodology for this estimation. We then present a summary of the results from the estimation. Finally, we discuss the parties' response to this analysis. Annex 1 contains details on the dataset we used and full results from the regressions.
7. The two key questions for this analysis are:
 - (a) Are each of the parties' content substitutes for one another?
 - (b) What is the relationship between the supply and views of UK and non-UK content?

8. In view of the limitations of our approach and our data, we must interpret the results with caution. However, at face value, the results we obtain suggest the following:
 - (a) The results are broadly consistent with the hypothesis that the content provided by each of the parties is a substitute for one another.
 - (b) The results are also consistent with the hypothesis that non-UK content is not a good substitute for the UK content available on the parties' websites.

Methodology

9. There are two key assumptions that are relevant to our estimation of viewing data for the parties' websites. First, we assume that at present the industry is unlikely to have reached equilibrium. Second, we assume that the supply of content is exogenous, and particularly that it is independent of demand. We discuss these two issues below.

Out of equilibrium estimation

10. Ideally we would estimate the relationships observed between providers in equilibrium. However, VOD is a nascent market for which both supply and demand are growing rapidly. To mitigate for this continuing growth, we use a time trend in order to estimate the trajectory towards an equilibrium steady state.

Exogenous supply

11. We assume that the supply of content by a provider is exogenous to the demand for content because we do not observe changes in supply in response to changes in demand. However, we note that in some cases we have referred to the 'supply of content' by reference to the list of titles or hours available. We note that this is effectively a measure of 'range' and not 'supply' in the usual sense of the word.
12. The pattern of [✂] supports this assumption (see Figures 1 and 2 below). We note that the supply of catch-up content is constrained by the hours of linear TV that were delivered in the preceding week; these will vary within certain bounds, as can be seen in Figure 2. The bounds are, however, stable over time. The only notable reductions in content are caused by technology outages. As such, we do not see any sign of reductions in supply of content that we would expect to see if supply were endogenous.
13. Similarly, archive content appears to be added on to the site in batches; neither [✂] shows signs of the supply of archive content being responsive to variations in demand.

FIGURE 1

[✂]

Source: CC analysis of BBC, ITV and C4C data.

FIGURE 2

[✂]

Source: CC analysis of BBC, ITV and C4C data.

Data

14. We collected data on the number of daily views (streamed or downloaded) on each of the VOD services offered by the parties since they were launched.
15. We also collected data in relation to the following explanatory variables:
 - (a) the daily supply of content for each of the VOD services;
 - (b) the number of views on BBC linear channels, ITV linear channels and C4C linear channels¹ on the previous day;
 - (c) days on which there occurred 'outages' on C4C ([REDACTED]);
 - (d) the age of the content: we separated out the views and supply of content by age band (0–7 days since broadcast, 8–30 days and 30+ days); we also split content by the method by which it is consumed (DTO, DTR and free); and
 - (e) throughout this analysis we have included a dummy variable to control for Sunday as this is a day when viewing might be expected to be higher as many consumers have more time to view content.

Substitutability by website and format

16. Overall, the results of the analysis presented in this appendix are broadly consistent with the view that the parties' VOD services are a substitute for one other. Table 1 shows the effect of the quantity of content that is supplied by one provider on the views of the content on the other parties' websites. [REDACTED]
17. [REDACTED]
18. [REDACTED]
19. Taking these comments together with the estimation results, we believe that the positive signs of these coefficients on the BBC supply are likely to reflect the market growth experienced as a result of the BBC's entry and marketing reach. We have been unable to separate this out from the competitive effect that the BBC has on C4C and ITV.
20. [REDACTED]
21. [REDACTED]
22. Thus if the BBC iPlayer is associated with market growth [REDACTED], we would not consider the positive coefficients on the BBC to be a rejection of our hypothesis that the three parties are good substitutes for one another.
23. [REDACTED] One possible interpretation of this coefficient is that it suggests that [REDACTED].² An alternative interpretation is that ITV.com might have had the same market-growing effect as the BBC iPlayer.

¹Including all time-shifted '+1' channels for each of the parties but excluding film4 and film4+1.

²[REDACTED]

TABLE 1 [X]

Source: [X]

[X]

24. [X]

TABLE 2 [X]

Source: [X]

[X]

25. [X]

TABLE 3 [X]

Source: [X]

[X]

Substitutability of content origin

26. In this subsection, we categorize the supply and views of content into UK- and non-UK-originated content in order to assess the effect of origin of content on the competitive constraints offered by UK and non-UK content upon one another.
27. The second column of Table 4 shows the effect of additional non-UK content on the viewing of UK content. The third column shows the reverse, that is the effect of additional UK content on the viewing of non-UK content.
28. The results are then split by the provider and format of the content. So, for example, the first row relates to UK and non-UK content broadcast on ITV 0–7 days ago.
29. The results presented in the third column of Table 4 suggest that, in general, the additional supply of UK VOD content has either no effect or a negative effect on the viewing of non-UK VOD content. This might indicate that UK content is a substitute for non-UK content. We focus on the results of catch-up on catch-up content and of archive on archive content.
30. However, the results presented in the second column suggest that, if there is a relationship, it is an asymmetric one. In particular, it would appear that the additional supply of non-UK content has either no effect or a positive effect on the viewing of UK content. This is consistent with the hypothesis that non-UK content is not a good substitute for UK content. It is not consistent with UK and non-UK content being good substitutes for one another.
31. The analysis would therefore appear to suggest that if there is any relationship between UK and non-UK content, it is an asymmetric constraint that is imposed by UK content on non-UK content. That is to say, at least in aggregate, non-UK content (previously screened on UK linear TV) does not offer a competitive constraint on the

UK content controlled by the parties,³ while the same UK content appears in one case to constrain the sale of non-UK content.

TABLE 4 [REDACTED]

Source: [REDACTED]

[REDACTED]

Parties' response to the CC's analysis of their viewing data

32. The parties raised two types of issue regarding our analysis: first, issues about identification of the econometric model; and second, issues about our interpretation of the results.

Identification of the econometric model

33. The parties raised a number of issues regarding the inherent limitations in our approach and the suitability of this approach to analysing market definition and competitive effects.
34. The parties noted the absence of a price effect in our analysis and questioned the extent to which it is possible to interpret this evidence in the context of the SSNIP test. Whilst we would agree that we cannot use this approach to assess consumers' substitution patterns in response to price changes, it allows us to look at the effects of relatively small changes in some non-price aspects of a VOD service.
35. They also argued that a robust approach would require data that is not available. First, the model does not include data on changes in the supply of non-VOD services, like PVRs. Second, the measure of supply does not account for changes in actual or perceived quality of content, nor in the mix of content over time. And third, the measure of demand does not capture duration of view, which may have changed over time.
36. In our view, the supply of free content and of non-VOD services like PVRs would not vary significantly on a daily basis. Moreover, we controlled for PVR viewing over the previous seven days as part of the BARB linear TV viewing figures. Therefore, we do not think that the lack of such information would result in misleading results. We agree with the second and third observations, but we have no reason to believe that there have been any such changes during the period of our analysis.⁴
37. In addition, the parties believed that some of the inconsistencies in our results were likely to have arisen because of the difficulties in controlling for the rapidly-changing nature of the VOD market. We agreed that it can be difficult to isolate substitution patterns in an evolving market and have been careful to point to the limitations of our approach in this regard. Specifically, we believed that the introduction of the BBC iPlayer has had a significant effect in expanding demand for other VOD services and that this effect may mask any substitution effects that we might otherwise observe.

³We note that it is likely that US content not previously broadcast on UK linear TV would offer still less a competitive constraint.

⁴Moreover, we have no reason to believe that any such changes would be correlated with the residual term in the respective regressions and therefore no reason to believe that they would cause biased or inconsistent statistical results.

38. The parties also doubted whether a linear time trend would adequately capture changes in viewing patterns from one year to the next and within each month and reflect the uptake in new technology and habit forming. However, from looking at the data we do not think that it is clear that a non-linear time trend, such as an exponential or quadratic time trend, would better capture changes in viewing patterns over the period.
39. The parties submitted that this analysis does not account for the activities of other rival VOD retailers.⁵ We were unable to collect data on all potential competing VOD retailers over this period. While we agree that this prevents us drawing conclusions on the impact of these other VOD retailers' activities on the parties' service, we are nevertheless confident that the relationships that we find to be statistically significant remain important results. Furthermore, we note that the iTunes data analysis does not have the same issue since, as the parties have told us, rival VOD retailers find it difficult to access the users of iTunes due to their loyalty to the store.

Consistency and interpretation of the results

40. The parties also raised a number of detailed points regarding the consistency of the results and our interpretation of those results.
41. The parties submitted that our conclusion that the evidence was broadly consistent with the hypothesis that the parties' content was a substitute for one another was incorrect and misleading. There were two aspects to this criticism. First, the parties made observations about our interpretation of the C4C outage dummy and argued that these results should be excluded. In our view, it is appropriate to include these results because they capture the effects of changes in supply, even if those changes are more significant than others.⁶ It is also our view that the results for the outage dummy are consistent with our conclusions. In particular, we think that [REDACTED]. In light of this information, [REDACTED].
42. The second aspect of the parties' concerns is that, once the C4C outage dummy results are excluded, they do not agree with our interpretation of the results. In particular, they argued that 12 of the 16 results were not consistent with substitution. In our view, we should look at the 8 results of the effect of catch-up on catch-up and of archive on archive rather than 16 results. In our view, it is not meaningful to look at the effect of catch-up on archive and vice versa, particularly given that at present ITV has a limited archive offer and the BBC iPlayer is only catch-up.
43. In our view, the results are clearest for [REDACTED] support our hypothesis and are statistically significant. [REDACTED] However, as we have been careful to note, it is our view that [REDACTED].
44. The parties also suggested that we were wrong to conclude that our results on the substitutability between UK and non-UK content were consistent with our conclusions.⁷ They said that six out of eight results point in the opposite direction. We note that, by examining the statistical significance of these results, it can be seen that not one of the results suggests that non-UK content is a substitute for UK content (see column 2 of Table 4).

⁵Parties' response to provisional findings, 24 December 2008.

⁶[REDACTED]

⁷Parties' response to provisional findings, 19 December 2008.

45. The parties raised concerns regarding the logic of our analysis and the credibility of the underlying theory of consumer behaviour. [REDACTED] We agree that this result is not consistent with a theory of consumer behaviour. However, it is our view that this result reflects the [REDACTED]. We also note that the ITV dataset contains a number of missing observations. The parties also pointed out that we did not find a statistically significant effect on viewing on the BBC iPlayer from changes in the supply of content on the BBC iPlayer. We agree that this result is inconsistent with our overall conclusions.
46. With regard to our analysis of the effect of non-UK content on UK content, the parties argued that the absence of a statistically significant effect might be incorrectly estimated if the model was incorrectly identified, and they raised two specific issues. First, as before, changes in the supply of ITV catch-up content had no statistically significant impact on any catch-up viewing. We agree that the ITV result is odd and refer to the discussion in paragraphs 34 and 23. Second, the parties argued that one plausible explanation for our result was that there had been an increase in the supply of free non-UK content available on the market, which may have increased total viewing of both non-UK and UK content. We looked at the viewing figures of 4OD, the only operational site at the time of an increase in US content, to see whether this could explain our result. We found no indication of a market expansion effect from non-UK content.
47. In conclusion, we recognize that there are limitations to our approach and inconsistencies in some of the results, and we set these out. We are particularly aware that, with the possible exception of C4C, the parties have only recently launched VOD services and that performance issues around the platforms are likely to affect our results. However, taking all the issues into account, it is our view that the results of our analysis are broadly consistent with the finding that the parties' content is a substitute for one another and that non-UK content is not a good substitute for UK content.

Section 2: Analysis of iTunes viewing data

48. In this section, we describe our analysis of iTunes viewing data. We first describe the analysis that we have conducted. We then present a summary of the results from the estimations. Annexes 3 and 4 contain further details on the dataset that we have used and the full results from the regressions.
49. The two key questions are:
 - (a) Are each of the parties' content substitutes for one another?
 - (b) What is the relationship between the parties' content and third-party content?
50. Since [JV party] has noted [REDACTED] after the introduction of films on iTunes, we also look at the effect of the launch of films on [JV party] and [JV party] revenue.⁸
51. In view of the limitations of our approach and our data, we must interpret the results with caution. However, at face value the results we obtain suggest the following:
 - (a) The results are broadly consistent with the hypothesis that the content provided by each of the parties are substitutes for one another.

⁸[REDACTED]

- (b) The results are also consistent with the hypothesis that non-UK content is not a good substitute for the UK content available on the parties' websites.
- (c) The analysis suggests that any [redacted] that may have occurred in the month after films arrived on iTunes is not attributable to the appearance of films.
52. We found some exceptions to these results: first, [redacted], a UK provider of children's TV content which the results indicate might be substitutable for the [JV party] children's content; and secondly, [redacted]. However, further analysis suggests that the negative impact of [redacted] content on [JV party] revenue is a result of the popular programmes such as [redacted] and [redacted] that it has supplied to [JV party] and others for broadcast on linear TV over the last decade.
53. We also found that [JV party] content was not a substitute for [JV party] content. However, we note that the [JV party] content made available on iTunes did not appear to be representative of the full [JV party] archive.

Methodology

Assumptions

54. There are two key assumptions that we make when conducting the estimation. First, we consider that at present the industry is unlikely to have reached equilibrium. Secondly, we assume that the supply of content is exogenous. We discuss these two issues below.

Exogenous supply

55. As with the analysis of the parties' data, we begin by making an assumption that the supply of content is exogenous, and particularly that it is independent of demand. In support of this assumption, we note that the supply of content is relatively inflexible.⁹
56. We note that in some cases we have referred to the 'supply of content' by reference to the list of titles available. We note that this is effectively a measure of 'range' and not 'supply' in the usual sense of the word. In relation to this dataset, because it covers the top 1,000 or so titles purchased, appearance in the list of titles is also a measure of 'popularity'.

Out of equilibrium estimation

57. Ideally we would estimate the relationships observed between providers in a stable equilibrium. However, VOD is a nascent market for which both supply and demand are growing rapidly. To mitigate for this continuing growth, we use a time trend in order to estimate the trajectory towards an equilibrium steady state.¹⁰

⁹iTunes negotiates contracts for the supply of content from each provider. [redacted] Each contract specifies the content to be supplied and the content provider does not have the ability to reduce supply in response to fluctuations in demand.

¹⁰In settling upon our specification of the regression analysis, we considered using an alternative time trend which started at the beginning of the dataset rather than the beginning of the data on the provider in question. However, since content from some providers only came online during the period of the dataset we decided against using this alternative time trend.

58. In contrast to the analysis of the parties' data, we note that the iTunes content is not free; rather, it has a price for each programme. [REDACTED]¹¹

Data

59. In the following paragraphs, we summarize the dataset provided by iTunes.

60. The dataset begins on 29 August 2007 when iTunes added TV content to the iTunes store.¹² It consists of daily information on the revenue of each programme that was downloaded from the iTunes store on every day between 29 August 2007 and 13 July 2008. For the first four months of the data, the only providers of content were MTV Networks Europe and ABC. However, in the first four months of 2008, a number of other content providers are added to the platform. These include:

- Columbia Pictures (7 January 2008);
- Aardman (22 January 2008);
- Warner Bros (23 January 2008);
- BBCW (18 February 2008);
- C4C (18 March 2008);
- ITV (2 April 2008);
- HIT entertainment (28 April 2008); and
- Universal (5 May 2008).

61. Each provider is launched with a relatively large selection of content.¹³ As time goes by, new content is added to this stock of existing content. For example, BBCW begins with [REDACTED] hours of content and adds more than [REDACTED] hours of content on [REDACTED] separate days over the following five months. Figure 3 shows the number of new programmes added to the iTunes store on each day between 29 August 2007 and 13 July 2008.

FIGURE 3

Number of new programmes added to iTunes by day

[REDACTED]

Source: CC analysis of iTunes data.

62. We note that the data we were able to collect from iTunes covers approximately the top 1,000 programmes purchased on each day of the dataset, rather than every programme that was available each day. For example, the dataset does not include programmes that were not purchased on the day in question. We are therefore unable to count the total number of titles that are available each day and the proportion of titles which are included in this measure. As a result, if we interpret this variable as

¹¹[REDACTED]

¹²The iTunes application is available to download free of charge from www.apple.com.

¹³[REDACTED]

a measure of the full range that is offered by a provider, there would be a risk that some variables would be mis-measured, which might lead to unpredictable bias in the results and limit the conclusions that could be drawn on the relationships between providers. We cannot rule out this possibility. Nevertheless, the conclusions that we draw do not depend on the variable being defined as the full range. In particular, we note the following three important points:

- (a) First, this form of analysis assesses customer switching in response to changes in price. Since [REDACTED] we instead observed the changes in the range of content available in order indirectly to identify customer switching in response to variations in that range.
 - (b) Second, while the variables are unlikely to constitute the full range that is available, they nevertheless remain a consistent measure of the ‘popularity’ of the content by reference to the number of titles of a particular provider’s content appearing in the top 1,000 or so titles. The dataset provided by iTunes includes at a minimum the top 1,000 programmes purchased that day. However, where the 1,001st programme sold the same number of units as the 1,000th programme ([REDACTED]), that programme is also included in the data. By this method, on certain days we register more than [REDACTED] programmes whilst in the early days when little content was available on the site we register as few as [REDACTED]. Therefore we observe substantial variation and the dataset is not constrained to 1,000 titles. The popularity of one piece of content therefore does not come at the expense of the popularity of another piece of content.
 - (c) Finally, if it can be assumed that an increase in the popularity of a provider’s content reflects an improvement in some element of the offer of that provider, we can interpret the effect that this has on the demand for rival providers’ content in the same manner as we would the effect of a price cut by the provider.
63. If we accept this, it follows that negative coefficients between providers X and Y in our results will reflect the fact that some improvement in the offer of provider X causes customers to switch demand away from provider Y. Importantly, this interpretation is not affected by potential mis-measurement issues, though we note that it does rely on the assumption that ‘popularity’ is a function of a provider’s offer.
64. The increasing number of programmes on iTunes is reflected in the revenue earned per day (see Figure 4). [REDACTED]¹⁴

FIGURE 4

iTunes daily revenue (including VAT)

[REDACTED]

Source: CC analysis of iTunes data.

Impact of content providers

65. Table 5 shows the impact of different content providers as they are added to the iTunes store. [REDACTED]

¹⁴[REDACTED]

TABLE 5 Share of programmes and share of revenue, by provider

[X]

Source: CC analysis of iTunes data.

Analysis

66. In the following paragraphs, we set out the econometric analysis of this dataset. This is intended to inform our view of substitution patterns between parties, as well as with other content providers.

67. However, we first note that this dataset does not allow us directly to observe substitution patterns themselves, since it is not at a customer-specific level of granularity.¹⁵ Rather, we are limited to examining indirect evidence which may imply the existence of certain patterns of substitution.

68. For example, the key question in our analysis will be whether an increase in the range, or ‘popularity’, of content from provider X will diminish the demand for the content of provider Y. Where such a result is obtained, we would interpret it as being consistent with the content of provider X acting as a substitute for the content of provider Y. This interpretation is based on the proposition that an increase in the ‘range’, or ‘popularity’, of content from provider X that is correlated with a decline in the revenue at provider Y is consistent with customers substituting between the two stores. This data does not allow us to question customers individually to ascertain whether they have switched, but we can observe that the results at the end of a day’s trading are consistent with them behaving in such a manner.

69. Moreover, in the absence of such correlations, we can say that the day’s trading results are not consistent with the hypothesis that customers are switching between the two. Finally, when interpreting the results by reference to ‘range’, we should note that all of the above will depend crucially on the assumption that consumers are better off with a broader, rather than smaller, range of programmes from which to choose. When interpreting the results by reference to ‘popularity’, no such assumption is necessary.

70. The estimation that we specify is therefore as follows:

$$Rv_{it} = \alpha_i + \beta_1 Sat_t + \beta_2 Sun_t + \beta_3 Bank_t + \beta_4 date_t + \beta_5 AvOwn_{it} + \sum_{j \neq i} \beta_{j5} Av_{jt} + \varepsilon_{it}$$

71. Here, Rv_{it} is the revenue taken by provider i on day t ; we use revenue rather than the number of downloads since certain content is on occasion provided on offer at zero price.¹⁶ We have not included price as an explanatory variable since the data does not specify the price of each programme on each day. [X]

72. The variables Sat_t , Sun_t and $Bank_t$ are dummies for the weekend and bank holiday respectively.¹⁷ $Date_t$ is a simple time trend.¹⁸ $AvOwn_{it}$ is the number of the provider’s own programmes that are available (or register as ‘popular’) on day t . Finally, Av_{jt} is

¹⁵[X]

¹⁶This happens on [X] occasions over the [X] months for which we have data.

¹⁷We experimented with using a different dummy variable for each day, or a single dummy for weekends. This did not appear significantly to alter the results obtained and therefore we included just the more important dummy variables.

¹⁸See footnote 10.

the number of programmes from each of the other providers that are available (or register as ‘popular’) on day t.

Results

73. In the following paragraphs we summarize our results. In view of the limitations of our approach and our data, we must interpret the results with caution. Table 6 shows the full set of cross-provider effects. This shows the direction and significance of the relationship (details of the coefficients and t-statistics are provided in Annex 4). In the first cell of the top left-hand corner we see that additional [JV party] content is, unsurprisingly, correlated with an increase in [JV party] revenue. Similarly, and as we would expect, for every provider on the diagonal we find a positive effect of their own content on their own revenue.
74. Staying with the first row, we see that increased [JV party] content is also correlated with a reduction in the revenue of [JV party] content. As outlined in paragraph 68, we would conclude therefore that the evidence is consistent with the existence of a pattern of substitution between [JV party] and [JV party].
75. In contrast, we find positive correlations between the supply of [JV party] content and the revenue of [US provider], [US provider], [US provider], [JV party] and [US provider]. This is inconsistent with customers substituting between the content of [JV party] and each of these providers.

TABLE 6 Regression results: the effect of availability on revenue, by provider

The effect of the
availability of On the revenue of ...
content from ...

[✂]

Source: CC analysis of iTunes data.

[✂]

Substitution between the parties

76. Focusing on the parties, we see that, as expected, the content of each one has a positive effect on its revenue.¹⁹ Moreover we find that:
- (a) more [JV party] content is associated with less revenue for [JV party] and [JV party];
 - (b) more [JV party] content is associated with less revenue for [JV party]; and
 - (c) more [JV party] content is associated with less revenue for [JV party], but more revenue for [JV party].

¹⁹Though for [JV party] this is not a statistically significant effect.

TABLE 7 **Regression results: the effect of availability on revenue, JV parties only**

<i>The effect of the availability of content from ...</i>	<i>On the revenue of ...</i>
	[REDACTED]

Source: CC analysis of iTunes data.

[REDACTED]

77. Therefore it would appear that this evidence is consistent with the proposition that customers substitute between the content provided by the parties.
78. The exception to this is the positive effect of [JV party] content on [JV party] revenue. We have examined the content that [JV party] supplies to iTunes. The list of series is at Annex 2. In general, this [JV party] content is relatively old. [REDACTED]

TABLE 8 **Summary statistics for the JV parties (22 April to 13 July)**

	<i>Average revenue per day (22 April–13 July) £</i>	<i>Programmes on iTunes on 13 July</i>	<i>Average revenue per day per programme £</i>
[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]
[REDACTED]	[REDACTED]	[REDACTED]	[REDACTED]

Source: CC analysis of iTunes data.

79. To the extent that [REDACTED], we may therefore not be in a position to draw strong conclusions about future substitution patterns.

Substitution between [JV party] and third parties

80. It would appear from column 1 of Table 6 that [JV party] revenue is adversely affected by [UK provider] and [US provider] (as well as [JV party]). [UK provider] is a UK children’s TV provider which provides content similar to the [JV party] children’s content. In the case of [US provider], we have examined three hypotheses which might explain the negative coefficient:
- (a) that [US provider] large stock of cartoon content might impact on the viewing of [JV party] children’s content;
 - (b) that [US provider] content as a whole might act as a substitute for [JV party] content; and
 - (c) since [US provider] highest-earning programme, by some distance, is, a long-running sitcom on [JV party], that it is the [US provider] content that has been shown on the parties’ linear TV channels that [JV party] customers may view as a substitute, in the same way that they regard [JV party]’s other programmes as a substitute (see paragraph 76).
81. We have found that the only significantly negative correlation that [JV party] has is with [US provider] content that has already been screened on the parties’ linear TV channels, particularly [JV party] and [JV party] (for example, [REDACTED]). This is consistent with hypothesis (c).
82. The effect of [JV party] content on other providers’ revenue is shown in the first row of Table 6. We note significant positive effects on [US provider], [US provider], [US

provider] and [US provider]. None of these results is consistent with customers substituting between these providers and [JV party].

Substitution between [JV party] and third parties

83. Column 2 of Table 6 shows that there are no negative impacts on [JV party] revenue other than from [JV party] and [JV party]. The only other significant correlation is a positive one with [US provider]. This might be attributable to the number of high-profile [US provider] programmes that are broadcast by [JV party].
84. The content of [JV party] is only negatively correlated with the revenue of [JV party], [JV party] and the [US provider] carried by iTunes. Of the other providers, none are statistically significantly affected other than [US provider], [US provider] and [US provider] which are all positively correlated. We note that these are all children's TV programmes and might therefore be to some degree complementary to the entirely non-children's TV content available from [JV party].

Substitution between [JV party] and third parties

85. With the exception of [JV party], no provider has a significantly negative correlation with the revenue of [JV party] (column 3 of Table 6). [JV party] content meanwhile has had a negative effect on [US provider] and [US provider] (row 3 of Table 6). This might imply an asymmetric substitution pattern between these providers and [JV party].

Substitution between TV and film

86. [JV party] said that the introduction of films on iTunes on 4 June had led to [X]. We tested the significance of this result. We found that the introduction of film had a positive but insignificant correlation with [JV party] revenue (we also found an insignificant effect on [JV party] revenue). This suggests that the introduction of films on iTunes had no effect on [JV party] revenue.
87. The explanation for this [X] is that, after controlling for other effects, the [X]. That is to say, the [X] is attributable to other factors and not to the introduction of film.

Parties' response to the CC's analysis of the iTunes data

88. The parties submitted that there were inherent limitations to our approach when applied to their own data (paragraphs 5 to 31). Our view of these limitations is discussed in detail in paragraphs 33 to 47. We recognize that similar points might be made in relation to this analysis. However, we believe that, for similar reasons, the results of the analysis are broadly consistent with the finding that the parties' content is a substitute for one another and that non-UK content is not a good substitute for UK content.
89. The parties said that when examining the relationships between UK and US suppliers they found that half the results were negative and half were positive. The parties told us that this meant that half of these results contradicted the conclusions that we draw from the analysis. We note that we have not examined in any detail the question of whether UK content constrains US content or specialist content providers (eg Kids TV). We examined whether the parties' content is constrained by US suppliers or by the other parties to the JV. We note that there is just one statistically significant nega-

tive effect in these results that suggests a constraint from a US provider on a UK provider. We discussed this in detail in paragraphs 80 to 82.

90. The parties submitted that the data should be tested to see if it is stationary.²⁰ They also submitted that the analysis should include lagged variables in order to capture learning effects. Finally, they submitted that the analysis should check that there is no pattern to the residuals of the estimation. They considered each of these to be tests of the robustness of the results obtained.
91. We note, first, that since this is a nascent market this is necessarily an out-of-equilibrium estimation and so will not be stationary; on that basis, we included a time trend. As noted in paragraph 16 of Annex 1, we have corrected for autocorrelation in the residual term and any heteroskedasticity which creates variations in the residual term. Finally, we note that VOD learning effects would be reflected in the time trend.
92. In conclusion, in view of the limitations of our approach and our data, we must interpret the results with caution. However, the results obtained are broadly consistent with the hypothesis that the content provided by the parties is a substitute for one another and that non-UK content is not a good substitute for the UK content available on the parties' websites. The results suggest that [X] may have occurred in the month after films arrived on iTunes is not attributable to the appearance of films.

Section 3: Analysis of Virgin Media viewing data

93. In this section, we describe our analysis of Virgin Media viewing data. [X] In the following paragraphs, we describe our analysis of data provided by Virgin Media to try to see what effect [X] has had on the views of [X]. The results should be interpreted with caution, since we are unable to control for other important factors that may have varied during this period.²¹
94. Figure 5 shows the views of all archive content between January and June 2008. [X]

FIGURE 5

Views of archive content, January to June 2008

[X]

Source: CC analysis of Virgin Media data.

95. Although we can place only limited weight on this evidence and although it does not provide conclusive evidence of any substitution, it is consistent with a pattern of substitution away from [X] content and towards [X] content as the [X].

Section 4: Analysis of [X] viewing data

96. In this section, we describe information provided to us by [X] on a 'natural experiment'. The following paragraphs describe our analysis of [X] data to determine the impact of the significant increase in supply of US content on demand for UK content.

²⁰The mean or variance of stationary data would not change over time.

²¹[X]

97. [redacted] explained that whereas it considered UK content to be valuable for driving 'reach' on its VOD service (in terms of increasing the number of unique users), US content was more helpful in driving 'depth' (in terms of the level of usage among a narrow group of users). It gave an example of the effect that the introduction of US content had on its existing UK content offering.
98. [redacted] explained that the launch of content from [redacted] brought: [redacted].
99. [redacted] also explained that: [redacted].
100. Figure 6 shows [redacted].

FIGURE 6

[redacted]

Source: CC analysis of [redacted] data.

Note: [redacted].

101. [redacted]

FIGURE 7

[redacted]

Source: CC analysis of [redacted] data.

102. [redacted]²² This could suggest that availability of US-originated TV content does not cause viewers to substitute away from UK-originated content. Rather, the increase in viewing of US-originated content may be explained by either new viewers or existing viewers that watch US-originated content in addition to watching UK-originated content. This would imply that the two sources of content are not close substitutes.
103. However, we note that this data only represents a snapshot in time and, in principle, this high-level data might conceal substitution between US and UK content, within a market that is growing overall. We therefore regard this evidence as supporting the proposition that US and UK content are not close substitutes, but not as providing conclusive evidence. One way we would have liked to test this would have been by looking at the impact of the launch of additional UK content on UK viewing. However, as all the providers of UK content that are currently available on [redacted] have been available since launch, [redacted] was unable to provide us with any examples where UK content has been added to the service.

²²[redacted]

Technical description of the estimation of the parties' data

1. In this annex, we describe in greater detail the estimation of the parties' data. In the first part, we describe the analysis, including the estimation and the construction of the dataset. In the second part, we present the full regression results.

Estimation

2. In addition to the time trend, the exogenous supply of content from both provider i and the other parties to the JV (providers j) and the viewing on the linear channels for each of the parties on the previous day,²³ we include two important explanatory variables.
3. First, we include dummy variables for two technological outages on the 4OD website.²⁴ These each lasted [X] full days and therefore prevented any viewing of content on 4OD over that period. These therefore each represent substantial reductions in the supply of content on 4OD.
4. Secondly, we include a dummy variable for Sunday. This is intended to control for any systematic variation in viewing that occurs as a result of consumers having more time for viewing on demand on Sunday.
5. We have excluded from the regressions the marketing activities of the parties. We do so because the nature of marketing on linear TV channels makes it difficult to construct a meaningful measure of the advertising spend. For example, a voiceover at the end of a popular show advertising the VOD service will reach a large audience but does not cost the channel anything.
6. Excluding this variable will only affect the results obtained from the regressions if the marketing is correlated with any of the other explanatory variables. If it is correlated, OLS estimation will lead to biased estimates of the coefficients. If it is not, OLS estimates will remain unbiased.
7. We believe it is reasonable to assume that marketing will be uncorrelated with the time trend, the Sunday dummy, the outage dummies and linear viewing figures. It is also reasonable to assume that it will be uncorrelated with the supply of catch-up content, as the catch-up content supply is limited by the availability of content that has been shown on linear TV. It may be less reasonable to assume that marketing is uncorrelated with the supply of archive content since more archive content might be added to coincide with the marketing promotion. However, we compared the pattern by which archive content has been added on to the site with the timing of the major marketing campaigns on which we have information. This would appear to suggest that there is, in fact, no correlation between marketing activity and the addition of archive content.
8. The estimation we specify is therefore as follows:

²³The broadcast of popular programmes such as [X] will increase the viewing of content on the following day. This happens each seven days as a new episode is released. We have previously used day of the week dummies to control for these variations. However, these programmes do not occur throughout the dataset. Hence to improve the specification we replaced them with a variable that measures the provider's linear TV viewing figures on the previous day.

²⁴These outages occurred on [X].

$$Demand_i = \alpha_1 + \beta_1 Supply_i + \beta_2 Supply_j + \beta_3 Time + \beta_4 Sunday + \beta_5 Outdum1_{4od} + \beta_5 Outdum2_{4od} + \beta_5 Linear\ viewing_i + \beta_6 Linear\ viewing_j + \varepsilon_i$$

Dataset

9. With a new market such as VOD, we would expect, and we observe, companies to experiment with different business and pricing models. In particular, companies may go through a process of learning how best to draw revenue from the product.
10. To prevent the analysis from modelling these experiments, and thereby failing to estimate the substitution patterns between content providers, we restrict ourselves to an analysis of content that is provided at no cost to the consumer.
11. Within the free-to-view content that is supplied by the parties, we additionally distinguish between:
 - (a) for all of our analysis: the format of the content, which is the period of time since it was screened on linear TV, ie 0–7 days; 8–30 days; or 30+ days; and
 - (b) for the second section of our analysis: the origin of the content, which is whether the content is produced in the UK or outside the UK.
12. We have daily data on the content available and the number of views on each of the four websites operated by the parties:
 - (a) BBC iPlayer;
 - (b) ITV.com;
 - (c) 4OD; and
 - (d) Channel4.com.
13. We note that [redacted].²⁵
14. We also note the existence of ‘spikes’ in the catch-up (0–7 days) content available on a website shortly after the launch of that website. These involve the content rapidly expanding before contracting and settling at an intermediate level. This spike soon disappears and a stable level of supplied content is observed thereafter. We view these as experiments in the content level and therefore remove them from the dataset.
15. We note that the data for the usage of the [redacted] service contains some gaps where the format of all the content that was viewed on a particular day cannot be classified. The parties said that this called into question the robustness of the entirety of the ITV results.²⁶ Figure 1 shows the number of views for each day for each format of content to illustrate how much of an effect this is likely to have on the results. Figure 2 shows the number of views for each day by format of content for UK content. Figure 3 shows the number of views for each day by format of content for non-UK content. These diagrams show that [redacted]. These views have been excluded from the regressions and therefore particular care should be taken in interpreting the results

²⁵[redacted]

²⁶Parties’ response to provisional findings, 19 December 2008.

for the regressions in which [✂] views is the dependent variable. These missing observations will have no effect on the regressions in which [✂] or [✂] views are the dependent variable.

FIGURE 1

[✂]

Source: [✂].

FIGURE 2

[✂]

Source: [✂].

FIGURE 3

[✂]

Source: [✂].

16. Finally, having considered the potential for autocorrelation and heteroskedasticity in the regressions, we used Newey-West standard errors to account for autocorrelation up to seven lags. This method of estimation corrects the standard errors that we have estimated for any autocorrelation and heteroskedasticity that might be present. This allows us to be confident in the significance of the results that we report. We note that neither autocorrelation nor heteroskedasticity would bias the results that we obtain, but, if present, would impact on the efficiency of estimation.





