

Nozick on Sunk Costs*

David Ramsay Steele

Historical costs have powerful sway over untutored minds.¹

I. INTRODUCTION

Robert Nozick challenges the view of sunk costs held by economists,² a view which has for about a century been considered an established part of economic theory. Nozick produces arguments which, he believes, show the untenability of “the economists’ doctrine that sunk costs should be ignored” (p. 22).

In this article I argue that Nozick’s criticisms fail and that the economists’ doctrine emerges unscathed.

II. THE ECONOMISTS’ DOCTRINE OF SUNK COSTS

The sunk costs doctrine was for many years stated by economists without the words ‘sunk costs’. The doctrine is one conclusion of the theory of opportunity costs and was understood as such from the early days of opportunity cost theory. Wicksteed’s classic account, written in 1910, clearly explains and defends what we now call the doctrine of sunk costs.³

According to the opportunity cost concept, the cost of an action is what is given up by taking that action. The only costs which should be weighed in making a decision are the avoidable and, hence, necessarily future costs entailed by that decision; these costs consist of the benefits which would have flowed from the next best option, had it been elected instead. The sunk costs doctrine follows: costs incurred in the past,

* I thank N. Scott Arnold, David Barker, David Gordon, Barry Smith, and the editors of *Ethics* for helpful comments on drafts of this article.

1. George Stigler, *The Theory of Price*, 4th ed. (New York: Macmillan, 1987), p. 111.

2. Robert Nozick, *The Nature of Rationality* (Princeton, N.J.: Princeton University Press, 1993), pp. 21–26. Where not otherwise stated, page and line references are to this work.

3. Philip H. Wicksteed, *The Common Sense of Political Economy and Selected Papers and Reviews on Economic Theory*, 2 vols. (London: Routledge, 1933), vol. 1, pp. 88–94, 373–91.

Ethics 106 (April 1996): 605–620

© 1996 by The University of Chicago. All rights reserved. 0014-1704/96/0603-2006\$01.00

historical costs, should not be counted as costs of present or future decisions.

For example, Bill decides to open a factory producing widgets. Bill's outlays include an expensive durable machine, for which he gets a loan to be repaid over several years. The machine (we stipulate) cannot be resold, and the loan repayments would still have to be kept up if the factory closed. The factory is built, and production starts. It turns out that the market had been misjudged, and the income statement shows a net loss, which is expected to continue as long as the loan payments have to be made. But if we subtract the loan repayments from expenses, the factory is making a profit.

The economist holds that, in deciding whether to keep the factory open, the loan repayments should be ignored (and, although this takes us beyond Nozick's discussion, will in fact, in a competitive market, generally be ignored). Since those repayments would still have to be made if the factory closed, they are not costs of continued operation.

In thinking about such issues, it is often helpful to consider hypothetical alternative scenarios by which the situation under discussion might have come about. We can compare the situation of Bill's widget factory, after production has commenced, with a scenario in which the machine had been donated as a free gift, and coincidentally Bill had an old, unrelated debt, amounting to the same as the loan repayments for the machine in the first scenario. According to economists, the two situations are alike in all relevant respects; therefore, the optimal decision in each case must be the same.

III. NOZICK'S FIRST ARGUMENT: USING ONE ERROR TO COUNTERACT ANOTHER

I can distinguish five mutually independent arguments in Nozick's case against the sunk costs doctrine. I will label these 1 to 5 and follow Nozick in allocating the great majority of my space to his argument 1. I briefly discuss arguments 2 through 5 at the end of this article. Nozick's argument 1 is presented via the following example: "If I think it would be good for me to see many plays or attend many concerts this year, and I know that when the evening of the performance arrives I frequently will not feel like rousing myself at that moment to go out, then I can buy tickets to many of these events in advance. . . . Since I will not want to waste the money I have already spent on the tickets, I will attend more performances than I would if I left the decisions about attendance to each evening" (p. 22).

The protagonist of this story, whom I will call *N*, buys a ticket at time *A* because he can foresee that at the later time, *B*, he will feel too lethargic to go to the theater, and since at *A* he considers this "lethargy" undesirable, he wants to do something at *A* to make it more likely that at *B* he will override the lethargy. (This terminology of *A*, *B*, and a subsequent time, *C*, is employed by Nozick.) Since *N* knows

that he is given to counting sunk costs, he uses this feature of his personality to manipulate his future decision. He knows at *A* that at *B* he will believe it to be an argument in favor of visiting the theater that he would otherwise have wasted the money he had spent, at *A*, on the ticket. Hence *N*'s propensity to count sunk costs can be used by *N* at *A* to make it more likely that at *B* he will do what he thinks at *A* that he should do. If *N*'s reasoning were along the lines envisaged by Nozick, *N* at *A* might turn down the opportunity of a free ticket and insist on paying \$40 for it, or he might walk to the box office in the rain, which he detests, when he could just pick up the phone.

Nozick sees the counting of sunk costs as offering the possibility of a technique for overcoming temptation. If *N* at *A* knows that *N* at *B* will be prone to commit the sunk costs fallacy, *N* has one more technique with which he can act at *A* to change the way he will behave at *B*. "We can knowingly employ our tendency to take sunk costs seriously as a means of increasing our *future* rewards. If this tendency is irrational, it can be rationally utilized to check and overcome another irrationality" (p. 23).

The reference to future rewards arises because Nozick holds that frames of mind like that at *A* will recur at *C*, so that *N* will then look back on the whole sequence of events and prefer that he had gone to the theater at *B*.

IV. AN ADVANTAGEOUS ERROR IS AN ERROR

The most natural understanding of Nozick's argument 1 is that he accepts that *N* at *B* commits an error (an "irrationality") by counting sunk costs but that this error has an outcome which *N* at *A* regards as advantageous or beneficial, because it counteracts or cancels another error (or, more broadly, another deficiency or shortcoming), that of lethargy.

Insofar as Nozick asserts that we can imagine ways in which committing a certain kind of error might turn out to be advantageous, he does not contradict the economists' doctrine. It is not disputed that errors can turn out to be advantageous: two immediately obvious cases are (a) that in which the agent has incomplete knowledge, and by chance a decision relying on an error turns out to be best, and (b) the case where one error cancels out the effect of another error. Showing that an error may be beneficial or advantageous does not show that it is not an error. Drawing that conclusion would be an instance of the well-known fallacy of inferring the quality of a decision from its outcome.⁴

4. Ibid., vol. 1, pp. 72, 120–21; Stephen Nathanson, *The Ideal of Rationality: A Defense, within Reason* (Chicago: Open Court, 1994), pp. 49–51. Nathanson cites Rawls (*A Theory of Justice* [Cambridge, Mass.: Harvard University Press, 1971], p. 72), who calls it "subjectively rational" to do the best one can with the available information. I

Nozick raises the possibility that an economist might say that according to Nozick's argument, the "irrationality" of counting sunk costs is desirable only for someone with another "irrationality" (p. 24). Nozick responds to this hypothetical objection by referring to his arguments 2 and 3 and then stating his argument 5. That is to say, Nozick pointedly does not contest the claim that his first and fourth arguments rely on one irrationality (error or deficiency) to counteract another.

Yet it is misleading for Nozick to make his hypothetical economist appear to concede part of Nozick's argument and try to save a smaller area of validity for the sunk costs doctrine. Real economists do not say that counting sunk costs can never have a desirable outcome. Given that the agent's information is imperfect, any miscalculation, any folly, can turn out for the best. The economists' position is that counting sunk costs is an error: on the most obvious interpretation, this is confirmed rather than challenged by Nozick's argument 1.

If Nozick denies that counting sunk costs is always an error, it seems to follow from his position that no type of error is always an error, for any type of error might be compensated for by some other error to yield a desirable outcome. We can equally well imagine some circumstance in which supposing that $2 + 2 = 5$ would have a welcome outcome and then pronounce that this error is not always irrational, but only some of the time.

V. THE TWO FORMS OF THE SUNK COSTS FALLACY

Although economists sometimes speak of *the* fallacy of sunk costs, there are in fact two different forms of the fallacy, which tend to impel decisions in opposing directions. That form which takes up by far the greater attention and space in economics teaching I will call the "main form" of the fallacy; it is exemplified in the story of Bill and his widget factory. Here, according to economists, current operations should often be viewed as more profitable than uninstructed common sense would indicate, and current projects should often be continued where common sense might suggest that they should be terminated. The economist's rejection of the main form of the fallacy always, if it affects

prefer to avoid such locutions by simply saying that counting sunk costs is an *error*. There may seem to be an air of paradox about the claim that a theory of successful action would classify a specific decision as wrong even though the outcome were successful. But this is superficial. The theory never recommends acting in a way known to be unsuccessful or failing to act in a way known to be successful. Erroneous actions with advantageous outcomes are either (a) flukes, like blowing the rent money on lottery tickets and winning \$20 million, or (b) cases where a mistaken theory in the agent's mind leads to a pattern of successes. This latter is not always due to one error's counteracting another but would generally be susceptible to an explanation which would help in disentangling the successful outcomes of the mistakes from actual or possible unsuccessful outcomes.

behavior at all, leads to the continued pursuit of projects which the fallacy would indicate should be abandoned. That is, the main form of the fallacy can never work in the direction required by Nozick's argument, but only in the opposite direction (if it affects behavior at all).⁵

The other form of the fallacy, which most textbook treatments, for example, do not bother to mention,⁶ is the view that because expenditures have been made in the past and have not yet been recouped, therefore special attempts should be made to recoup those expenditures in the future. I will call this the "Concorde" form of the sunk costs fallacy.⁷ Thus, if Hillary has paid to commence the building of a canal, which is now half completed, this is sometimes believed to provide a reason for Hillary to complete the canal, even if, in an alternative scenario, the half-completed canal existed as a natural geographical feature, and, knowing what Hillary now knows, she would not think it worthwhile to "complete" the canal. The economist says that these two scenarios are alike in all relevant respects: past expenditures do not justify future expenditures. The economist does not deny that, looking ahead, we ought to try to make income exceed costs. But we ought not to look back at costs irretrievably incurred in the past and make our behavior different from what it would have been if those past costs had not been incurred, in an attempt to recoup those costs in the future. Because those costs are now unavoidable, they are irrelevant, *qua* costs, to any current or future decisions.⁸

5. The fuller opportunity cost doctrine from which the sunk costs doctrine is derived might push in the opposite direction because of implicit costs. I follow Nozick in confining the discussion to those cases where a decision is made about whether to continue with some project. Some discussions of sunk costs relate instead to optimal pricing: whereas noneconomists often suppose that what was paid in the past to produce or acquire something should influence what is charged for it now, economists say that this is immaterial.

6. I tested my impressions of what economists teach by looking at the relevant passages in forty introductory economics or intermediate price theory textbooks. Some failed to unambiguously mention sunk costs. Most that did do so mentioned only the main form, while others mentioned both forms. I found none where the Concorde form was given greater emphasis or where only the Concorde form was mentioned, except Stigler (pp. 111–12), which, however, gives a passing mention of the Concorde form and whose other examples are all about price setting rather than project continuation. I count as a "mention" an example which fits one or the other form—most texts don't explicitly distinguish them. Of the texts I looked at, only David Friedman's *Price Theory: An Intermediate Text* (Cincinnati: South-Western, 1986) clearly explains the distinction between the two forms, calling them "opposite mistakes" (pp. 279–80).

7. See Richard Dawkins, *The Selfish Gene*, 2d ed. (Oxford: Oxford University Press, 1989), p. 150.

8. On the economist's view of costs, see Wicksteed, pp. 373–91. For more recent accounts, see Armen Alchian, "Costs," in *International Encyclopedia of the Social Sciences*, ed. David L. Sills (New York: Macmillan, 1968), pp. 404–14; James M. Buchanan, *Cost and Choice: An Inquiry in Economic Theory* (Chicago: Markham, 1969).

The second, or Concorde, form of the sunk costs doctrine is alone treated by Nozick. He does not mention the more commonly discussed main form of the doctrine and leaves the impression that the Concorde form (the rejection of the Concorde fallacy) is the whole doctrine. Only the Concorde form can help Nozick's argument.

VI. WHY *N* MIGHT GO TO THE THEATER, IRRESPECTIVE OF SUNK COSTS

People quite reasonably take steps to change their future circumstances in such a way that they will be more likely at some future time to choose to do one thing rather than another. It is not unusual for a person to buy exercise equipment in order to "make herself" exercise. This well-known stratagem doesn't necessarily involve counting sunk costs and may be free of error from the economist's standpoint.

There are reasons not involving sunk costs why *N*, on the evening of the performance, might not go to the theater if he had not bought tickets, even though he would go if he had bought tickets. An unfocused awareness of these other motivations helps to make *N*'s stratagem of committing himself by buying tickets in advance look plausible. This intuitive plausibility does not rest on counting sunk costs.

The most obvious consideration—given a passing mention by Nozick (p. 22, lines 19–20)—is that if *N* has already bought a ticket, he does not have to pay for a ticket at *B*. Since the payment for the ticket has already been made and (we will suppose) cannot be unmade, the theater effectively has free admission for *N* at *B*. If he had not bought the ticket, on the other hand, admission would cost the price of the ticket, say, \$40. There would additionally be a nonpecuniary cost, the time and effort spent making the purchase, which *N* might negatively value at \$9. (Further, although Nozick assumes that *N* at *B* can costlessly be certain that seats of the relevant quality and price are not sold out, this is rarely the case.)⁹

N's plan at *A* makes sense without having *N* at *B* count sunk costs. *N* at *A* anticipates that *N* at *B* might not be sufficiently motivated to go to the theater, so *N* at *A* provides the motivation of a free ticket (free both monetarily and in terms of the nonpecuniary expense of the transaction of acquiring it) and in practice also dispels any uncer-

9. Theater tickets are often cheaper when purchased in advance. From the standpoint of Nozick's argument, this is unfortunate, since it means that the anticipated sunk cost, and therefore the ability to manipulate one's own future behavior by incurring a cost, is lower. On the other hand, an individual might well think he shouldn't "waste" the theater ticket since its acquisition was such a bargain, and this erroneous thought might prompt him to go—another illustration of the fact that once we start relying on our intellectual confusion to get us to do things, any type of intellectual confusion may do just as well.

tainty at *B* about availability and price of tickets by buying the ticket in advance.

This stratagem is most dependable where *N* at *B* does *not* fall for the sunk costs fallacy in its main form. If *N* at *B* is in the grip of this main form, then he will be inclined to rate the cost of going to the theater as \$49 higher than he should (or perhaps some positive amount less than \$49 higher, for the error exists if sunk costs are given any weight at all as present costs). *N* will fail to fully take on board the fact that the ticket, once paid for, is free.¹⁰ And this error may dissuade him from going to the theater.

Not only does the above rationale offer an alternative to Nozick's reliance on sunk costs, but an understanding of this rationale makes the rational utilization of the Concorde fallacy appear far less promising. Suppose that *N* pays \$40 for the ticket and would have paid up to \$56 (a consumer's surplus of \$16). When buying the ticket, *N* anticipates the visit to be worth \$65 (\$56 plus the nonpecuniary outlay valued at \$9). Suppose that on the night, the visit is worth \$5 to him. This is a considerable drop in *N*'s valuation of the theater visit, a reduction of \$60, yet *N* will still go to the theater—even if its value falls to one thousandth of a cent, he will go. (There are further costs of going, of course, but the valuation in question is already net of those costs.) Before there is any scope for the Concorde fallacy to make a difference, *N*'s valuation of the theater visit must have fallen by at least \$65.

There is another distinct motivation for *N* at *B*: he may not want to waste the ticket he has available at *B*. This is different from counting sunk costs, where *N* at *B* doesn't want to waste the money he paid at *A* for the ticket. If *N* at *B* possesses a ticket because he received it as a gift, there are no historical costs for *N*—he never bought the ticket—but he may still feel that he doesn't want to waste it. (No doubt this, too, is an error, but it is not an error of counting sunk costs.) *N* at *A* might buy the ticket to take advantage of *this* foreseen motivation on the part of *N* at *B*. In that case, there would be a sunk cost, but it would have nothing to do with *N*'s reason at *B* for going to the theater or with *N*'s reason at *A* for buying the ticket.

In Nozick's story, *N* buys a ticket to induce himself to behave in a certain way at a future time. This intrapersonal stratagem has an interpersonal counterpart: a person might buy someone else a ticket in order to change that other person's behavior. Notice that if this or

10. This assumes that the ticket has no other worthwhile employment. Suppose that, shortly before the performance, a friend offers *N* \$2 for the ticket. Then at that point the cost to *N* of using the ticket to go to the theater is \$2. Similarly, if the friend credibly offers \$20,000 for the ticket, then \$20,000 is the ticket cost (there are other costs) to *N* of going to the theater that night.

any other ploy works interpersonally, it cannot be by utilizing a sunk cost error; hence, if it works intrapersonally for the same reason that it might work interpersonally, sunk costs are not intrinsic to that reason.

Another possibility is that *N* at *B* intrinsically values the completion of projects already commenced or merely decided upon. (One of these projects might be attending the theater no less than eight times a year.) It is quite wrong to suppose, as Nozick sometimes seems to do, that this phenomenon in any way conflicts with the economists' view of sunk costs. I look at this further in Section XI below.

VII. THE SWITCH IN *N*'S GOALS

N at *B* has goals (and underlying preferences) that are not only different from those of *N* at *A* but opposed, so that frustrating *N*'s pursuit of his goals at *B* can help *N* pursue his goals at *A* and vice versa. This feature, the mutual antagonism of *N*'s goals on two different occasions, is essential to Nozick's first argument. The argument is therefore based on a sleight of hand, a virtual equivocation, for *N*'s counting of sunk costs at *B* is held to help *N* pursue the goals he holds at *A*, when he does not count sunk costs (or is not required by the argument to count sunk costs). Counting sunk costs is not claimed to help *N* pursue the goals he pursues at *B*, when he counts sunk costs. *N*'s goals have changed from *A* to *B*, so it is no surprise that an error he makes in pursuing his goals at one time, while this error does make him less effective at pursuing those goals, may very well help to make him more effective in pursuit of his goals at a different time, when he has contrary goals. At *B*, the point where it is claimed that *N* counts sunk costs, counting sunk costs does look like an error, since there is nothing to suggest that it helps *N* in effectively pursuing his goals at *B*, if we look at these goals in isolation from his goals at *A*.

Naturally, Nozick might dispute the "present aim" view of rationality implicit in the above discussion, but then there would be no need for Nozick's particular story. It would be enough for him to point out that a person may blunder in pursuit of (say) an objectively wrong goal, and, hence, a blunder may serve the agent's best interests. But it would then be evident that such a line of attack was pointless, for all that economists have claimed is that counting sunk costs is an error given the agent's actual goals at the time of the decision. Economists have not claimed that these goals were always wise.

But, it might possibly be said, *N*'s counting sunk costs at *B* does effectively serve *N*'s goals at *A*, and these are more important than *N*'s goals at *B*, as they represent the authentic *N*, when he is in full command of his faculties, when his will is strong, and when he has a superior grasp of his own true interests. But this is evidently not persuasive enough to convince *N* at *B*. If counting sunk costs is an error, if it gives *N* at *B* a misleading picture of his alternatives, then

N at *A* can be seen as trying to frustrate the actual wishes of *N* at *B*. This becomes instantly clear when we replace *N* at *A* and *N* at *B* with two different individuals (the interpersonal parallel). That individual *P* can thwart the desires of individual *Q* by capitalizing on *Q*'s propensity to believe fallacies or commit errors is hardly surprising or interesting.

Counting sunk costs remains, in Nozick's example, an error, just as economists say. What Nozick has done is to identify a hypothetical case where the error of counting sunk costs can have welcome consequences—welcome to an outside observer or to the individual in a different, and contrary, frame of mind. The desirable quality of the consequences depends upon preferences contrary to those of the agent at the time of decision.

Nozick's discussion gives us no reason not to impute to him the following view: the economist's advice not to count sunk costs is unfortunate in some cases, because it is best in those cases for an individual to commit the error of counting sunk costs and thereby to be less effective at gratifying his preferences of the moment. If this is Nozick's view, it does not contradict the economist's doctrine.

Consider *N* at *A*, or at any other *A*-like, strong-willed time. It can never be to his advantage to count sunk costs while in this frame of mind, for any reason remotely like that in Nozick's example. ("Never" except in the oblique sense that it may be a concomitant of counting sunk costs at *B* that he also counts sunk costs at *A*.) Nozick's discussion in his argument 1 (and possibly his arguments 4 and 5) does not in the least go against the view that, relevant to a particular set of goals or preferences, counting sunk costs is always an error, in that it is not optimally conducive to attainment of those goals or satisfaction of those preferences. Nozick disapproves of *N*'s goals and, hence, his preferences at *B*. Judged by a standard external to those goals, Nozick believes that we do not want people like *N* at *B* to be good at fulfilling their actual desires.

VIII. PRIVILEGING STRONG-WILLED FRAMES OF MIND

Against my position, it might be contended that I am talking as though there were two agents, but there is truly only one. *N* really has the same preferences and goals all along: it's simply that, at *B*, his "real" preferences and goals are inoperative because of his weakness of will. (The terminology of strong and weak will is not employed by Nozick here, but it seems to be in the spirit of his discussion.)

Any such defense is foredoomed, since in the relevant sense, *N*'s goals at *B* are real. *N* at *B* is both capable of counting sunk costs and of learning not to count sunk costs—precisely this is crucial to Nozick's account. It is only at *B* that *N* is stated to count sunk costs. If we are going to deny the reality of *N*'s goals at *B*, then (aside from the

implausibility of this denial) we must deny that *N* at *B* counts sunk costs and, hence, that *N* ever advantageously counts sunk costs.

We can be impartial with respect to the two sets of goals, *A* and *B*, or we can privilege one (Nozick would vote for *A*) at the expense of the other. In the first treatment, we simply observe that *N* wants different things at different times. We regard *N*'s wants at *B* as every bit as legitimate, in every sense, as *N*'s wants at *A*. In the second treatment, we side with *A* against *B*. Economic theory does not dispose us to either of these treatments; it is neutral on the matter (although no doubt economists often find the first more congenial).

Under the first treatment, it is uncontroversially evident that *N* at *B* is committing an error if he counts sunk costs: he will be poorer at gratifying his desires. It is not relevant that he has different and incompatible desires on other occasions. He just changes his mind from time to time.

Under the second treatment, we side with *N* at *A* against *N* at *B*, and perhaps in some sense we suppose that *N*'s goals at *A* are his true goals, his goals at *B* a betrayal of his true goals. But this merely means that we agree with *N* at *A* that an error by *N* at *B* may be welcome, because we judge the goals of *N* at *B* to be reprehensible.

IX. COMMITTING "ERRORS" ON PRINCIPLE

If an error leads to good results in some cases, we can, as it were, knowingly commit the error in order to get the good results. But then the "error" is no longer an error. It is a rule that superficially looks like an error when seen out of its proper context. An example is "With this rifle, you should aim to the right of the target." If the rifle has a bias to the left, and you shoot to the right to compensate for the bias, you are not thereby committing an error (even if you miss the target). If you know nothing of the left bias and shoot to the right because of poor aim, thereby unwittingly compensating for the left bias, you commit an error (even if you hit the bull's-eye).

Nozick writes of "rationally" utilizing "our tendency to take sunk costs seriously as a means of increasing our future rewards" (p. 23). But is it feasible for *N* at *A* to make *N* at *B* believe something fallacious, while *N* at *A* clearheadedly sees through the fallacy? *N* at *B* can hardly be expected to say to himself: "Ah, now it's time to start accepting a fallacy which will make me want to do what I wouldn't otherwise want to do." If *N* at *B* falls for the Concorde fallacy, then *N* at *A* must fall for it, too (if the occasion arises). So it's doubtful whether what Nozick seems to be recommending—that *N* at *A* rationally utilizes the proneness to error of *N* at *B*—could ever be accomplished.

X. IS THE CONCORDE ERROR ADVANTAGEOUS?

Nozick holds that "taking sunk costs into account sometimes is desirable (so the economists' general condemnation is mistaken) and some-

times is not" (p. 23). As we have seen, this evades the question of whether counting sunk costs is always an error, as economists maintain, regardless of the desirability of the outcome. Nozick and the economists can agree that counting sunk costs will sometimes be advantageous and sometimes disadvantageous. Does Nozick hold that it will so often be advantageous that the best policy is to count sunk costs rather than not? While this would not, as Nozick apparently supposes, refute the economists' doctrine, it would certainly form a fascinating rider to it, and one that perhaps ought to be mentioned in economics teaching. Nozick's view is unclear. At one point he seems to commit himself definitely to an agnostic position (p. 23, lines 34–36), but his assertion that the economists' rule "is not an appropriate general principle of decision" (p. 22) is troubling.

The citing of a hypothetical example where counting sunk costs turns out to be beneficial does not take us very far. Nozick makes no attempt to show that committing this kind of error will always, or typically, or more often than not, or in any appreciable proportion of instances be beneficial. We would scarcely be impressed by the demand that we reject the rule that you should not whimsically kill the next person you meet, because, after all, that person *might* secretly be a serial killer who not only richly deserves death but is getting ready to kill again; hence, the prohibition of random homicide is sometimes desirable and sometimes not. For that matter, in the theater example, almost any kind of error will do. If *N* is in the habit of losing his way, he might intend to go to a bar and end up by mistake at the theater. (Sometimes it's desirable to lose your way and sometimes not. And if you know that you are prone to lose your way, you can utilize this propensity by calculating in advance that you may be unable to find your way to the bar.)

Does Nozick perhaps hold that the agent can distinguish a subset of cases where counting sunk costs is likely to be advantageous, and count sunk costs on those occasions, failing to count them on all others? Nozick does not tell us of any general policy rule that he recommends in place of the economists' rule never to count sunk costs. If he recommends the rule always to count sunk costs, it seems immediately plausible that the disadvantageous applications will vastly outweigh the advantageous ones. If, at the other extreme, he recommends that we count sunk costs only in those cases where doing so is likely to be advantageous, this presupposes that we have some way to distinguish these cases from others. In that event, to save his argument Nozick would have to avoid any mechanism by which the agent perceives the advantages and counts sunk costs because he sees the advantages. If he can do this, he does not need to count sunk costs at all: pursuit of the advantages would be sufficient.

Nozick cannot be recommending counting sunk costs as a general rule. He states that the economists' doctrine "may be a correct rule

for the maximization of monetary profits" (p. 22). And, as we have seen, the main form of the sunk costs fallacy always tends to impel the agent in the opposite direction from that which Nozick finds desirable.

I have suggested that the persuasiveness of Nozick's theater ticket example derives from the assumption that *N* at *B* does not commit the main form of the sunk costs error. Nozick could take the position that he is referring only to the Concorde fallacy and that the main form of the fallacy is irrelevant to his argument. But it is doubtful that we could rely on an individual to be prone to the Concorde fallacy while immune to the main fallacy. The main fallacy can be quite subtle and may require protracted analysis to identify, whereas seeing through the Concorde fallacy makes no great demand on the intellect.

Even if we confine the recommended cases for counting sunk costs to the Concorde fallacy in noncommercial situations, counting sunk costs is going to lead to many disadvantageous outcomes. If *N* follows the rule at *B*, he is going to follow it at *A*. Furthermore, he is going to follow it at *A*-like times with respect to costs incurred at *B*-like times. And there is no reason to suppose even that following it at *B*-like times with respect to costs incurred at *A*-like times will always, or often, be advantageous.

In Nozick's example, *N*'s lethargy explains his reduced inclination to go to the theater. But this reduced inclination might be a result of bad reviews, a tornado watch, *N*'s being implored to help a niece with her homework, *N*'s having a cough and being unwilling to distract other members of the audience, or *N*'s receiving a free ticket for another show which is better in every way. Counting sunk costs must counteract these motives as effectively as it would counteract lethargy.

Even in the theater example, it's by no means clear that the Concorde fallacy will impel *N* to go to the theater. For *N* has expended resources on, say, comfortable furnishings in his home and has spent many an evening assiduously developing his human capital in the direction of becoming a virtuoso couch potato—every time he spends an evening away from home, he loses a chance to recoup some return on this past investment. Overeating or overdrinking can easily be facilitated by the erroneous thought that it would be a waste not to finish the bottle or what's in the fridge, because of the money or effort already expended in acquiring these. In such instances, proneness to the Concorde fallacy would make succumbing to temptation more likely.

XI. WANTING TO STAY THE COURSE IS INDEPENDENT OF SUNK COSTS

The sunk costs doctrine doesn't prohibit any attention to historical costs or any guidance of present decisions by looking at past costs. For example, someone may examine the record of past costs and incomes

to suggest what might happen in future analogous situations. Or decisions taken and costs incurred in the past may have had repercussions which result in relevantly modifying the actual situation as it obtains later, at the moment of decision.

It is also possible that the agent may have a preference for finishing what she has started or for honoring her commitments. An individual may simply have a preference for making and honoring commitments; this may not, and generally does not, have anything to do with sunk costs.

At times Nozick seems to confound some such rationally impeccable motive with a sunk costs error (p. 22, lines 34–41), a tendency illustrated by his use of the grotesque phrase “honoring sunk costs” (p. 23). Costs qua costs cannot be honored, although commitments can. In honoring costs, if anyone would ever want to do anything so strange, one would be treating them as something other than costs, as, for example, in religiously venerating a recipe for Dundee cake, one would not be employing it as a recipe. As a matter of practical relevance, it will often be most important to maintain a commitment at an early stage when few costs may have been incurred.

If an agent has a preference or motivation inclining her to keep commitments or finish what she has started, this intrinsically has nothing to do with historical costs and may indeed occur where there are no historical costs. We should also be clear that even where someone is more inclined to persist with some course of action because of attention to historical costs, this is not necessarily an instance of what economists would identify as a sunk costs fallacy.

A person might imaginably (although somewhat bizarrely) have an independent preference for pursuing actions where there are sunk costs—“independent” in the sense that this preference does not arise from a sunk cost fallacy but is one of the array of preferences with which the agent is endowed prior to any calculation of the benefits and costs of particular actions. The possibility of such an independent preference was recognized by Wicksteed, who remarks that a man setting his selling price by reference to past costs of production “is either allowing an irrelevant consideration to affect his judgment or else is deliberately taking a commercial risk to gratify a personal feeling.”¹¹

In the case of Bill’s machine considered above, Bill might have promised his great-aunt, when she was on her deathbed, that he would close down the factory if it failed to recoup the cost of the machine. Any number of such motivations could be imagined. Perhaps in a stressful moment, Bill swore a vow to himself that he would recoup

11. Wicksteed, p. 386.

the cost of the machine or close the factory and now feels obliged to comply with that vow. Such cases are unusual—typically, Bill will not have sworn any such peculiar oath. However, even in those few cases where such a motive arises, the sunk costs doctrine is not contradicted. For in such cases, an additional goal of counting historical costs has been added to the conditions of the problem. As a result, a historical cost is used to derive a goal; in the stated conditions, the historical cost does not influence decisions because it is a cost but rather because it generates a goal. In such a case, the agent does not commit the mistake of supposing that the historical cost of the machine is a cost of keeping the factory running.

XII. NOZICK'S FOURTH ARGUMENT

Nozick's fourth argument occupies the paragraph beginning near the bottom of page 23. As with argument 1, argument 4 has the agent deliberately incurring costs because he reckons that in the future he will count sunk costs and thus will be motivated to behave differently. But unlike argument 1, argument 4 relates not to a specific action (such as going to the theater on a particular night) but to the choice of a *principle*. Nozick conceives a principle as a rule for grouping actions so that they are treated uniformly (pp. 3–4, 17–18). For the most part, my criticisms of Nozick's argument 1 also apply to his argument 4, but there are additional difficulties, of which I will mention two.

First, what is meant by choosing one principle rather than another? One would think that the only principle in question were that of not succumbing to temptation. I surmise that Nozick is implicitly referring back to his earlier discussion (p. 19, lines 13–17), and that the sort of thing he has in mind here is the desirability of a motive for preferring (a) the principle never to succumb to temptation to (b) the principle never to succumb to temptation except on a single imminent occasion. If this is Nozick's intention here, then it is mistaken. If past costs would afford a reason for maintaining principle a, they would equally afford a reason for maintaining principle b. The past costs provide no means of discriminating between the two alternative principles.

Second, the Concorde principle is irrelevant here, because it is in essence the notion that we should do something to recoup a past loss. In contrast, anyone can see that it is senseless to say: "I expended costs, yet these were amply compensated by the ensuing benefits; nonetheless, I must do something to recover those costs." Why should we suppose that the individual who has successfully resisted temptation on a number of occasions will now regard himself as being, so to speak, behind on the deal? Unless he does regard himself as being in a losing position so far, the Concorde fallacy is inapplicable.

Nozick's discussion of principles points to the possibility of a more straightforward and thoroughly rational approach: by grouping actions together, the agent attributes more significance to each action by virtue of its membership in the group. Thus, the agent derives more utility from an action because it belongs to the group,¹² and, hence, the cost of failing to perform that action is increased. This account is admittedly limited in that it leaves open the process by which the agent comes to ascribe enhanced significance to actions as members of a group, but some such method has decided advantages over relying on the Concorde approach, for example: (a) it does not depend upon the inherently undesirable practice of relying on one's own muddleheadedness to make one do the right thing; (b) it has more direct relevance to the important notion of commitment, for the degree of commitment can vary independent of total past costs expended; (c) it avoids perverse consequences of the Concorde approach (such as abandoning continued pursuit of a goal *because* one discovers that it is turning out to be easier to achieve than one had believed, or deliberately incurring unnecessary costs only because they will affect one's later behavior through the Concorde fallacy); and (d) it avoids the indeterminacy of the Concorde approach (which urges one to keep allocating more resources to a losing project, without specifying at what point, if ever, this throwing of good resources after bad is to be discontinued).

XIII. NOZICK'S SECOND, THIRD, AND FIFTH ARGUMENTS

Nozick's second argument is framed thus: "We do *not* treat our past commitments to others as of no account except insofar as they affect our future returns" (p. 22). The immediate context shows that Nozick believes that this statement contradicts the sunk costs doctrine. But economists have never been in any doubt that individuals sometimes *do* erroneously count sunk costs.¹³ And the mere fact of having a preference for keeping commitments to others does not entail any counting of sunk costs.

Nozick's third argument is that "we do *not* treat the past efforts we have devoted to ongoing projects of work or life as of no account

12. But not as much utility as is derived from the entire group. Nozick claims that "the penalty for violating the principle this time becomes the disutility of violating it always" (p. 19). We do not, I take it, want to view getting drunk today as *equivalent* to getting drunk every remaining day of our lives.

13. Wicksteed, pp. 91–94, 387; Robert H. Frank, *Microeconomics and Behavior* (New York: McGraw-Hill, 1991), pp. 13–14, 226–36. For a typical example of the way in which economists acknowledge that businesspeople are prone to count costs and exhort them not to do so, see Thomas H. Nagle, *The Strategy and Tactics of Pricing: A Guide to Profitable Decision-Making* (Englewood Cliffs, N.J.: Prentice Hall, 1987), pp. 21–23.

(except insofar as this makes their continuance more likely to bring benefits than other freshly started projects would). Such projects help to define our sense of ourselves and of our lives" (p. 22).

Here, too, what we do is not conclusive with respect to what we should, rationally or efficiently, do. And at the point of decision, defining her sense of herself and of her life is preferred by the agent to other desiderabilia. This does not conflict with anything in economic theory and has nothing to do with sunk costs. There is not even any claim here that the agent values defining her sense of herself and of her life just because she has expended resources on this in the past—although that would not amount to a sunk costs fallacy, as we have seen.

Nozick's fifth argument (p. 24, lines 34–40), draws upon a well-known argument by Schelling¹⁴ that it may be useful to convince others that we will "irrationally" stick to our guns even in the face of threats which make it disadvantageous for us. The possibility that commitment to certain behaviors, a commitment that disregards momentary estimations of advantage, may be advantageous in the long term, because of the way in which awareness of such a commitment may modify the behaviors of other people, has been familiar to many economists for at least some decades.¹⁵ It has nothing intrinsically to do with sunk costs, although there has been discussion of the advantages of behaving "irrationally,"¹⁶ and if the propensity to maintain such commitments were made more likely because of a propensity to count sunk costs, this would be one more conceivable way in which an error could, on occasion, have advantageous consequences.

I conclude that of Nozick's five arguments against the economists' doctrine, arguments 1 and 4 are faulty in several fatal respects, while arguments 2, 3, and 5 embody simple mistakes. My narrow concern has been Nozick's claim that he has refuted the sunk costs doctrine as economists understand it, and my argument should not be taken to imply that there is nothing of value in Nozick's discussion or that parts of it could not profitably be reformulated to avoid the unwarranted allegation that the economists' doctrine is at fault.

14. Thomas Schelling, *Arms and Influence* (New Haven, Conn.: Yale University Press, 1966), pp. 35–91.

15. See Robert H. Frank, *Passions within Reason: The Strategic Role of the Emotions* (New York: Norton, 1988), pp. 47–70, and the references cited in p. 47 n.

16. Schelling, *Arms and Influence*, and *The Strategy of Conflict* (Cambridge, Mass.: Harvard University Press, 1960); Frank.