The Production and Consumption of the Arts: A View of Cultural Economics

By David Throsby
Macquarie University, Sydney, Australia

I am grateful to Will Baumol, Tom Bradshaw, Tony Bryant, Graham Madden, Sir Alan Peacock, Pasquale Lucio Scandizzo, Mark Davidson Schuster, and two referees for helpful comments, and to numerous colleagues for supplying me with copies of their work.

I. Introduction

By most criteria the arts comprise a significant area of economic activity. In 1990, American consumers spent $5 billion on admissions to theater, opera, galleries, and other nonprofit arts events (more than on admissions to spectator sports), $4.1 billion on movie admissions, and $17.6 billion on books. Because of difficulties in defining boundaries around the arts industry, statistics on its contribution to GDP are problematical, but available data suggest that the arts (theater, music, opera, dance, visual arts, crafts, literature, community, and folk arts) account for a little under one percent of the United States GDP and a little over one percent of the civilian labor force. If the net is cast somewhat wider, defining the "cultural industries" as including the arts, motion pictures, radio and television, and printing and publishing, an aggregate value of output can be measured for 1988 of about $130 billion or 2.5 percent of GDP (National Endowment for the Arts 1992).

Likewise, support for the arts and culture in the U. S. through government and voluntary contributions amounts to a significant annual commitment of funds. Combined federal, state, and local government expenditure on the arts and museums in 1987 amounted to about $0.8 billion, and in 1990, 6.4 percent of charitable giving was channeled to arts, culture, and the humanities, yielding a total level of voluntary contributions in these areas of $7.9 billion in that year. Private markets in the arts, too, are of significant size. Looking at the international art trade, for example, we can note that the worldwide net sales of the two major art auction houses (Christie’s and Sotheby’s) amounted to $6.6 billion in 1989–90.1

Yet, despite the fact that production and consumption of art have been elements of human activity for longer than most of the phenomena that have en-

1 The year in which van Gogh’s “Portrait of Dr. Gachet” sold for $82.5 million, in the following year the downturn in the market more than halved the value of sales. Data from Policy Studies Institute (1991, pp. 52–53).
gaged the attention of contemporary economists, it is only relatively recently that serious work has begun to be undertaken in the area that has come to be known as “cultural economics,” or more particularly the economics of the arts. Major economists of the past have largely ignored this field. There have been some who have pursued a scholarly interest in art in addition to their economics, such as Adam Smith, who wrote serious essays on music, painting, dancing, and poetry, or John Ruskin, who was a leading nineteenth century critic of art and architecture. Others have been (and are) artists themselves of one sort or another, as Peacock (1980) has pointed out. It is also well known that Keynes was an active and passionate devotee of painting, the theater, and ballet, and was centrally involved in establishing the Arts Council of Great Britain, the principal vehicle for public support for the arts in that country to this day. But although Keynes spoke and wrote often on the importance of the arts in society, he never produced a major work in the field (James Heilbrun 1984).

It was not until 1959 that John Kenneth Galbraith tackled the confrontation between economics and art, in a lecture published in The Liberal Hour (1960). He looked at the economic situation of the artist and at the potential for good design to promote exports of American manufactures. At around the same time across the Atlantic, Lionel Robbins (1963) was the first British economist of modern times to analyze the economic role of the state in support for the arts and in financing public museums and galleries, followed soon after by Peacock’s (1969) interpretation of arts subsidies within the framework of traditional welfare economics.

But if contemporary cultural economics has a point of origin, it would lie in the pages of the book by William J. Baumol and William Bowen, Performing Arts—The Economic Dilemma, published in 1966. For the first time a major branch of the arts was subject to systematic theoretical and empirical scrutiny. To those economists who cared to read it at the time, it showed the extent to which their discipline could illuminate a new and challenging area of interest, using the familiar tools of economic inquiry.

Since that time the field of cultural economics has acquired an Association (founded in 1973), a journal (first published in 1977), an international conference (first held in 1979), and a growing body of literature. At a fundamental level it has had to begin by facing the question of defining art within the context of received economic theory. When asked to define jazz, Louis Armstrong is reputed to have replied, “If you gotta ask, you ain’t never going to know.” A similar conundrum has beset economists such as Galbraith (1974) and Kenneth E. Boulding (1985) who have grappled with definitional issues in seeking to bring art within the economic calculus.

The aim of the present article is to consider how far economics has gone over the last thirty years, or how far it might go, in exposing and analyzing the peculiar problems posed by production and consumption of the arts. The paper is not a literature review; hence, many significant works in the burgeoning literature of cultural economics have not been cited. Rather, it is an attempt to draw together and interpret some principal lines of research. In Part II of this article we look at the role of tastes as a fundamental driving force in the demand for and supply of art. Parts III and IV consider the markets for art works and for performing arts services respectively. In Part V the characteristics of labor markets in the arts are discussed, and in Part VI we look at the broad area of public policy...
toward the arts; questions of rationalizing state support for artistic activity have been among the major preoccupations of economists who have delved into this field since the writings of Robbins referred to above. Finally Part VII speculates on some lines of future development.

II. Taste for the Arts

The neoclassical view of tastes as given and as differing in some systematic but unmeasured way between individuals says nothing in particular about the arts. But theories of demand that consign taste to a residual status shed no light on the formation of tastes or on their profound influence on life-cycle consumption patterns.

The new consumer theory, as is now well known, suggests that tastes are similar between individuals, with variations in behavior caused by differing shadow prices of commodities produced according to household production functions in which material goods and services, including the arts, enter as inputs. But the arts can be further distinguished in this theory by their being addictive, in the sense that an increase in an individual's present consumption of the arts will increase her future consumption. Such a view can in fact be traced back to Alfred Marshall, who recognized that the taste for "good music" was an acquired taste that would increase over time with exposure. In the household production model, the relative consumption of the arts will rise over time, not because of a shift in tastes, but because the shadow price of the arts falls as experience, understanding and other human capital attributes associated with the arts are acquired. Following the work of George J. Stigler and Gary S. Becker (1977) and of Becker and Kevin Murphy (1988), a theory of rational addiction covering goods ranging from heroin (bad) to the arts (good) can now be seen as consistent with more general theories of rational choice.

Regardless of the theoretical underpinnings, it is clear that the endogenization of tastes in economic models is likely to be essential if any progress is to be made in explaining demand for the arts. Whether one calls it addiction or the cultivation of taste, the most relevant first step is to make taste for the arts dependent on past consumption, providing thereby a plausible explanation for the rightward shifting of the long-run demand curve. In turn, further elaboration of these ideas will enable a sharper definition of "cultural goods" and their relationship with other goods in the life-cycle consumption and investment patterns of individuals and households. As we have noted, cultural consumption can be interpreted as a process leading both to present satisfaction and to the accumulation of knowledge and experience affecting future consumption. Extension of the standard lifetime utility maximization models to include cultural as well as material goods can demonstrate the cumulative nature of the utility generated by consumption of cultural goods. In contrast to ordinary goods, the current consumption of cultural goods can be seen as adding to rather than subtracting from the process of capital accumulation over time, with obvious implications for growth theory and its predictions for optimal growth rates. Within such a model Scandizzo (1992) finds growth rates for cultural consumption that are chaotic and unstable, due to the fact that it is rational to respond to the peculiar cumulative features of culture in a seemingly irrational way.

2 In Chapter III of Book III of his Principles, Marshall wrote: "It is therefore no exception to the Law (of diminishing marginal utility) that the more good music a man hears, the stronger is his taste for it likely to become" (Marshall 1891, p. 151)
But, at the heart of the matter regarding tastes, both art lovers and artists themselves will argue that the inconsistencies, spontaneity, and unpredictability in behavior in the arts will always defy rational explanation, because these responses derive from notions of mystery, imagination, and the unfathomable creative impulse. While this is doubtless true of some individual actions and reactions in both production and consumption of art, aggregate data suggest otherwise. Tastes for the arts do seem to be moved by systematic phenomena, such that the aggregate behavior of consumers and of artists can be modeled in ways that are mostly consistent with economic theory. In the remainder of this article, we explore how much economics can say about such matters. In the following two sections we turn from a consideration of cultural goods in general to the properties of the specific artistic goods and services exchanged in the market place, looking first at art objects, and then at the performing arts.

III. Markets for Art Works

Original art objects (paintings, pieces of sculpture, and other artifacts) are, as a generic commodity group, characterized by a set of attributes that distinguish them from all other goods. They are created only by individuals. Every unit of output is differentiated from every other unit of output, an extreme case of a heterogeneous commodity. For the work of artists no longer living, supply is nonaugmentable. Art works can be copied but not reproduced, in the sense that ultimately there is only one unique original of every work of art. Paintings and sculptures provide clear consumption benefits to purchasers through their utilitarian characteristics as durable private goods. At the same time, artworks form part of the cultural capital of a nation or of the world (some more so than others), and thus have, to a greater or lesser degree, public-good characteristics, especially when they are acquired by galleries or collections for public showing. Because art works can be resold, and their prices may rise over time, they have the characteristics of financial assets, and as such may be sought as a hedge against inflation, as a store of wealth or as a source of speculative capital gain.

For the purposes of economic analysis of market behavior, the essential features of artworks that animate the utility functions of buyers can be captured by distinguishing between art as decoration, i.e., art providing immediate consumption services through its aesthetic qualities, and art as asset, i.e., art providing financial services through its potential for price appreciation. A Lancastrian interpretation of demand is clearly appropriate, where “artistic” characteristics of works such as size, color, and other aesthetic values enter the utility function for art as decoration, and riskiness, expected rates of return, and other financial variables influence demand for art as asset. In such a construction of utility and demand for art objects, the reputation of the artist can be seen as playing a significant role. It is the primary correlate with expected rate of return in assessing the attractiveness of the artist’s work as a financial instrument; at the same time, the place accorded an artist by historical or contemporary aesthetic judgment is indicative of the contribution the artist’s work makes to the appeal of art for its own sake, as decoration, or as a store of cultural value. Any plausible theory of demand for art works must evidently capture the role of an artist’s real or imagined position in the firmament of artistic creation in linking the various motives of buyers seeking particular works in the marketplace.

The structural features of the art mar-
ket in which these demands are expressed vary somewhat between countries. But generally the art market can be characterized as comprising a series of linked submarkets. At the lowest level, sometimes referred to as the "primary" market, unorganized individual artists provide works to galleries, local art fairs and exhibitions, small dealers, and private buyers; this market is highly decentralized. At the "secondary" level, in markets located mostly in significant cities where art is traded, such as New York, London, Paris, and Sydney, established artists, dealers, and public and private collectors circulate works by live artists who have managed to make the transition from the primary market (for example, by succeeding in having their work taken up by a recognized commercial gallery or dealer, or purchased by a respectable public art museum), and works by dead artists whose names are still recognized. Finally, at the highest level, an international market exists in which the major auction houses are the main players, and where the works of artists of the highest reputation are traded at prices that frequently make headline news.

Market structure varies between these levels. At the lowest level, competition is widespread. Universally there are more artists and would-be artists than there are buyers interested in acquiring their work, so prices are low. Although most serious painters undergo significant periods of training to qualify as professional artists, as a group they lack the credentialing mechanisms of doctors and attorneys, and are thus unable to exert any supply-side power in this market in order to restrict competition or to raise prices.

The secondary market is considerably more concentrated on both buying and selling sides. Only a relatively small proportion of artists make the transition from the primary market, so that in any one center there is only a restricted number of artists represented in the market at any one time. Given the resources required to enter the market on the buying side and the somewhat arcane nature of the product, the numbers of buyers, both individual and institutional, is similarly constrained. Considerable market power can be exercised by the galleries and dealers who handle most of the sales in this market. Not infrequently a gallery owner or art dealer can tie up the work of a particular artist so that he can behave effectively as a monopsonist in dealing with the artist and as a first-degree monopolist in dealing with buyers. The tendency towards increased concentration at the center of this market is heightened by the fact that, over time, galleries and dealers who handle the most successful artists attract both artists and buyers from other allegiances to their stable; such key players in the market can exert a powerful influence on the rise and fall of artists' reputations, and hence on the future price expectations for their work (Leslie Singer 1981).

With some market participants holding considerable power, it is not surprising that markets for artworks can show evidence of information asymmetry. A particular manifestation of this phenomenon is the possibility of fakes or forgeries, which are akin to "lemons" in the used car market, where the quality of goods is known to sellers but not to buyers. For example, in the market for nonfigurative painting, following Tibor Scitovsky's discussion of pattern-complex art in The Joyless Economy (1976), Roger McCain (1980) has suggested that the inability of some buyers to see the higher-order patterns of such art opens the way to fraudulent operators to misrepresent patternless work as something better, i.e., to create a market for lemons. The effect of such dishonesty may be, as George
Akerlof (1970) showed in the case of automobiles, to drive out business altogether. While such extreme outcomes have not occurred in the arts, the tendency exists, and it might then be argued that correction of the informational shortcomings in such markets could provide a rationale in principle for subsidy of pattern-complex art through some form of collective intervention.

Ultimately, the major focus in looking at the economics of the art market is likely to be on prices. Two questions are of interest. How do rates of return on investment in art compare with returns elsewhere? and What are the main determinants of the prices of art works?

Evidence on the comparability of rates of return between art and other forms of asset holding suggests fairly decisively that the average rates of return to investment in the secondary and tertiary art markets are below those yielded by comparable financial assets. For example, W. Baumol (1986), using Gerald Reitlinger's (1961) data spanning more than 300 years, computed a real rate of return of 0.6 percent per year for a sample of paintings resold twice or more over the period 1652 to 1961, compared with a real annual rate of return of about 2.5 percent on safe British government securities over the same period. Similarly, for 1946–1968, John Picard Stein (1977) estimated the nominal appreciation of paintings at 10.5 percent per year, compared with a nominal annual return in the stockmarket of 14.3 percent. Bruno S. Frey and Werner W. Pomperehne (1989b) found that, while certain works of artists such as Cezanne, Gaugin, van Gogh, Monet, and Renoir have performed much better than others over the long term, the average real rate of return on paintings was only 1.6 percent per year in the period 1950–1987, based on a sample of 415 transactions; they estimated that the real annual rate of return on financial investments in the U.K., the U.S., Germany, or France over that period was around 2.4 percent. Finally, for a portfolio of contemporary U.S. art held over the period 1987–1989, Singer (1990) found a negative price differential against alternative financial instruments of just over two percent.

Most of these writers point to the greater variability in the return to holding paintings rather than financial assets. Most also suggest that the observed spread between average rates of return on art and other assets, after adjustment for risk, is a measure of the consumption value of art works to their owners. This hypothesis implies that if those who buy art for nonpecuniary reasons such as aesthetic pleasure, status, or prestige value also hold stocks and bonds, there cannot be market equilibrium unless the difference in return, adjusted for risk, equals the value of the nonpecuniary benefits provided by the art. Because the motives of individual buyers in acquiring art are likely, as noted earlier, to range across a spectrum from demand for art purely as decoration to demand for art purely as asset, such a hypothesis about average differentials is likely to cover a wide variation in individual behavior.

To answer the second question posed above would require a properly articulated model of both demand and supply behavior in the market. Frey and Pomperehne (1989a, Ch. 6) report a price equation that could be interpreted as a reduced form of such a larger model, wherein demand for artworks is explained by price, aesthetic quality, consumer income, financial market characteristics, and other variables, and supply is determined by prices and costs of production. In their estimated equations derived from data on sales of work by 100 top American and European artists over the period 1971–1981, price is shown to be influenced significantly from the sup-
ply side by production costs and size and type of work, and from the demand side by consumer income, aesthetic evaluation, and rates of return on other assets.

Intrinsic to their estimates is the derivation of a measure of aesthetic quality. Frey and Pommerehne adopt a measure that assigns points to the evaluation by art experts of various aspects of an artist’s work. The resulting measures can be used to rank artists in order of the “quality” of their output as assessed by the cognoscenti of the art world. Following the earlier work of Friedrich Schneider and Pommerehne (1983a), these authors show that these aesthetic evaluations are not random, but are systematically related to measurable attributes of the artist’s work history, including school of work, career stage, and past achievement.3 Even so, there is no convincing evidence to date that art “experts” can consistently outperform the market in judging the potential appreciation of currently traded art.

Artists in turn have an interest in advancing their careers, if only for the recognition that such progression provides for the quality of their work. Artist’s careers can be seen as a series of stages related to the stratified market structure discussed earlier, whereby advancement can be seen as a step up from one market level to the next. Buyers, in turn, might be seen as trying to pick those artists who are most likely to make a transition to a higher stage, in the expectation that their prices will rise as a result. W. Baumol (1986) and Singer (1990) have examined the probabilities that a given artist will make a successful transition; not surprisingly the odds are very low. The consequence of this for buyers is that playing the art market is little more than a game of chance. For artists, it means that expected returns from art practice, on average, are likely to be low, a matter to which we return in Section V below.

IV. The Performing Arts

A. Demand

A demand function for attendance at live events in theater, opera, dance, and music would be expected to contain own price, price of substitute entertainments, consumer income, and quality characteristics of performances as explanatory variables. The diversity of the product, and the discrimination of consumers in deciding their attendances at particular performances, suggests that the qualitative characteristics of events (who is appearing, what the critics have said, what work(s) are being performed) are likely to dominate price in determining demand. Furthermore, consumption of the live arts is highly time-intensive, indicating that the price of leisure time is likely to be more influential in determining demand than the ticket price itself.

In examining the determinants of demand in the performing arts, a distinction can be drawn between demand for immediately accessible entertainments such as popular musicals, live entertainers, circuses, and so on, and demand for what are sometimes referred to as the “higher” performing arts, including opera, “serious” drama, classical music, jazz, classical and modern dance, and performances in any art form that are experimental or avant-garde. In the former case, the less discriminating nature of demand means that substitutes are more readily available, and hence own-price responsiveness is likely to be greater. In the latter case, where consumption reflects the sort of acquired taste discussed in Section II

3 Nevertheless, in a dynamic world, it is not apparent that aesthetic judgments are made independently of the price of an artist’s work; Thorstein Veblen pointed out as long ago as 1899 in The Theory of the Leisure Class that a person’s assessment of the beauty of an object is likely to be influenced by its cost. If so, the simple price determination models referred to above will need to be respecified.
lower price elasticities might be expected among established consumers, for whom qualitative characteristics of performances are likely to be decisive. With regard to income, on the other hand, demand for the performing arts would be expected to be somewhat more responsive. In the first place, some performing arts can be seen as luxury items, associated with social status and the desires of the wealthy for conspicuous consumption; the phenomenon of “first nights,” at which attendees are more interested in looking at each other than at what is on the stage, is common in many countries. Second, increasing consumption of the arts over the long term is linked with education, which in turn is a significant determinant of income. Hence a positive association can be established via this path between attendance at the performing arts and the income of consumers.

Empirical studies of demand for the performing arts undertaken over a number of years have been broadly consistent with the above observations. At the outset, several early studies identified performing arts audiences as being of significantly higher educational, occupational and income status than the community at large (W. Baumol and Bowen 1966; Ford Foundation 1974). Subsequently time-series and cross-section demand studies have built up a clearer picture of the importance of various factors affecting demand for the performing arts. Most interest has centered on price. One of the earliest studies, that of Thomas Moore (1966), is typical; he found a price elasticity of demand for Broadway theater of \(-0.33\) to \(-0.63\), and this result has been broadly confirmed by a number of subsequent studies in different art forms and different countries. It is always true, however, that demand for groups of companies is likely to be less elastic than demand for the output of a single firm or even for a single performance. Thus what is true for the industry may not be true for the individual enterprise. Marianne Victorius Felton (1992), for example, found price elasticities for the major U.S. orchestras of around \(-0.6\) and for major ballet and opera companies ranging from about \(-0.1\) to about \(-0.6\), whereas the elasticities facing individual companies within these groups were found to be significantly larger, with some greater in absolute terms than \(-1\).

Although most (but not all) studies have identified a significant positive coefficient on consumer income in estimated demand equations, the corresponding elasticity estimates have varied above and below \(1\). Because, as noted above, live arts consumption is time-intensive, gains in attendances over the long run due to increasing incomes are likely to be offset to some extent by the increasing price of leisure. Glenn Withers (1980) confirmed this proposition using data covering all U.S. performing arts for the period 1929–1973; he found a “pure” income elasticity of around unity, composed of a “full” income effect (imputing leisure time as part of income) of \(2.7\), offset by a leisure price effect of about \(-1.6\).

Cross-elasticities, too, have been found to be significant, given the wide range of possible substitutes for the performing arts, not just from outside the industry (movies, reading, watching television) but also from within. In the latter regard, James H. Gapinski (1986) examined competition between theater, opera, dance, and symphony orchestra performances in London through the decade of the 1970s, and found demand elasticities with respect to substitute price within this group of companies to be always positive and ranging up to 2 and above.

Finally, because of the difficulties of measuring the quality of performances,
most demand studies in the performing arts have had to consign the effects of variations in quality to residual status. An exception is the experimental study by Throsby (1983) which identified several quality characteristics of live theater performances, including standard of script, acting, and production, and found consumer demand to be inelastic with respect to ticket price but strongly responsive to variations in expected quality.

B. Production and Cost

Performing arts firms, like other productive enterprises, combine labor and capital with given technology to produce output. While the inputs employed by the firm are clearly enough defined, the specification of output is not so straightforward. Superficially, the output of a symphony orchestra is a concert, or that of a theater company is a performance of a play. But there is a difference between the manufacture of a production—the setting up, the rehearsing, the integration of various elements into a single interpretation of a score or a script—and a performance, which is a repetitive process that reproduces the same “output” over and over again. Furthermore, interpretation of output in these terms does not capture the fact that the purpose of a performance is to provide a “cultural experience” for an audience which itself may be thought of as the final product; hence, an alternative measure of the output of a performing company can be taken as the number of attendances over a given time period, an approach similar to that used in defining the output of a university as the number of enrollments, or that of a hospital as the number of patients. Nevertheless, a distinction should perhaps be made in the performing arts between output produced, measured as the number of seats available for sale over a given period (equivalent as an output variable to the number of performances), and output sold, measured as actual numbers paying to attend.

When performances are given in closed venues of fixed capacity, output of the performing arts can be seen as an excludable local public good, nonrival in consumption up to the point where a capacity constraint is met. At a more general level, output of the performing arts can be characterized as a mixed good, with joint production of a private component enjoyed by individual attendees and a public-good component deriving from the value of the arts and culture to society at large.

Interpreting output as number of attendances, and ignoring administrative and other inputs not related directly to artistic production, a simple model of the performing arts production process can be set up as follows. Suppose a performing company (theater group, opera company, musical ensemble, etc.) can vary the number of productions and the season length for each production within a given accounting period, say one year, over which time venue capacity \( v \) is fixed. Let \( y_{ij} \) = attendances at the \( i \)-th performance of the \( j \)-th production, \( i = 1, \ldots, m_j; j = 1, \ldots, n; y_{ij} \leq v \). Define \( L^s \) and \( K^s \) as the set-up labor and capital required to mount a production (rehearsals, construction of sets, etc.) and \( L^r \) and \( K^r \) as the operating labor and capital employed per production (actors’ and attendants’ labor, consumables, etc.). Then, assuming output price is constant, we can write

\[
y_j = \sum_i n_{ij} = y_j(L^s_j, K^s_j, m_j, q_j) \tag{1}
\]

with

\[
m_j = m_j(L^r_j, K^r_j) \tag{2}
\]

where \( q_j \) summarizes quality variables describing the \( j \)-th production; note that quality in this context may be measured...
in terms of the lavishness of the production, in which case it is clearly not independent of \( L_s \) and \( K_s \), or in terms of other characteristics such as type or period of work being performed, in which case a strong influence on \( y \) would be likely. In this system \( m_j \), the number of performances of the \( j \)-th production, may be seen as an intermediate product, being interpreted as an output in (2) and an input in (1). For a given production, the relationship portrayed in (2) is likely to be homogeneous of degree one, with fixed factor proportions, although these proportions will differ between productions. In (1) it would be expected that \( \delta y_j/\delta m_j > 0 \), \( \delta y_j/\delta m_j^2 < 0 \), that is, extending the season length of a production is likely to cause attendances at the margin to decline, indicating diminishing marginal products of operating labor and capital in (2) with respect to eventual output in (1). However, the marginal products of set-up labor and capital are less easy to predict; it does not immediately follow that plays with more complicated sets, or orchestras that have had more rehearsals, will draw larger audiences. For the short period represented by (1) and (2), possibilities for factor substitution are very limited.

The relationships depicted in (1) and (2) could be estimated from data for one or a group of companies in a particular art form for a single period of time. Over the longer run, time-series data for a given company, or pooled data covering a group of companies, might be used to estimate an extended version of this system wherein venue capacity is also variable. In such a situation the scope for factor substitution between productions, or between more/fewer productions and shorter/longer runs, is considerably greater. A long-run production function might be specified as:

\[
y_t = y_t(L_t', K_t', q_t, v_t)
\]

where \( y_t \) = total attendances for a company in period \( t \), and where the factor inputs are measured as aggregate quantities in the different categories over the time period considered. A proxy for \( q_t \) capturing at least some aspects of aggregate quality in the period would be required.

A simpler alternative to (3) for which data might be more readily obtainable is provided by

\[
y_t = y_t(n_t, \tilde{m}_t, q_t, v_t)
\]

where \( n_t \) is the number of productions for a company in period \( t \), and \( \tilde{m} \) is the mean number of performances per production. The venue size \( v \) in (3) and (4) could be varied within a given location (for example by installing new capacity) or by performing in a series of venues of different sizes.

Cost functions can be derived for the above production model at given factor prices. In the very short run, relating to a single performance, or to a fixed season length for a single production, the interpretation of output as a local public good means that virtually all costs are committed to the first unit of output, and thereafter marginal costs are close to zero. In the short run the firm will typically incur relatively high fixed costs and relatively low variable costs, indicating declining average total costs. In the longer period, performing companies can be expected to have access to significant scale economies in long-run production and cost functions.

Empirical evidence on the production and cost relationships of firms in the performing arts is sparse. Several production functions consistent with the above propositions were obtained for some Australian companies by Throsby (1977), while Gapinski (1980, 1984), using Cobb-Douglas and transcendental functional forms with U.S. and British data, has demonstrated the basic regularity of pro-
duction processes in the major performing arts. More attention has been paid to costs, partly because a direct relationship between costs and output that bypasses the underlying production function is simpler to estimate, and partly because data have been more readily available. For example, studies by Steven Globerman and Sam H. Book (1974) and Mark Lange et al. (1985) found evidence of scale economies with respect to attendances and performances among theater, opera, and dance companies and symphony orchestras.

Further research into production and cost relationships in the performing arts will require more carefully articulated models than the simple constructions outlined above. It will be necessary to search for more appropriate measures of output and to make better use of existing measures, for example by controlling for demand effects when paid attendances are used as output; to account for the peculiarities of some labor contracts in the performing arts; to allow for the non-substitutability between some types of labor (a tenor cannot stand in for an ailing soprano, nor can he be retrained as one); to find quality measures that go beyond crude budgetary proxies; and to model explicitly the differences in production and cost structures between performing media, from small-scale drama to grand opera. Equally, better data will be essential to further empirical progress. If advances can be made in these respects, one possible benefit could be that future work in the economics of the performing arts may be more useful to managers of performing companies than has been the case in the past.

C. Firm Structure and Behavior

A look at the performing arts industry in most countries shows a range of firm types operating side by side in the marketplace. The three common forms observed are for-profit proprietary companies, private nonprofit firms, and publicly owned and operated firms organized on a nonprofit basis. In the U.S., data from the Census of Service Industries in 1987 indicate that 22 percent of performing arts organizations were nonprofit (tax exempt) in that year, accounting for 27 percent of revenue, with the remainder being for-profit (taxable) enterprises (National Endowment for the Arts 1992). These figures indicate a growth in the proportion of activity in the performing arts accounted for by nonprofit firms; corresponding figures for 1982 are 19 and 25 percent respectively.

What explains this observed distribution of firm types in the performing arts? In particular, why is the nonprofit form prevalent in this area, especially in certain subsectors of the industry? Several theories have been put forward to explain the appearance of nonprofit firms in a market economy. Burton Weisbrod (1977) has suggested that voluntarily established and financed nonprofit organizations represent a response to an unsatisfied demand for public goods, where consumers consider the level of government supply of such goods to be inadequate. Alternatively, Henry Hansmann (1980) has argued that consumers may prefer nonprofit to profit-seeking firms when buyers lack information or cannot adequately judge the quality of goods or services offered, such that ordinary contractual mechanisms do not provide buyers with protection against exploitation by producers. Other hypotheses relate the spread of nonprofit firms to the availability of government grants in certain areas or to the growth in charitable giving encouraged by tax exemptions to donors.

A nonprofit firm can be characterized as one which (a) is forbidden by law to distribute any surplus to its owners; (b) is exempt from corporate income taxes; and (c) may receive contributions that attract tax exemptions for donors.
While each of these explanations can be applied individually in greater or lesser degree to the performing arts, the most direct explanation of the existence and distribution of different types of firms in this area comes from an examination of market characteristics and consumer motivation. Specifically, for some areas of the performing arts, the demand curve facing the firm is above average cost over a reasonable range of output and hence profits are possible, though by no means secure ex ante because of demand uncertainty. These areas of the performing arts correspond broadly to the popular entertainment end of the spectrum described earlier. Firms in this sector tend to run a single production over a long season and hence exhibit relatively low average fixed costs, with little difference between average and marginal cost. Requirements for third-degree price discrimination are met (separable markets, differing elasticities, resale not possible between markets), and increased profits can be pursued by these firms via differential ticket prices for different groups of seats. In these conditions the profit-seeking proprietary form arises as the most appropriate to this sector of the market.

In other areas of the performing arts, however, the average cost facing the firm is everywhere greater than demand, with no ticket price that can cover costs, even with differential ticket pricing for different parts of the house. Such a circumstance is characteristic of the “serious” or “high” arts as earlier defined.5

Under these conditions, if production is to occur at all, additional revenue will be required to close the gap between income and expenditure. These funds will have to come from voluntary contributions, or from government subvention, or both. Because neither private nor public donors are likely to contribute to a firm where there is a possibility that funds provided may simply add directly or indirectly to profits taken by the firm’s owners, it is most appropriate for enterprises operating in this market sector, whether private or public, to be incorporated on a nonprofit basis.

Nevertheless, it still remains to be explained why people will donate to such firms, because as a rule neither consumers nor governments are willing to support unprofitable enterprises in other areas. The answer in the case of the arts appears to lie chiefly in the social worth of the firm’s output as perceived both by voluntary contributors and by governments. We return to this question when we consider the rationalization of public support for the arts in Section VI below.

Modeling the behavior of nonprofit performing arts firms raises some intriguing questions. If the major argument of the utility function of such a firm or its managers is not profit, what is it? In fact, consistent with our earlier discussion, it can be proposed that the firm’s utility will be positively related to the level of attendances and to quality characteristics of output over a given period, and that the utility function will be maximized subject to a budget constraint requiring zero net revenue. This model assumes no separation of ownership and control of the firm, such that the “firm’s utility function” is identical with that of its owners and managers. A simple model of such a firm can be put forward as follows.

Imagine a performing company characterized by the production conditions described in (1) to (4) above. Let the inverse demand function be \( p = p(y) \). Without debating the means of measuring quality, let us simply assume a scalar \( q \) is available representing, in the eyes of the firm and its supporters, a composite index of out-

---

5 The dividing line between the two groups is not entirely clear cut; occasional examples may be found of profit-seeking firms in the high arts (e.g., some literary publishers), or of nonprofits producing popular culture (e.g., some local theater groups).
put quality such that increasing $q$ unambiguously denotes higher quality. Suppose that the level of donations and grants received by the firm, $g = g(q)$, is an increasing function of output quality, i.e., donors will be prepared to give more to assist the firm achieve higher quality levels. Further, the firm gains additional gross revenue $h$ from merchandising activities such as provision of restaurant and bar services, sales of books, T-shirts, etc.; such revenue would be expected to be an increasing function of $y$. Let $c = c(y,q)$ be the total cost function of the firm covering all activities, artistic, administrative, and merchandising. Then the firm's decision problem is to select the price of output, $p$, and the quality of output, $q$, to maximize

$$U = U(y,q)$$

subject to

$$p(y)y + g(q) + h(y) - c(y,q) = 0. \quad (6)$$

The first-order conditions describing the constrained optimum can be written:

$$\frac{U_y}{\lambda} + p_y + y + h_y = c_y$$
$$\frac{U_q}{\lambda} + g_y = c_q$$
$$p(y)y + g(q) + h(y) = c(y,q) \quad (7)$$

where subscripts indicate partial derivatives and $\lambda$ is the multiplier on the constraint (6).

The first optimality condition in (7) indicates that under these assumptions the utility-maximizing nonprofit performing company will set its average seat price where marginal revenue derived from ticket and merchandise sales is less than marginal cost, the difference being accounted for by the extra utility that the firm gains from increasing output. Similarly, as shown in the second condition in (7), the firm gains additional utility through increasing quality, and will therefore devote resources to increasing quality beyond the point where marginal grant income brought in by a unit increase in quality equals the marginal cost of securing that increase.

Several writers have modeled the behavior of nonprofit performing companies, with particular emphasis on pricing decisions, including Throsby and Withers (1979) and David Austen-Smith (1984). The most thoroughgoing theoretical treatment is due to Hansmann (1981), whose model permits analysis of the polar cases of the firm as audience maximizer, quality maximizer, and budget maximizer. Hansmann argues that nonprofit performing arts firms engage in a process of voluntary price discrimination, whereby they induce patrons to pay over voluntarily some or all of the consumer's surplus they enjoy through attending a performance. This implies that, insofar as the firm's pricing behavior is concerned with the maximization of revenue (given the achievement of other objectives), it would set prices to maximize the sum of ticket revenue and voluntary donations, rather than just ticket revenue alone.

Much scope remains for further theoretical and empirical work on the behavior of nonprofit firms in the performing arts. Some potentially fruitful lines of inquiry include the following. First, casual empiricism suggests strongly that as we move closer to the "serious" arts end of the spectrum, quality motives of firms become more important. A number of firms in the high arts appear to operate by setting ambitious quality goals in terms of repertoire, casting, etc. for a forthcoming season, and then assessing whether expected ticket and grant revenue will be sufficient to meet these goals. If not, the quality target is revised downwards until a feasible proposal is reached.

Note that most of this discussion of nonprofit performing companies is also applicable to art museums; see Martin Feldstein (1991).
A model which gave primacy to quality and which explored further the interdependencies between financial variables and quality would be likely to have some explanatory power; a prerequisite to such a model, however, would be some progress in resolving outstanding problems in interpreting and measuring quality in the performing arts.

Second, the essentially deterministic nature of most models so far proposed is a long way short of a reality where the occurrence of success or failure is subject to substantial uncertainty and seems often to be merely quixotic. The development of a stochastic approach to decision making in this field, where firms might be seen as trying to maximize the probability of success of meeting certain objectives subject to fuzzy constraints, would appear to hold promise. Third, in a related vein, dynamic extensions to existing models would be useful to capture the many interperiod and developmental questions that are ignored in static formulations. For example, when the level of grants is dependent on certain strategies adopted by the firm, there may be a lag between firm behavior and an induced donor response. Such a proposition is embodied in a model developed by Austen-Smith and Stephen P. Jenkins (1985), in which grant income in period \( t + 1 \) is dependent on output and financial performance in period \( t \), and the zero profit constraint is not binding in every period.

Finally, we have referred above to the “firm” as if it were, like the neoclassical construct, indistinguishable from its manager(s). In fact, nonprofit firms in the performing arts may be structured in a range of ways, from loose cooperatives to complex corporate systems involving separation of policy-making and executive functions. Managerial utility models of firm behavior of the Williamson type, in which managers pursue their own goals of personal advancement rather than objectives more identifiable with corporate well being, have been applied in some nonprofit areas such as health care, but have found little favor in the arts. Paul J. DiMaggio (1987) argues that economists have accepted somewhat too readily the notion of artists and arts managers as being entirely unselfish and driven only by an ethic of cultural achievement on behalf of the enterprise, which suggests that analysis of conflicts in goal formulation and pursuit within nonprofit arts organizations could open up a richer prospect than has been hitherto expected.

One proposition worth pursuing is that the individual welfare function of artists and of arts managers in nonprofit companies might include a variable denoting peer group reputation. Maximizing one’s standing among one’s peer group, which is clearly easier to achieve when profit maximization is not the sole or main object of the firm, is, however, not necessarily inconsistent with long-run wealth maximization, because prestigious awards given by artists to artists (such as the Booker Prize and other literary awards that are judged largely by writers) are often treated by consumers as indicators of the quality of the product.

D. Technology and the “Income Gap”

So far in this section there has been no mention of technological progress. In a pure sense, productivity improvements in the live performing arts are not possible in the same way as they are in, say, manufacturing. The physical output per unit of labor input in the latter sector has risen by several thousand percent over the last two hundred years, whereas the labor required to give a live performance of a Mozart string quartet has not changed in the slightest over the same period. This fact, first remarked formally in 1965 by Baumol and Bowen, has given
rise to one of the longest-running debates in the area of cultural economics: do the performing arts suffer from a “cost disease”?

In their original exposition, Baumol and Bowen (1965) used a two-sector model comprising a productive (manufacturing) and an unproductive (arts) sector to argue that wage rises in the economy as a whole could be offset in the productive sector by productivity gains, but would simply result in rising relative costs in the unproductive sector. If demand were static, or rising too slowly, the outcome would be an inexorably widening gap between revenues and costs for firms in the performing arts, requiring ever-increasing subventions from governments or private philanthropists to keep such firms alive. In the 25 years following its initial presentation, this proposition has been widely seized upon in a number of countries as spelling doom for the live arts unless governments intervened, and both government funding agencies and the companies they support have made much of the cost-disease hypothesis in pressing for ever more generous subsidies.

Over this period, too, there has been some clarification and reassessment of the original proposition and the assumptions on which it was based, and some empirical work has looked more closely at what cost trends have actually occurred. Essentially it can now be said that, while the basic logic of the cost disease is, in its own terms, unarguable, the causal chain linking certain characteristics of production of the live arts to a widening income gap for performing companies is by no means as inexorable as many have supposed, for a number of reasons.

First, technical change can occur within the performing arts. New venue design, improved sound systems, better lighting, and so on can bring live performances to larger immediate audiences. More importantly, broadcasting and recording technology can extend consumption of a single performance thousands or even millions of times. While there are certainly qualitative differences between live and reproduced performances for both performers and audiences, the scope for media dissemination to extend audiences and to improve the financial position of many performing companies is clear.7

Second, there have been substantial opportunities for adjustment in factor use by firms in response to cost pressures, for example by improving technical efficiency at existing levels of output; by reducing inputs of some factors, such as performing plays that have simpler sets or smaller casts; by other quality changes, such as performance of fewer contemporary works for which copyright fees have to be paid; and by adopting other revenue-improving strategies, such as merchandising activities and better marketing and packaging. Of course, it is clear that such actions by a firm may have significant short- and long-run implications for the quality of its output, a variable which, as noted earlier, is likely to be one of the primary arguments of its utility function.

Third, the factor price adjustments at the core of the cost-disease hypothesis have not occurred to the full extent suggested by the two-sector model. The realities of labor markets in the arts indicate that performing arts firms may not have to match increases in wages generated elsewhere in the economy in order to attract or retain an adequate supply of

7 However, Hilda Batumol and William Baumol (1984a) point out that the media themselves are subject to a similar cost disease; for example, the costs of production of television drama have risen substantially relative to the costs of transmission over the last thirty years as a result of similar pressures to those affecting the live arts.
labor. As a result, a gradual erosion in the relative earnings position of workers in the performing arts is likely, rather than a full transmission of wage rises in other sectors.

Fourth, there are grounds for hope among performing companies that rising consumer incomes will go some way towards offsetting the negative effects of ticket price rises forced through cost pressures, despite the somewhat ambiguous conclusions drawn earlier as to the magnitude of the long-run income elasticity of demand for the performing arts. Although an inexorable rightward movement in the demand curve cannot be hailed as the unequivocal savior of the performing arts, at least the shifts in expenditures and tastes induced by secularly rising incomes and educational levels within the community are in the right direction.

A number of empirical analyses of trends in costs and revenues for performing companies have examined aspects of the cost-disease phenomenon following Baumol’s and Bowen’s original research, including studies by Dick Netzer (1978), Peacock, Eddie Shoesmith, and Geoffrey Millner (1982), Hilda Baumol and Baumol (1984b), and Samuel Schwarz (1986). These studies have generally pointed to production-side adjustments by performing companies over time and have found little evidence of differential rates of inflation in the performing arts compared with other sectors of the economy. In fact, Baumol and Baumol (1980) have noted the relatively slower rate of cost increase in the live arts that has occurred during periods of high general inflation such as the 1970s. They attribute this outcome, which is consistent with earlier historical periods, to the fact that in inflationary times there is an erosion in philanthropic support, together with a heightened money illusion that restrains rises in ticket prices, forcing performing companies to concede lower wage rises, reduce labor inputs, change the repertoire, and adopt other cost-reducing strategies such as those mentioned above. These studies have also shown that the combined impacts of production adjustments, increased demand, and generally rising levels of unearned revenue have countered any tendency towards a secular rise in deficits among performing companies, suggesting that although the cost disease will doubtless continue to present the performing arts with difficult problems, it is unlikely to be terminal.

V. Labor Markets for Artists

Nowhere in cultural economics might the differences between the arts and other industries be expected to be so marked as in the labor market area, especially with regard to the functions of artists as economic agents. The popular image of the artist, whether actor, musician, painter, or poet, as a flamboyant bohemian devoted only to realizing a creative dream and oblivious to financial concerns is a portrayal far removed from the philistine economic man who lies at the heart of conventional economic models of labor market behavior. Nevertheless, despite these apparent differences from the supposed norm, routine economic analysis can be readily applied to study the supply of artistic labor, the levels of income earned by artists, and the distribution of income across the artistic labor force.

A prior question, however, is a definitional one: Who is an artist? For the purposes of theoretical economic analysis, a categorization based simply on the willingness or capacity of individuals to sell their artistic labor or its immediate products should provide a sufficient initial definition that could then be further refined to suit specific contexts. However, for empirical work, the identification of artists or of the arts labor force presents
greater difficulties. Census classifications of “artists” are often at odds with the categories used by researchers in the field, a problem compounded by the fact that many artists hold more than one job, and their employment outside the arts may provide the basis for their census designation. Empirical work based on survey instruments must determine appropriate filters to identify artists in a labor market where external screening devices such as formal credentialling are largely absent (Gregory Wassall and Neil Alper 1985).

A. Labor Supply

Acknowledging these difficulties, what can be said about professional artists as a group and about how they supply their labor to the market place? Census data indicate that U.S. artists are predominantly male (57% in 1991) and white (92%)8, and have comprised an increasing proportion of the workforce over the last forty years, rising from 0.73 percent in 1950 to 1.31 percent in 1990 (National Endowment for the Arts 1992). A distinction can be made between the working conditions of performers (actors, musicians, dancers), who mostly are temporary employees without job security when they work for money, and creative artists making things for sale (painters, sculptors, craftspeople, writers, composers), who work mainly as independent self-employed contractors. These employment characteristics mean that the performing arts is the only area of the industry where labor unions have had a significant impact. Overall, about one-quarter of artists are members of a labor union. Artists spend long hours at their arts work, some part of which may be unpaid (performers practicing to maintain their skills, other creative artists spending time producing work that is not for sale). For example, in Australia about 70 percent of artists work longer in total than the standard full-time working week. In the U.S. only 20–25 percent of artists work more or less full-time at arts work, and it would appear that a majority of those working less than full-time at their chosen profession would prefer to spend more time at the arts but are deterred from doing so by the need to earn an income elsewhere.

These characteristics suggest that theories of labor supply in the arts will need to account for multiple job-holding by artists, and in particular for the differences in their motivations in supplying work to the arts and nonarts labor markets. The primary desire to create art as a principal occupation must be recognized as the essential driving force behind an artist’s labor supply decisions. In this respect artists may be seen as similar to academics, researchers, and other professionals where nonpecuniary motives relating to work satisfaction exert a significant influence on patterns of time allocation. Nevertheless, artists as a group differ by virtue of the fact that their professional creative work alone is, in the majority of cases, unlikely to generate a living wage over a reasonable period of time, either because the hourly earnings are too low and/or because remunerative work opportunities are not available.

Empirical work on artists’ employment, based mostly on sample surveys, has yielded results consistent with these observations. For example, Joan Jeffri (1991), in a survey of over one thousand U.S. painters, found that pursuit of art at the expense of income was widespread; of her sample, 70 percent had turned down lucrative opportunities that were not artistically fulfilling on more than one occasion. At the same time, the need for artists to seek income-supporting work

8 These proportions are found to be somewhat smaller in surveys targeted specifically at practicing professional artists.
outside the arts is apparent; Wassall and Alper (1992) reviewed a number of surveys which document the extent of multiple job-holding by artists, including their own 1981 survey of over 3,000 New England artists which found only 24 percent of artists holding no nonartistic job. They observed that as artists’ careers solidify as they grow older, more time is devoted to arts work, though it is not clear whether this results from greater career success of existing artists or from the less successful dropping out. Throsby (1992) estimated labor supply functions for Australian artists with arts and nonarts wage rates as explanatory variables; this study found evidence of artists supplying the nonarts labor market only up to the point where an adequate return was received to support their primary artistic work, with attention being switched thereafter to supplying primarily the arts labor market.

B. Earnings Functions and Career Choice

If artists’ labor supply decisions must be modeled with due regard to the peculiarities of artistic work, it follows that the construction of earnings functions and models of career choice in this area will also need to account for these features. To begin with, it will be essential for earnings functions to distinguish clearly between arts and nonarts sources of income; indeed, within income derived from the arts, it may be desirable to differentiate further between income from work as a pure creative artist, such as from acting or selling paintings, and income from other arts-related work, such as from teaching within an artist’s area of practice. Earnings equations can be readily postulated along the lines suggested by human capital theory, with education, on-the-job experience, and other factors as explanatory variables alongside the usual sociodemographic indicators. But the influence of the various determinants of earnings will be expected to differ between different types of work.

One question to be asked is: if human capital variables do in fact help to explain earnings from arts work, what is the mechanism by which this influence operates? In the performing arts, for example, the amount of training that an artist undergoes might affect the average hourly earnings she receives, because better trained performers can command higher fees. Alternatively, or in addition, there may be a relationship acting through the number of hours worked, for example because more established artists can obtain more engagements.

Several empirical studies have thrown light on artists’ earnings and on the influence of expected income and other variables on career choice. The first and simplest question has been whether the popular notion of the penniless artist bears any resemblance to reality: are income levels in the arts lower than in comparable occupations? Although comparison of artists’ incomes with those of the average worker using census data may not reveal a substantial differential, especially if total (arts and nonarts) income is used as the measure of artists’ earnings, comparisons with more specific occupational groups do indicate an appreciably lower level of mean and median earnings among artists than among other workers of similar educational and professional standing. Thus, in 1990 the median weekly earnings of all full-time wage and salary earners in the U.S. was $415, and that of all full-time artists was $499. However, in the same year the median weekly earnings of all full-time managerial and professional workers, a group broadly comparable with artists in terms of educational attainment, was $608. While these crude comparisons do not control for other systematic differences between
categories, as well as being affected by the sorts of definitional problems noted earlier, they are at least consistent with analyses of survey data which indicate that artists suffer a significant earnings penalty. It appears that this earnings differential is due in part to lower hourly earnings and in part to the fact that as a group artists devote less time to arts work than do other workers to their particular occupations.\(^9\) Wassall and Alper (1992) conclude that artists do not choose to work less; their extensive multiple jobholding indicates that they work in other jobs to supplement artistic incomes deemed by them to be inadequate. (p. 191)

Second, age-earnings profiles are steeper for artists than for other workers. Randall K. Filer (1986) suggests that this is caused not so much by differential drop-out rates, but rather by high rates of employment growth in the arts generating the possibility of rents for older, more established artists for whom younger artists are an imperfect substitute.

Third, the role of education, a mainstay of the human capital model, has been found consistently in the above studies to be not as influential in determining income levels in the arts as is the case in other occupations. In other words, if annual earnings of different types of workers are regressed against a range of explanatory variables in a human capital framework, the schooling variables generally turn up with lower and less significant coefficients for artists than they do for other occupational groups. For instance, using such a model estimated from 1980 Census data, Filer (1990) calculated that an additional year of schooling would add about $1900 to the incomes of managers, professionals and technicians, but only about $1150 to the incomes of professional artists. He suggested that the abilities required for a successful artistic career are not easily taught in formal education systems, and that artistic skills contributing to earning capacity may be more readily acquired through on-the-job experience.

Fourth, it is readily seen that artists' incomes are more variable than those of other groups, both across time for an individual artist, and across artists at a given point in time. Such an observation accords with the general view of an artistic career as being risky. Does this riskiness deter new entrants to the arts labor market, or does the lure of performing in Carnegie Hall, having a work shown in the Museum of Modern Art, or winning an Oscar or a Pulitzer Prize, exert an irresistible attraction, no matter how minuscule the probabilities of such success? The evidence on this question is unclear. Although a number of models of career choice in the arts have been proposed, sound data to quantify them are in short supply. Nevertheless, whatever the eventual empirical verdict on whether or not artists are risk averse, there seems no doubt at least that the level of earnings risk in the arts is one of the factors causing artists to take non-arts jobs, where a more stable return can be expected.

Fifth, consideration of the skewness in artists' earnings raises the question of the role of talent in determining success in an artistic career. While it is quite plausible to take estimated earnings functions and to attribute at least some of the (often large) unexplained residual to differences in talent, such a hypothesis remains untestable when no independent measure of talent is forthcoming. An alternative approach has been to consider how con-

\(^9\) A further explanation of artists' low incomes could lie in the fact that some part of their work often escapes the market process, and is appropriated by users who do not pay. Although copyright revenue forms a significant component of some artists' incomes, there are many others whose creative work remains largely unprotected, despite the growing volume of national legislation and international agreement covering intellectual property.
ditions of artistic production, consumption, and exchange have promoted some talented individuals to the top of the income tree. Sherwin Rosen's (1981) "superstar" model construes an artist's revenue as a convex function of talent; there is imperfect substitution in consumption between different sellers, because several mediocre performances do not substitute for one good one. Hence small differences in talent among performers are magnified into large earnings differentials. Furthermore Rosen sees consumption technology in the arts and elsewhere as enabling scale economies in consumption. These two characteristics taken together mean that a few talented persons can command large markets and the highest returns. The same result can be obtained without reference to differences in talent; Moshe Adler (1985) suggested that the costs of acquiring knowledge about a range of performers prompts consumers to concentrate their demand on a small number of individuals who will thereby be elevated to stardom.10

Finally, some indications for policy may be able to be drawn from these sorts of analyses. For example, Felton (1980) used estimated labor supply, revenue, and production functions to evaluate the most cost-effective means of inducing composers to spend more time at composition. She found that patronage through commissions, prizes, and grants that increased temporary income would be more effective than subsidizing publication, recording, or performance in bringing forth the required increase in labor supply. The question of why a government might have an interest in stimulating the output of artists is addressed in the next section.

VI. Public Policy Toward the Arts

Finding a rationale and guiding principles for government support of the arts was one of the major areas of concern of the earliest postwar writings in cultural economics, and these issues have continued to recur in the literature ever since. Consideration of these matters can conveniently be divided into positive and normative aspects.

A. Positive Aspects

Governments in all democratic countries assist the arts in some way or another, using a variety of instruments including: subsidies to companies and individuals; direct provision of artistic goods and services through state-owned enterprises; tax concessions to individual and corporate donors to the arts; tax exemptions to artists and arts organizations; regulation, such as local content requirements for television drama; the provision of information; support for arts education and training; and legislation affecting the economic rights of artists, for example through copyright laws. Levels of support vary markedly between countries. International comparisons of government expenditures in this field are tricky because of differences in definition of "the arts" between countries, difficulties in defining comparable funding sources, and problems in quantifying indirect aid. Acknowledging these difficulties, Table 1 assembles data on direct public expenditure on arts and museums for eight countries for the year 1987.

The relatively low level of expenditure in the United States is explained in part by the smaller extent of direct provision (state ownership) of arts facilities in

10 In a similar vein, Glenn MacDonald (1988) explains the skewed income distribution among artists by the presence of large numbers of poorly paid young hopefuls trying to gain a foothold in the marketplace; for an assessment of all these models in the context of the market for singers, see Ruth Towse (1992).
TABLE 1
PUBLIC EXPENDITURE ON ARTS AND MUSEUMS, VARIOUS COUNTRIES 1987

<table>
<thead>
<tr>
<th>Country</th>
<th>All public expenditure on arts(a) as proportion of GDP</th>
<th>Public expenditure on arts(a) per head</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>%</td>
<td>Central govt.</td>
</tr>
<tr>
<td>United States</td>
<td>0.05</td>
<td>1.7</td>
</tr>
<tr>
<td>Canada</td>
<td>0.34</td>
<td>12.1</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>0.79</td>
<td>7.4</td>
</tr>
<tr>
<td>W. Germany</td>
<td>0.77</td>
<td>0.4</td>
</tr>
<tr>
<td>France</td>
<td>0.77</td>
<td>7.2</td>
</tr>
<tr>
<td>Netherlands</td>
<td>0.45</td>
<td>16.3</td>
</tr>
<tr>
<td>Sweden</td>
<td>0.42</td>
<td>29.2</td>
</tr>
<tr>
<td>Australia(b)</td>
<td>0.39</td>
<td>6.0</td>
</tr>
</tbody>
</table>

Notes: (a) Comprising recurrent expenditure on arts and museums; this is essentially a “British” definition of what is included in the arts, and differs from some American studies (Schuster 1988).
(b) 1989-1989.

American compared with the European countries, and in part by the fact that the U.S. places by far the greatest reliance on voluntary support to the arts through charitable giving. Estimating the public sector component of these latter resource flows is problematical. Schuster (1985) estimated that in 1982-83 private donations to the arts and humanities from individual, corporate, and foundation sources in the U.S. was about four times as great as the level of direct government support, and that the public cost of these voluntary donations in terms of taxation revenue forgone was about half of the aggregate amount. Applying these proportions in broad terms to the data for the U.S. in Table 1 suggests that, even after allowance is made for indirect support for the arts through the tax system, overall levels of public subvention in the U.S. are still likely to come out lower than in the other countries in the table. This ranking of the U.S. relative to the rest of the world has apparently not changed greatly over the last few decades. Somewhat similar data for the late 1960s painted a similar picture (Scitovsky 1972). Scitovsky interpreted it as indicating differences in American and European tastes which would only be altered by education.

One noteworthy characteristic of Table 1 is the variation between countries in the proportions of support contributed by central and by regional governments. France and Germany provide the majority of their arts assistance through regional and local channels, for example, whereas Sweden is more centralist. In the U.S., the 1980s have seen a growth in fiscal decentralization across the board, in which arts funding has been caught up. Appropriations for the National Endowment for the Arts, the principal vehicle for direct central government financing of the arts, fell by about 30 percent in real terms over the decade, while appropriations for State arts agencies nearly doubled in real terms over the same period. The relative incidence of the costs of indirect arts support has
also been shifting away from the center in recent years. Netzer (1992) sees these developments as desirable in terms both of the theory of fiscal federalism and the theory of public support for cultural activities. Further, DiMaggio (1991) points to the benefits to artists and organizations of a coexistence of federal and state grant-making, through increased flexibility and a lessening of the costs of wrong decisions, as well as a means of risk pooling in times of political vulnerability.

The basic issue to be considered in the positive economics of arts funding is the question of why governments, whether federal, state, or local, have intervened in the market to the extent indicated in Table 1. What has been the principal motivation behind these significant levels of government expenditure? A major consideration influencing legislators has undoubtedly been a sense of the appropriateness of a government role in supporting the cultural life of the community. Countless examples can be drawn from the rhetoric of politicians of all persuasions around the world of the "duty" and "responsibility" of the public sector in fostering the arts. Such sentiments are consistent with theories of the state arising from outside of neoclassical economics, or indeed from outside of economics altogether, but within the conventional economic paradigm they sit somewhat less easily. They might be construed in one of two ways.

If in fact the community were indifferent or hostile to the arts, the imposition of government preferences for the arts through subsidies would be dictatorial or, at best, paternalistic. On the other hand, governments may simply be acting on a belief that voters share their view of a proper role for the public sector in providing support to the arts; if so, and if this belief were correct, the actions of governments could be argued to be broadly in line with individual utilities. Occasional ad hoc evidence from opinion polls in several countries lends some support to the proposition that a majority of voters approve of government involvement in this area, notwithstanding the occasional outcry when the role of the artist as social critic creates a tension between public funding and freedom of artistic expression.

An additional positive explanation of government subsidies to the arts is that they represent the outcome of rent-seeking behavior by individuals and enterprises in the arts industry. Certainly studies that have sought to uncover the determinants of variations in public expenditure on the arts across regional jurisdictions, or the factors influencing voting behavior in referenda on levels of arts support, have found evidence of a demand for private benefit through arts subsidies (Withers 1979; Bruce Seaman 1981; Schneider and Pommerehne 1983b; Throsby 1984). Furthermore, if an indication of rent-seeking is provided by an excess of factor payments in the arts over the levels that a free market would provide, then a prima facie case for the existence of rent-seeking would appear to be made (William Grampp 1989). However, the bidding up of factor rewards by self-interested agents in the arts is but one of many possible explanations for observed factor-price differentials. Such effects would also be seen as a result of governments acting on the sort of perceived mandate described above, even in the absence of rent-seeking behavior. Furthermore, there may be efficiency or other reasons for payments to factors that exceed private market levels, a consideration which brings us to the normative aspects of these questions.

B. Normative Aspects

A concern for optimal allocation of private and public resources within a social-welfare-maximizing framework in a free
Throsby: The Production and Consumption of the Arts

exchange economy leads inevitably to the question: Is there an economic rationale in normative terms for spending tax revenue in support of the arts, regardless of what governments actually do? We examine this question initially in the context of a standard competitive model wherein resource allocation is guided by the free and independent choices of sovereign consumers. The textbook market-failure grounds for government intervention in such an economy are well known. Many of these grounds can be applied to the arts, including possibilities that the arts give rise to external benefits in production and consumption, that there are nonmarket demands for the arts for option, existence, and bequest values, and that the arts exhibit public-good characteristics alongside the private benefits conferred by individual consumption. Of course such speculations, if valid, would provide only a prima facie case for corrective government action. Before such action would be warranted in normative terms, it would need to be shown that at the margin the social benefits gained from intervention would outweigh the direct costs involved in comparison to alternative means of achieving the same ends. Further, an assurance is required that obstacles such as political corruption or bureaucratic inefficiency in delivery mechanisms will not prejudice an optimal outcome.

Moving beyond standard efficiency considerations raises the possibility that the arts might be deemed a merit good in Richard Musgrave's (1959) original terms, and that, if so, this would provide normative grounds for collective action. At first glance, the arts would seem to fit the "merit want" description rather closely: society apparently sees the arts as "meritorious," yet people do not demand them in private markets to the extent that such a view would suggest, providing a presumptive case for corrective intervention. Closer examination, however, suggests that a number of the characteristics that might be ascribed to the arts as a merit good can actually be explained as generalized externalities or social goods. For instance, a belief that the arts are socially beneficial when held by people who do not themselves consume the arts directly, or an acceptance by some individuals of the desirability of others' consumption, can be accounted for in this way. In such cases what appears at first sight to be "imposed choice" turns out to be ultimately consistent with the principle of consumer sovereignty.

Nevertheless, this may not be the complete story. Are there aspects of a normative case for intervention in arts markets that still lie beyond the standard welfare analysis based on rational action in accordance with well-informed individual preferences? Four aspects of this question may be considered. First, the efficient operation of market processes requires fully-informed consumers. If individuals lack the necessary information on which to base their market choices, or at a more fundamental level are ignorant of their own welfare, then they may take decisions that are not in their own best interests, and corrective action, at the least through provision of information and education, might be justified. Although data are lacking, a plausible intuitive case can be made that this situation applies in some measure to the arts, especially in the light of discussion in Section II above concerning the development of tastes for cultural goods.

Second, a scrutiny of the relationship between preference and action, interpreted unquestioningly in conventional welfare analysis as a direct linkage, indicates that there may be significant cases where the observed behavior of individuals is inconsistent with their underlying values, for reasons such as misperception, weakness of will, or fluctuations in
preferences over time. Such cases of apparent irrationality present a challenge to the accepted definition of consumer sovereignty, suggesting that a broader and more carefully articulated interpretation of consumer sovereignty might be warranted (Alan Hamlin 1990). Government action which appears initially to restrict consumer choice might then be seen to be normatively consistent with a broader sovereignty notion.

Third, it is possible that a traditional social welfare function that admits only individual utilities as its arguments may be too restrictive in the context of socially meritorious goods such as the arts. The suggestion of an “augmented” social welfare function, where society expresses desires that are additional to those of its members as individuals, has been around for some time. More recently attention has been paid to the specific possibility of “irreducibly social goods” (Charles Taylor 1990), that is, goods which contain some element of benefit that cannot ultimately be attributed to some individual. In the general discussion of this proposition, the arts can be advanced as a significant case, but in order to do so the boundary of conventional economic thinking must be extended to encompass ideas of culture and civilization drawn from philosophy, aesthetics, and political and social theory.

Finally, distributional issues must be addressed. Within the standard welfare economics framework, it is of course generally assumed that any adverse equity impacts of measures designed to improve allocative efficiency will be dealt with by lump-sum transfers. This matter is of some concern in the arts, because the benefits of subsidies to encourage artistic activity will almost certainly have a regressive incidence on consumers, though not necessarily on producers. However, suggestions for dealing with this problem directly through adjustments to policy instruments, for example by channeling some subsidies to performing firms through consumer vouchers issued to low-income customers, have generally been judged ineffective, unworkable, or both. Ultimately the verdict yielded by the standard model must be that any un-toward distributional implications of arts assistance should be seen not so much as a reason for modifying support procedures but as an added impetus to ensure that general redistributive policy is operating to the desired extent.

Distributional questions, too, are an important aspect of the concept of merit goods, providing a rationale for in-kind transfers to the disadvantaged in areas such as housing and education. The arts do not fit comfortably alongside these examples, because they can scarcely be construed as a social or economic necessity, no matter how convinced artists and others might be of the central importance of art in life. Nevertheless a distributional motive from time to time influences the behavior of the managers of performing companies and art museums, for example, in their efforts to keep prices down as a means of enabling access to the arts by all, and not just by the rich.

Many economists over the past twenty years have put forward and analyzed market failure, merit, and distributional arguments for and against public support for the arts, such that by now there are few theoretical stones left unturned within the confines of the competitive model, and the focus of further enquiry in this area of the field must be primarily empirical. All of the effects noted above are in principle measurable, and it remains for well-designed research to put quantitative flesh on the theoretical

\[11\] Discussions of vouchers, such as that of Mark Blaug (1976, pp. 140–42), emphasize the difficulties of targeting such schemes on the “right” recipients; nevertheless, there is at least one scheme that has been judged a success (see W. Baumol 1979).
bones. This task should be relatively straightforward in respect to immediate neighborhood effects such as the specific spillovers in production and consumption arising, for example, from arts festivals or from the presence of a dynamic arts sector within a regional economy. However, studies in this area have tended to focus on direct expenditures and their multiplier effects, rather than on genuine instances of market failure that might warrant government intervention (Seaman 1987). More generally, the arts can be seen as a potentially leading sector in central city renewal and in urban economic development. There is scope for further research in this area, both conceptually (to specify more clearly the linkages between the arts, local demand, tourist expenditures, and urban growth patterns, and their implications for public policy) and empirically (to estimate the magnitudes involved).

But the task of quantifying broader public-good benefits of the arts, if they exist, has proved more daunting. While hedonic methods may be applicable at local level (David E. Clark and James R. Kahn 1988), the evaluation of more diffused community demand for the arts as a social good is likely to prove more amenable to contingent valuation approaches. Little work has been undertaken to date, apart from exploratory studies such as those of Throsby and Withers (1986) and William Morrison and Edwin West (1986), which have used survey techniques to estimate individual willingness to pay for the public-good component of the arts under conditions controlling for free-rider and informational biases. These studies have indicated that respondents not only perceive benefits such as definition of national identity and social criticism arising from the arts, but also are willing to pay for these benefits out of taxation at levels that are somewhat higher than existing government subventions. Although these results are consistent with some of the theoretical propositions mentioned above, and indeed could be taken so far as to suggest that market failure grounds for government support for the arts may be empirically more robust than some of the skeptics have supposed, they are only partial in their coverage, are inevitably location- and time-specific, and may not control for all sources of bias affecting such investigations. They should be seen as pointers to further needed research in this area.

We noted above that further work on the efficiency grounds for public assistance to the arts would be likely to be primarily empirical. By contrast, extension of the standard welfare analysis into the broader domains noted earlier offers rich possibilities for both theoretical as well as empirical advance in its application to the arts. Such efforts are likely to be more fruitful if they incorporate ideas and models drawn from other disciplines.

Finally, once the question of whether governments should assist the arts has been resolved, or if an affirmative answer to it is taken for granted, the next consideration is how might a given volume of support best be raised and distributed among competing avenues of expenditure? These issues have of course been central questions of public finance across all areas of social policy over the years, among which the arts have received some share of attention, especially in regard to the incidence of the costs of support and the effects of tax policy on resources available to major cultural institutions (Alan L. Feld, Michael O’Hare, and Schuster 1983; Charles T. Clotfelter 1991). The distribution of public support among competing claimants, whether disbursed as direct government grants or through tax revenues forgone, comprises the principal dimension of a govern-
ment's arts policy. Such a policy is likely to have multiple objectives, including the expansion of access to the arts, the expression of national or regional cultural identity, the promotion of excellence, the encouragement of regional growth, and so on. For this and other reasons, quantitative evaluation of the "productivity" of different types of support presents difficult theoretical and empirical problems, most of which have yet to be addressed. Furthermore, there are manifest social and political dimensions to these questions that suggest they should not be studied in isolation from their institutional context.

VII. Conclusion

Several reviewers of the progress of cultural economics over the years have observed that many writers, including themselves, have begun their books or papers with an apology for presuming that economics might have anything useful to say about art. In the light of the wide range of imaginative and resourceful work in applying economics to problems in the arts that we have drawn attention to in this article, henceforward such disclaimers may no longer be thought necessary. Three lines of future development in cultural economics can be foreseen.

First, despite progress to date, there are, as we have noted throughout this paper, numerous nontrivial theoretical and empirical problems yet to be explored in this area that are susceptible to the powerful tools of positive and normative economic analysis. Interest is likely to continue to arise particularly from existing fields of economic enquiry where applications to the arts and culture occur as a special case, such as in the theory of consumer behavior, the economics of nonprofit enterprise, the economics of charitable giving, urban economics, labor market economics, and other areas.

Second, an essential element in future work will be the provision of better data. At several points in this article attention has been drawn to the serious constraint imposed on research in cultural economics by the lack of comprehensive statistics on the arts industry or its subsectors. Once-off sources such as the Ford Foundation data tape for the performing arts (1974) have been wrung dry and have become obsolete. Census data often cannot be provided on a sufficiently detailed definitional basis to be useful for research. Specialist arts authorities such as the NEA and the state arts agencies have suffered budget cuts which tend to fall more heavily on apparently expendable areas such as data gathering than on giving grants to artists. The U.S. experience in these respects is mirrored to a greater or lesser degree in most other countries. Cultural economists will need to pay greater attention to the collection of new data in future if they wish their work to be taken seriously by other researchers or to be useful to policy makers, organizations or individuals working in the field.

Finally, while theoretical and empirical developments within the conventional paradigms of economics will doubtless continue, the arts and culture do also challenge tradition-bound economists to focus their eyes on a wider horizon. The very breadth of the subject area of the arts, embracing as it does a range of issues in philosophy, aesthetics, history, sociology, politics, and many other disciplines, would appear to make cultural economics a natural area for eclectic theoretical and methodological advance. Indeed, economists in this area have recently been grappling with ideas from cognitive psychology (Frey 1991), aesthetics (Gianfranco Mossetto 1992) and other fields, work that is illustrative of the scope for a broader methodological
foundation in this field. While efforts such as these of necessity remain highly speculative, further work in cultural economics may well point to new lines of development in the methodology of economics more generally.

REFERENCES


Grampp, William D. "Rent-Seeking in Arts Policy,"


Taylor, Charles. "Irreducibly Social Goods," in...


