Jeanette D. Snowball

Measuring the Value of Culture

Methods and Examples in Cultural Economics

Springer

2 Using Economic Impact Studies to Value the Arts

Economic impact studies can be used for valuing all sort of cultural goods, but especially those that attract large numbers of tourists (and thus tourist spending) from outside the impact area. This means that, while there are certain exceptions, like the "superstar" museums discussed by Frey (1998) and "blockbuster" art exhibits (Skinner 2006), the economic impact method is most commonly applied to cultural events, but can also include not-for-profit cultural sectors within a particular region.

For example, economic impact studies conducted on cultural goods have included opera, subsidized theatre, orchestras, the impact of local cinema, art and other museums, music institutes, not-for-profit arts and culture organizations, heritage sites and so on. A very useful resource for these kinds of studies is the Impact Database, hosted by the Centre for Cultural Policy Research at the University of Glasgow. While the method can be easily extended to just about any cultural good or service, this chapter gives special prominence to arts festivals because of their spectacular recent growth.

Since the 1980s there has been an explosion of the number of festivals of all types, not just arts festivals, but folk festivals, harvest festivals, food festivals, family festivals, carnivals, literary festivals – the list is endless. It is estimated that there are more than 300 festivals in the UK (British Federation of Festivals 2004), 1300 in Australia (Johnson et al 2005) and more than 5000 in the US (Blumenthal 2002). Quinn (2005) and Liu (2005) suggest that this change occurred because of a change in the way the service sector of the economy was viewed – not simply as a support for, or parasite on, manufacturing any longer, but as an important, wealth-creating sector in its own right. The implication of such a change for festivals, especially arts festivals, has been that many cities see them as a way to put "even mundane places permanently onto the lucrative tourist map" (Liu 2005:3). In other words, arts festivals are increasingly being regarded primarily as generators of financial benefits – the "just add culture and stir

approach to urban regeneration" (Gibson and Stevenson 2004 in Quinn 2005).

As the number of arts festivals has increased world-wide, so too has the interest in valuing them through the use of economic impact studies. As Frey (2005) points out, this presents cultural economists with an interesting dilemma: on the one hand, various commentators (Seaman 1987; Madden 2001) have expressed their concern with using such a purely market-based measure to "value" the arts. On the other hand, arts organizations and practitioners have shown little interest in non-market valuation methods (like willingness-to-pay studies) and a growing reliance on economic impact data as a means of arguing for public and private sponsorship.

It is not that the economic impact method itself is particularly complex, as Tyrrell and Johnson (2006) point out. However, the number and extent of the assumptions and decisions made about such crucial inputs as visitor numbers, the size of the multiplier and the data collection time and method leave economic impact results vulnerable to manipulation and bias (Crompton 2006). Non-specialist readers of economic impact reports are seldom equipped to evaluate such methodological niceties and tend to accept the final figure as a "scientific" outcome.

Economic impact studies attempt to answer the question, "If the event had not taken place, what would the loss of revenue to the impact area have been?" In other words, they attempt to calculate all the additional economic activity that takes place in the impact area as a result of the event or facility being studied. These sorts of studies concentrate on the private good aspect of the arts, captured by market transactions, rather than the merit or public good aspects.

As such, commentators are deeply divided on the usefulness of economic impact studies in valuing cultural goods. On the one side are mostly academic arguments that hotly contest the use of economic impact studies, arguing that, in the case of arts advocacy, they are worse than useless and may even be harmful to the cause by encouraging inappropriate comparisons with other sectors and downplaying the whole purpose of culture (Seaman 1987; Madden 1998, 2001 amongst others). On the other side are mostly practitioners and arts organizations who argue that economic impact studies can provide useful information about culture and cultural events and are, moreover, very effective in helping to lobby for public and community support (Vogelsong et al 2001, Heaney and Heaney 2003, Crompton 1995, 1999, 2001 amongst others). Both dissenters and promot-

ers of the method, however, recognize that there are potentially dangerous methodological issues as well.

This chapter introduces the theoretical debates around using economic impact studies in cultural valuation. Chapter three offers practical advice and examples of how to go about conducting a reliable study.

2.1 The Benefits of Using Economic Impact Studies

The case in favour of the use of economic impact studies in arts advocacy rests mostly, as Cohen et al. (2003) suggest, on the pragmatic rather than the ideal. In other words, the fact that such studies can produce a "bottom line" figure, which can be easily understood and compared, is perhaps the most important positive argument put forward. The numbers are important, as many commentators show, because they are the basis upon which funding decisions are often made, "Public officials, boosters and the media accept the quantifiable which appears to represent reality in order to justify a desired project" (Johnson and Sack 1996:370).

Goldman and Nakazawa (1997) agree, stating that, when "hard choices" about which of a number of desirable projects to fund have to be made, economic impact figures can play an important part. When the funds are provided, or partly provided by community residents, "they [expect to] receive a return on their investment in the form of new jobs and more household income" (Crompton 1999:143) and this return can be shown in economic impact figures. For this reason, a vast number of economic impact studies on the arts have been conducted. Madden (2001) reports that, from 1973 to 1993 more than 200 arts economic impact studies were conducted in the United States alone. Since then, the number has continued to grow and, while few academic journals publish the results, the websites of many arts councils provide a long list of arts impact studies and even "Do-it-yourself" kits for arts managers and events organizers wanting to use the technique (Jackson et al. 2005). A number of computerized input-output models are also available, such as IMPLAN and RIMMS II.

On a more specific level, economic impact studies can provide information about how money can best be spent to improve an event. Such improvements can be both in terms of financial gains for the community, for ex-

ample improving areas in which visitors are shown to spend most (Vogelsong et al. 2001), and in terms of improving event quality and composition in order to attract new visitors and to keep regular visitors loyal (Heaney and Heaney 2003).

Heaney and Heaney (2003) conducted and economic impact analysis on a two week summer music institute in Stevens Point, USA. They argue that direct impact figures of participants can be used to expand or improve those areas of the event that visitor spending flags as important. For example, in the case of the music institute, it was found that visitor spending on travel was large, since Stevens Point is fairly remote and the authors thus suggest that information on travel routes, maps and websites could be improved since this might be an important "decision-making determinant" for visitors and also impact on "customer satisfaction" (Heaney and Heaney 2003:260). They also argue that indirect and induced economic impact figures could be used to gain community support and sponsorship, especially from those industries (like accommodation and food) that are shown by the economic impact analysis to attract significant visitor spending. In addition, they suggest that economic impact figures are useful in increasing the "stature and validity" of the institution and in lobbying for local government support on the basis that the event increases economic activity in the region.

A number of economic impact practitioners (Herrero et al. 2004; Cohen et al. 2003) recognize that they are only measuring a partial value of the good or event in question, that is, that the arts generate other significant benefits as well. Seaman (2003) points out that the arts generate three types of impact: (i) consumption values, made up of use and non-use values best measured by contingent valuation methods, (ii) long run increases in productivity and economic development, best measured by hedonic pricing models, and (iii) short run net increases in economic activity, best measured by economic impact studies. Guetzkow (2002) points out that the impact of the arts on communities occurs in many ways (for example, through direct involvement, audience participation and in simply having artists and arts organizations present) and on both an individual and community level. Economic impact studies, while one of the most popular forms of measuring value, capture only a part of the "impact" of the arts on communities and individuals.

Herrero et al. (2004) conducted an economic impact analysis of the "European Capitals of Culture" festival event in Salamanca in 2002. They found that the festival generated 556.1 million euros for Castilla y Leon, 247.2

million euros for the rest of Spain and 803.3 million euros in total (2004:15). They argue that this is an important way of valuing the festival because a city nominated as a "European Capital of Culture" must be financially sustainable (profitable) in the long run since, "along with the cultural organization itself, there is a need for a remarkable effort in the form of creating new cultural facilities, urban redesign, tourist equipment and communication in the city" (Herrero et al. 2004:3).

A study conducted on the economic impact of the non-profit arts organization in ninety-one American communities by Cohen et al. (2003) showed that, through the spending of audiences and the organizations themselves, they added $134 billion per annum to the American economy. While Cohen et al. (2003) question the validity of using such financial figures to justify government support for the arts, they report that the study was cited numerous times in political debates and led to a new resolution encouraging the support of non-profit arts organization being adopted by the conference of mayors as well as a $10 million increase in funds to the National Endowment for the Arts being passed by the House of Representatives in 2002 – the largest such increase in nearly twenty years. They conclude that, "At this time in history, economic development is perhaps the most persuasive message when making the case for support [of the arts] to local, state and national leaders" (Cohen et al. 2004:31).

There is no doubt that tourism is big business: including transport, it is estimated to be the largest world industry at $4.4 trillion per year (Liu 2005) and to be growing at around 5% per year (McGuigan 2005). In such a context, festivals are seen as a way for cities to raise their profile and differentiate themselves from other places in a competitive market. The unfortunate consequence is that measuring the value of arts festivals has come to be seen as simply an offshoot of "urban tourism research" (Visser 2005), which focuses only on the potential of festivals as economic development strategies, rather than on the value of the product they produce – the arts.

The process was helped along considerably by the work of Richard Florida and his influential book *The rise of the creative class* (2002). Simply put, his thesis was that the concentrated presence of the "creative class" in cities was positively correlated with higher levels of human capital and with faster economic growth. Despite a number of criticisms, his work has remained important and has been shown to be robust in more recent studies (Rushton 2006). On the one hand, as Rushton (2006:10) points out, this offers arts proponents an alternative approach to the "economically dubious economic impact studies" based on visitor spending. However, it rein-

forces the idea that art and artists mean money and this link, once made in the minds of promoters and policy makers, is difficult to redefine.

A general review of the press reporting of both increases in funding and in funding cuts seems to verify the argument that the arts in many countries are being valued primarily as drivers of economic growth and development, rather than in terms of their intrinsic cultural value. For example, *Back Stage* (2003) published a highly critical report of a big cut in funding to the California Arts Council. The basis for the criticism was almost entirely the expected fall in economic impact as a result of the cut, rather than the loss of any aesthetic or qualitative values that the arts might provide. Similarly, in South Africa, when announcing provincial government sponsorship for the National Arts Festival, spokespeople concentrated their remarks on the economic benefits, increased tourist attraction and job opportunities offered by the event.

A recent article by Crompton (2006) provides a wide range of examples of the successful use of (often inflated) economic impact studies to argue for increased funding of various tourism projects, including cultural events. Ironically, it is the very success of such studies that has encouraged what Crompton refers to as "mischievous" practices (a less diplomatic word would be "dishonest"), discussed in more detail in the next chapter. However, he concludes with a warning that such hugely inflated economic impact figures are beginning to lose their power as advocacy tools because of the skepticism with which they are coming to be viewed. If economists want these studies to be taken seriously, they will have to be conducted with greater integrity.

2.2 The Dangers of Using Economic Impact Studies

Criticism of economic impact studies in arts advocacy can be divided into methodological issues (dealt with in chapter three) and conceptual problems. The latter generally argue that even the most sophisticated impact study would not be a good way to motivate for public funds and that such focus on financial indicators, rather than helping, may harm the arguments of arts advocators.

The first problem is that any economic impact study is highly sensitive to the impact area or regional delineation of the research question. Since eco-

nomic impact relies on the spending of visitors from outside the region and, in some cases, additional spending of locals within the region, the question of opportunity cost must inevitably arise. That is, where is the money coming from and what other sectors or regions have suffered because this one has gained? As Seaman (1987:731) puts it, "when enquiring as to the source of these quasi-mysterious 'exogenous' increases in overall spending, one often discovers that they may not constitute net increases, but merely changes in the composition of spending demand". The point is not a new one and was also made by Baade and Dye (1988:41) in their analysis of the rationale for the public subsidization of sports stadiums, where they argue that "net new activity" often includes "a reallocation of the preexisting level of local residents' spending".

Madden (2001) elaborates, pointing out that it is not enough to show that there are multiplier effects within the impact region, rather one must show that these effects are larger in the benefiting region or sector than in those industries or areas from which the event has diverted funds. The same argument holds for diverting government funds towards art:

> "Increases in government expenditure must ultimately come from somewhere – either diverted away from alternative policy expenditures, or away from the expenditures of citizens through their taxes. The net effect depends on the 'inverse' impacts of the area from which the extra money is diverted" (Madden 2001:167).

Madden's (2001:172) point is that lobbying for public funds based on projected financial gains for one region by diverting spending away from other regions is not a politically neutral game. In fact, he argues that it smacks of protectionism and "is an invitation to war – event war". Using financial figures to lobby for arts support can be dangerous in that such studies encourage numerical comparisons with other industries or events whose purpose is entirely different from the arts. This is verified by Crompton's (2006:79) personal experience in what he calls "the economic impacts arms race": When the council was presented with an economic impact figure of $16 million for a three week festival, they rejected the results on the grounds that the impact of a 3 day rodeo event was calculated (by another consultant) to be nearly $60 million.

Both Madden (2001) and Seaman (1987) point out that it is unlikely that the arts will ever be shown to have the impact of a "basic industry" like, for example, petroleum and coal products, and that such comparisons are in any case spurious. Even if the arts could be shown to have a comparably

large economic impact, this would still not be a good reason to lobby for public support. Gazel and Schwer (1997) show that the impact of three *Grateful Dead* rock concerts on the Las Vegas economy was between $17 and $28 million, but no one would dream of using these figures to argue for the public support of the rock band.

Another problem with focusing on the financial impacts of the arts in regions where there is a large difference between rich and poor residents is that most of the money accrues to the wealthier residents of the community who have some means of capitalizing on the presence of visitors. For example, in the Grahamstown economy (a small host city in South Africa), it was found that household income from the arts festival held there was vastly greater amongst wealthier European-origin (white) residents than in the poorer African-origin (black) communities, although the percentage of retained income in poorer communities was higher (Antrobus et al. 1997). Seaman (1987:746) concludes, "Arts proponents, therefore, are involved in a dangerous game when they resort to impact studies. In a sense, they are choosing to play one of their weakest cards, while holding back their aces". In other words, it is the positive spillovers provided by the arts to society, the primary cause of market failure, which should be used to motivate for public support to the arts and not the more frequently cited economic benefits.

An example of a case where the financial benefits of arts festivals are regarded as more important than the social, is given in McGuigan's (2005) discussion of the selection of cities for the European Capital of Culture competition. His argument is that, although the criteria for the competition might have started out as being cultural, they have come to focus more and more on the potential financial benefits the prize can provide. "It is not at all surprising that Liverpool won the British competition for designation as European Capital of Culture 2008. It is a perfect test case for the efficacy of culture-led regeneration" (McGuigan 2005:239).

Quinn (2005 and 2006) makes an eloquent argument against seeing the arts and arts festivals primarily as drivers of economic growth on two grounds. Firstly, the focus on arts events as commodities for sale to tourists undermines the traditional role of festivals as community gatherings where people could come together to establish new, or re-establish old, networks and connections. Festivals can be places set apart from everyday life where alternative views and values can be shared and considered, where resistance to the dominant regime can be expressed and where cultural capital (Throsby 2001) can be valued and shared. Of course, arts events can also

be about "vibe" and atmosphere, good food and entertainment, but as Peter Aspden (Aspden and Clark 2004) suggests, this should not be their *only* or main function.

Quinn's (2006) second argument is that, if arts festivals are treated primarily as commodities, they are likely to lose their capacity to interest and excite and thus to attract tourists and generate economic or cultural benefits. Liu (2005:4) talks about the threat to traditional cultures of "McDonaldization" and the spread of uniform American cultures:

> In the summer of 2002 I was witness to the first time that paraders in Disney costumes appeared at the end of the traditional Nadaam festival in Ulaanbaatar, Mongolia. No doubt, Mickey, Minnie and Donald will become regulars at the annual celebrations, alongside with Genghis Khan's cart, the wrestlers, acrobats and horses. The warning is quite clear. Unless traditional culture is valued and nurtured, it will not survive the next generation brought up with MTV, videogames and survival shows.

Aspden (Aspden and Clark 2004) agrees that many modern arts festivals simply present works already performed elsewhere at higher prices, "it is a sort of supermarket, where the paying public is persuaded to bulk-buy processed culture". Along with the increased competition for audiences and artists and crowding out effects experienced by some festival managers (Blumenthal 2002), it is not difficult to see that a focus on the monetary gains from such events may defeat both cultural and financial ends.

In such a context, Quinn (2006) argues that arts festivals may fail to fulfil the host community's cultural or economic expectations. Preoccupation with short-run financial benefit excludes longer term values, such as the festival enhancing the identity of the place and its development potential (in terms of infrastructure, increases in demand for artistic products year round, increases in social welfare and so on). To illustrate her point, Quinn (2006) reports the results of a study of two festivals in Ireland, Wexford Festival Opera, geared from the start towards providing high quality opera and the attraction of foreign tourists, and the Galway Arts Festival, a much more community based enterprise. The findings, based on an opinion survey of both visitors and local residents, show that at both festivals participants were aware of the non-market values the festivals provided, like creating arts awareness, developing arts venue infrastructure, helping to build the reputation of the host town and creating more arts activities throughout the year.

However, the Galway Arts Festival became more interested in the financial benefits of the festival and focused more attention on tourists, thus moving away from their original aim of being a community-based festival. Local communities began to resent this move, feeling excluded and dissatisfied because the original relationship between the festival and the community was not being maintained. Quinn (2006) argues that, for a festival to be sustainable in the long run, its aims must be clear from the outset and that it needs to be recognized that different aims will generate different returns. For example, a festival that balances local – international orientation is likely to be very useful in developing cultural infrastructure in the host city and increasing appreciation of and participation in the arts, but is not going to be as successful in attracting visitor demand as a festival which focuses on international (rather than local) visitors and artists.

Madden (2001) goes so far as to say that government intervention based on economic impact figures could do more harm than good, since the objectives of government are seldom aligned with those of the arts. Cohen and Pate (2000:109), talking from the perspective of artists, support this view: "Artists have asserted (in conversation with us) that they feel it is absurd to make decisions on their future funding without fully recognizing the aesthetic worth of their product". In a similar vein, Tusa (1999 cited in Reeves 2002:36) states that,

> "Mozart is Mozart because of his music and not because he created a tourist industry in Salzburg or gave his name to decadent chocolate and marzipan Salzburg kugel. Picasso is important because he taught a century new ways of looking at objects and not because his paintings in the Guggenheim Museum are regenerating an otherwise derelict northern Spanish port…Absolute quality is paramount in attempting a valuation of the arts; all other factors are interesting, useful but secondary".

McCarthy et al. (2004), identify two types of cultural value: "Intrinsic" value, which has a direct relationship with the artistic product itself (like pleasure and cognitive growth); and "instrumental" value, including economic impact, which is a by-product of artistic production and can also be generated by substitute goods. They argue that the "culture wars" in America in the early 1990s necessitated "proving" the value of the arts, which increased focus on the instrumental values disproportionately because they are more easily measured than the intrinsic benefits. Matarasso (1997) found the same pattern in the UK, stating that studies of the impact of the arts have tended to focus on the financial, rather than the social, benefits to the exclusion of "the real purpose of the arts".

In addition to such categorical dismissals of the method, there are other, interpretive problems, arising largely from the attempt to extract information from economic impact studies not designed for the purpose. Personal experience of the disbelief and disappointment of event organizers when presented with final reports has prompted further thought about the ways in which such figures are viewed and used. The first anomaly occurs when stated event aims are compared to the use of financial figures in declaring the festival "a success". The more public good characteristics an event has, the smaller will be the benefits captured by an economic impact study and the larger will be the consumer surplus and the value of non-market goods, which has not been measured by an economic impact study at all. Madden (2001) argues that economic impact studies measure spending on the event, in other words, *costs*, not benefits and that if the arts were available for free, they would undoubtedly increase our well-being or utility by a greater amount because of decreased opportunity costs. The usefulness of economic impact studies in these cases, however well they are conducted, is thus questionable.

Another problem area of economic impact study interpretation arises when arts event organizers try to draw conclusions about the relative importance of various activities from data on visitor spending. In 2001, South African National Arts Festival organizers commissioned a consumer research study chiefly in order to establish consumer spending patterns and opinions with a view to using this information to lobby for further sponsorship of the event. Despite the generally good opinion of festival-goers of Main and Fringe shows – an average of 4 out of 5 for quality and price - festival organizers expressed disappointment that spending on tickets was only the third highest expenditure category, accommodation being first and shopping second (Snowball and Antrobus 2001:18). At first glance, this result does appear to be contradictory or to indicate that festival visitors are interested in other aspects of the event more than in shows. However, if one considers that shows at the festival were highly subsidized - on average, ticket prices for the exact same shows outside of the festival were found to be 41% higher (Snowball 2004) - while accommodation and shopping are not, the errors that can be made by using only expenditure data to indicate interest or value become apparent.

Even those who criticize impact studies as a tool for arguing for the public support of the arts, however, do recognize that they can be useful. Seaman (1987) points out that, if economic impact studies are conducted with methodological soundness, they can be used successfully to examine the relationships between various parts of the economy and to make predictions about income and output changes. Madden (1998) agrees, adding

that, in addition to financial flows, economic impact studies can provide important information about the effects of demand and supply shocks on regional economies and a way of comparing the financial redistribution that results from different projects. However, both Madden and Seaman point out that economic impact studies are seldom put to only these uses and they both argue vehemently that, in the majority of cases, economic impact studies of arts and culture are an "abuse of economic analysis" (Seaman 1987:725).

2.3 Conclusions

Considering both the advantages and disadvantages of using economic impact studies in arts valuation, one can make two general observations. Firstly, accepted as a *partial* analysis, and especially if conducted in conjunction with some other sort of study better suited to capturing the aesthetic or cultural values of the event or facility, impact studies can be useful to a certain extent, especially in lobbying for funds. Secondly, economists requested to conduct such studies find themselves in a particularly uncomfortable position. On the one hand are commentators like Madden (2001:174) who argue that undertaking such studies is a "prostitution of economics", since economists must know how questionable the reasoning behind such studies is. On the other hand, attempts to convince arts managers of the usefulness of the conceptually more complex contingent valuation methodologies is, as Madden (2001) also admits, slow work.

It is likely that arts practitioners will continue to commission economic impact studies, despite their theoretical problems, because of their effectiveness in arguing for public and private support. In such an atmosphere, economists who argue against their use are likely to be regarded rather like Cassandra of ancient Troy – doomed to speak the truth, but not to be believed. In this situation, practitioners are better served by refining the method and making sure that methodological issues are kept high on the research agenda rather than by rejecting economic impact as a valuation method entirely.

References

Antrobus, G., Williams, V., Fryer, D., Khumalo, B., Streak, J. & Webb, A., (1997) The economic impact of the 1996 Standard Bank National Arts Festival. Department of Economics, Rhodes University: Grahamstown

Aspden, P and Clark, A. (2004) High brow, low blows. Financial Times Magazine 19/06/04

Blumenthal, R. (2002) Festivals, festivals everywhere: Summer arts events multiply, testing the limits of growth. New York Times 30/07/02

Cohen, C. and Pate, M. (2000) Making a meal of arts evaluation: can social audit offer a more balanced approach? Managing Leisure 5:103-120.

Crompton, J. (1995) Economic impact analysis of sports facilities and events: Eleven sources of misapplication. Journal of Sports Management 9:14-35.

Crompton, J. (1999) The economic impact of sports tournaments and events. Parks and Recreation 34,9:142-151.

Crompton, J., Lee, S. and Schuster, T. (2001) A guide for undertaking economic impact studies: The Springfest example. Journal of Travel Research 40:79-87.

Crompton, J. (2006) Economic impact studies: Instruments for political shenanigans? Journal of Travel Research 45:67-82.

Florida, R. (2002) The Rise of the Creative Class. Basic Books, New York

Frey, B. (2005) What values should count in the arts? The tension between economic effects and cultural value. Institute for Empirical Research in Economics, University of Zurich. Working paper no. 253.

Goldman, G. and Nakazawa, A. (1997) Determining economic impacts for a community. Economic Development Review 15,1:48-52.

Gazel, R. and Schwer, K. (1997) Beyond rock and roll: The economic impact of the Grateful Dead on a local economy. Journal of Cultural Economics 21:41-55.

Guetzkow, J. (2002) How the arts impact communities. Taking the Measure of Culture conference, Center for Arts and Cultural Policy Studies, Princeton University Working Paper Series.

Heaney, J. and Heaney, M. (2003) Using economic impact analysis for arts management: An empirical application to a music institute in the USA. International Journal of Nonprofit and Voluntary Sector Marketing 8,3:251-266.

Herrero, L., Sanz, J., Devesa, M., Bdate, A. and Del Barrio, M. (2004) Economic impact of cultural macrofestivals: the case study of Salamanca 2002, European capital of culture. Presented at the 13th International Conference on Cultural Economics, Association for Cultural Economics International.

Johnson, A. and Sack, A. (1996) Assessing the value of sports facilities: the importance of non-economic factors. Economic Development Quarterly 10,4:369-382

Liu, J. (2005) Tourism and the value of culture in regions. The Annals of Regional Science 39:1-9.

Madden, C. (1998) "Discussion paper: The economic benefits of art. Creative New Zealand [On line] Available: www.creativenz.gov.nz [Accessed 13/08/04].

Madden, C. (2001) Using economic impact studies in arts and cultural advocacy: a cautionary note. Media International Australia incorporating Culture and Policy. Feb,98:161-178.

Matarasso, F. (1997) Use of Ornament? The social impact of participation in the arts. Comedia [On line]
Available:
http://www.comedia.org.uk/pages/pdf/downloads/use_or_ornament.pdf
[Accessed 18/05/07]

McCarthy, K, Ondaatje, E., Zakaras, L. and Brooks, A. (2004) Gift of the Muse: Reframing the debate about the benefits of the arts. Rand Corporation: Santa Monica, Arlington and Pittsburgh.

McGuigan, J. (2005) Neo-liberalism, culture and policy. International Journal of Cultural Policy 11:229-241.

Quinn, B. (2005) Arts festivals and the city. Urban Studies 42:927-943

Quinn, B. (2006) Problematising 'festival tourism': Arts festivals and sustainable development in Ireland. Journal of Sustainable Tourism 14:288-306

Rushton, M. (2006) The creative class and urban economic growth revisited. Paper presented at the 14th International Conference of the Associated for Cultural Economics International, Vienna.

Seaman, B. (1987) Arts impact studies: A fashionable excess. Towse, R. (Ed.) 1997 Cultural economics: the arts, the heritage and the media industries Vol. 2 Edward Elgar: Cheltenham.

Seaman, B. (2003) The economic impact of the arts in Handbook of Cultural Economics. Towse, R. (ed) Edward Elgar.

Snowball, J. D. and Antrobus, G. (2001) Consumer Research: A survey of visitors at the 2001 National Arts Festival, Grahamstown. Department of Economics, Rhodes University: Grahamstown. Commissioned by the Grahamstown Foundation.

Tyrrell, T. and Johnson, R. (2006) The economic impacts of tourism: A special issue. Journal of Travel Research 45:3-7.

Visser, G. (2005) Let's be festive: Exploratory notes on festival tourism in South Africa. Urban Forum 16:155-175

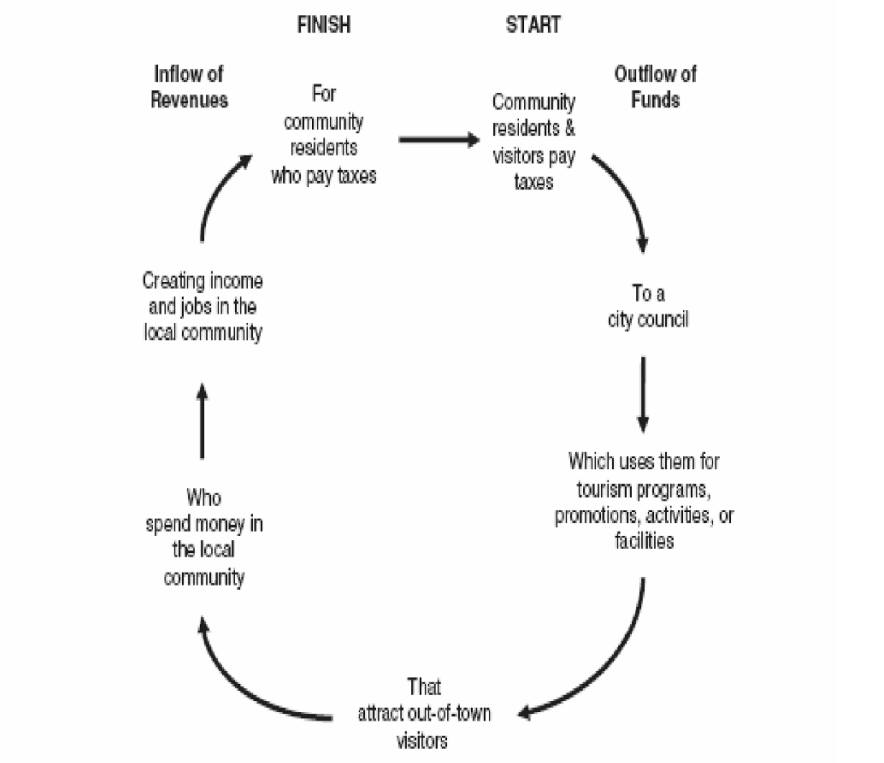
Vogelsong, H., Graefe, A. and Estes, C. (2001) Economic impact analysis: A look at useful methods. Parks and Recreation 36,3:28-33.

3 Calculating Economic Impact

In the introduction to a recent special edition of the *Journal of Travel Research* (Volume 45, 2006), Tyrrell and Johnson define economic impact analysis as seeking "to estimate changes in regional spending, output, income and/or employment associated with tourism policy, events, facilities or destinations". As discussed in the previous chapter, there are a number of good reasons that economic impact data may be very useful in valuing an arts event or cultural facility. As long as it is always presented as a partial value, that is, as long as it is always acknowledged that the economic impact figure is not capturing *all* value, such a figure can be a powerful way to demonstrate the worth of a particular project.

Crompton (2006) argues that it is important not to see economic impact studies as a cost-benefit analysis that applies only to the host city council. For example, if the council spends more on the event than it generates directly in revenue, this is not a good argument for discontinuing because it does not take into account the benefits to community residents, who pay the council taxes in the first place. Below (fig. 3.1) is the useful diagram that he provides showing the rationale for conducting studies of cultural events or facilities that attract tourists from outside the impact region.

As mentioned in the previous chapter, while economic impact studies can also be used to value permanent local cultural facilities or arts organizations, they are most often used when "new" spending occurs in the region because of tourists. When the method is used to value the cultural industries or not-for-profit arts organizations of a whole region, it is more often the local economic activity stimulated by the organizations themselves (for example, their spending on equipment and labour) than the, often small, impact of tourists from outside the region that is mainly considered. This is because, as further discussed below, the spending of local residents in the impact area is not usually included, since this spending may have taken place in any case.



(Crompton 2006:68)

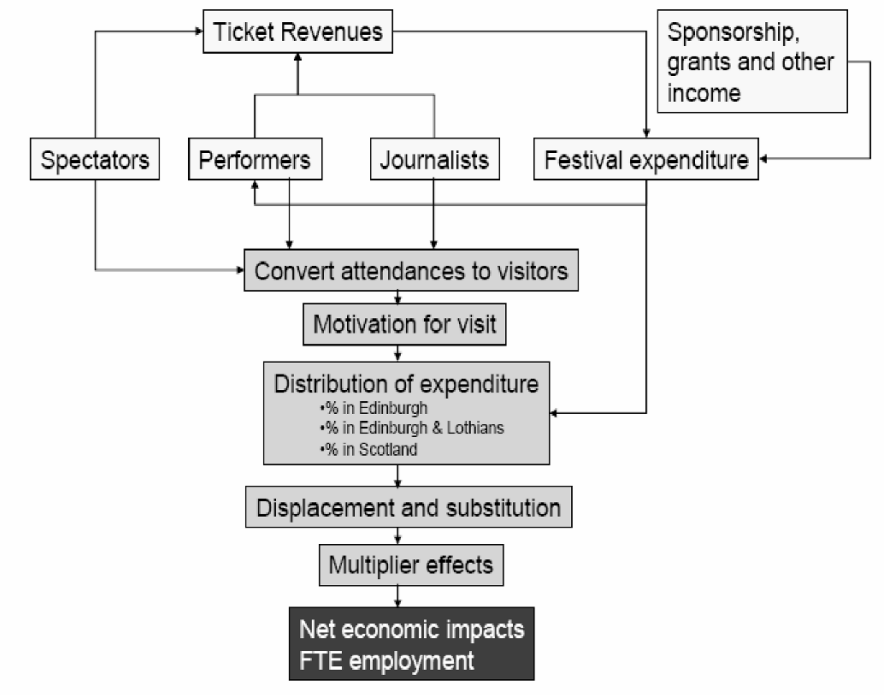
Fig. 3.1. The Economic Impact Cycle of Costs and Benefits

The implication of the above diagram for economic impact studies of cultural events is that it is not enough to track spending and income by organizers or the host city council in question – one also has to take into account the indirect impacts in terms of job creation and income that the event has on local residents. This chapter offers practical advice on how to conduct such as study, with particular reference to cultural events and festivals that draw tourists into the impact region.

3.1 Direct Net Economic Impact

The first step in any economic impact study is to determine the net injections into the impact area as a result of the event, often referred to as direct impact or first-round spending. One way of doing this is to collect visitor

spending data using quota or probability sampling and interview or self-completion questionnaires. However, Tyrrell and Johnston (2001) argue that, in addition to visitor spending, spending by producers, the value of the time donated by volunteers and media spending should also be included. The identification of important spending groups will depend on the type of event being valued. For example, an economic impact study of the Edinburgh Festivals in 2004-2005 identified four "strands" of spending, from spectators, performers, journalists and the festival organizers, as shown in the diagram below.



(Edinburgh Festivals Study 2004-2005:6)

Fig. 3.2. Calculating Economic Impact at the Edinburgh Festivals

Net direct spending should not include spending that would have occurred in any case, for example, spending by locals that could be regarded as "diversions of spending" from other goods in the area (Seaman 2003) and spending by "casual" visitors or "time switchers" who would have come regardless of the event. This is why taking into account "motivation for visit" and "displacement and substitution effects" are part of the process of calculating economic impact shown in figure 3.2.

Since short term events sometimes include foreign producers, gross direct impact may be very large, but net direct impact may be negative. This is because of the diversion of local spending away from local goods towards "foreign" producers who take most of their profits away with them. Thus, in the case of arts festivals or events which rely on many performers and vendors from outside the region, it is likely that net and gross direct economic impact will be considerably different. Results are also likely to be affected by how the impact area is defined ("distribution of expenditure" in figure 3.2), as well as the way in which data is collected. The following section discusses methodological issues in calculating net direct impact.

3.1.1 Data Collection Methods and Sampling

The way in which visitor spending data is collected can have a huge effect on the calculation of economic impact, especially if the festival population is highly differentiated in terms of language, number of trips and length of stay. Most studies rely on either interviews with attenders using random or stratified sampling, or self-completion questionnaires returned by mail or at collection points at the event. The advantage of using interviews is that sampling is easier to control and data is generally more reliable and of a better quality. However, this method is expensive and can be supplemented with self-completion questionnaires handed out at, for example, festival shows, restaurants and other facilities near or associated with the cultural resource, art exhibitions, souvenir shops, craft markets and so on. While a cost-effective way of increasing the sample size, combining the two methods (or relying on only one method) can introduce bias.

An example of the difference that data collection methods can make is in the comparison of the economic impact of the South African National Arts Festival (NAF) calculated by two different teams using different methods. One team of researchers had conducted economic impact studies on the NAF for a number of years (Antrobus et al. 1996; 1997; Snowball and Antrobus 2001, 2003; Antrobus and Snowball 2004; Snowball and Antrobus 2006), referred to as the Antrobus method. A second, unrelated team (Saayman et al. 2005), referred to as the Saayman method, conducted a similar study in 2005.

The 2005 Saayman study was based on self-completion questionnaires offered in English. The 2006 Antrobus study (and those preceding it) used

both self-completion questionnaires in English and face-to-face interviews with festival goers on a quota basis. The quota was based on past studies and was put in place to prevent interviewers from introducing any unintentional bias. While the Saayman method produced a visitor profile showing that 73% of festival audiences were European-origin (white), English speaking people, the Antrobus method showed only about 64% in this category. It is probable that this is because people who read and write English easily are much more likely to fill in a self-completion questionnaire than second-language English speakers are. In addition, if income is split along racial lines, as it is in South Africa, self-completion data submitted by a higher proportion of wealthier, white people is likely to produce visitor spending data that is upwardly biased. This was shown to be the case in the 2006 study, where the average visitor expenditure per group for the entire stay was R3 204 (about 320), while it was only R2 167 (about 217) on average using interview data. An important finding is thus that, in the situation where groups of attenders show a strong correlation between language and wealth characteristics, self-completion data is likely to be biased unless questionnaires are offered in more than one language where literacy levels are high (Bragge and Snowball 2007).

Loomis (2007) identifies two types of sampling bias that can inflate the final economic impact figure. The first is "avidity bias", that is, people who visit an area where the sample is being selected (like a craft market area or other facilities associated with the event) more often, are more likely to be included in the sample. The second sort of bias is "length of stay" bias – that is, that visitors who stay longer (and who thus are likely to spend more) are more likely to be included in the sample than short duration visitors.

Depending on the type of event being valued, sampling procedures can be used to overcome these forms of bias. For example, using exit intercept interviews may reduce the length of stay bias, but might increase avidity bias (Loomis 2007:43). A way of reducing bias might also be to sample visitors at multiple points. For example, in the case of a festival, sampling could be done at craft markets, show and ticket queues, art exhibitions and other facilities in order to include as many types of visitor as possible. However, those who spent more time on such activities and who stayed longer are still likely to be over-represented.

Loomis (2007:43), adapting from Nowell et al. (1988), suggests correcting for length of stay and avidity bias by weighting visitors sampled by the re-

ciprocal of their length of stay (in days) and number of trips respectively. So length of stay could be corrected by:

$\mu_{LS} = N/\Sigma\ (1/LS_i)$ (1)

where μ_{LS} is the corrected length of stay, N is the total number of visitors sampled, and LS_i is the length of stay in days of visitor i. Similarly, avidity bias could be corrected for using

$\mu_{Trips} = N/\Sigma\ (1/Trips_i)$ (2)

where μ_{Trips} is the corrected number of trips and $Trips_i$ is the number of trips for person i (Loomis 2007:43).

In a case study of rafting trips on the Snake River in Wyoming, Loomis (2007:45) found a significant difference in the corrected and uncorrected estimates of the number of trips per rafter (a 42% overstatement without the correction), the trip length (17% overstatement) and thus the number of visitor days, calculated by multiplying trip length and number of trips (52% overstatement). The cumulative effect of the corrections was that the economic impact of the rafters was 58% lower than the uncorrected figures.

Visitor spending data collection at the South African National Arts Festival

Please estimate how much you intend to spend in total at the Festival on the following items.
(Grahamstown locals, please report only spending in addition to your normal monthly expenses).
Important!
Are these figures for: □ only yourself? □ your whole traveling group/family?
□ Accommodation: R _____ □ Shows: R _____
□ Food & drinks: R _____ □ Shopping: R _____
□ Other (please specify)
How many people are you paying for at the Festival?
(Snowball and Antrobus 2006)

The accuracy with which visitors report their spending is also dependent on how the data is collected. For example, it has been shown that visitor spending data is more accurate if it is collected soon after the spending has taken place. A long time lapse between the spending and the reporting can result in bias because of inaccurate recollection and "telescoping" – reporting spending outside of the study area that may have occurred on the same trip (Wilton and Nickerson 2006). These problems are exacerbated for impact studies of permanent tourist attractions, like museums, opera houses and theatres, but are unlikely to occur in event studies because data is typi-

cally collected from visitors during their attendance. However, if data for the whole trip is being asked for (as opposed to spending on the day of the interview only), some acknowledgement of "prediction bias" needs to be made, since it is quite likely that visitors will be required to project total spending before their trip is complete.

Crompton et al. (2001) argue that group spending figures are likely to be more accurate than individual spending, especially for accommodation. This is also likely to apply to group ticket sales, especially at events where bookings can be made far in advance. However, it is unlikely that group members would be able to account for other spending categories for all their members (like souvenir spending and spending on food and drinks), so caution needs to be exercised. One way of dealing with the problem is to give visitors the choice of reporting group or individual spending as in the South African National Arts Festival study illustrated above. However, if this is done, it needs to be very carefully recorded which sort of data (group or individual) is being reported.

Visitor spending data collection at Art in the Park, Blackburn

Please can you estimate how much your group/family spent on the following items during your outing to Art in the Park.
(Include all expenditure related to attending the event e.g. staying in Blackburn for a meal/drink afterwards etc.)

Items	Spent of Saturday £	Spend on Sunday £
A. Food (restaurants, take-aways, snacks etc)		
B. Drink (soft drinks, alcohol, tea/coffee etc)		
C. Transport (bus fare, taxi, train, petrol, parking etc)		
D. Gifts/Souvenirs (toys, crafts, programmes etc)		
E. Other items (please specify)		
F. Estimated total cost of day out		

Please can you estimate how much your group/family would spend in Blackburn on a *normal weekend*?

Saturday	£
Sunday	£

(Wood 2005:42)

Finally, the number of spending categories needs to be thought about carefully. Leaving out a potential major category or categories, could result in

an underestimation of visitor spending in the impact area. For example, in the 2006 South African National Arts Festival study, a frequently reported "other" category was transport. How much of this spending actually took place in the impact area, however, is doubtful. On the other hand, having too many categories can quickly lead to interviewee fatigue and the possible over-estimation of total spending due to rounding up and imperfect recall. For shorter festivals, like the Art in the Park event in Blackburn, Lancashire (Wood 2005), daily spending figures can aid recall and provide useful information on spending patterns and times to organizers.

3.1.2 The Area of Study and Local Spectators

Crompton (1995; 2006) argues that a failure to define accurately the impact area of the study could lead to widely differing results and is one of the most frequent "mischievous" procedures used to inflate economic impact estimates. As would be expected, the larger the area under consideration, the less would be the "leakages" from the system and thus the greater the multiplier effect and the reported economic impact. Crompton (2006:73) also points out that when one expands the area of study from, for example, one city to the whole province, a great many more "visitors" will count as "local spectators", much of whose spending should be excluded from the analysis.

Despite researchers carefully defining the area of study, media reports may persist in misreporting results. For example, the impact figure for the 2003 South African National Arts Festival (R33 million) was reported as being applicable to the whole Eastern Cape Province in the headline, "Grahamstown festival earns Eastern Cape R33m" (*Daily Dispatch* 2003). This is simply untrue, since the figures were calculated for Grahamstown only, not the whole province. Provincial impact figures would have included fewer "visitors", but also fewer leakages and would have produced substantially different figures. However, since the Eastern Cape government is now a major Festival sponsor, reporting the impact as applying to the whole region made better political sense, but fallacious economics!

Since economic impact studies are attempting to measure the financial difference between a situation with the event versus a situation without the event, only spending that occurs directly because of the event should be included (Tyrrell and Johnson 2006). Spending that would have occurred in any case should be excluded. Thus, only attenders from outside the

study area should be included in the study, since the spending of local residents does not represent injections of new money, but merely the "recycling" of money already in the area (Crompton 1995). In small cities and towns, local visitors are unlikely to make up a large proportion of festival goers, but in larger centres, the effect of including local residents can be quite marked. For example, a study of the Melbourne International Festival of the Arts (1996:13) showed that local visitors made up 73% of festival goers and spent $3.14 million, but were excluded entirely from the study on the basis that the spending would have occurred even if the festival had not taken place.

There are two possible exceptions to this rule: firstly, spending by local residents who have stayed in town specifically to attend the festival could legitimately be regarded as new money entering the region. One could estimate this figure by asking local residents to report spending in addition to their normal monthly expenses. "However, these types of estimates are very tenuous and economists invariably recommend that all expenditure by local residents should be disregarded" (Crompton 1995:27). Getz (1991:303) also suggests that if the event keeps local residents of the host community at home, rather than spending their money outside the community, local spending could be included in the impact analysis. Another alternative would be to report economic impact both with and without the additional spending of local attenders.

If additional spending by local residents is going to be included in the impact study, one needs to make careful provision for this in the collection of spending data. Referring to the two examples in section 1.1, note that the South African National Arts Festival survey asks local residents ("Grahamstown locals") to report only spending in addition to normal monthly expenses; the Art in the Park event in Blackburn (Wood 2005:42) asks visitors to "estimate how much your group/family would spend in Blackburn on a *normal weekend*", so that the extent of additional spending as a result of the event can be estimated.

Examples of arts impact studies that did include additional spending by local residents are the 1990 - 1991 Edinburgh Festival study (1991:9) and the Adelaide Festival study (1990:13), which included spending by residents of the region "which is additional to normal economic activity" (Edinburgh Festival study 1991:9). The argument here is that, especially for events of short duration, like a festival, local residents may choose to stay in town and "holiday" at the festival instead of spending entertainment funds outside the area (Adelaide festival study 1990:13).

In a household survey of local residents of the South African National Arts Festival host city (Grahamstown) conducted in 2003 (Snowball and Antrobus 2003), about 87% of the sample reported some spending at the festival. However, of these, 48% claimed that they spent no more than they did in a normal week, while 52% reported some extra spending as a result of the festival. In order for local spending to have a net positive effect on the region, it has to be shown that there is import substitution, in other words, that spending which would have occurred outside the region now occurs in the impact area as a direct result of the event (Seaman 2004). Results show that, for those with additional spending, 53% would have spent it in Grahamstown anyway, 31% would not have spent it (saved) and only 15% would have spent it outside the area.

Furthermore, Seaman (1987:732) points out that it is important to ask how local spending is being funded. "If it is from savings at a local bank, the secondary effect would be a reduction in the available pool of loanable funds for, perhaps, local investment or consumption projects far removed from the arts". In other words, the opportunity cost of diverted local spending should also be considered. All this verifies Crompton's (1995; 2006) point that counting local spending as part of economic impact is problematic.

It is also argued (Crompton 1995, Crompton et al. 2001, Tyrrell and Johnston 2001) that the expenditure from visitors who would have come to the area regardless of the event being measured should not be counted as contributing to the economic impact of the event, since they would have spent money in the area anyway. Crompton (1995) defines "time switchers" as people who may have been planning to visit the area for some time (to visit family and friends or to experience some other feature of the area, like museums and game parks), but have switched the time of their visit to coincide with the festival or event. "Casuals" are visitors who may already have been in the area for whatever reason, and decide to attend rather than do something else. In both cases, these visitors would have spent money in the area regardless of whether there was a festival (or other event) or not.

Another issue is whether visitors to the impact region come primarily for the cultural event or whether they are combining their visit with other reasons for visiting the area. In the latter case, not all spending can be attributed directly to the cultural good in question. For example, the study of the economic impact of the Southern Festival of Books in Memphis (Wallace and Honey 2005) found that, in addition to attending the book fair, a significant percentage of visitors also went to other local attractions like:

Graceland (17.5%), the Memphis Zoo (2.5%), Beale Street (30%), the Peabody Hotel (35%) and other things. Strictly speaking, one could argue that the income from visitor spending (in this study, all attributed to the book fair) should be partly discounted to take into account other attractions. On the other hand, the survey did find that 85% of out-of-town visitors came primarily for the festival and only used these visitor numbers to calculate economic impact, excluding all spending by local residents and the spending of visitors not primarily there for the festival.

The 2004-2005 Edinburgh Festival study found that, after taking into account displacement (spending that would have occurred in the impact area in any case) and leakages (spending outside of Edinburgh), the percentage of total visitor spending that could be attributed directly to the Festivals decreased to 61% of the total.

Table 3.1: Impact Area, Visitor Numbers and the Employment Multiplier

	Grahamstown	**Melbourne**	**Edinburgh**	**Adelaide**
Area definition	Somewhat vague: Albany district included small farming villages, but no bigger cities.	Well defined: separate studies for "city of Melbourne and suburbs" & "greater Melbourne".	Well defined: separate studies done for several areas.	Well defined: 80km radius & separate regional impact.
Including locals	Tickets bought by locals excluded: 21 per cent locals.	Excluded all spending of locals: 73 per cent locals.	Included spending by locals "additional to normal economic activity": 1 million pounds.	Excluded most local's spending except residents "holidaying at the festival".
Including "time switchers" & "casuals"	No check. Did ask foreigners main reason for visiting SA.	Did check: only 24 per cent gave festival as main or only reason for visit.	No check.	Did check: 51,4 per cent of visitors there primarily to attend festival. "Extended stay" also counted.
Employment multiplier	No permanent jobs claimed other than the organizers.	No permanent jobs claimed.	Probable overstatement. 1300 full-time jobs claimed.	No mention of employment created.

(Snowball and Antrobus 2002:1305)

A comparison of economic impact studies conducted at four arts festivals (Snowball and Antrobus 2002) showed that, in general, the area of study was fairly well defined, but that only two of the four studies examined excluded local spending entirely or checked for time switchers and casuals (see table 3.1). However, while these are important considerations, deter-

mining visitor numbers, especially for un-gated events with free shows, can be even more problematic and crucial.

3.1.3 Determining Visitor Numbers

Even the best estimates of net economic impact will depend to a great extent on the accuracy of visitor number calculation because, once visitor spending has been determined, it is multiplied by the estimated number of visitors. However, very little on how this figure is arrived at has been written. Crompton et al. (2001) discuss the example of Springfest (a 4 day annual cultural festival held in Ocean City) where they used an hourly and bi-hourly counting method at various access points to arrive at an estimate. Wilton and Nickerson (2006) used counts at highway entry points (traffic counts) and airports to estimate tourist numbers – a useful method particularly when the impact of a long running event or a permanent tourism attraction is being calculated. Estimates from aerial photographs can also be used as estimates when the event in question is held mainly outside.

The National Arts Festival studies in South Africa (Antrobus et al. 1996 and 1997) used two methods for visitor number calculation that relied on having accurate ticket sales data and accurate data of at least one type of accommodation. The first method ("ticket sales method") was to collect data on the average number of ticketed events attended by each respondent for the whole festival. This average number included festival visitors who had attended no ticketed shows and were mainly concerned with shopping at the craft markets and/or attending free shows, street theatre and art exhibitions. The average number of ticketed shows attended was then divided by the total number of tickets sold, excluding those sold to local residents, to arrive at a total number of visitors (21 662 in 1996 and 20 700 in 1997).

The second method ("accommodation method") used the visitor questionnaire to determine the percentage of visitors using university accommodation and the average number of nights festival visitors stayed. Data on the number of bed nights sold during the festival was then obtained from Rhodes University and divided by the average length of stay to give the number of visitors in the university residence accommodation category. Since the percentage of visitors in this category was known from the visitor questionnaire, total visitor numbers, excluding local residents and including day visitors, could then be calculated (25 808 in 1996 and 19 822 in 1997). In both the 1996 and 1997 surveys comparable results were ob-

tained using the two methods. The weak point is that both methods rely heavily on a representative sample of festival goers being drawn from the population to avoid over- or understatement of important figures, like the number of shows attended and the length of stay.

Length of stay bias, discussed in section 1.1, is also mentioned as a problem by Wilton and Nickerson (2006). In additional to a lower probability of being interviewed, short-stay visitors are also less likely to be willing to complete questionnaires because they are generally in more of a hurry than longer-stay attenders (Antrobus et al. 1997). In addition to having a direct effect on spending data and economic impact, such bias may also affect visitor number calculations. At the 1998 South African National Arts Festival (Antrobus and Snowball 1998), a specific attempt to collect data from day and short-stay visitors was made by using a "sixty second interview". The motivation behind it was that, since the interview would take only one minute of the visitor's time, even those who were only staying for a day or two might be willing to help.

Table 3.2. Data Collection and Visitor Numbers at the South African Festival

Category	**1996**	**1997**	**2003**	**2004**	**2006**
Method (percentage interview versus self-completion)	84	42	100	41	74
Average number of ticketed shows attended per person	5.2	6	4.9	6	9
Total number of tickets sold	184 761	157 380	95 913	104 617	111 776
Percentage of local respondents	20	21	33	17	7.5
Visitor numbers (* excluding locals)	25 000* (31 250)	20 000* (25 300)	20 000	20 000 (16 600*)	20 700 (17 000*)

Of those interviewed, 12.6% were day visitors. Longer, self-completion questionnaires were also used at the 1998 Festival, but of these respondents, only 5.6% were day visitors. The implication for the average length of stay and thus for the calculation of visitor numbers for use in economic impact studies is great. In particular, the method of data collection (self-completion versus interviews) and the length of the questionnaire are likely to have important effects. In general, visitor number calculations using the ticket sales and accommodation methods are likely to be downwardly biased, especially if only self-completion questionnaires are used and the questionnaire is long. The 2003 and 2004 studies used the ticket sales method of visitor number calculation in conjunction with counting methods (particularly at free events and street shows) done by festival organizers in an attempt to compensate for this possible underestimation (see table 3.2).

As discussed earlier, the economic impact of the South African National Arts Festival (NAF) had been calculated by one team of researchers for a number of years, referred to as the Antrobus method, and was then calculated for one year by an unrelated team referred to as the Saayman method. The results were, unsurprisingly, remarkably different: the 2004 Antrobus method estimate being R38.5 million (about 4 million Euros), while the 2005 Saayman method was R53.5 million (5.5 million Euros). After careful examination of results, it became apparent that, while the Antrobus method was based on estimated visitor numbers of about 20 700, the Saayman method estimated 70 000 visitors. It is not known how the Saayman method calculated the number of visitors, but given reliable ticket sales data for 2005, their figure implies that, over the entire average stay of 4.5 days, visitors attended less than 2 ticketed shows each, which seems implausible given attendance figures from previous studies (Bragge and Snowball 2007).

Another interesting component of table 3.2 is that, when a method of reducing the number of local residents interviewed was introduced in 2006, the percentage of locals dropped from around a 20% average to only 7.5%. However, as can be seen from the visitor numbers excluding locals, the number of outside tourists calculated using the ticket sales and accommodation methods remained roughly the same between 2004 and 2006.

The reporting and interpretation of even very robust visitor number calculation is also far from obvious. Crompton (1999) points out that one of the most contentious parts of any economic impact study is the calculation of visitor numbers because average spending per visitor is multiplied by the number of visitors in order to determine the first round, or direct, economic impact of the event. However, visitor number estimates are also important because there is a prestige component to being able to claim to have many visitors, particularly where the emergence of a number of cultural events in recent years has led to considerable competition.

It is obviously in the interests of the organizers, therefore, to report as many people as possible attending and visitor number calculations are thus quite often based on a certain amount of wishful thinking. Reporting of actual discreet visitors versus visitor days thus needs some clarification since the term “visitor numbers” is not as unambiguous as it sounds. Confusion can lead to many recriminations and disbelief, since it is possible that an obviously smaller event can claim a larger number of visitors, when referring to visitor days, than a much larger one which refers to actual individual people (Snowball 2004).

The confusion arises because of the differing length of time that visitors spend at an event. For example, a festival that is located near large cities is more likely to attract day visitors and short-stay visitors than one that is located in a more isolated area. For example, in 2003 the average visitor at the South African National Arts Festival stayed for about 6 days and there were 20 000 visitors (using the ticket sales and counting methods). However, this figure refers to individual people, so the number of visitor days (i.e. the addition of the number of visitors who were at the festival each day) was in fact around 121 000. The latter figure was, of course, much more acceptable to organizers and also gives a better idea of the size of the event, since it takes into account the characteristics of the particular festival location and makes comparison with other events more meaningful. While a festival located closer to large cities may thus claim to have a larger number of different individuals attending, it is the comparison of visitors per day that is most revealing. However, it must be emphasized that the reporting *method* makes no difference to the economic impact, but (as Loomis 2007 suggests) there may be differences in the spending patterns of shorter-stay and longer-stay attenders that would be reflected in final impact figures.

3.1.4 Producers, Sponsors, Vendors and the Media

Tyrrell and Johnston (2001) argue that, in addition to calculating the spending of visitors or spectators, spending in the impact area by producers, sponsors, vendors and the media should also be included. However, they also point out that, in order to avoid double counting, the source, starting point, destination and reason for the expenditure also needs to be tracked if this method is to be followed.

The implication is that economic impact studies should also include producer surveys – both visiting and local business people and performers – in order to fully capture the economic effects of the event. A survey of local businesses at the South African National Arts Festival showed increased spending of R2.3 million on wages to local residents, materials, site fees, electricity, fuel and living expenses (like accommodation, food and general consumer spending). Visiting traders spent about R38 million in Grahamstown during the 1996 Festival, while visiting performers spent about R1.3 million in 1997, mostly on the provision of accommodation and food for cast members and their families. However, as Tyrrell and Johnston (2001) point out, counting producer expenditure as well as the

full value of sponsorship (often used to pay for the costs of artistic productions on the Main program) could result in double counting (Antrobus et al. 1997a).

Surveys of local businesses were also conducted at the 1996 (Antrobus et al. 1997a) and 2003 (Snowball and Antrobus 2003) South African National Arts Festivals. Despite the long time interval, the two business surveys showed very similar results. In both cases, local businesses who provided food and drinks, services (travel agencies, banks) or goods related to festival activities (photography, florists, pharmacies) experienced a significant increase in their monthly income during the event. Both surveys also found that for hardware stores, building contractors and the media (including printing services) business increased directly before the Festival as households and producers prepared for the event. An interesting point from the 2003 survey was that businesses who reported no change in monthly income admitted that, since the festival takes place in the Rhodes University holidays, they would, without the festival, have experienced a fall in income. In other words, events which take place out of season could have an important role in smoothing the cyclical nature of the earnings of local businesses.

Another way of looking at the effects of visitor spending on a local economy is to consider the forward and backward linkages of businesses supplying tourists. Cai et al. (2006) compiled linkage indices for the tourism industry in Hawaii using Leontief and Ghosh supply-driven multipliers. While the backward linkage multipliers are the same for tourist and non-tourist production, there are some interesting differences in the forward linkages. Except for industries specifically geared towards tourist supply, like hotels and air transport, "the web of forward linkages tends to be greater when producing for tourism than for non-tourism consumption" (Cai et al. 2006:43). In other words, tourist related sales may be more beneficial and important for economic activity than non-tourist sales and this could also be a consideration when valuing a tourist event.

However, at most festivals, money-making opportunities are not limited to local traders and the presence of traders from outside the area might have negative implications for local business people. Seaman (1987) argues strongly that, to the extent to which visiting traders provide competition to local stores, the receipts from these activities should be accepted as a substitution for the earnings of those in the local community and thus subtracted from the first round spending in economic impact calculations.

At the South African National Arts Festival the presence of visiting traders was perceived to increase competition in some sectors (particularly amongst clothing retailers) and a cause of congestion and overcrowding that discouraged regular local customers from shopping. If one accepts the Crompton et al. (2001) view that municipal spending on events is expected to generate a financial return for local tax payers, one can understand the frustration of local businesses regarding visiting traders who have paid little or nothing towards the event, but are reaping the benefits, in some cases in direct competition with local businesses.

The value of media coverage, especially in the longer term, could also have a significant effect on the impact area. At the SA National Arts Festival, the Rand value of media coverage of the festival (including newspapers, magazines, online articles, radio and TV) has increased steadily from about R38 million (3.8 million Euros) in 2002 to nearly R80 million (8 million Euros) in 2004. It could thus be argued that the additional publicity has a significant impact on local businesses, private schools, the university, estate agents, game lodges in the surrounding area and other related industries. Such effects would fall into the "long run increases in productivity and economic development" category mentioned by Seaman (2003) and are typically not included in a short run economic impact study.

3.1.5 Supply Constraints and Other Costs

In addition to demand side errors, Seaman (2004) points out that few, if any, cultural economic impact studies have referred to supply side constraints. He argues that if event visitors displace or crowd out visitors who might otherwise have come to the region, then not all event visitor spending can be included as a benefit since, if the event had not taken place, other visitors would have spent in the impact area.

A study of the impact of the Spoleto festival in Charleston, USA actually found that the festival had a negative impact on the local hotel industry (Litvin and Fetter 2006). The authors compared hotel occupancy and daily hotel rates during the festival with those outside of festival times. Despite the fact that the festival is quite long (17 days) and attracts a fairly affluent audience, they found that, on average, Charleston hotels were less full during the festival than in prior and following weeks and that rates were also not higher. They conclude that the festival causes more potential tourists to stay away than it attracts.

The Blackburn study also took into account losses that might have occurred because people who would normally have been in the impact area on non-festival days might have spent more on their normal shopping trip than they did at the festival. However, it is unlikely that such differences would have been large, because 92% of respondents reported that they would not normally have been in the area (Wood 1995:46).

Supply constraints certainly do exist in cities and towns, particularly with regard to accommodation, but Seaman (2004) also gives reasons why this may not occur in certain cases. Firstly, if the timing of the event is annual and well advertised, it is almost certainly known in advance and non-festival tourists could easily reschedule their visit. In fact, some events are specifically scheduled for tourist "low-season" to take advantage of excess capacity. Secondly, some (fairly) small amount of visitor displacement may occur, but this could be offset by additional local spending that would have occurred outside the area, were it not for the festival. In addition, even if there is some crowding out of non-festival visitors, those attending may spend more money than the former group. Finally, it may be that there is enough excess capacity available in the town to cater for both groups.

Crompton (1995; 2006) argues that economic impact analysis should also take into account the opportunity costs of public or private spending on the event as well as any negative impact that the event may have. He points out that local government spending on the event should not be counted as an injection of new funds into the area because the money has come from local residents in the form of taxes, in other words the original "investment" (Crompton et al. 2001). Public funding from outside the region can likewise only be counted as new money (and thus included in the economic impact) if it would not otherwise have been spent in the impact area.

In discussing the economic impact of sports facilities, Johnson and Sack (1986:376) agreed that one needs to ask, "Would a similar or larger amount of state support now be available for a project with more direct economic impact if the tennis complex had not been built?". Although they conceded that this question may be unanswerable, respondents in their study agreed that some of the city's "political capital" had been spent in lobbying for state funds for the project, which, the authors argued, should be acknowledged as a cost (Johnson and Sack 1996:376).

Table 3.3. Benefits and Costs at *Arts in the Park* in Blackburn

Benefits	**Problems**
Local Residents	
Money to the town	Crowds
New visitors to the town	Strangers
Opportunities for family fun	Traffic
Good image for Blackburn	Parking
Bringing community together	Safety on streets
Pride in the area	Litter
Opportunity for new experiences	Vandalism
Something to look forward to	Other crime
Investment in the local area	
Local Businesses	
Money to the town	Demand business can't meet
New visitors to the town	Traffic congestions
Good image of Blackburn	Parking problems
Bringing community together	Litter
Pride in area	Vandalism
New customers to your business	Other crime
Non-local customers for your business	Putting-off regular customers
Greater awareness of your business	
Investment in the local area	
Community Groups	
Community development	Funding
Good image for Blackburn	Traffic congestion
Bringing community together	Parking problems
Pride in the area	Litter
Good publicity for your organisation	Vandalism
Funding for your organisation	Other crime
More involvement in your organisation	Cost of involvement
Awareness of your organisation	
Investment in the local area	
Something to look forward to	

(Wood 2005:43)

While public funds to events might make up a fairly small percentage of the economic impact of the event, studies of whole cultural sectors and non-profit arts organizations in general show that a large percentage of the operating budget comes from government sources. For example, in a study of the economic impact of Florida's arts and cultural industries it was found that 51% of income for these organizations was unearned. Of the total income, federal, state and local government grants made up nearly 19%. In this case, it would be very important to ask questions about how else the money could have been spent and whether it might have been spent in the impact area in any event, especially in the case of local and state grants.

Cost to the host community, in terms of contributions by local government in the form of increased police presence, refuse removal, etc, are also often not subtracted from the total impact figure. An ideal study would consider a wide variety of costs, or negative impacts, such as the pressure on infrastructure, traffic flow problems, overcrowding of the town centre, increased crime, inconvenience to local residents, like lack of parking, crowds, noise and litter, increased competition to local stores and a feeling of antagonism by local store holders to visiting traders (Antrobus et al. 1997, Wood 2005; see table 3.3). Another possible opportunity cost is that of "festival refugees" - those local residents who deliberately leave town when the festival is in progress to avoid the inconvenience. However, in some cases, these residents then let their houses to festival visitors at a profit, so the cost imposed by the loss to the town of the spending of "refugees" during festival may not be significant.

While putting monetary values on longer term benefits and non-market or opportunity costs is often difficult, they could at least be mentioned in the economic impact report in a qualitative way, which would further serve to alert readers to the partial nature of the economic impact figure.

3.2 Indirect Impact

In addition to direct effects, injections of new money into a region generate indirect impact as a result of successive rounds of spending that occur within the impact area: this is known as the multiplier effect. Second round or indirect spending is then added to the estimate of direct spending in order to calculate total economic impact. The following section discusses the use of multipliers, including the importance of the employment multiplier, in calculating this figure.

Once visitor numbers and expenditure have been determined and the direct impact worked out, a multiplier size must be determined in order to calculate indirect impact. The size of the multiplier and thus the effects of successive rounds of spending, will depend on the "leakages" from the economy being considered. Leakages represent the amount of money that is taken out of the host economy in the form of spending by local earners outside the host economy and savings. The amount of leakage (and thus the size of the multiplier) is determined by the level of imports into the impact area, which depends on the size and nature of the host city. For ex-

ample, in 2003 a business survey was conducted (Snowball and Antrobus 2003) as part of the South African National Arts Festival economic impact study. The impact area, Grahamstown, is a small town with little manufacturing. Local businesses were surveyed to determine the extent to which stock sold in Grahamstown is sourced from outside the region. It was found that, on average, 87% of stock was bought from larger cities in the province. The result suggests that leakages from the area can be expected to be large in second round spending, even if local residents spend festival earnings at local businesses.

Crompton (1995:29) states that, "It is not desirable to take the [multiplier] results of an economic impact assessment from similar studies in other communities and apply it, because the combinations of business interrelationships in communities are structured differently so linkages and leakages will be different". Seaman (2003) reiterates this, but since it is time consuming and expensive to calculate a multiplier from first principles, the tendency in many studies seems to be exactly that, i.e. to use multipliers that have been derived for the region, or for other events, or simply to use an estimate. However, some countries have developed regional input-output models, like Minnesota IMPLAN Group and RIMS II developed by the USA Bureau of Economic Analysis, which can be adjusted to the specific region and help with calculating indirect impact.

One way to estimate multiplier size is to use general rules from meta-analysis. A great deal of literature exists regarding the determinants of regional multipliers, and fortunately, results tend to agree with one another. One very important factor is the size of the region in question and its location. A number of studies (Shahidsaless et al. 1983; Baaijens and Nijkamp 2000 and Greenberg et al. 2002) all found the size of the multiplier to be positively related to both population size and the physical area of the region. Shahidsaless et al. (1983) used the well-known Central Place Theory to further expand on this, finding that the marginal propensity of a region to import is negatively related to the distance to another, larger market. Thus, the closer the impact region to a larger market, the more likely it is to rely on that market for goods and services, the greater the leakages will be and the smaller the multiplier. Greenberg et al. (2002) also find that smaller, rural towns have fewer opportunities for the creation of backward and forward linkages, thus also decreasing multiplier size.

Table 3.4. Variables Included in the Rough-Set Analysis Model

Geographic variables	
Population size (POP)	1. <10,000 2. 10,000 – 100,000 3. 100,000 – 200,000 4. 200,000 – 1,000,000 5. 1,000,000 – 5,000,000 6. no available information
Surface (in km^2) (SUR)	1. <500 (VS = very small) 2. 500 – 10,000 (S = small) 3. 10,000 – 100,000 (M = medium) 4. >100,000 (L = large)
Geographic feature of the area (GEO)	1. island 2. archipelago 3. country (no island) 4. region (no island)
Political autonomy (POA)	1. independent (I) 2. not independent (NI)
Tourist variables	
Number of tourist arrivals (TOA)	1. <10,000 (S) 2. 10,000 – 700,000 (M) 3. >700,000 (L)
Share of arrivals from most important country of origin in total number of arrivals "largest country share" (LCS)	1. <45% (S) 2. 45-75% (M) 3. >75% (L) 4. no available information
Type of tourist attractiveness (ATR)	1. "sun, sand and beach holidays" 2. nature and culture holidays (NC) 3. "mixed" i.e. both "sun, sand and beach holidays" and nature and culture holidays (MI)
Decision variable	
Average tourist income multiplier (TIM)	1. <=0.5 (S) 2. 0.5 – 1.0 (M) 3. 1.0 – 1.5 (L) 4. >1.5 (VL = very large)

(Baaijens and Nijkamp, 1997: 843)

Baaijens and Nijkamp (2000) undertook a very useful meta-analysis of eleven economic impact studies from different countries and developed a "rough-set analysis model. As shown in table 3.4, they provide a means of estimating a multiplier for any region or event based on geographic variables (population size, surface area of the impact area, geographical features of the area and political autonomy) and tourist variables (number of tourists, origin of tourists and tourism type). For example, a small region (less than 500km square) with a population of between 20 000 and 150 000 is likely to have a multiplier of less than 0.5, except where there is a very large (10 000 to 70 000) inflow of people from outside the region.

Table 3.5. Determining the Multiplier Value for Four Arts Festivals

Festival	Year	Multiplier Determination	Direct Impact	Indirect Impact	Implied multiplier value[1]
South African National Arts Festival	1996	"Using a factor of 0,6 for leakages" (1997a:15)	R25.9 million	R4 million	0.18
Edinburgh Festivals	1990 1991	Carried out by Leisure Research Services No discussion of how the "multiplier" was calculated.	43.9m Pounds (Edinburgh and the Lothians)	9.2m Pounds (Edinburgh and the Lothians)	0.22
	1996	Scottish Tourism multiplier study (1996:6)	122m Pounds (Edinburgh)	30m Pounds	0.24
Adelaide Festival	1990	Input-output multipliers generated by Centre for South Australian Studies (1998)	$8.8 million (S. Australia)	$10 million (S. Australia)	1.15
Melbourne Festival	1995	"The derived multipliers for the 1992 Adelaide Formula One Grand Prix have been used to undertake a formal economic evaluation of the 1995 MIFA" (1996:25)	$18.37 million	$20.11 million	1.1

(Snowball and Antrobus 2002:1305)

A second and related method of multiplier estimation is to examine multiplier sizes of similar studies. For example, comparing multipliers used at other arts festivals and taking into account the source of the multiplier (what research it is based on), the size of the impact area and the varying nature of the event can provide a useful starting point. Table 5 below shows such a comparison and can suggest at least upper and lower bounds for other studies. A complication of any such study is that multipliers and their sources are often not explicitly discussed in economic impact reports, leaving them open to misrepresentation and exaggeration, especially where sales multipliers and employment multipliers are used (Crompton 2006).

[1] Except for the South African National Arts Festival, no actual multiplier values were given in any of the festival reports. The above values are thus obtained by inference (comparing direct and indirect income) and may thus be subject to error or over-simplification. However, as a rough comparison, they are still interesting.

Crompton et al. (2001:81) suggest that a sales multiplier, that is, the effect of an extra unit of visitor spending on economic activity, is not actually a very useful measure of economic impact and that it is the personal income multiplier that should be used since "it enables the economic benefits received by residents to be related to the cost they invested". Given the extent of the leakages from local business, turnover in terms of the value of sales is unlikely to be related closely to increases in income. However, specific data on household and business income changes, as a result of the event, is difficult to collect, since respondents are often reluctant to give accurate figures, making the estimation of an income multiplier, in the absence of a pre-existing input-output model, difficult to estimate.

Table 3.6. Edinburgh Festival 2004-2005 Multipliers

	Edinburgh		**The Lothians**		**Scotland**	
	Output	Income	Output	Income	Output	Income
Accommodation	1.52	0.33	1.53	0.35	1.74	0.51
Food and drink	1.70	0.42	1.71	0.45	1.94	052
Entertainment	1.55	0.50	1.56	0.53	1.78	0.67
Shopping	1.54	0.33	1.56	0.35	1.93	0.34
Transport	1.39	0.31	1.40	0.33	1.53	0.36

(Edinburgh Festivals Economic Impact Report 2005:10)

More sophisticated studies provide both sales and income multipliers. For example, the very comprehensive 2004-2005 Edinburgh Festivals report used multipliers from the Scottish Tourism Multiplier Study (1991), which provided sales (output) and income multipliers for various sectors (accommodation, food and drink, entertainment, shopping and transport). As can be seen in table 3.6, the multiplier values vary between sectors, but are generally far smaller for income than for output (taking into account leakages of various sorts from the impact area). In addition, as the impact area increases, so does the size of the multiplier – an expected change, since a larger area generally has fewer leakages.

The employment multiplier shows how many full-time jobs are created as a result of the event or festival. However, as Crompton (1995:22) and Crompton et al. (2001) point out, the employment multiplier is most unreliable because it assumes that, "all existing employees are fully utilized so an increase in external visitor spending will inevitably lead to an increase in the level of employment". Especially for a once a year event, like a festival, it is unlikely that many new jobs would be created. Rather, existing employees would work overtime or "casual" short term help would be hired for the period of the event. An interesting point that other commentators do not seem to take into account, is that not all the jobs created will

necessarily be filled by people from the impact area. For example the 1996 South African National Arts Festival study found that, of the 387 temporary jobs created by visiting traders, only 36% were taken by people from the impact area (Antrobus et al. 1996:15). This seems to be largely as a result of the lack of technical skills amongst unemployed local residents, which encourages artists and vendors to bring in their own personnel.

Skinner (2006) argues that the only really reliable way to find out whether the event in question has attracted additional tourists from outside the region (and thus created additional jobs) is to conduct an *ex post* study of the region. Skinner conducts such a study to determine whether three "block buster" art exhibitions in Jackson, Mississippi had any impact on regional income using a time series analysis. The measure of economic growth is the change in employment figures. Taking into account seasonal changes and the business cycle, Skinner (2006:123) finds that, for at least two of the three exhibitions, the estimated job creation from the *ex ante* economic impact studies (721 jobs) is "virtually indistinguishable from the 699 job creation forecast of the filtered time series model". The findings support the idea that, at least in some cases, employment multipliers can give reliable information on the job creation potential of cultural events.

When the cultural good in question is a permanent feature, like the Columbia Museum of Art (CMA), using impact figures to calculate increases in jobs is more feasible. For example, the CMA generates 61 jobs directly from its own operation and 98 jobs indirectly from visitor spending (Miley Gallo and Associates 2006). Since the CMA operates all year round, jobs created directly are likely to be stable, but since tourism is seasonal, the indirect jobs may be an overstatement. As in short event studies, existing staff might be utilized to deal with temporary tourist influxes and so not create "new" jobs.

3.3 Total Economic Impact

The total economic impact of the event can be calculated quite simply by adding the net direct and indirect impact figures. As the preceding discussion has suggested, the method itself is not very complex, but the underlying assumptions that lead to the net direct and indirect figures used in the final calculation are crucial.

Table 3.7 shows the calculation of economic impact at the South African National Arts Festival over a number of years. While the study is not very sophisticated (only one multiplier is used and, owing to data unavailability, an assumption of the size of immediate outflows was made based on past studies), it does serve to show the mechanics of the calculation and give an idea of trends. Total visitor spending (row A) is calculated by multiplying the average spending per visitor by the estimated number of visitors. This figure included only *additional* spending by local residents. The presence of "time switchers" and "casuals" was investigated, but found to be negligible – expected, given the relative remoteness and small size of the host city. The potential negative effects of the "crowding out" of local suppliers and "festival refugees" was not included, but was expected to be small.

To this is added sponsorship (assumed not to have been spent in the region if not for the festival) and rental earnings by local organisations holding shares in the Village Green craft market. In row C, an estimate of the immediate outflows (earnings by non-local traders and performers) is made by extrapolating from the 1996 data (the last time a producer survey was conducted), which showed outflows as approximately 30% of total visitor spending. Row D shows the net direct impact and E the indirect impact, calculated by multiplying direct impact by the estimated multiplier.

Table 3.7. Economic Impact Calculations at the South African Arts Festival

Category (All figures in millions of Rands)[2]	**1996**	**1997**	**2003**	**2004**	**2006**
A. Total visitor spending	25.9	23.5	27	30	33.6
B. Sponsorship and craft market	1.5	1.5	13	13	13
C. Immediate outflows (30% of A + B)	7.6	7	12	13	14
D. Net direct impact (A + B – C)	19.8	15	28	30	32.6
E. Indirect impact (D x 0.18)	3.4	2.7	5	5.5	5.9
F. Total economic impact	23.3	17.7	33	35.5	38.5
Total impact at 2000 prices (Rands)[3]	30	21	27	28.6	25.5

While not as accurate as an ideal study, impact figures over time can provide useful information on the evolution of an event, as well as an indication of its financial value. Noteworthy points of the above table are the fall in ticket sales, visitor numbers and real economic impact from 1996 to 1997, with a slow recovery being reflected in the 2003 and 2006 figures. However, ticket sales of 1996 levels (184 761) have not been seen again, with 2003 and 2004 figures hovering around 100 000, but showing an up-

[2] Where one Rand is equal to about ten Euros

[3] Using the Consumer Price Index from Statistics South Africa

ward trend. The same is true of real economic impact (at 2000 prices), which is now starting to approach 1996 levels again.

The major reason posited for the decline is the emergence in the New South Africa of many other arts festivals, the greatest rivals at present being the *Klein Karoo Nationale Kunstefees* (KKNK), started in 1995 and catering for a mostly Afrikaans speaking audience, and the *Aardklop* Arts Festival that began in 1998. Perhaps the decline is also partly due to the freedom of expression and speech in the New South Africa. During the apartheid era, the National Arts Festival played a very important role as an outlet for otherwise repressed expressions of political and social protest. Once the elections had been held and the New South Africa was underway, its role necessarily had to change and the falling off of ticket sales and economic impact may have been a reflection of this period of uncertainty (Snowball and Webb 2007). The stabilization and growth of the festival in recent years may thus indicate that this tricky period has been successfully negotiated, even in the face of increased competition from other festivals, and that it is finding its place in the new democracy.

3.4 Conclusions

A thoughtfully designed economic impact study can produce credible results that can be a useful way of valuing a cultural event or resource. However, the opportunities for misrepresentation are quite considerable and care should be taken at each stage of the process. While economic impact figures can be useful in arguing for the public and private support of the arts, this relies on their being regarded as believable, particularly where comparisons over time, or with other tourist events, are made.

Quite often, economists may feel pressured to come up with the “right”, that is, large, economic impact figure, especially if the study has been commissioned by event organisers or sponsors. Sometimes, it can be very difficult to present figures that may not show such a big financial impact and it is then useful to be able to suggest other ways of evaluating the cultural resource. One of the most popular methods is contingent valuation, or willingness to pay studies, which will be discussed in the next chapter.

References

Adelaide Festival Study, McDonald, S. (1990) The Adelaide Festival: the economic impact. Centre for South Australian economic studies: University of Adelaide, Adelaide.

Antrobus, G., Williams, V., Fryer, D., Khumalo, B., Streak, J. & Webb, A., (1996) The economic impact of the 1996 Standard Bank National Arts Festival. Department of Economics, Rhodes University: Grahamstown

Antrobus, G., Webb, A. and Mather, D. (1997) The economic impact of the 1997 Standard Bank National Arts Festival. Department of Economics, Rhodes University: Grahamstown.

Antrobus, G. and Snowball, J. (2004) National Arts Festival 2004: Festino Survey. Department of Economics, Rhodes University: Grahamstown.

Baaijens, S. and Nijkamp, P. (2000) Meta-analytic methods for comparative and exploratory policy research: An application to the assessment of regional tourist multipliers. Journal of Policy Modeling 22,7:821-858.

Bragge, B. and Snowball, J. (2007) Why arts proponents love economic impact studies and economists hate them. Paper submitted to AIMAC Cultural Policy conference, Valencia, Spain.

Cai, J., Leung, P. and Mak, J. (2006) Tourism's forward and backward linkages. Journal of Travel Research 45:36-52.

Cap Gemini Ernst & Young; Milward Brown Ulster (2005) Study of the economic and social impact of subsidised theatre in Northern Ireland. [On line] Available:http:///www.artscouncilni.org/departs/all/reports/research/subsidised%20theatre.pdf [Accessed 11/05/07].

Crompton, J., Lee, S. and Schuster, T. (2001) A guide for undertaking economic impact studies: The Springfest example. Journal of Travel Research 40:79-87.

Crompton, J. (1995) Economic impact analysis of sports facilities and events: Eleven sources of misapplication. Journal of Sports Management 9:14-35.

Crompton, J. (2006) Economic impact studies: Instruments for political shenanigans? Journal of Travel Research 45:67-82.

Edinburgh Festivals Study (1991) Final Report: Edinburgh festivals study 1990/91. Scottish Tourist Board: Edinburgh.

Edinburgh's Year Round Festivals, 2004-2005 Economic Impact Study (2005). SQW Economic Development Consultants and TNS Travel and Tourism: Edinburgh.

Getz, D. (1991), Festivals, Special Events and Tourism, Van Nostrand Reinhold, New York, NY

Greenberg, M., Lewis, D., Frisch, M., Lowrie, K., and Mayer, H. (2002) The US department of energy's regional economic legacy: special dimensions of a half century of dependency. Socio-economic Planning Sciences 36:109-125.

Johnson, A. and Sack, A. (1996) Assessing the value of sports facilities: the importance of non-economic factors. Economic Development Quarterly 10,4:369-382.

Litvin, S. and Fetter, E. (2006) Can a festival be too successful? A review of Spoleto, USA. International Journal of Contemporary Hospitality Management 18,1:41-49.

Loomis, J. (2007) Correcting for on-site visitor sampling bias when estimating the regional economic effects of tourism. Tourism economics 12,1:41-47.

Melbourne International Festival of the Arts (1994). A report on the economic and social impacts. Ernst and Young, Melbourne.

Melbourne International Festival of the Arts (1996) Evaluation of attendance and economic impact. Commissioned by the City of Melbourne and MIFA: Melbourne

Miley Gallo and Associates (2006) The economic impact of the Columbia Museum of Art on the city of Columbia [On line] Available: http://www.colmusart.org/assets/files/ecoimpact.pdf [Accessed 15/05/07].

Saayman, M., Slabbert, E. and Saayman, A. (2005) The economic impact and profile of visitors to the Grahamstown National Arts Festival 2005. Institute for Tourism and Leisure Studies, North-West University, Potchefstroom.

Seaman, B. (1987) Arts impact studies: A fashionable excess. Towse, R. (Ed) 1997 Cultural economics: the arts, the heritage and the media industries Vol. 2 Edward Elgar: Cheltenham

Seaman, B. (2003b) The economic impact of the arts in Handbook of Cultural Economics. Towse, R. (Ed) Edward Elgar.

Seaman, B. (2004) The supply constraint problem in economic impact analysis: An arts/sports disparity. Presented at Lasting effects: assessing the future of economic impact analysis conference, New York.

Shahidsaless, S., Gillis, W. and Shaffer, R. (1983) Community characteristics and employment multipliers in non-metropolitan counties, 1950 – 1970. Land Economics 59,1: 84-93.

Skinner, S. (2006) Estimating the real growth effects of blockbuster art exhibits: A time series approach. Journal of Cultural Economics 30:109-125.

Snowball, J. and Antrobus, G. (2001) Consumer Research: A survey of visitors at the 2001 National Arts Festival, Grahamstown. Department of Economics, Rhodes University: Grahamstown.

Snowball, J. and Antrobus, G. (2003) Economic valuation of the 2003 Grahamstown National Arts Festival: economic impact, business and household surveys. Department of Economics, Rhodes University: Grahamstown.

Snowball, J. and Antrobus, G. (2002) Valuing the arts: Pitfall in economic impact studies of arts Festivals. South African Journal of Economics 70,8:1297-1319.

Snowball, J. and Antrobus, G. (2006) National Arts Festival 2006 Festino Survey. Department of Economics, Rhodes University: Grahamstown.

Tyrrell, T. and Johnston, R. (2001) A framework for assessing direct economic impacts of tourist events: Distinguishing origins, destinations and causes of expenditure. Journal of Travel Research 40:94-100.

Tyrrell, T. and Johnson, R. (2006) The economic impacts of tourism: A special issue. Journal of Travel Research 45:3-7.

Wessex Group (2000) Virginians for the Arts: The economic impact of arts and cultural organisations in Virginia. [On line] Available: http://www.vaforarts.org/tempdocs/Wessex.pdf [Accessed 23/05/07]

Wilton, J. and Nickerson, N. (2006) Collecting and using visitor spending data. Journal of Travel Research 45:17-25

Wood, E. (2005) Measuring the economic and social impacts of local authority events. International Journal of Public Sector Management 18,1:37-53

4 The Contingent Valuation Method

In addition to the economic or financial benefits that the arts and heritage provide, they also supply us with non-market goods, like pleasure, national pride, building and expressing community identity, etc. (as discussed in chapter one). Such values relate to the public good nature of the arts and are sometimes referred to as positive externalities because they are *external* to the market. Expressing these values in monetary terms is not always possible (or desirable, as Klamer argues), but, as in environmental economics, such value estimates can be a useful indication of public (non-expert) opinions and reveal values not associated with purely financial gains.

There are a number of ways in which the value of externalities provided by public goods may be measured and they can broadly be divided into revealed and stated preference methods. Revealed preference methods involve gathering market data on actual spending, such as how much it cost to travel to a heritage site or how much more the property in a particular culture-rich area costs. Stated preference methods rely on hypothetical scenarios and directly ask respondents what value they place on a good.

The travel cost method measures the access costs of users to a specific site or event, i.e. what consumers are willing to pay in travel costs in order to attend, for example, an arts festival, a heritage site or a museum (Mundy and McLean 1998:292). The hedonic price method assumes that households "migrate into or out of geographic regions based on tradeoffs between quality of life in those regions and differences in housing prices" (Ready et al. 1997:439), i.e. what consumers are willing to pay in increased housing prices (or decreased wages) in order to live in an area that has certain cultural amenities. The contingent valuation (CV) method, so-called because the valuation is contingent upon the given scenario, asks respondents directly what they would be willing to pay, or willing to accept, in a hypothetical market situation to conserve or expand some public good (Ready et al. 1997:439). The CV method also includes stated choice or choice experiment surveys, which ask respondents to choose between

various scenarios in which the levels of the attributes making up the good are varied (further discussed in chapter six)[1].

Although not a focus of this book, the travel cost method is becoming more popular as a valuation method in cultural economics and a number of recent studies have used it. For example, Poor and Smith (2004) valued the historic city of St. Mary's, considered to be one of the most significant historical sites in the US, using a zonal travel cost method. Visitor benefits are shown to be extensive, but the authors point out that they are measuring only use values and that non-use value may also be important. Bedate et al. (2004) use travel cost to estimate consumer surplus for cultural goods (two cathedrals, a music event and a museum) in the Castilla y Leon region in Spain. Other examples are Boter et al. (2005) who use the method to value Dutch Museums, and Alberini and Longo (2006) who combine travel cost and contingent valuation to estimate domestic visitors' use and non-use values for cultural heritage sites in Armenia.

The hedonic pricing valuation method has proved not to be very useful in the case of cultural goods because it has been shown that the arts and cultural activities are by no means the most important determinants of either individuals' or firms' decisions to move to a particular area. Also, the particular cultural attraction would have to be very large and well known in order to enter into the decision making process at all (Bille and Schulze 2006). Although the work of Florida (2002) has done much to forge a link between the "creative classes" and economic growth and development, causality has not been conclusively proven. As a result, the hedonic pricing method has not been much used in cultural good valuation.

Contingent valuation (CV) is by far the most popular valuation method, both in environmental economics (where it started) and in cultural economics, because it is currently the only method of measuring non-use values. The National Oceanic and Atmospheric Administration (NOAA) review on the use of CV studies published in the Federal Register in 1993, states that,

> "This [CV] approach allows the valuation of a wider variety of non-market goods and services than is possible with any of the aforementioned techniques (site-specific valuation methods such as travel cost, factor income approach, or hedonic pricing models) and is the only

[1] Both willingness to pay and choice experiments (also referred to as conjoint analysis) are part of contingent valuation or stated preference methods, but the term "contingent valuation" is often used to refer to willingness to pay studies, while "stated preference" often means choice experiments.

> method currently available for the express purpose of estimating passive use values" (NOAA 1993).

Passive use values (also referred to as non-use values) are defined by Carson et al. (1999:100) as "those portions of total value (measured by WTP or WTA) that are unobtainable using indirect measurement techniques which rely on observed market behaviour". Thus, while methods like travel cost and hedonic pricing will capture the non-market value of the good to users, this does not include people who, for whatever reason, might be non-users, but who are still willing to pay to preserve or support the public good. In other words, passive use values require no direct involvement of the user with the good at all and "as a result, economists are fond of saying that passive use leaves no behavioural trace" (Carson et al. 2003:258). Diamond and Hausman (1993) outline three types of non-use values: the value of one's own possible future use of the good, the value of one's enjoyment of the use of the good by others (also called bequest value) and finally, values unrelated to human use of the good.

A number of models can be used to estimate willingness to pay (WTP) or willingness to accept (WTA) functions (Johansson 1993; Carson et al. 1999), but the general theory can be presented as follows: if initial utility (U_0) is a function of certain levels of income, prices, private goods and public goods, and an increase in the amount of public good supplied increases utility (U_1), then WTP represents (or exactly offsets) the difference between U_0 and U_1, such that the final level of utility is unchanged. In the WTA case, the decline in utility is exactly equal to the increase in utility due to the compensation amount.

The rest of this chapter discusses the use of WTP studies in cultural economics and some of the problems associated with the method.

4.1 Examples of WTP Studies in Cultural Economics

Willingness to pay studies have been used in the valuation of a wide variety of cultural goods. The appendix table to this chapter gives some examples of studies that have been conducted on the arts, heritage, historical and archaeological sites, museums, libraries, broadcasting and theatre. While it is not possible to discuss all these case studies here, the following

section gives a brief summary of some them and others are used as examples in later methodological discussions.

WTP Studies of the Arts and Theatre

One of the earliest arts WTP studies was conducted by Thompson, Throsby and Withers (1983) in Australia. The survey asked a random sample of 827 Sydney residents detailed questions about the arts, including their interests, attendance, participation, perception of benefits or costs, views on public funding and WTP to support the arts under various conditions. Eighty per cent of respondents were interested in at least one of the art forms mentioned and there was wide-spread agreement with the idea that the arts provided "community public benefits", such as national pride, assistance in understanding and interpreting "our country and its culture", as well as general educational value (Throsby and Withers 1985:591). Just under three-quarters of respondents voted for increased arts subsidy at a significantly higher level than government support at the time. The authors conclude that "The notion of the arts as a luxury and as only an elite pleasure foisted on an unknowing or resentful public is simply wrong" (Throsby and Withers 1985:594).

Another early WTP study on the arts was conducted by Morrison and West (1986) in Ontario, consisting of a sample of 463 households, using telephone interviews. In particular, they wished to investigate whether increasing arts subsidies would benefit mainly the wealthy arts attenders, or whether what they refer to as the "external" benefits from the performing arts, justified tax spending even for non-attenders. The Morrison and West (1986:68) survey posed the following question to the relevant portion of their sample:

> "Since you personally do not go to live performances of drama, dance, classical music or opera, what benefits, if any, do you feel you get from these things in return for paying taxes towards them?"

Responses (not suggested by the interviewer) covered many of those proposed by the theory, such as national pride, welfare to future generations and educational importance. However, 40% of respondents claimed that they gained no benefit. It should be noted that, unlike the Thompson et al. (1983) study, which found that public arts funding in Australia was significantly below what the public was willing to support, Morrison and

West (1986:69) found that the majority of Canadians were not in support of an increase in arts funding, suggesting that, "all external benefits have already been internalised or, in other words, that no marginally relevant externalities remain".

Studies which combined willingness to pay with economic impact methods were conducted by Thompson et al. (1998 and 2002) on the impact of the arts on the Kentucky economy. In addition to money spent on tickets, Kentucky households stated that they were willing to pay $21.8 million (1998) and $16.9 million (2002) in order to avoid a 25% decline in the number of arts performances in Kentucky. "The value of these donations [WTP] give a minimum estimate of the difference between the value that Kentuckians place on the arts and the amount they pay for arts performances through ticket prices" (Thompson et al. 1998:7), and thus provide an indication of the monetary value of the public good benefits provided by the arts. The economic impact of the arts in Kentucky was calculated to be $22 million and provided 1 324 jobs (Thompson et al. 1998:18). By including the contingent valuation measure of the value of the arts as a public good, the estimated value of the arts to the Kentucky economy was found to be almost double what it would have been if only the economic impact figures had been used (Thompson et al. 1998:3).

An often-cited study by Bille Hanson (1997) estimated the value of the Royal Theatre in Copenhagen to a random sample of the Danish population using 1843 telephone interviews. Although only about 7% of the population actually use the Theatre, non-use values were found to be quite substantial, with a median WTP of 60 DKK and a mean value of 154 DKK. She concludes that "the citizenry's WTP is of a magnitude which at the least measures up to the public subsidy which is given to the theatre" (1997:18).

Glass et al. (1999) surveyed 515 households in Kansas asking a dichotomous choice WTP question for "a substantial increase in the amount of arts activity in your local area" (Glass et al. 1999:28) to be paid for by an increase in tax. It was found that Kansas households would be WTP $19 million in increased taxes for the rather vaguely defined "substantial increase" and that, as in other studies, participants in arts events or education programmes were more likely to favour the increase. In addition to this, an economic impact study found that expenditure by the Kansas Arts Commission, which provides grants to arts organizations, resulted in an impact of just over $1 million in the region.

In a developing country example, Snowball (2005) conducted a WTP study of two South African arts festivals using telephone interviews of the local populations of the towns in which the festivals take place. In a country with wealth still partly divided along racial lines, a major point of investigation was whether low income African-origin residents perceived the arts festivals as providing as much value as the high income, mostly European origin residents. Results show that, despite lower average attendance at festival events, a higher percentage of respondents from the low income areas were willing to pay to prevent the festivals from getting smaller, although their average WTP amount was lower.

Museums

Sanz et al. (2003) conducted a survey of 1108 visitors to the National Museum of Sculpture in Valladolid, Spain, to find out if they would be willing to contribute to a preservation fund for the museum. They also conducted telephone interviews with 1041 Valladolid households to determine non-use values for the museum. The mean WTP of direct users was between 25 and 40 , and passive use, option, bequest and existence value (i.e. non-use value) was between 27 and 36 .

Maddison and Foster (2003) asked 400 visitors to the British Museum whether they would be willing to pay to reduce congestion (indicated by photographs showing varying numbers of people in the museum). They find that there is a "congestion externality of £8.05 per person" (2003:186) and suggest that, rather than the current free entry, a charge should be made at busy periods. This is particularly feasible, since most visitors are foreign with high household incomes and only 1% of visitors interviewed said that they came to the Museum because it was free.

Historical Sites and Heritage

There have been a large number of willingness to pay studies conducted on historical sites and heritage in general. For example, Willis (1994) used a payment card (listing various willingness to pay amounts) to elicit visitor's WTP an entrance fee to Durham Cathedral, as opposed to a voluntary contribution payment method being used at the time. However, he found that the revenue from voluntary contributions was actually slightly larger than the revenue that could have been collected by charging one price based on

average WTP. Using photographs to illustrate various scenrios, Pollicino and Maddison (2001) asked a random sample of residents of Lincoln whether they would be willing to pay to decrease the time between cleaning their Cathedral from 40 to 10 years. Although Lincoln Cathedral is acknowledged to be one of the most beautiful cathedrals in the UK, it suffers significant damage due to air pollution. They found an average household WTP of between £15 and £23 and then calculated the annual damage to the Cathedral as a result of air pollution to be between £0.4 million and £0.6 million.

Examples of other heritage or historical sites that have been valued using WTP studies are Nidaros Cathedral in Norway (Navrud and Strand 1992), the Fez Medina in Morocco (Carson et al. 1997), Warkworth Castle in England (Willis and Powe 1998), Rivoli castle in Italy (Scarpa et al. 1998), the Surrey Archival History Centre in the UK (Ozdemiroglu and Mourato 2001), the historic centre of the town of Noto in Sicily (Cuccia and Signorello 2002), Bulgarian Monasteries (Mourato et al. 2002), historical shipwrecks off the coast of North Carolina (Whitehead and Finney 2003), the restoration of an old Arab pirate tower in Valencia (Del Saz Salazar and Marques 2005) and many others.

A recent interesting advance is the use of WTP studies to place a value on cultural heritage in developing countries, where resources for heritage preservation are typically limited. Kim et al. (2007) use a WTP study to value the Changdeok Palace, a World Cultural Heritage site in Korea. Interviews were conducted with tourists visiting the site, who were given information about the historical and cultural value of the Palace and then asked their WTP to improve services (tourist guides), research and upkeep of the site. They find that the mean WTP of between $5.70 and $6 is 2.5 times higher than the current admission price. Dutta et al. (2007) report on a similar study of Prinsep Ghat in Calcutta, India. A random sample of 181 Calcutta residents was interviewed and asked whether they would be willing to pay to renovate the site and develop it as a tourist attraction. They find that the cost of renovation is far below the WTP estimate and conclude that using market figures (i.e. profits from tourism) to value historic sites significantly underestimates their total economic value.

Archaeological sites

Related to studies of heritage sites and historical buildings are those valuing aspects of archaeological sites. For example, Ringanti et al. (1998) interviewed visitors to the Campi Flegrei Archaeological Park in Naples to determine their willingness to pay to conserve parts of, and the whole, site. Maddison and Mourato (1999) used a WTP study to calculate the welfare effect of removing the roads around Stonehenge (placing one in an underground tunnel and closing another). Interestingly, while a greater percentage of those surveyed chose the current scenario, because there were significant positive externalities to drivers on current roads being able to see the sight, the average WTP to provide the tunnel was higher, suggesting greater intensity of preference in the latter case. Mourato et al. (2004) interviewed visitors to the site of Machu Picchu historic sanctuary in Peru to determine their maximum WTP for access to the site. They then show how such data can be used to estimate a demand curve and to analyse the likely effects of various proposed entry fee changes.

Broadcasting

In one of the very first WTP studies, Bohm (1972) conducted some experiments on WTP for a half-hour popular TV programme in Sweden. While the good itself was not very interesting, experimentation with subsamples on the effects of incentives to over- and understate WTP was the foundation of the "Bohm interval" method of detecting hypothetical bias (further discussed later in this chapter). Papandrea (1999) investigated WTP for mandatory local television programming in Australia. The majority of those interviewed were positive about the benefits of Australian programming, but only 12% of the sample thought it should be increased, while 65% felt it should stay at current levels.

Libraries

Willingness to pay studies of libraries have demonstrated their public and private good characteristics – that is, that it is not necessarily only those who use the library directly who benefit. Harless and Allen (1999) surveyed members of the Virginia Commonwealth University to establish their WTP for extended reference desk services. Mean values for 18.5 ad-

ditional service hours (excluding 10% of the highest and lowest values) were $5.59 per semester for students and $45.76 for faculty. In this case, the benefits exceeded the cost of providing the additional service by a ratio of 3.5 to 1. Holt et al. (1999) conducted a survey of the St Louis Public Library to determine how much users would be willing to pay to avoid library closure and, in a separate sub-sample, how much they would be willing to accept in compensation for library closure. 88% of respondents were unwilling to accept any reasonable level of compensation for library closure, indicating significant non-market values.

Finally, Aabo (2005) attempted to measure both use and non-use values of public libraries in Norway. Findings showed that the benefits are four times higher than the costs. In addition, only 3% of the sample indicated that closing down their local library in order to spend the money on some other local project would increase their welfare.

4.2 The WTP Method and the Exxon Controversy

Contingent valuation surveys were, until about 1990, a little-studied branch of environmental economics. However, on the 24th of March 1989, the Exxon Valdez oil tanker, misjudging its position, crashed into Bligh Reef in Prince William Sound off the coast of Alaska (Carson et al. 2003). The tanker released 11 million gallons of crude oil into the Sound, hugely damaging what had been one of the most environmentally pristine areas on earth. The local newspaper, *Anchorage Daily News* (1999), in a retrospective of the oil spill ten years after the event, termed the accident "one of the worst human-caused environmental disasters in history". The newspaper also reports the slowness of the response to the disaster as a result of not having disaster management programmes in place.

However, all this was to change when it became known that the State of Alaska and the Federal Government intended to use a contingent valuation study to litigate against the oil company, arguing that they were liable for more than just the clean-up costs and lost income, but also for the non-market value of the huge environmental damage caused to the Sound. As a result of this, "the attention paid to the conceptual underpinnings and estimation techniques for passive use value changed rather abruptly" (Carson et al. 2003).

Early in 1990, Carson et al. (2003) were commissioned to run a huge WTP study to determine the value to the American public of the Prince William Sound, the details of which have recently been published in *Environmental and Resource Economics* (2003). After being given a large amount of information, including maps and photographs, respondents were asked to vote for or against a proposal to provide an escort ship programme for oil tankers entering the sound. The results of the survey showed that American households were willing to pay about $2.8 billion for the programme. Although the case was settled out of court, the State Attorney General was reported as saying that the CV study "was the heavy artillery in our briefcase" (*Anchorage Daily News* 1993).

In an attempt to discredit the method, the oil company appears to have funded quite a lot of research into the method. Many of these papers were presented at a conference organized by Cambridge Economics, Inc., held in Washington in 1992 and were published with transcriptions of discussions (Hausman (editor) 1993). Although there are few direct references to it, traces of a violent debate and doubt over the bias of Exxon sponsored research remain. For example, in a 1996 article, Diamond adds a footnote to the effect that, while earlier research may have been funded by the Exxon company, the current article (also on CV criticism) was not. Carson et al. (1993:258-9) comment that "Much of the recent criticism of CV is contained in the Exxon-sponsored conference volume …" At the conference itself, discussion was often heated, with one non-economist referring to the process during a discussion as "a milieu of duelling economists" (Hausman 1993:458). Despite this, however, the papers presented at this conference set the agenda for contingent valuation research and many of the issues raised are still being debated.

In their 2003 article, Carson et al. conclude that the out of court settlement between the State of Alaska and the Exxon Company approximated the CV value quite well. They also note that, since 1989, there have been no major oil spills in American waters, perhaps suggesting that the inclusion of non-use values in possible litigation, while not uncontested, is still acting as a significant deterrent to oil companies taking risks that may result in pollution.

One of the results of this heated debate was the National Oceanic and Atmospheric Administration review on the use of CV studies (NOAA 1993), chaired by, amongst others, Kenneth Arrow and Robert Solow. The report concluded that, if its guidelines and recommendations were followed, "CV studies can produce estimates reliable enough to be the starting point of a

judicial process of damage assessment, including lost passive-use values" (NOAA 1993:24). Another result was that, although CV studies have now been applied in many fields, most of the methodological research has been conducted in the field of environmental economics.

4.3 Criticisms and Defence of the CV Method

Ever since the Exxon case in the early 1990s, there has been a storm of sometimes quite vehement criticism of the CV method (mostly represented by willingness to pay studies) and equally heated defence. The following section reviews some of the major areas of criticism, focusing on the nature of hypothetical markets and the ways in which CV responses have violated the assumptions of neoclassical economic theory.

4.3.1 Hypothetical Markets and the "Free Rider" Problem

One of the first questions many critics of the willingness to pay method ask is how one can be sure that respondents will tell the truth. The NOAA panel (1993:7) raised the issue of "implausibly large responses" and it has since been found in many studies that hypothetical markets tend to overestimate WTP. The current debate has largely shifted from whether or not hypothetical bias, defined as the difference between hypothetical and real WTP, exists to a discussion on how the problem can be detected and controlled. There are two major directions in this field. Firstly, there are those who, by comparing real and hypothetical valuations for the same good, seek to find some method of calibration for hypothetical responses that will bring them more in line with real responses. Secondly, there is ongoing research into how questionnaire design can be used to moderate this form of bias.

There are at least two possible reasons that respondents may overstate their willingness to pay for a good in a hypothetical scenario. Firstly, they may be "free riders" – that is, they are deliberately over-stating their true value of the good (knowing they won't really have to pay the amount they nominate) in order to ensure that it is provided. Secondly, respondents may not be making a realistic market decision, taking into account their budgets. In

this case, they are expressing a positive attitude towards the good, rather than an accurate valuation. The following section provides a historical background to the hypothetical bias and related "free rider" problem in contingent valuation and reviews the current debates regarding calibration and questionnaire design.

The Bohm Test for Hypothetical Bias

The liability or free-rider problem in estimating the value to individuals of public or mixed goods was first formally acknowledged by Samuelson (1954). He pointed out that, so long as all goods were private and operated in a perfectly competitive market, there was no incentive for individuals to misrepresent their demand since they all aspired to their highest indifference curve within the constraints of a budget. However, for a public good, voting or signalling other than that of the market must be used: "Now it is in the selfish interest of each person to give false signals, to pretend to have less interest in a given collective consumption activity than they really do" (Samuelson 1954:388). This is the central problem with a social economy or a socially shared public good. By pretending that it is less valuable than it in fact is, any individual may hope to escape personal liability for their consumption of the good.

Samuelson (1954:389) added, however, that the failure of the market did not mean that there was no optimal solution to the provision of public goods:

> "Given sufficient knowledge, the optimal decisions can always be found by scanning over all the attainable states of the world and selecting the one which according to the postulated ethical welfare function is best. The solution exists; the problem is how to find it".

Some of the earliest and seminal research on the existence and control for free riding and hypothetical bias in WTP studies was conducted by Peter Bohm (1972). He argued that the theory of the free rider motive to conceal demand preferences had not been empirically tested and that other, perhaps stronger, motives could result in honest demand revelation. Bohm (1972) thus conducted a series of WTP studies on small groups of paid volunteers. The public good under discussion was a new TV comedy series. Each respondent was asked to state the highest admission fee that they would pay to see a half-hour programme. If the stated cost of showing the programme was smaller or equal to the amount that the group was prepared to pay, the

programme would be shown and each person would have to pay something.

The first group were told that, should the programme be shown, they would have to pay the amount they had stated; the second group would have to pay some percentage of the amount they stated; group three, a variable amount; group four, a flat rate; and group five, nothing. The results revealed that there was no significant difference of the stated WTP between any of the groups. This is particularly surprising between groups one and five, since conventional free rider theory would postulate that group one people would have an incentive to understate their WTP, while group five would have an incentive to overstate. A later non-hypothetical study for an actual public good (Bohm 1984) verified these results.

Bohm (1972:125) did not, however, rule out the possibility that differences in stated WTP existed when different liability methods were used, but suggested that the results of the experiment showed that these differences were not very large. Rather than using the conventional cheating strategies which economic theory postulates, he suggested that group one revealed a true WTP because, "people tend to regard their impact on total demand, however small it may be, as important" and that group five did likewise because "they feel morally obliged to do so".

A recent study, using students in a computer-based experiment by Murphy et al. (2005b), also suggests that the "free rider" phenomenon in WTP studies is not as severe as one might think. Using post-decision questions, such as how much the decisions of others affected the respondent's WTP, they conclude that "most respondents were neutral with respect to the likelihood that others would free ride and report that free riding had little influence on their own payment decision" (2005:336).

There is also a stream of recent research that argues that there is no evidence of hypothetical bias at all. In a laboratory experiment with 264 students divided into smaller groups, Vossler and McKee (2006) tested the difference in response to real and hypothetical willingness to pay situations. They found no evidence of hypothetical bias since "voting decisions do not vary systematically when payment is hypothetical versus when it is real". In a similar sort of experiment, Guzman and Kolstad (2006) argue that, in hypothetical markets, respondents are unlikely to incur the opportunity cost of gathering more information about the good in question. Since willingness to pay tends to increase as information about the good

increases, this suggests that contingent valuation estimated in hypothetical markets might actually be *downwardly* biased. Murphy et al. (2005a and 2005b) find that hypothetical bias does exist, but is much more prevalent at higher willingness to pay amounts.

Morrison and West (1986:63), in their study of the performing arts in Canada, agree that other motives may counteract the free rider problem, but argue that it is the very *unimportance* of individual answers that would ensure honesty; "The fact that respondents may perceive their views to have little or no weight is likely to induce true preferences and to reduce strategic bias rather than encourage false or inaccurate responses". They also argued that those who do not want to be part of the study could more easily refuse the interview than provide untrue answers.

What Bohm (1979) suggested was that economic motives to over- or understate demand and WTP may (or may not) be overcome by other, perhaps moral, motives, and that these motives may change depending on factors such as the issue under discussion and questionnaire design. Given that the motives for over- or understating WTP may be different in each situation, he proposed what came to be called "the Bohm interval method" as a way of verifying the results of the study.

Bohm (1979) argued that if two similar sample groups were given the same questionnaire, but differing liabilities, and that the directions of the likely misrepresentation of demand for the two groups were known, then they could act as controls for each other. This effect could be further enhanced because there may be unknown reasons for the participants to reveal their preferences accurately, as suggested above. For example, two population samples were asked to reveal their preferences for a particular public good project. If the project was to be carried out (i.e. the WTP for the project exceeded or was at least equal to its cost) then the first group would have to pay the amount that they stated (or some related sum), while the second group would be asked to pay only a nominal amount or nothing at all. In this case, there was a clear economic incentive for the first group to understate, and for the second group to overstate their WTP.

If the average WTP for both parties was not significantly different, then one could assume that no serious misrepresentation had taken place and that motives to reveal true preferences had dominated. Even if there was a (fairly small) difference between the WTP of the two groups, the responses of the first group could be regarded as a bottom limit and those of the sec-

ond group as a top limit - the true WTP occurring in the interval between the two. However, the larger the interval, the less accurate the study could be (Bohm 1979).

A number of studies since have used some version of the Bohm interval with largely similar findings to Bohm's original tests. The Thompson et al. (1983) study of the arts in Australia used the Bohm interval method to control for the free-rider problem, but differed from the Bohm method by asking both the liable and non-liable questions in the same sample. Respondents were asked the following two questions (amongst others) directly after each other:

> (a) What is the maximum you would want paid out of your taxes each year to support the arts at their current level, if your taxes were adjusted so that you would actually have to pay the amount you nominated?
>
> (b) Now, suppose that there would be no change in your total taxes. What is the maximum you would want paid out of your taxes each year to support the arts at their current level?
> (Throsby and Withers 1985:32)

In order to check that the question order was not significant, half the sample were asked these questions in the above order and the other half in the reverse order. In any event, no significant difference in response was noted. Where some difference between the liable and non-liable response was found, it was not as great as would have been expected. For example, 14.5% of the sample nominated a WTP of $61 - $100 per year when not liable which decreased to 13.4% when liable.

Throsby and Withers (1986) estimated that about 65% of respondents to their survey gave honest WTP answers, i.e. their WTP did not differ between the liability and non-liability questions. Of the remaining 35%, only one third were "strong" free riders (those who give a positive WTP under non-liability and zero WTP under liability), the rest being "weak" free riders (those who give non-zero responses, but lower responses under liability). A difficulty which Throsby and Withers (1985) considered was the hypothetical nature of the WTP questions, both liable and non-liable. This differs from both the Bohm experiments (1972 and 1984) in which some real payment had to be made.

The Morrison and West (1986:66) study followed the Bohm method more closely. Two separate samples were asked the following question: "Do you think that $3,35 a year in taxes is too little, too much or just right?". The first sample was asked the question on the understanding that their taxes would not change, while the second sample's taxes would increase or decrease accordingly. Again, no significant difference between the stated WTP of the two samples could be found.

A survey at the South African National Arts Festival (Snowball & Antrobus 2001) used the Bohm interval method in a very similar way to Throsby and Withers (1985) except that a once-off WTP amount of R5 was used (dichotomous choice yes/no response), rather than a bidding system asking for a maximum WTP amount. Respondents were told that, "The Festival costs between R8 and R10 million each year which is paid by the sponsors, mainly the Standard Bank. 2001 will be the last year that the Standard Bank will pay for the Festival. If no new sponsor is found, the Festival will end". They were then asked the following questions:

> "Supposing that there would be **no change in your monthly taxes**, would you be willing to allow R5 of your taxes each month to be spent on supporting the Festival?
>
> Now suppose that **you would actually have to pay** the R5 extra in taxes each month. Considering your monthly expenses, would you still be willing to pay R5 a month towards supporting the Festival?" (Snowball 2000:92).

Of the 80 telephone interviews conducted, 9.4% of the Grahamstown West sample, (largely populated by European-origin residents who have high income and education levels), and 20.6% of the Grahamstown East sample (which has a largely African-origin, low income and education population) responded "yes" and then "no" or "don't know" to the non-liability and liability questions. This suggests that, on average, only 13.6% of the sample were free riders.

Table 4.1. Differences between Liable and Non-liable Willingness to Pay

Study	Result
Bohm (1972)	No significant difference
Morrison and West (1986)	No significant difference
Throsby and Withers (1985)	35% (12% "strong" free riders; 23% "weak" free riders)
Snowball and Antrobus (2001)	13.6% (9.4% in high income area; 20.6% in low income area)

However, in the Grahamstown study, not all those who answered "yes" to the non-liable questions and "no" to the liable question can automatically be regarded as free riders, because of the extremely low levels of income recorded (less than R500 a month) for some respondents in Grahamstown East. In these cases, it is quite possible that the refusal to pay extra is the result of honesty about an income constraint, rather than because of free rider behaviour. This would also explain why there appear to be more than twice as many free riders in the low-income, township area compared to the more affluent west side of town. This conclusion is supported by Booysen's (2001:695) article on the major reason for the non-payment for services in South Africa. He concludes that, "the problem of non-payment of services appears to be a problem of ability-to-pay, rather than willingness-to-pay."

Although there have been few studies that have used the Bohm interval as such, his experimental work gave rise to two distinct streams in hypothetical bias research. One stream consists of a growing literature comparing real with hypothetical values in an attempt to find some general rule of calibration, in other words, the ratio of mean hypothetical to mean actual WTP (List and Gallet 2001; Murphy et al. 2005a). Studies range in size and type from tiny samples conducted in laboratory conditions (Botelho and Pinto 2002) to larger mail surveys (Nestor 1998) and face-to-face interviews (List and Shogren 1998).

Experiments with Calibration between Real and Hypothetical Bids

A study in cultural economics comparing real and hypothetical bids was conducted by Willis (1998) who asked visitors to an historical site, Warkworth Castle, who seemed to have decided not to enter the site when they discovered the ticket price, what their WTP for a ticket would be. The tickets were then offered to respondents at their stated WTP. Willis argued that the use of a private good to test the validity of the WTP method, usually used for public or mixed goods, is justified because, if the WTP method cannot closely approximate the real value of private goods (which are often easier to define and better known) then it is unlikely that the method will be more accurate for public goods.

Willis (1998) found that, of the 43 respondents who considered themselves potential visitors, only 17 accepted the tickets to the site at their stated

WTP price. Responses to questions about why respondents would not accept the tickets at their nominated price were not clear. This seems to suggest that, "a substantial proportion of CVM [contingent valuation method] values are neither very robust nor reliable and are subject to considerable ambivalence" (Willis 1998:616).

List and Shogren (1998) conducted first a hypothetical and then a real within-sample WTP study at a collectable baseball card auction. Respondents were first asked to bid hypothetically for a card or a group of cards and the sealed bids placed in a voting-style box without the interviewer seeing the amount. They were then asked to bid in reality for the cards, the highest bidder to be notified by telephone and, on sending the money, receiving the cards. The results showed a ratio of hypothetical-to-actual overbidding of between 2.2 to 3.5, suggesting that the NOAA Panel report proposed "divide by two" rule might be justified (List and Shogren 1998:203). However, they also found that hypothetical bias was highly context specific (so that adding or subtracting substitute cards from the set being bid on changed calibration significantly) and that overstatement was significantly higher for non-dealer respondents who were less familiar with the good and its usual market price.

Nestor (1998) argues that there are great advantages in combining hypothetical or stated preference data with real market or revealed preference data, such as reducing multicollinearity between site or good characteristics and extending the data analysis beyond the range available, using one sort of study only. Using a study on waste disposal options, where some households were already paying an optional tariff for sorted recycling, while others did not yet have the option available to them, she showed that, although hypothetical bias was present, it was not of a very high magnitude. While results were "mixed", she did not find, in contrast to List and Shogren (1998), that experience with the "good" being valued (in other words, more information) had any effect on WTP amounts or hypothetical bias.

Whitehead et al. (2000) conducted a similar analysis, combining revealed preferences from travel cost methods with stated preference data for quality improvements in outdoor recreational facilities. They also found that, by including stated preference techniques, they are able to comment not only on the value of such quality improvements to current users, but also to potential users. Botelho and Pinto (2002) conducted a laboratory experiment comparing real and hypothetical WTP for the provision of an information leaflet on the otter. While hypothetical responses were higher

than real responses in a split-sample survey, the hypothetical data could be used to predict, given the demographic characteristics of the respondent, "what they would have responded if they had been placed in the real treatment" (Botelho and Pinto 2002:995). They thus conclude that, while hypothetical bias is present, the valuations of the hypothetical survey still convey useful information, especially if they can be calibrated to real data.

There is thus growing evidence that, although hypothetical bias probably exists, the data from such surveys can still be used to inform decisions. However, the search for a general calibration ratio does not appear promising for various reasons. Firstly, it is only really possible to conduct such tests where a real market situation for the good exists – mostly limited to private goods (Lusk 2003). Carson (1997) argues that, for some classes of goods, in particular, public goods to be provided by voluntary contributions and private goods to be purchased, it is impossible to design a survey in which the dominant incentive strategy will be to tell the truth because, "the survey response must provide the possibility of strictly altering the respondent's choice set without strictly expanding or contracting it …" On the other hand, WTP surveys on public goods provided through the use of coercive payment mechanisms and quasi-public goods appear to equal or underestimate real market valuations. "Thus, a single private good, rather than representing the "best" case for a successful CV, represents one of the worst cases" (Carson 1997:1503)

The second problem with calibration experiments is that they appear to be good and context specific. A meta-analysis of 29 real and hypothetical WTP calibration experiments, covering a wide range of goods, methodologies and elicitation techniques, comes to this conclusion, as well as arguing that, in general, calibration is relatively small: 1.26 to 1.30 (List and Gallet 2001). The study also finds evidence, as in the baseball card example, that the more subjects know about the good, usually a private good, the smaller will be the hypothetical bias. A related study, a recent meta-analysis of 28 WTP studies by Murphy et al. (2005a) found a calibration factor (median) of only about 1.35, with 70% of their 83 observations having a factor below 2. However, they note that the sample shows significant positive skewness, with the calibration factor for the top ten observations being a mean of 10.3, compared to 1.54 for the rest. Research in this area thus seems to be moving toward the NOAA panel's conclusion that "No automatic or mechanical calibration of responses seems to be possible" (NOAA 1993:25).

"Cheap Talk" Questionnaire Design

The other proposed way of controlling for the hypothetical bias is through questionnaire design. Research in this area has mostly focused on the "cheap talk" design first suggested by Cummings and Taylor (1999). They experimented with questionnaire design that explicitly made respondents aware of hypothetical bias before asking them the WTP questions.

> "The cheap talk script makes three general points: it describes the hypothetical bias phenomena; it discuses possible explanations for this phenomena; and it requests that subjects vote in the upcoming hypothetical referendum as if it were a real referendum" (Cummings and Taylor 1999:651).

While the cheap talk design produced results for hypothetical surveys that were statistically indistinguishable from surveys that required some real payment, the authors admitted that the method had some problems. The most important problem is that the length of the cheap talk explanation is unrealistic for use in most surveys and particularly in telephone surveys. Tests using much shorter versions of the format were unsuccessful (Cummings and Taylor 1999:656).

However, Aadland and Caplan (2003) conducted a 1000 household WTP study of curbside recycling using a cheap talk questionnaire design significantly shorter than that of Cummings and Taylor (1999), but conveying similar information. They find that the design significantly reduces willingness to pay amounts and, by comparing WTP amounts with real market data, are able to conclude that it reduces hypothetical bias.

A WTP mail survey was conducted by Lusk (2003) for genetically modified "golden rice" using a longer cheap talk script for half the sample. He found that the cheap talk design did significantly reduce WTP bids, but that it was most effective amongst respondents who were inexperienced and had no knowledge of the subject (genetically modified foods) prior to the survey. For those with prior knowledge, the cheap talk design reduced WTP, but not by a statistically significant amount.

A study conducted by List (2001) using a cheap talk design in a split-sample field experiment at a collectable sports card auction found very similar results. While List found that the cheap talk design significantly reduced hypothetical bias amongst inexperienced bidders, it had no effect

on traders and experienced bidders. He comments that this result represents a challenge for future CV research design in eliminating hypothetical bias for all subject types, but that recognizing the importance of the link between "reliability of stated values and the experience with the good" is an important step forward (List 2001:1504-5).

In a recent study, Murphy et al. (2005b) conducted experiments which included both hypothetical and real WTP scenarios. They found that including a cheap talk section in a hypothetical WTP questionnaire reduced the number of positive responses, compared to the hypothetical situation without the cheap talk design. However, there was still a significant difference between hypothetical and real results, especially at the lower bid amounts ($3 or $6). At higher payment amounts, the cheap talk design did appear to be able to align the hypothetical and real bids.

Related to this is an innovative paper by Champ and Bishop (2001) who use both stated and revealed preference data from two large studies on the purchase of wind-generated electricity. They found that, by asking respondents how certain they were that their WTP responses reflected their true values, where 1 represented "very uncertain" and 10 "very certain", they could identify those respondents likely to be overbidding in the hypothetical market. Simply by recoding "yes" WTP responses from respondents who rated themselves less than 8 on the certainty scale to "no", they found that WTP results replicated those of real market data for the same good and statistically significant hypothetical bias vanished.

Despite the different directions that research into hypothetical bias in contingent valuation studies has taken, it does seem possible to draw some general conclusions. Firstly, the better informed respondents are, or the more familiar they are with the good, perhaps also reflected by their certainty that their WTP bids reflect their true preferences, the less likely it appears that they will be prone to overstating WTP. Secondly, it appears that the Diamond and Hausman (1994) claim that calibration could lie anywhere between 1 and 10 is overly pessimistic and that Bohm's original idea, that hypothetical bias in properly designed surveys is not likely to be large, stands. The crucial issue in controlling for hypothetical bias, therefore, appears to be the information that is provided to respondents (along with their personal previous experience of the good, where applicable) and the way in which the questionnaire is designed.

4.3.2 The Embedding Effect and "Warm Glow" Hypothesis

One of the major ways in which hypothetical bias is detected in WTP surveys of goods without direct market parallels is in inconsistency of the results with what economic theory predicts. Foremost amongst these is the insensitivity of willingness to pay bids to the scope or amount of the good being valued.

> "Usually, though not always, it is reasonable to suppose that more of something regarded as a good is better so long as an individual is not satiated. This is in general translated into a willingness to pay somewhat more for more of a good … Also, if marginal or incremental willingness to pay for additional amounts does decline with the amount already available, it is usually not reasonable to assume that it declines very abruptly" (NOAA 1993:6).

Initially, a number of WTP studies were cited as demonstrating such insensitivity to scope. For example, Kahneman (1986 cited in NOAA 1993:4) found that WTP for the cleanup of all the fishing lakes in Ontario was "only slightly more" than the WTP to clean the lakes in one region. A similar case is the now infamous "birds" study (Desvousges et al. 1993), which showed that the WTP to preserve 2000, 20 000, or 200 000 birds was the same. Diamond's later (1996) analysis in fact argues that, given certain Neoclassical assumptions about the utility function, WTP to save 100 000 birds, for example, must be at least 100 times bigger than the WTP to save 1000 birds in order for the study to pass the scope test.

Positive Attitudes and "Warm Glow"

The explanation for scope insensitivity and related phenomena put forward by critics of the WTP method, like Diamond and Hausman (1994) and Kahneman and Knetsch (1992), was that respondents were not really valuing the good in an economically rational way, but were merely expressing a positive attitude to the good in question, or "purchasing moral satisfaction" from giving to some worthy cause. This "warm glow" hypothesis argues that, since individuals are only showing support for some composite good (like the arts or the environment), there is no reason to suppose that their WTP would vary much with the amount of the specific good in question. The NOAA report (1993:8) notes that, if warm glow does occur, the WTP results "should not be taken as reliable estimates".

Diamond (1996) argues that, if the process of, for example, giving money to save 1000 birds, provides utility in addition to the actual birds saved (the good), then the WTP estimate is, at least in part, capturing some moral satisfaction or warm glow. In other words, 1000 birds are not a perfect substitute for 1000 birds *saved*. This could lead to a situation in which the government could, for example, propose the development of a wilderness area where no development is really intended, and increase welfare by announcing that the results of a WTP study to block the development had been successful. "In other words, if we are not willing to restrict analysis to preferences defined over the state of the environment, rather than also including the process resulting in that state, then we need a new welfare economics …" (Diamond 1996:346).

Interestingly, this ties in well with the work of Sen (1985) who argues that neoclassical utility theory and welfare economics are flawed in that they do not take into account the process by which the final outcome is reached or consider the role of the choice set available to the individual who made the choice. Sen would undoubtedly agree with Diamond that we need a new welfare economics, but in Sen's case it would be a theory that included utility derived from the process (taking into account the limitations of the other available choices) rather then excluding it. Carson et al. (2001:177) agree with this view, stating that, "It is utility, whatever its source, that matters to total value" and that the motivations behind increases in utility are irrelevant as far as economic theory is concerned.

Some of the most interesting theoretical work on CV has been done by Kahneman et al. (1999). They argue that attitudes are much more closely linked to WTP bids than the economic concept of preferences and that this poses a problem because attitudes violate many of the assumptions of rational choice theory. Preferences are assumed to exist already and imply a choice of one good over another that can be linked to a budget constraint and quantified. However, attitudes do not imply choice and are thus not easily fitted into choice theory. For example, preferring bagpipes to opera implies a choice of one over the other. Having a positive attitude to bagpipes does not imply a negative attitude towards opera. Diamond and Hausman (1993:27) comment that, "Our conclusion is that people do indeed care about preserving wilderness areas, but we infer that standard CV questionnaires do not generate a description of preferences but, rather, elicit responses that generally express concern about preserving wilderness". That is, they argue that CV responses capture attitudes ("general concern") not preferences.

In several experiments, Kahneman et al. (1999) find that attitudes, expressed on rating scales, are highly correlated with valuations expressed in dollar terms. For example, ratings on numerical scales were elicited in the case of a proposed intervention to protect the peregrine falcon being threatened by pollution. Measures used were willingness to pay, degree of political support for the proposed intervention, personal satisfaction from contributing to the scheme and rating of the importance of the problem as a public issue. All the rating systems showed a high degree of correlation (Kahneman et al. 1999).

One of the expected results of the theories of Kahneman et al. (1999) is that WTP studies will inevitably be insensitive to the size or amount of the good being valued because respondents will be using "judgment by prototype" to make decisions and to express an attitude, rather than focusing on the specific situation described. Using a number of examples and with reference to well-known psychological theories, they show that "judging by prototype" is inevitably prone to "extension neglect", that is, unless the attention of respondents is specifically drawn to issues of size or number, they will have little or no effect on the valuation of the good. For example, in the Desvousges et al. birds study, Kahneman et al. (1999:212-3) suggest that the scenario brings to mind a mental image of "an exhausted bird, its feathers soaked in black oil, unable to escape" and that this image is likely to dominate expressions of the respondent's attitude to the problem, including their willingness to pay for a solution, which will not be much affected by the *number* of birds saved. They apply a similar theory to the "add-up test" (or individual valuation and summation).

Adequate Responsiveness and Studies that do Show Scope Sensitivity

The issue of scope sensitivity is a vital one, since the NOAA report (1993:5) sets this measure of internal consistency or rationality as a minimum requirement of proof that the WTP responses "corresponded to some reality". As such, it falls into the panel's "burden of proof" subset, "… if a CV survey suffered from any of the following maladies, we would judge its finding unreliable" (NOAA 1993). However, the NOAA panel, after the initial report was published, noted that the "adequate responsiveness" mentioned in the report did not necessarily mean a statistically significant responsiveness to scope, which may depend crucially on the context and was

ultimately a "judgment call" (Arrow et al. 1994, reported in Smith and Osborne 1996:288).

While critics of the method argue that WTP studies will never be able to produce responses consistent with the economic theory of rational choice, however responsiveness to scope may be measured, this is hotly contested by a number of other commentators, whose main argument is usually that it is the survey design and/or method that is at fault.

Hanemann (1994) is the foremost of such defenders. In terms of the case studies often cited in the case for scope insensitivity, he argues that the Kahneman (1986) lakes study actually found a 50% difference in WTP for lake cleanup in one region as opposed to all of Ontario and that the survey itself was not well run, consisting of only a brief telephone interview using an open ended WTP question (1994). In the case of the "birds" study, he argues that the results *were* consistent with what economic theory would predict, given that changes in bird population were described to respondents as "much less than 1 percent" (2000), "less than 1 percent" (20 000) and "about 2 percent" (200 000). Presented in this manner, it would have been surprising if such tiny percentage changes in bird population size had elicited much difference in WTP. However, when this was pointed out to William Desvousges at the Exxon conference, he countered that focus group studies had shown that respondents understood the questions and that respondents were not faced with more than one scenario, so that the relative change in bird population size between scenarios may not have been significant (Desvousges (discussion session) in Hausman, 1993:161):

> "Now, what we were trying to do in the focus groups- keep in mind that we're not asking the same respondent 2000 versus 20 000 versus 200 000 – so I think that that is a difference here. What we did was to take the three versions and give a respondent one of those three versions".

Hanemann (1994) also argues that if one takes into account that birds or lakes may be regarded as substitutes for each other and that saving or cleaning them has a declining marginal utility, one would not expect WTP to increase proportionately with an increase in the good. Randall et al. (1981 in Dupont 2003) agree that sequencing can be understood in terms of economic theory, that is, as the list progresses, budget constraints become increasingly limiting and substitution effects set in.

In a follow up study by Schkade and Payne (1992, cited in Carson and Mitchell 1993), using the Desvousges et al. survey instrument, respondents were asked to "think aloud" as they made their WTP decision. While the

study revealed that respondents thought of a strikingly large number of irrelevant factors, indicating the difficulty of the task (as Diamond and Hausman (1994) point out), it also forced respondents to concentrate more and results showed a 100% difference in median WTP to preserve 2000 and 200 000 birds (Carson and Mitchell 1993).

Carson (1997:1505) supports this view, arguing that "specific, avoidable survey problems can mimic the impression of insensitivity to scope" and that these are mostly associated with studies using mall intercepts and short telephone interviews and which have vague descriptions of the good, the provision and/or the payment mechanism.

> "If there is a central problem with contingent valuation it is that people will try to answer whatever question is put to them. The quality of the response is crucially dependent on the information provided to and perceived by the respondent, and the seriousness with which the respondent takes the survey interview" (Mitchell and Carson 1993: 1266-7).

An example of a cultural economics study that did show scope sensitivity is the WTP to preserve the Campi Flegrei conducted by Ringanti and Scarpa (1998). Using face-to-face interviews, they asked respondents' WTP to preserve the Campi Flegrei, an ancient Roman archaeological site near Naples. The WTP questions were designed specifically to test whether bids were significantly different for the whole site versus parts of the site. The WTP to conserve the whole site was $28.81, while conserving only parts of the site had significantly lower WTP results: $10.18 for the part of the Campi Flegrei not open to the public and $6.99 for the Bagnoli. The authors conclude that, with proper questionnaire design, WTP bids are sensitive to the scope of the good being valued.

A growing number of WTP studies conducted in accordance with the NOAA panel guidelines are showing sensitivity to scope. Further research into the theoretical reasons behind those studies showing scope insensitivity is also advancing. Carson et al. (1996) conducted a study looking at four large-sample environmental quality willingness to pay studies, including the original Exxon study, done in 1991, and a follow up study in 1993, all of which conformed to the NOAA panel's guidelines. All studies showed sensitivity to scope, with studies valuing more inclusive public goods producing higher WTP estimates than those valuing "smaller" goods. In a later review (2001) Carson et al. reinforce their claim that the vast majority of modern CV studies do show sensitivity to scope, and that

those that don't generally did not follow NOAA guidelines on questionnaire design and survey administration.

Using a meta-analysis of five studies conducted on the WTP for improvements in visibility (air quality) at national parks, Smith and Osborne (1996:290) tested for scope sensitivity and the context-specific a priori expectations dictated by economic theory. They found that, "Regardless of the sample composition or model specification, there was a statistically significant, positive relationship between willingness to pay and proportionate improvement in the visibility range" (1996:295). They also found that when differences between studies in the economic commodity being valued are included, for example, the base visibility conditions, the amount of time the changes in visibility were to occur for, etc., the differences in WTP results were economically plausible.

Possible Reasons for Apparent Scope Insensitivity

On a more theoretical level, Bateman et al. (2004) show that scope sensitivity is significantly affected by whether the "visible choice set" is revealed to respondents before they are asked to bid, or not and that this, rather than the warm glow hypothesis, can account for the variation in scope sensitivity in WTP studies. When testing for scope within a single interview, the amount of the good is varied (quantitative nesting) or the type or quality of the good is changed (qualitative nesting) and respondents are asked their WTP for each scenario. In an experiment and field study, Bateman et al. (2004) found that, where respondents were told in advance what all the possible choices were, the studies were far more sensitive to scope. Furthermore, in the studies with visible choice sets, the order in which the good was varied, bottom-up or top-down, was not a significant factor – in contrast to many studies reporting the so-called sequencing effect, where the order of questions makes a vast difference to the WTP bid.

Bateman et al. (2004:89) offer some suggestions as to why removing the element of surprise should have such a dramatic effect on scope sensitivity. Firstly, they point out that they are not claiming to have solved the issue of strategic bidding, but that by revealing all options before a vote, they have held "strategy space constant" and achieved internal consistency, in that responses are sensitive to scope and do not vary with list direction. This result is as predicted by choice theory: "different strategy spaces result in different valuations". A second suggestion is that reported

sensitivity to list direction is as a result of the gains/losses asymmetry (also relevant to the difference between WTP and willingness to accept), which postulates that gains are worth less than losses because of loss aversion, also known as the "endowment effect". Finally, they suggest that the element of surprise can produce negative reactions in respondents, resulting in a lowering of stated values.

Related to this, Dupont (2003) found strong evidence that sequencing effects depend very much on how familiar respondents were with the good in question. In a WTP study used to value improvements to swimming, fishing and recreational boating in Hamilton Harbour in Ontario, she found that WTP bids for passive users and potentially active users were significantly affected by question ordering, while those for current users were not. She concludes that, "If respondents are not familiar with the environmental good/activity (passive users) or are uncertain as to the enjoyment they will obtain from it (potentially active users) then they may be more subject to order effects than active users" (Dupont 2003:336).

One of the few cultural economics WTP studies designed to test specifically for scope sensitivity was conducted by Whitehead and Finney (2003). In a study of the willingness to pay to protect submerged maritime cultural resources like shipwrecks, Whitehead and Finney (2003) find that WTP bids are insensitive to scope, but suggest that this is because of the way the WTP question was asked, rather than unrealistic valuations. Using telephone interviews of residents of North Carolina, respondents were asked their WTP to protect shipwrecks off the North Carolina coast from treasure hunters. In the first version, a single bounded WTP question was asked, for protecting 50 or 100 shipwrecks at a particular price, to which respondents could vote yes or no. If they voted yes, the number of shipwrecks protected and the price was increased and if they voted no, a lower number of shipwrecks and a lower price was suggested.

An interesting aspect of the study was that, no matter what their first vote (yes or no) was, respondents were reluctant to change in the second round. Whitehead and Finney (2003:323) suggest that this is because of "incentive incompatibility": If, after voting yes, respondents are offered the good at a higher price, they may assume that the higher priced good is a waste of money, and vote "no" in the second round, regardless of real willingness to pay. If, after voting no, respondents are offered the good at a lower price, they may assume that the good is of lower quality and still vote no in the second round.

Despite changing the amount of the good offered as well as the price, Whitehead and Finney (2003) found significant incentive incompatibility. That is, less that 10% of respondents voted "no" on the first question and "yes" for a lower price and quantity, while less than 20% voted "yes" for the first question and "no" for a higher price and quantity. As a result, willingness to pay did not show sensitivity to the size of the good being valued. They also suggest that an increase from 50 to 100 shipwrecks out of a total of 5000 (information that was supplied to respondents) represents an increase from 1% to 2%, which may not have been perceived as significant – similar to the Desvouges et al. "birds" study discussed earlier.

Although not the focus of their research, this sort of incentive incompatibility is also evident in the Alberini et al. (2005) study which uses a double-bounded willingness to pay question to place a value on the conservation of S. Erasmo Island in the Lagoon of Venice. As the table below shows, only 5.95% of respondents changed a "no" vote to "yes" in the second round (offered at a lower price) and 17.17% changed a "yes" vote to "no". The evidence thus suggests that, in some studies at least, apparent scope insensitivity may in fact be as a result of questionnaire design issues.

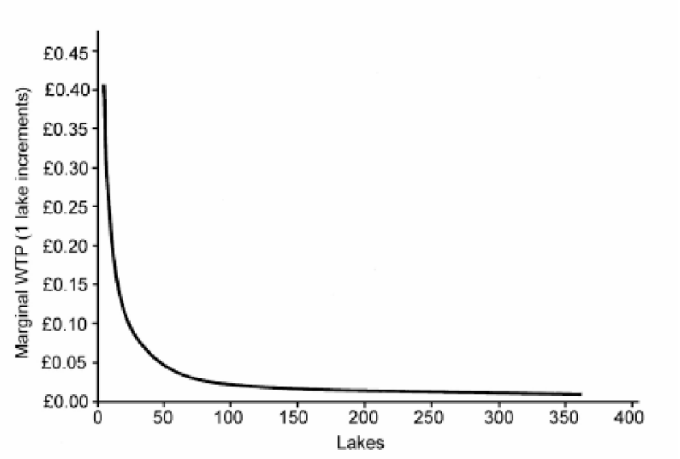
Table 4.2. Double-bounded dichotomous choice responses in two cultural goods studies

WTP	Whitehead & Finney: % WTP to protect 50 wrecks	Whitehead & Finney: % WTP to protect 100 wrecks	Alberini et al. % WTP to conserve S. Erasmo Island
No-no	46.14	46.40	53.54
No-yes	6.59	7.88	5.95
Yes-yes	28.64	29.05	23.34
Yes-no	18.64	16.67	17.17

(Whitehead and Finney 2003:235; Alberini et al. 2005:165)

Further evidence of this is found in the Bateman and Brouwer (2006) survey conducted to measure the WTP to reduce the risk of skin cancer in New Zealand. They specifically tested for scope sensitivity by first eliciting individual WTP and then broadening this to household WTP. They also tested question formats: open-ended (OE) versus dichotomous choice (DC). They find that the DC format WTP question shows statistically significant sensitivity to scope of a magnitude that conforms to *a priori* expectations, while the OE format question does not. "Consequently, we cannot reject the hypothesis that the DC responses are consistent with standard theory" (2006:210).

Another general finding that appears to be emerging is that WTP studies are not sensitive to changes in scope when these changes are small or seen to be too small to matter. A study by Bateman et al. (2005) measuring the WTP to change the acidity level in a large sample of mountain lakes illustrates this point. They find that respondents' WTP did vary depending on the number of lakes being valued, but decreased significantly once a threshold level had been reached. Thus, while willingness to pay to protect five lakes was estimated to be between £16.39 and £20.29, the maximum values for protecting all 400 lakes were only between £22.40 and £30.18. Bateman et al. (2005) argue that this finding does not indicate scope insensitivity, but rather that, once the threshold has been reached, marginal WTP to protect additional lakes is very small.



(Source: Bateman et al. 2005:289)

Fig. 4.1. Marginal WTP for Lakes Protected from Further Degradation

Sensitivity to scope will continue to be the acid test of internal consistency in WTP studies. Kahneman et al. (1999:217) conclude that, "Insensitivity to scope is the inevitable result of general rules that govern human judgment. It is naïve to expect broad psychological laws to be overcome by minor methodological adjustments". However, the general consensus seems to be that insensitivity to scope is as a result of poor survey design, rather than proof that contingent valuation itself does not conform to economic theory.

4.3.3 Disparities Between WTP and WTP

Another important way in which contingent valuation surveys violate the standard assumptions of economic theory is in the large difference in the outcome of the survey depending on whether respondents are asked their willingness to pay to avoid a decline in some good or service or if they are asked their willingness to accept (WTA) compensation for the same proposed decline. The Expected Utility Hypothesis, put forward by Willig in 1976 and extended by Randall and Stoll in 1980, implies that, when income effects are small (which most studies suggest they are), then differences between WTP and WTA measures will be small (Inder and O'Brien 2003; Hanemann 1991; Boyce and Brown 1992:1367).

In other words, the theory suggests that, when goods are sold in a competitive market with no transactions costs, they take on the properties of money. In this case, not only can WTP be shown to be equal to WTA, but these can also be shown to approximate the average market price of the good, with straight line indifference curves resulting (Shogren et al. 1994). However, as several authors have noted (Boyce and Brown 1992; Morrison 1997; Hanemann 1991; Shogren et al. 1994) this has not proved to be the case - WTA usually exceeds WTP by a factor of anything from two to ten and sometimes more than this.

Given that income effects have usually been shown to be small, two distinct theories have emerged as possible explanations for the WTP-WTA disparity, namely the substitution effect first put forward by Hanemann (1991) and the loss aversion or so-called endowment effect, proposed by Kahneman et al. (1990). The following section reviews these two theories and their possible combination (Morrison 1997a and 1997b) as well as some other later suggestions.

The Substitution Effect

Hanemann (1991) argues that the implications of the Randall and Stoll model have been misunderstood. He points out that, for changes in the quantity of a good, there is no presumption that WTP must equal WTA, especially if the good has no close substitutes. He develops a model showing that, if income effects are held constant, WTP and WTA measures will become increasingly divergent as the number of substitute goods is re-

duced. “In the limit, WTP could equal the individual’s entire (finite) income, while WTA could be infinite” (1991:625-6). Shogren et al. (1994) use the example of the trade-off between health and wealth to illustrate this, showing that, with no perfect substitutes, indifference curves will take the usual shape, which is convex to the origin. This argument is supported by Amiran and Hagen (2003) who show that, if utility is asymptotically bounded, full compensation for the loss of a public good, which has no perfect or close substitutes, may be impossible because the extra income from compensation (WTA) cannot purchase an equivalent good. It is thus possible for WTA to be infinite without violating any fundamental neoclassical assumptions.

Adamowicz et al. (1993) used two private good experiments to test for the presence of the substitution effect in WTP-WTA disparity. The first experiment obtained WTP and WTA responses for a movie ticket, where the film was also available on video. The second regarded tickets to a live hockey game that would also be shown on TV. The sample was split in each case, one half being informed of the substitutes (the video and TV coverage) the other, not. While the movie/video experiment (using open-ended questions) showed no sensitivity to the presence of substitutes, the hockey experiment (using close-ended questions) did. “Using the mean value welfare measure, the difference between the WTP and WTA was $13.40 (40 percent) smaller for the substitute sub-sample than for the no-substitute subsample” (1993:425). The authors conclude that the availability of substitutes does have a significant effect on the difference between WTP and WTA.

Brown (1994) however, disputes the findings of Adamowicz et al. (1993). He points out that, while both WTP and WTA declined in absolute terms when substitutes were offered (as one would expect), the ratio of WTP to WTA did not change much at all, the WTA/WTP ratio being about 1.86 with or without the substitute.

An interesting test for the effect of substitute goods in cultural economics (and one of the few studies of a public good) was undertaken by Kling et al. (2004). Their study investigated the value of preserving the Northern Hotel in Fort Collins, Colorado. The hotel is a local landmark, originally built in 1866 and host to a number of famous personalities in the past, including President Franklin Roosevelt. However, the hotel had deteriorated seriously and, if not restored, might have become irreparable. Kling et al. (2004) set out to examine the WTP-WTA disparity specifically by splitting the sample between two valuation scenarios: one in which respondents

were offered a refund of tax revenue surplus or the building restoration (WTA), and one in which building restoration was coupled with an added tax payment (WTP). They found that WTA bids were significantly higher than WTP amounts and concluded that, in the case of such cultural heritage, there is a significant non-substitutability effect, since there are no close substitutes for unique heritage.

The same seems to apply to a study of the St Louis Public Library (Holt et al. 1998, 1999), which asked library users how much they would be willing to accept to give up some library services and also their willingness to pay ("dollar value") for these services. The findings show that 88% of respondents refused to consider closing the library "at any price", with WTA bids beyond any reasonable level and far exceeding WTP - $136 million (WTA) compared to $15 million. Reasons included that the library was "Important / needed / priceless / too valuable" (Holt and Elliott 1998:43).

The Endowment Effect

The other main argument put forward to explain why WTP and WTA can differ so much is a form of loss aversion or the "endowment effect". In a number of experiments, Kahneman et al. (1990) found that initial ownership, or endowment, affected significantly what subjects would be willing to accept to part with the good. For example, in one of the experiments, one group of subjects was given redeemable tokens to trade while another was given consumer goods (coffee mugs or chocolate bars) and told that, if they could arrive at a mutually agreeable price, they could trade. The expected volume of trade (V*) was about 50% in both markets, in other words, with randomly assigned goods or tokens, about half the goods would change hands. However, in the token market, where tokens had no value, other than that imputed by the researcher, the ratio of actual trade (V) to expected trade was 0.91, while in the consumption good market, V/V* was only 0.31. This pattern persisted even when a number of iterations of the market simulation were run with the same subjects. All eight experiments produced robust results, implying that "the observed undertrading of consumption goods may be largely due to a reluctance to part with entitlements" (Kahneman et al. 1990:1339).

The result contradicts the Coase theorem, which argues that, where there are no transactions costs, the allocation of resources amongst bargaining individuals should not be affected by initial property rights or endow-

ments. “However, if the marginal rate of substitution between one good and another is affected by endowment, then the individual who is assigned the property right will be more likely to retain it” (Kahneman et al. 1990:1340). Kahneman et al. (1990) conclude that the market asymmetry observed in the WTP-WTA disparity is thus not likely to be some kind of cognitive mistake that will disappear with the use of a different survey method or repeat iterations, but that theorists need to accept that utility depends strongly on starting point. In a later study Thampapillia (2000:510) illustrates that “the shape of an individual’s underlying utility function, and hence the indifference curves, changes when an individual’s endowment changes”.

Thompson et al. (1998 and 2002), in their WTP studies of the value of the arts to Kentucky households, also recognise “loss aversion”, in that mean WTP per household to avoid a 25% loss was found to be significantly more ($11.44) than the WTP to increase arts performances and exhibitions by the same amount ($6.21). Considering that the costs of increasing arts performances and exhibitions in Kentucky by 25% would cost $12.6 million, a WTP to avoid this loss of $16.9 million indicates a net benefit of $4.3 million. Thompson et al. conclude that the relatively small net benefit indicates that current levels of support are relatively efficient.

Aabo (2005) used a split sample to value public libraries in Norway. One sub-sample was asked their WTA compensation in the form of a tax rebate for closing the library down and the other was asked their WTP to prevent library closure. WTA estimates were found to be five times higher than WTP. However, Aabo argues that, where respondents feel that they have “an inherent right to the good, WTP measures will underestimate welfare impacts” (2005:492). In this study, it was found that 94% of Norwegians felt that they had such rights and Aabo thus suggests that the WTP estimate will be biased downwards, the “true” value being closer to the WTA estimate.

In stark contrast to this, Shogren et al. (1994) ran a number of Kahneman-like tests (coffee mugs and chocolate) and discovered no significant difference in WTP and WTA after the first iteration. In other words, for private goods, with no transactions costs, they found that respondents do learn what reasonable values to place on goods as they gain more market experience and that this equalizes WTP and WTA. However, when the same set of tests were conducted for a good for which there are few substitutes (health in the form of reduced risk of food contamination), they

found that WTA is always greater than WTP and conclude that Hanemann's (1991) argument is correct.

One of the major differences between the Shogren et al. (1994) experiments and the Kahneman et al. (1990) tests, however, is that Kahneman et al. used a Becker-DeGroot-Marschak auction, while Shogren et al. used a Vickery auction. Shogren and Hayes (1997) argue that the Vickery auction is far more realistic and like a real market, thus encouraging market-like learning, while the Kahneman et al. method is not. "Is the endowment effect a fundamental part of choice or simply an artefact of a weak exchange environment? The weaker the exchange institution, the weaker the socialization of rational behaviour and the stronger the potential hold of asocial anomalies on choice" (Shogren and Hayes 1997:243).

The idea that market discipline encourages learning is supported by a later study by List (2003) using information from two field studies – a tradable sports cards auction and a collectable pin auction. Using a method similar to that of Kahneman et al. (1990), List finds that, while a significant endowment effect does appear in pooled data (indicated by the low percentage of respondents who were willing to trade their "endowed" good for an alternative good of similar value), the picture changes dramatically when the sample is divided up according to trading experience. In this case, professional traders and experienced non-traders show no endowment effects, while inexperienced traders show a huge effect.

Morrison (1997a and 1997b), however, points out that the possibility that both the income and substitution effects are present, and their possible joint effect on differences between the two contingent valuation methods (WTP and WTA), has not yet been adequately addressed. On conducting an experiment, first allowing both endowment and substitution effects to vary, and then controlling for substitution effects, Morrison (1997a) finds first no significant difference between WTP and WTA and then, in the latter experiment, a significant difference. She thus argues that the Shogren et al. (1994) results, which show that no endowment effect exists, are unjustified.

Psychological Theories for the WTP-WTA Disparity

In addition to these two main streams of thought (that is, the substitution and endowment effects), a few commentators have also suggested rather more psychological theories as to why WTP and WTA often differ so

much. Boyce and Brown (1992) suggested that "intrinsic" values may be important in valuing a complex good such as the environment or some aspect of culture. "If an environmental commodity has intrinsic value, we argue that kinked or inflected indifference curves between the commodity and other money expenditures result when intrinsic values are included in WTA measures of value but (at least partially) excluded from WTP measures of value" (Boyce and Brown 1992:1367). In other words, the moral responsibility of preserving a species or an art form or cultural heritage may be captured in a WTA measure, thus inflating it beyond WTP.

Biel et al. (2006) confirm this hypothesis with their WTP-WTA experiment, using a real payment situation amongst students who were paid to participate in the study. Using a split sample, respondents were asked whether they were willing to pay, or willing to accept a reduction in payment, to support a World Wildlife Fund otter protection project. Only 9 out of 48 participants were willing to pay to support the project, while 23 out of 51 were willing to accept a payment reduction. In addition, follow-up questions exploring the emotional responses to the WTP or WTA questions found that, for those who chose not to donate, emotions such as annoyance, self disgust, discomfort, self dissatisfaction and shame were of a much higher intensity in the WTA scenario. They conclude that WTP-WTA disparities "are largely a result of asymmetric emotional experiences and moral perceptions" (2006:10)

Inder and O'Brien (2003) suggest that the unpleasant psychological uncertainty caused by having to make a choice, either as a seller (WTA) or a buyer (WTP), may cause the wide range of disparities observed. They argue that loss aversion is not adequate to explain differences in WTP and WTA because it looks at only one side of the equation – the seller who is willing to accept compensation for a loss – and does not take into account the psychological effect of the sale on the buyer, who might be unsure about what the asset is worth to them, the utility it will generate and what price the buyer will accept. Band and O'Brien suggest that, to compensate for such uncertainty, sellers (WTA measure) will tend to overstate their values, while buyers (WTP measure) will understate theirs, resulting in wide and varying disparity between WTP, WTA and market price.

The NOAA Panel (1993:18) recommended that only the WTP method should be used, "because [it] is the conservative choice". Many commentators (Cummings in Shogren et al. 1994 and Boyce and Brown 1992, amongst others) agree that only the WTP method of contingent valuation studies should be used when attempting to value a non-market good with

possible "moral" values because of the likelihood of eliciting infinite WTA even where WTP is small or zero. "It is difficult to see how such infinite individual WTA values can be incorporated into benefit-cost analysis..." (Boyce and Brown 1992:1371). While some commentators, like Mansfield (1999), still argue that compensating value is the correct measure of welfare loss rather than WTP, in practice most studies have stuck to the WTP recommendation.

4.3.4 The Mixed Good Bias

The arts, as a mixed good, have both private and public good characteristics. Some parts of the arts, like concerts and exhibitions, are excludable and have private good characteristics in that those who do not pay an entrance fee are prevented from consuming the good. However, the arts also have public good characteristics, such as the externalities which they generate - usually not excludable and in most cases, non-rival. The following section discusses the difficulty of separating the pure non-market externalities from the market or financial benefits of a mixed good in WTP studies.

Attenders versus non-attenders

Throsby (1984) suggests that there should be separate demand functions for the public and private components of the good and that this may reveal an incentive for the misrepresentation of demand preferences. He is supported by Morrison and West (1986:69) who point out that "horizontal inequity" (differences in benefit between people of similar income levels) may make the idea that it is the wealthy, well educated people that benefit most from arts subsidies too simplistic: "The theatre attender [as opposed to the non-attender of similar income] will enjoy the benefits of subsidy to the direct use of theatres as well as the external benefits".

This idea was later articulated as the distinction between total value and passive use value that Carson et al. (1999) made. They point out that WTP consists of two parts, passive use value (externalities that accrue to users and non-users) and direct use value (that accrues only to users). The main point of Throsby's (1984) argument is that those people who attend the live performing arts will thus have a (logical) reason to overstate their

WTP for the service, even where conventional free rider behaviour is absent. Consider an economy which has a mixed good (x) with two users, A and B. User A demands only the public parts of the good (non-attender) and user B demands the private good component as well (attender). Looked at simply, if there were an increase in subsidy to good x, B would gain more from the decline in price than A (since B is an attender), even though the tax shares of A and B would be equal. There is, therefore, an incentive for B to overstate WTP, since an increase in subsidy would benefit B more than A, while the costs would be shared equally (Throsby 1984:280 - 282).

> "The essence of this analysis is that an incentive for overstatement of preferences exists if the individual's gain from the increase in his or her consumption enabled by subsidisation of the private [part of the mixed] good exceeds his or her perceived share of the tax necessary to finance this increase in output" (Throsby 1984:282).

In applying this theory to their study on the level of public finance for the arts in Australia, Throsby and Withers (1986) found a strong correlation between users (i.e. attenders) and higher WTP; "A downward adjustment of about 40% in average stated WTP is required to eliminate this source of bias". Once this had been done, the results of the study could be used to provide information on the determinants of demand for the public part of the good, such as income, taste and education (Throsby 1984).

These results are confirmed in a number of later cultural economics valuation studies, as shown in table 4.3. Without exception, all the studies show that users or potential users of the resource are willing to pay more than non-users. The mixed good bias fits well with the Thompson et al. (2002) suggestion, which states that, for those who frequently attend the highly sponsored shows, willingness to pay may represent some sacrifice of consumer surplus, or an avoidance of travel costs, rather than purely a valuation of positive externalities.

A particularly interesting study is that of the Surrey Historical Centre (Ozdemiroglu and Mourato 2001), which proposed two valuation scenarios: In the first case, the Centre would remain operational and open to the public as it is presently. In the second case, the Centre would remain operational, but would not be open to the general public. The findings show that, for users, the two scenarios have significantly different values; non-use value being £24, use-value (open to the public) being £34. However, for current

non-users, the average WTP for both scenarios was £13, clearly illustrating differing values depending on use.

Table 4.3. Use-value Bias in WTP Studies

Source	Case Study	Findings
Carson et al. (1997)	Preservation of the Fes Medina	Visitors to the site are willing to pay $70 each on average, while visitors to Morocco who did not visit the Fes are willing to pay about $30 each. WTP amongst the European population in general was $2.15
Ozdemiroglu and Mourato (2001)	Surrey Historical Centre archive	An average user's WTP to prevent closure and collection loss was £34 per person per year and £24 to prevent loss of access. Non-users were willing to pay £13 per year for both scenarios.
Snowball and Antrobus (2001)	National Arts Festival, South Africa	56% of low/no attendance respondents were willing to pay, 76% of medium attenders and 100% of frequent attenders.
Snowball and Antrobus (2003)	Oudtshoorn Arts Festival, South Africa	Average WTP of respondents who attended 0-3 shows was R3.33; 4-7 shows was R11.60; 8-12 shows was R13.80; 13 or more shows was R15.22
Del Saz Salazar and Marques (2005)	Old Arab Tower restoration, Valencia	Mean WTP higher for those with "high consumption" of cultural goods
Alberini et al. (2005)	S. Erasmo Lagoon, Venice	Percentage of "no" WTP responses is 49% amongst recent visitors, 50% amongst potential future visitors and 79% amongst non-users.
Del Saz Salazar and Menendez (2007)	El Parque Central, Valencia	Respondents who lived closer to the proposed urban park were willing to pay more than those who lived further away.

Such results are also found in studies of arts festivals (Snowball and Antrobus 2001; 2003) and heritage sites (Del Saz Salazar and Marques 2005; Alberini et al. 2005 and Carson et al. 1997), where willingness to pay was positively correlated with the number of shows attended or visits to the site. In fact, the Del Saz Salazar and Marques study (2005) found that WTP to restore an old Arab tower in Valencia was 133% higher for those respondents who were categorised as "high consumers" of cultural goods in general than those who were low/average cultural good consumers. This suggests that appreciation of cultural goods in general, as well as use of the specific good being valued, could be biasing WTP results upwards. When generalising WTP results to a whole population, therefore, it is very important to examine the extent to which users, or "high" cultural goods consumers, may be influencing the average values.

Mixed Good Characteristics and Earnings

In addition to the consumer surplus of attenders, Seaman (2003a:13) suggests that what a WTP study may also be capturing is a willingness to pay for expected economic benefit. This could be reflected not only in present earnings, but also in the long-term growth of the area. "Would not a very real component of the "bequest" motive of a parent of a 13 year old child be the possibility that that child could possibly manoeuvre into a position to be a direct beneficiary of the higher income and expanded job opportunities that the event would provide?" Such benefit might particularly apply to a cultural event or attraction that would result in an increase in visitors to the area, like an arts festival, a major museum or the restoration of a major heritage site.

Studies conducted by Thompson et al. (1998 and 2002) added WTP estimates to economic impact figures to measure the total value of the arts to the Kentucky economy ($43.8 million a year). However, this method was criticized by Seaman (2003a) on two grounds. Firstly, he argues that part of what the WTP estimate may be picking up is the projected current or future earnings of the respondent or their household due to the arts. In other words, the WTP estimate includes, to some extent, economic impact and that simply adding them leads to double counting. Secondly, he argues that, rather than giving a total value, the WTP figure measures only one scenario (that of a 25% decrease in arts events), not the value of the externalities provided by 100% of Kentucky arts events.

In the study conducted at two South African arts festivals (the Grahamstown National Arts Festival and the *Klein Karoo Nationale Kunstefees* in Oudtshoorn), respondents were asked a follow-up question to the WTP question, probing the reasons for being willing or not willing to support the Festival (Snowball 2005). Respondents, who were local residents in the town where the festival took place, were also asked about their earnings from the festival as a result of additional employment or providing accommodation to tourists. Both the amount of money earned at the festival and the respondent's views on the economic benefits that the festival can or could provide (often cited as a reason for positive willingness to pay) were positive and highly significant determinants of willingness to pay.

Despite much lower average festival earnings, citing economic benefits as a reason for positive WTP was particularly statistically significant in the lower income areas. This suggests that, in areas where cultural events are

associated with increases in tourism and thus earnings, where there is a high level of unemployment or low average income, willingness to pay figures will also be capturing some of the value of expected future financial gains. In the higher income areas of the two South African festival towns, actual earnings from festival tourists were statistically significant positive determinants of WTP, and are thus also likely to be biasing the WTP figure upwards (Snowball 2005).

Table 4.4. Festival earnings and spending at two arts festivals in South Africa

	Grahamstown		**Oudtshoorn**	
	High income	Low income	High income	Low income
Average current festival earnings	R390 ($60)	R106 ($16.31)	R833 ($128.15)	R341 ($52.46)
Percentage of respondents who cited financial gains (current or future expected) as a reason for willingness to pay.	19%	53%	35%	40%

(Source: Snowball 2005:118)

If one could establish the extent to which the WTP figure reflects the value of non-market externalities versus market benefits, the ideal situation, as suggested by Seaman (2003a), would be to run both a WTP and accurate economic impact study at the same event. One could then combine the two results, appropriately weighted, in order to arrive at a true economic value for the festival. However, simply adding WTP and economic impact figures would result in some double counting.

4.3.5 Other Categorical Critiques of the CV Method

Sunstein (2002) and Sunstein et al. (2001) argue strongly that the contingent valuation method, especially in the case of cultural goods, will produce incoherent valuation estimates. Their point is that WTP will vary sharply with the category of goods being considered, that is, estimates will be "category bound". For example, for cultural amenities being considered in isolation, it is quite likely that a significant WTP amount will be elicited, but that if other categories were included, like some aspect of health care, the WTP for the cultural amenity would drop sharply. "When people explore particular problems in isolation, they normalize them by comparing them to a cognitively accessible comparison set consisting of cases

from the same basic category. When cases from other categories are introduced into the picture, people's judgments can be greatly affected, because the process of normalization is disrupted" (Sunstein 2002:1). Thus, WTP estimates obtained in isolation, particularly for cultural goods, which appear fairly low on most people's public goods rankings, are judged unreliable.

Sunstein (2002) poses two solutions to the problem, both of which he rejects. Firstly, one could include as wide a range as possible of other public goods (health care, education, environment, etc) in the WTP questionnaire. Sunstein rejects this on the grounds that the task would be incredibly complex and result in "cognitive overload". An alternative would be to include at least some other categories, but this would allow for serious manipulation, since the categories chosen would certainly affect the WTP result. Epstein (2003:267-8) supports this position, arguing that, even for private goods, if removed from a specific context where price comparison with similar articles is possible, WTP responses are likely to be "haphazard and of limited reliability".

A second problem, raised by Sunstein et al. (2001), is that of the validity of assuming that WTP responses are a monetary expression of utility. Like Kahneman et al. (1999), they argue that even when people can produce "coherent and consistent moral intuitions", they are not easily translated into numbers. They consider the example of a jury deciding on the penalty (number of years in jail) for a particular crime, which when considered in isolation, appears just, but when considered along with penalties for other crimes, is not coherent. Sunstein (2002) concludes that, while the market mechanism is not suitable for making judgments about the provision of cultural goods, WTP surveys are not a solution either, particularly when one considers income constraints.

Recalling somewhat the endowment theory discussed above, Epstein (2003) argues that, even for private goods, value is not well defined, consisting of a market value and (an often divergent) subjective value. He argues that, even if subjective value is ignored, market values (for example, quotes from various builders for adding to a house) can vary quite a lot. If market values for private goods are unreliable, how much more unreliable will be stated preferences for public goods in hypothetical markets?

Epstein (2003) also points out that there are a number of biases introduced into WTP studies by, for example, assuming that non-use or existence values are always positive and that the lower bound of the WTP response is

zero. There may be cases where existence value may in fact be negative. A second point is that existence value can apply to many goods, some of which may exclude others. "To ask whether people attach existence value to redwood trees does not tip the scales in favour of their preservation. It is necessary to pose a second question: do they also attach existence value to redwood furniture. But one cannot have both" (Epstein 2003:272).

In fact, there is one early WTP study that does acknowledge both the possible positive and negative externalities of the good. Maddison and Mourato (1999) conducted a study of the WTP to run one of the roads around Stonehenge underground and to close the other in order to restore the peace of this ancient mystical site. However, after being shown maps and pictures of the current and alternate scenarios, respondents were asked, without financial considerations, which scenario they preferred. This allowed those who preferred the current two roads scenario to the tunnel scenario to state their positive value. Regardless of which option they chose, respondents were then asked their WTP in order to secure their preferred option.

While about half the respondents considered the site to be "disfigured" by the roads, a similar number attached a positive value to being able to see the stone circle when driving past or not liking to drive through long tunnels. This suggests that, at least for some respondents, there is a cost associated with running the road through a tunnel. However, WTP for the tunnel scenario was still considerably higher (£12.80) than the current scenario (£4.80), suggesting a much stronger preference amongst those who chose the tunnel scenario. The net heritage benefit of building the tunnel was calculated to be £149m.

Even the most ardent supporters of CV studies do however recognize their limitations. Carson et al. (2001), while concluding that many of the problems of CV can be overcome, mention two areas in which the method is limited. Firstly, WTP studies are limited by wealth. In developed countries, this is less of a problem, but in the developing world, where government policies are often designed specifically to provide public goods to the poorer parts of the population, this can be a significant handicap. The second limitation is that, like all real or hypothetical markets, only the preferences of the current generation and perhaps those imputed for future generations are taken into account, leaving the true preferences of future people unknown and essentially unknowable.

Throsby (2003) also identifies areas where WTP will not capture the full value of the good, particularly if the good is cultural. For example, if the respondent does not personally value the good, but recognizes that value may accrue to others, either as individuals or society as a whole, and is thus willing to pay for it. Throsby (2003:279) argues that such societal values are particularly relevant to cultural goods because of their nature:

> Culture can be defined as the set of beliefs, traditions, customs, etc. which identify a group and bind its members together … So the value of cultural goods that I am discussing here is a value identifiable in relation to the group rather than to the isolated characteristics of individuals. It is apparent that the perception of this sort of value is not going to be captured by an expression of individual WTP".

Thus, Throsby (2003) argues that, while WTP studies can certainly capture some of the economic value of a cultural good, the relationship will never be perfect and that even the best studies will tend to undervalue the cultural good in question.

4.4 Conclusions

The above review has given examples of how willingness to pay studies have been used to value cultural goods. Using case studies, the chapter outlines the debate around the major ways in which the results of contingent valuation studies have been criticized and shown to be inconsistent with neoclassical economic theory, that is, by being implausibly large, by showing insensitivity to scope and by demonstrating a large disparity between WTP and WTA measures, as well as other general criticisms. As the NOAA Panel (1993) suggests, many of these problems can be controlled for through rigorous and conservative questionnaire design and much of the CV debate has shifted to this arena. The NOAA recommendations, as well as the results of later research, are discussed in the following chapter, focussing specifically on the practical problems of questionnaire design, result interpretation and validity testing.

Appendix 4.1

Examples of WTP Studies in Cultural Economics[2]

YEAR	AUTHOR	AREA	GOOD	SOURCE
1972	Bohm	Broadcasting	TV programme	*European Economic Review* 3:111-130
1980	Throsby, O'Shea	Arts	Mildura Arts Center, Australia	Research Paper 210, Macquarie University
1983	Thompson et al.	Arts	Tax support for Australian arts	Research Paper 261, Macquarie University
1983	Scheider, Pommerehne	Theatre	Swiss municipal theatre	Book: *Economics of Cultural Decisions,* Hendon and Shanahan (Eds)
1986	Morrison, West	Arts	Performing arts and culture, Ontario	*Journal of Behavioural Economics* 15:57-72
1989	Frey, Pommerehne	Museums	Purchase of two Picasso paintings by Swiss city	Book: *Muses and markets*
1990	Ehrenberg, Mills	Broadcasting	TV viewing	Supplement to *Broadcast* (November)
1992	Navrud, Strand	Historical site	Nidaros Cathedral, Norway	Book: *Pricing the European Environment* Navrud (Ed)
1994	Willis	Historical site	Durham Cathedral, England	*Journal of Environmental Planning and Management* 37,3:267-278
1994	Grosclaude, Soguel	Historical site	Building maintenance, Neuchatel, Swiss	*Journal of Environmental Planning and Management* 37,3:279-287
1994	Martin	Museums	Quebec museums	*Journal of Cultural Economics* 18,4:255-270
1995	Schwer, Daneshvary	Broadcasting	PBS TV in Las Vegas	*Journal of Media Economics* 8,3:95-109
1996	Garrod et al.	Historical	Building renovation, Newcastle	*Cities* 16,6:423-430
1996	Longwood et al.	Heritage	Grazing, Australian alps	*Journal of Environmental Planning and Management* 39,3:357-370
1996	Beltran, Rojas	Archaeological sites	Mexican archaeological sites	*Annals of Tourism Research* 23,2:463-478
1997	Carson et al.	Historical site	Fes Medina, Morocco	World Bank Report, Washington
1997	Bille Hansen	Theatre	Royal Theatre in Copenhagen	*Journal of Cultural Economics* 21, 1:1-28

[2] The above table is largely adapted from, with some updating and additions, the very useful annotated bibliographies of Noonan (2002) and EFTEC (2005). While it includes some actual referenda (which are not part of hypothetical market valuation) the major focus is on the willingness to pay method, not including conjoint analysis (choice experiments) that is discussed in chapter six.

1998	Willis, Powe	Historical site	Warkworth Castle, England	*Journal of Environmental Planning and Management* 41,5;611-619
1998	Chambers et al.	Historical site	Building in St. Genevieve, MO	*Public Finance Review* 26:137-154
1998	Scarpa et al.	Historical site	Rivoli Castle, Italy	Book: *Environmental Resource Valuation: Applications of the CV method in Italy* Scarpa, Sirchai , Bravi (Eds)
1998	Pagiola	Historical site	City of Split, Croatia	World Bank website
1998	Bravi, Scarpa	Museums	Italian museums	Book: *Valuing cultural heritage* Bravi, Scarpa, Sirchia (Eds)
1998	Riganti, Scarpa	Archaeological sites	Campi Flegrei, Italy	Book: *Economic resource valuation* Bishop, Romano (Eds)
1999	Glass et al.	Arts	Local area arts in Kansas	University of Kansas Institute for Public Policy and Business, Research report 257
1999	Maddison, Mourato	Historical site	Road options for Stonehenge	CSERGE Working Paper GEC 99-08 (Available on internet)
1999	Papandrea	Broadcasting	Domestic TV	*Journal of Cultural Economics* 23,3:147-164
1999	Harless, Allen	Libraries	Services at University Library	*College & Research Libraries* 60:56-69
1999	Holt et al.	Libraries	St Louis public libraries	*Public Libraries* 38,2:98-108
2000	Santagata, Sigorello	Historical site	Napoli Musei Aperti, Italy	*Journal of Cultural Economics* 24,3:181-204
2000	Creigh-Tyte	Heritage	Historical environment in Britain	Paper presented at the *Association of Cultural Economics International* conference in Minneapolis
2000	Begin et al.	Broadcasting	Filmmaking in Quebec and France	*International Journal of Cultural Policy* 7,2:355-368
2001	Pollicino, Maddison	Historical Site	Cleaning Lincoln Cathedral, England	*Journal of Cultural Economics* 25,2:131-148
2001	Ozdemiroglu, Mourato	Heritage	Surrey History Centre	Paper presented at Economic Valuation of Cultural Heritage Conference, University College, London
2001	Jennings	Broadcasting	Irish public radio	M. Lit. submitted at Trinity College, Dublin University
2001	Snowball, Antrobus	Arts	Arts Festival in South Africa	*South African Journal of Economics* 69,4:752-766
2002	Cuccia, Signorello	Historical site	Visits to Noto, Italy	Book: *The economics of heritage* Rizzo, Towse (Eds)
2002	Willis	Historical site	Bosco di Capodimonte	*Journal of Cultural Economics* 26:307-324
2002	Aabo, Strand	Libraries	Norwegian libraries	*Journal of Librarianship and Information Science* 34,1:5-15
2002	Thompson et	Arts	Arts in Kentucky	*Journal of Cultural Economics*

	al.			26:87-113
2002	Mourato et al.	Heritage	Bulgarian monasteries	Book: *Valuing Cultural Heritage* (Navrud, Ready Eds.)
2003	Maddison, Foster	Museums	Congestion in British Museum	*Oxford Economic Papers* 55:173-190
2003	Whitehead, Finney	Heritage	Submerged maritime heritage, North Carolina	*Journal of Cultural Economics* 27, 3-4:231-240
2003	Sanz et al.	Museum	National Museum of Sculpture, Valladolid, Spain	*Journal of Cultural Economics* 27:241-257
2004	Mourato et al.	Historical site	Access to Machu Picchu	*World Economics* 5, 3
2004	Kling et al.	Historical site	Hotel in Fr Collins, CO	*Urban Studies* 41,10:2025-2041
2004	Alberini et al.	Heritage	Value of S. Erasmo in the Lagoon of Venice	*Journal of Environmental Planning and Management* 48,2:155-175
2004	Brown	Heritage	National Trust properties in England	PhD dissertation, Imperial College London
2005	Del Saz Salazar, Marques	Historical site	Restoring old Arab tower, Spain	*Journal of Cultural Heritage* 6:69-77
2005	Aabo	Libraries	Norwegian public libraries	*New Library World* 106, 1218/1219:487-495
2005	Snowball	Arts	Arts Festival in South Africa	*Journal of Cultural Economics* 29:107-125
2006	Jim	Heritage	Heritage trees in cities	*Environmental Monitoring and Assessment* 116:53-80
2006	Tuan, Navrud	Historical site	World Heritage site in Vietnam	*Environmental and Resource Economics*
2007	Kim et al.	Historical site	Changdeok Palace, Korea	*Tourism Management* 28:317-322
2007	Dutta et al.	Heritage	Prinsep Ghat, Calcutta	*Tourism Management* 28:83-95

References

Aabo, S. (2005) Are public libraries worth their price? A contingent valuation study of Norwegian public libraries. New Library World 106, 1218/1219:487-495

Aadland, D. and Caplan, A. (2003) Willingness to pay for curbside recycling with detection and mitigation of hypothetical bias. American Journal of Agricultural Economics 85,2:492-502.

Adamowicz, W., Bhardwaj, V. and Macnab, B. (1993) Experiments on the difference between willingness to pay and willingness to accept. Land Economic 69,4:416-427.

Alberini, A., Rosato, P., Longo, A. and Zanatta, V. (2005) Information and willingness to pay in a contingent valuation study: The value of S. Erasmo in the Lagoon of Venice. Journal of Environmental Planning and Management 48,2:155-175.

Alberini, A. and Longo, A. (2006) Combining the travel cost and contingent behaviour methods to value cultural heritage sites: Evidence from Armenia. Journal of Cultural Economics 30:287-304.

Anchorage Daily News (1999) Hard Aground: Disaster in Prince William Sound Special feature published form 21-24 March. [On line] Available: www.and.com [Accessed 18/09/02].

Bateman, I.J., Cole, M., Cooper, P., Georgiou, S., Hadley, D. and Poe, G.L. (2004) On visible choice sets and scope sensitivity. Journal of Environmental Economics and Management 47:71-93.

Bateman, I., Cooper, P., Georgiou, S., Navrud, S., Poe, G., Ready, R., Riera, P., Ryan, M. and Vossler, C. (2005) Economic valuation of policies for managing acidity in remote mountain lakes: Examining validity through scope sensitivity testing. Aquatic Sciences 67:274-291.

Bateman, I. and Brouwer, R. (2006) Consistency and construction in stated WTP for health risk reductions: A novel scope-sensitivity test. Resource and Energy Economics 28:199-214

Bedate, A., Herrero, L. and Sanz, J. (2004) Economic valuation of the cultural heritage: application to four case studies in Spain. Journal of Cultural Heritage 5,1:101-111.

Biel, A. (2006) Emotions, morality and public goods: The WTA-WTP disparity revisited. Department of Economics, School of Business, Economics and Law, Goteborg University, Working Papers in Economics, no. 193.

Bille Hansen, T. (1997) The willingness to pay for the Royal Theatre in Copenhagen as a public good. Journal of Cultural Economics 21:1-28.

Bille, T. and Schulze, G. (2006) Culture in urban and regional development. Handbook of the Economics of Art and Culture. Ginsburgh, V. and Throsby, D. (Eds.) North Holland:Amsterdam.

Bohm, P. (1972) Estimating demand for public goods: An experiment. Scandinavian Journal of Economics 3:113-130.

Bohm, P. (1979) Estimating willingness-to-pay: Why and how? Scandinavian Journal of Economics 81:142-153.

Bohm, P. (1984) Revealing demand for an actual public good. Journal of Public Economics 24:135-151

Booysen, F. (2001) Non-payment of services: A problem of ability-to-pay. South African Journal of Economics 69,4:674-697.

Botelho, A. and Pinto, L. (2002) Hypothetical, real and predicted real willingness to pay in open-ended surveys: experimental results. *Applied Economics Letters* 9:993-996.

Boter, J., Rouwendal, J. and Wedel, M. (2005) Employing travel time to compare the value of competing cultural organizations. Journal of Cultural Economics 29,1:19-33.

Boyce, R. and Brown, T. (1992) An experimental examination of intrinsic values as a source of the WTA-WTP disparity. The American Economic Review 82,5:1366-1372.

Brown, T. (1994) Experiments on the difference between willingness to pay and willingness to accept: Comment. Land Economics 70,4:520-522.

Carson, R. and Mitchell, R. (1993) The issue of scope in contingent valuation studies. American Journal of Agricultural Economics 75,5:1263-1268.

Carson, R., Mitchell, R., Conway, M. and Navrud, S. (1997) Non-Moroccan Values for Rehabilitating the Fes Medina. World Bank Report: Washington.

Carson, R.T. (1997) Contingent valuation: Theoretical advances and empirical tests since the NOAA panel. American Journal of Agricultural Economics 79,5:1501-1507.

Carson, R., Flores, N. and Mitchell, R. (1999) Theory and measurement of passive use value in Bateman, I. and Willis, K. Valuing Environmental Preferences: Theory and Practice of the CV method in the US, EU and Developing Countries. Oxford University Press: New York.

Carson, R., Flores, N. and Meade, N. (2001) Contingent valuation: Controversies and evidence. Environmental and Resource Economics 19:173-210.

Carson, R., Mitchell, R., Hanemann, M., Kopp, R., Presser, S. and Ruud, P. (2003) Contingent valuation and lost passive use damages from the Exxon Valdez oil spill. Environmental and Resource Economics 25;257-286.

Chambers, C., Chambers P. and Whitehead, J. (1998) Contingent valuation of quasi-public goods: validity, reliability and application to valuing a historic site. Public Finance Review 26,2:137-155.

Champ, P. and Bishop, R. (2001) Donation payment mechanisms and contingent valuation: An empirical study of hypothetical bias. Environmental and Resource Economics 19:383-402.

Cuccia, T. and Signorello, G. (2002) Heritage Conservation in Sicily. Rizzo I. and R. Towse (eds.), Methods for Measuring the Demand for the Arts and Heritage: Theoretical Issues. Edward Elgar Publishing 2002

Cummings, R. and Taylor, L. (1999) Unbiased value estimates for environmental goods: A cheap talk design for the contingent valuation method. American Economic Review 89,3:649-666.

Del Saz Salazar, S. and Marques, M. (2005) Valuing cultural heritage: the social benefit of restoring an old Arab tower. Journal of Cultural Heritage 6:69-77

Del Saz Salazar, S. and Menendez, L. (2007) Estimating the non-market benefits of an urban park: Does proximity matter? Land Use Policy 24:296-305.

Desvousges, W., Johnson, F. Dunford, R. Boyle, K., Hudson, S. and Wilson, K. (1993) Measuring natural resource damages with contingent valuation: Tests of validity and reliability. In Hausman, J. (Ed.) Contingent Valuation: A critical assessment. Elsivier Science Publishers: Amsterdam.

Diamond, P. and Hausman, J. A. (1993) On contingent valuation measurement of nonuse values. In Hausman, J. (Ed.) Contingent Valuation: A critical assessment Elsivier Science Publishers: Amsterdam

Diamond, P. and Hausman, J.A. (1994) Contingent valuation: Is some number better than no number? Journal of Economic Perspectives 8,4:45-65.

Diamond, P. (1996) Testing the internal consistency of contingent valuation surveys. Journal of Environmental Economics and Management 30:337-347.

Dupont, D. P. (2003) CVM embedding effects when there are active, potentially active and passive users of environmental goods. Environmental and Resource Economics 25:319-341.

Dutta, M., Banerjee, S. and Husain, Z. (2007) Untapped demand for heritage: A contingent valuation study of Prisep Ghat, Calcutta. Tourism Management 28:83-95.

Eftec (2005) Valuation of the historic environment: Annotated bibliography of heritage valuation studies. Economics for the Environment Consultancy Ltd. London.

Epstein, R. (2003) The regrettable necessity of contingent valuation. Journal of Cultural Economics 27:259-274.

Glass, R., Clifford, N., Harris, B., Wollsey, C. and Krider, C. (1999) Economic scope, impact and marketing study of the Kansas Arts Commission. The University of Kansas, Institute for Public Policy and Business*:* Research report number 257.

Guzman, R. and Kolstad, C. (2006) Researching preference, valuation and hypothetical bias. Environmental and Resource Economics. Springer Online Publication.

Hanemann, M. (1991) Willingness to pay and willingness to accept: how much can they differ? The American Economic Review 81,3:635-647.

Hanemann, M. (1994) Valuing the environment through contingent valuation. Journal of Economic Perspectives 8,4:19-43.

Harless, D. and Allen, F. (1999) Using the contingent valuation method to measure patron benefits of reference desk services in an academic library. College and Research Libraries January:56-69

Hausman, J. (Ed.) (1993) Contingent Valuation: A critical assessment. North-Holland: Amsterdam.

Holt, G. and Elliott, D. (1999) Proving your library's worth: A test case. Library Journal November:42-44.

Inder, B. and O'Brien, T. (2003) The endowment effect and the role of uncertainty. Bulletin of Economic Research 55,3:289-301.

Johansson, P. (1993) Cost-Benefit Analysis of Environmental Change. Cambridge University Press: London.

Kahneman, D., Knetsch, J. and Thaler, R. (1990) Experimental tests of the endowment effect and the Coase theorem. *Journal* of Political Economy 98,6:1325-1348.

Kahneman, D and Knetsch, J. (1992) Valuing public goods: The purchase of moral satisfaction. Journal of Environmental Economics and Management 22:57-70

Kahneman, D., Ritov, I. and Schkade, D. (1999) Economic preferences or attitude expressions: an analysis of dollar responses to public issues. Journal of Risk and Uncertainty 19, 1-3:203-235.

Kim, S., Wong, K. and Cho, M. (2007) Assessing the economic value of a world heritage site and willingness to pay determinants: A case of Changdeok Palace. Tourism Management 28:317-322

Kling, R., Revier, C. and Sable, K. (2004) Estimating the public good value of preserving a local historical landmark: The role of non-substitutability and citizen information. Urban Studies 41,10:2025-2041.

List, J. and Shogren, J. (1998) Calibration of the difference between actual and hypothetical valuation in a field experiment. Journal of Economic Behavior and Organization 37:193-205.

List, J. (2001) Do explicit warnings eliminate the hypothetical bias in elicitation procedures? Evidence from field auctions for sportscards. The American Economic Review 91,5:1498-1507.

List, J. and Gallet, C. (2001) What experimental protocol influences disparities between actual and hypothetical stated values? Environmental and Resource Economics 20:241-254.

List, J. (2003) Does market experience eliminate market anomalies? The Quarterly Journal of Economics Feb: 41-71.

Lusk, J. (2003) Effects of cheap talk on consumer willingness to pay for golden rice. American Journal of Agricultural Economics 85,4:840-856.

Noonan, D. (2002) Contingent valuation studies in the arts and culture: An annotated bibliography. The Cultural Policy Center at the University of Chicago: Working Paper.

Maddison, D. and Mourato, S. (1999) Valuing different road options for Stonehenge. CSERGE Working Paper GEC 99-08

Maddison, D. and Foster, T. (2003) Valuing congestion costs in the British Museum. Oxford Economic Papers 55:173-190.

Mansfield, C. (1999) Despairing over disparities: Explaining the difference between willingness to pay and willingness to accept. Environmental and Resource Economics 13:219-234.

Mitchell, R. and Carson, R. (1993) Current Issues in the Design, Administration, and Analysis of Contingent Valuation Surveys, University of California at San Diego, Economics Working Paper Series 93-54, Department of Economics, UC San Diego

Morrison, W. and West, E. (1986) Subsidies for the performing arts: Evidence of voter preference. Journal of Behavioral Economics 15, Fall:57-72.

Morrison, G. (1997a) Willingness to pay and willingness to accept: some evidence of an endowment effect. Applied Economics 29:411-417.

Morrison, G. (1997b) Resolving differences in willingness to pay and willingness to accept: Comment. The American Economic Review 87,1:236-240.

Mourato, S., Kontoleon, A. and Danchev, A. (2002) Preserving cultural heritage in transition economies: A contingent valuation study of Bulgarian monasteries. Navrud, S. and Ready, R. (Eds.) Valuing Cultural Heritage. Edward Elgar Publishing.

Mourato, S., Ozdemiroglu, E., Hett, T. and Atkinson, G. (2004) Pricing cultural heritage: A new approach to managing ancient resources. World Economics 5, 3:1-19.

Mundy, B. and McLean, D. (1998) Using the contingent valuation approach for natural resource and environmental damage application. Appraisal Journal 66,3:290-298.

Murphy, J., Allen, G., Stevens, T. and Weatherhead, D. (2005a) A meta-analysis of hypothetical bias in stated preference valuation. Environmental and Resource Economics 30:313-325.

Murphy, J., Stevens, T. and Weatherhead, D. (2005b) Is cheap talk effective at eliminating hypothetical bias in a provision point mechanism? Environmental and Resource Economics 30:327-343.

Navrud, S. and Strand, J. (1992) The preservation value of Nidaros Cathedral, in Pricing the European Environment (Navrud, S. (ed.) Oxford University Press: New York.

Nestor, D. (1998) Policy evaluation with combined actual and contingent response data. American Journal of Agricultural Economics 80,2:264-277.

NOAA (1993) Arrow, K.J., Solow, R., Leamer, E., Radner, R. and Schuman, H. Report of the NOAA Panel on contingent valuation. National Oceanic and Atmospheric Administration Federal Register 58,10

Ozdemiroglu, E. and Mouraton, S. (2001) Valuing our recorded heritage. Paper presented at the Economic Valuation of Cultural Heritage conference, University College, London.

Papandrea, F. (1999) Willingness to pay for domestic television programming. Journal of Cultural Economics 23:149-166

Pollicino, M. and Maddison, D. (2001) Valuing the benefits of cleaning Lincoln Cathedral. Journal of Cultural Economics 25:131-148.

Poor, J. and Smith, J. (2004) Travel cost analysis of a cultural heritage site: The case of historic St. Mary's city of Maryland. Journal of Cultural Economics 28,3:217-229.

Ready, R., Berger, N. and Blomquist, G. (1997) Measuring amenity benefits from farmland: hedonic pricing versus contingent valuation. Journal of Urban and Regional Policy 28,4:438-458.

Ringanti, P. and Scarpa, R. (1998) Categorical nesting and information effects on WTP estimates for the conservation of cultural heritage in Campi Flegrei. Bishop, R. and Romano, D. (Eds.) Environmental and Resource Valuation. Kluwer: Boston.

Samuelson, P. (1954) The pure theory of public expenditure. Review of Economics and Statistics 36:378-389.

Sanz, J., Herrero, L. and Bedate, A. (2003) Contingent valuation and semiparametic methods: A case study of the National Museum of Sculpture in Valladolid, Spain. Journal of Cultural Economics 27:241-257.

Scarpa, R., Sirchia, G. and Bravi, M. (1998) Kernel vs logit modelling of single bounded CV responses: Valuing access to architectural and visual arts in Italy. In Bishop, R. and Romano, D. (Eds.) Environmental Resource Valuation: Applications of the Contingent Valuation Method in Italy. Kluwer: Boston.

Seaman, B. (2003a) Contingent Valuation vs. Economic impact: substitutes or complements? Paper delivered at the Regional Science Association International Conference, North American Meetings: Philadelphia

Sen, A. (1985) Well-Being, agency and freedom: The Dewey lectures 1984. The Journal of Philosophy 82,4:169-221.

Shogren, J., Shin, A. Hayes, D. and Kleinbenstein, J. (1994) Resolving differences in willingness to pay and willingness to accept. The American Economic Review 84,1:255-270.

Shogren, J. and Hayes, D. (1997) Resolving differences in willingness to pay and willingness to accept: Reply in The American Economic Review 87,1:241-244.

Smith, K. and Osborne, L. (1996) Do contingent valuation methods pass a 'scope' test? A meta-analysis. Journal of Environmental Economics and Management 31:286-301.

Snowball, J. (2000) Towards more accurate measurement of the value of the arts to society: Economic impact and willingness to pay studies at the Standard Bank National Arts Festival. Unpublished Masters Thesis: Department of Economics, Rhodes University, Grahamstown, South Africa.

Snowball, J. and Antrobus, G. (2001) Measuring the value of the arts to society: the importance of the value of externalities to lower income and education groups in South Africa. South African Journal of Economics 69,4:752-766.

Snowball, J. and Antrobus, G. (2003) Results of the WTP household survey at the Klein Karoo Nationale Kunstefees. Unpublished research paper, Rhodes University: South Africa.

Snowball, J. (2005) Art for the masses? Justification for the public support of the arts in developing countries – two arts festivals in South Africa. Journal of Cultural Economics 29:107-125.

Sunstein, C., Kahneman, D., Schkade, D. and Ritov, I. (2001) Predictably Incoherent Judgments. Paper presented at the Cultural Policy Center at the University of Chicago workshop. [On line] Available: http://cutural policy. Uchicago.edu/workshop/sunstein.Html [Accessed: 19/02/03].

Sunstein, C. (2002) Coherent and incoherent valuation: a problem with contingent valuation of cultural amenities. The Cultural Policy Center at the University of Chicago: Working paper.

Thampapillai, D. (2000) Willingness to pay and willingness to accept: a simple conceptual exposition. Applied Economics Letters 7:509-511.

Thompson, B., Throsby, D. and Withers, G. (1983) Measuring community benefits from the arts. Research paper 261 School of Economic and Financial Studies: Macquarie University.

Thompson, E., Berger, M. and Allen, S. (1998) Arts and the Kentucky economy. Center for Business and Economic Research, University of Kentucky.

Thompson, E., Berger, M. and Blomquist, G. (2002) "Valuing the arts: A contingent valuation approach" *Journal of Cultural Economics* 26:87-113.

Throsby, D. (1984) The measurement of willingness-to-pay for mixed goods. Oxford Bulletin of Economics and Statistics 46,4 279-289

Throsby, D. and Withers, G. (1985) What price culture? Journal of Cultural Economics 9,2:1-33.

Throsby, D. and Withers, G. (1986) Strategic bias and demand for public goods: theory and application to the arts. Towse, R. (Ed.) Cultural economics: The arts, the heritage and the media industries. Vol 2 Edward Elgar: Cheltenham.

Throsby, D. (2003) Determining the value of cultural goods: How much (or how little) does contingent valuation tell us? *Journal* of Cultural Economics 27:275-285.

Vossler, C. and McKee, M. (2006) Induced-value tests of contingent valuation elicitation mechanisms. Environmental and Resource Economics 35:137-168.

Whitehead, J., Haab, T. and Huang, J. (2000) Measuring recreation benefits of quality improvements with revealed and stated behavior data. Resource and Energy Economics 22:339-254.

Whitehead, J. and Finney, S. (2003) Willingness to pay for submerged maritime cultural resources. Journal of Cultural Economics 27:231-240.

Willis, K. (1994) Paying for heritage: What price for Durham Cathedral? Journal of Environmental Planning and Management 37,3:267-278.

Willis, K. and Powe, N. (1998) Contingent valuation and real economic commitments: A private good experiment. Journal of Environmental Planning and Management 41,5:611-620.