

Balancing the Costs of Forecasting Errors in Parole Decisions*

Richard Berk
Department of Statistics
Department of Criminology
University of Pennsylvania

11/30/2010

1 Introduction

Parole decisions are can be based on a variety of factors. Some are automatically considered and closely prescribe the actions to be taken. Among these are mandatory releases after a full prison term has been served. Other factors can be discretionary. They may or may not be considered and their use is subject to interpretation.

Among the discretionary factors commonly weighed are two fundamental kinds of risk. One risk is that an individual will be released only to commit new crimes, sometimes very serious ones. The other risk is that an individual will not be released although behavior on parole would have been exemplary. Since the 1920s, parole boards in the United States have tried to minimize both kinds of risks by constructing forecasts of how individuals will fare under supervision in the community and then using those forecasts to inform parole decisions (Borden, 1928; Burgess, 1928). The results of this enterprise are mixed (Dean and Dugan, 1968; Farrington, 1987; Gottfredson and Moriarty, 2006; Berk, 2008a).

*Thanks go to Shawn Busway and Jim Acker for very helpful comments on an earlier draft of this paper.

One reason for the mixed performance is a failure to appreciate that there are inevitable tradeoffs between the two kinds of parole risks. If one is more likely, the other is less likely. A related reason is that the consequences of the competing risks have been almost universally been treated symmetrically. The costs of releasing an individual who then commits crimes are implicitly assumed to equal the costs of failing to release an individual who would have succeeded in parole. An important consequence is that the forecasts that follow build in that equivalence (Berk and Cooley, 1987). If the costs are not really equal, the forecasts can be misleading, often badly so (Berk, 2010).

In this paper, I will review ways to think about forecasting errors in parole decisions. The key issue will be how to properly address the consequences of the two kinds of forecasting errors, known more formally as false positives and false negatives. In the parole context, false positives can be individuals incorrectly projected to be poor parole risks. False negatives can be individuals who are incorrectly projected to be good parole risks. Over many parole decisions, both kinds of errors are virtually inevitable and when improperly considered, adversely affect the quality of those decisions.

2 False Positives, False Negatives, and Forecasting

For a didactic discussion of false positives and false negatives, it is helpful to work with archetypes. Popular culture can provide them. For this paper, therefore, one of the ongoing appeals of the Star Wars movies is the presence of good and evil starkly represented respectively by Luke Skywalker and Darth Vader. Darth Vader vaporized an entire planet populated by millions of peace-loving humanoids. Luke Skywalker saved the Federation of such planets.

	Predict Darth Vader	Predict Luke Skywalker
Darth Vader	True Positive	False Negative
Luke Skywalker	False Positive	True Negative

Table 1: Parole Forecasting Typology

2.1 False Positives and False Negatives

In the extreme at least, parole boards may need to ascertain whether individuals whose cases are being reviewed are more like Darth Vader or more like Luke Skywalker. Parole boards are often especially sensitive to violent crimes committed by individuals on parole so their risk assessment instruments and procedures may be tuned to identify the Darth Vaders. When they succeed, the term “true positive” can be applied. When they fail, the term “false negative” can be applied. In a parallel fashion, if the risk instrument correctly identifies a Luke Skywalker, one can use the term “true negative.” And when that identification is in error, the term “false positive” can be applied. The four possibilities are shown in Table 1, with the two kinds of errors in bold type. A case can fall in any one of the four cells.

Looking at Table 1, there is a simple way to avoid all false negatives: just proceed as if all cases were like Darth Vader. That is, use only the “Predict Darth Vader” column. One can see that all of the true Darth Vaders will be correctly identified. However, all of the true Luke Skywalkers will automatically become false positives. In a parallel fashion, one can see in Table 1 that there is a simple way to avoid all false positives: just proceed as if all cases were like Luke Skywalker. That is, use only the “Predict Luke Skywalker” column. One can see that all true Luke Skywalkers will be correctly identified. However, all true Darth Vaders will automatically become false negatives. Both solutions are simple because there is no need to do any forecasting. One rule or the other can be applied without any consideration of risk. But in practice, both solutions are likely to be unsatisfactory in part because all differences between individuals are ignored. For example, first-time offenders are treated the same as habitual offenders. One response is to introduce forecasts of risk based on individual differences.

2.2 Parole Decisions as Parole Forecasts

At the time parole decisions are made, the consequences of parole decisions cannot be known. That is why forecasts are required even when they are not acknowledged as such. A real forecast is being made, often implicitly. If the forecasted outcome is favorable, the individual is more likely released. If the forecasted outcome is unfavorable, the individual is less likely released.

The basis of parole forecasts is necessarily historical. In effect, past cases serve as “training data.” For these cases, the parole decision and subsequent

outcomes can be known. Why cases fall in any of the four cells in Table 1 can be considered. Sometimes that information is used to construct systematic risk instruments. Sometimes it is used anecdotally. In both instances, the historical information informs current decisions that are also shaped by various legal requirements and often combined with behavioral evaluations derived from formal clinical training, research findings and theory from the social sciences, and a range of craft lore. For example, an inmate's "motivation" can be an important factor.

Because at the time parole decisions are made, one cannot know into which of Table 1's cells a case actually belongs, Table 1 is no help with the instant case. Rather, Table 1 is an essential part of the process by which *forecasting procedures* are constructed from historical data. We turn to that next.

2.3 The Relative Costs of Forecasting Errors

In an ideal world, the forecasts would be perfect. There would be no false positives and no false negatives. In practice, there can be a substantial number of both. Moreover, there are tradeoffs between the two. If the net is cast to catch a larger number of true negatives, it will almost certainly catch a larger number of false positives as well. If the net is cast to catch a larger number of true positives, it will almost certainly catch a larger number of false negatives as well. This means that as false negatives decrease, false positives increase, and likewise, as false positives increase, false negatives decrease. How then should the tradeoff between false positives and false negatives be undertaken? The answer is that it depends on their relative costs.

When a case becomes a false positive or false negative, a forecasting error has been made, and there costs associated with each. A false negative often means that a serious crime has been committed. There are one or more victims and various criminal justice agencies become involved. A false positive often means that an individual spends additional time in prison, which not only can be costly in monetary terms, but can have negative consequences for the individual and the individual's family. Obtaining credible quantitative costs for such forecasting errors can be difficult. Fortunately, it is also not necessary. All one needs is the relative costs: how much more or less costly a false negative is compared to a false positive.

One can arrive at the relative costs by deciding how many false positives

one will accept for every false negative or equivalently, how many false negatives one will accept for every false positive. This is really not much different from more familiar pricing mechanisms. If a consumer in a supermarket is indifferent between four boxes of cereal and one gallon of milk, a gallon of milk is worth four times more to that consumer than a box of cereal. Note that the absolute value of the cereal and the milk need not be determined to arrive at their relative value.

In the parole setting, if one can accept 10 false positives for every false negative, false negatives are 10 times more costly than false positives; one false negative counts the same as 10 false positives. Conversely, if one can accept 10 false negatives for every false positive, false positives are 10 times more costly than false negatives; one false positive counts the same as 10 false negatives.¹

2.4 Building in Relative Costs

The relative costs of forecasting errors should be consciously built into the forecasts, and the reasoning rests on common sense. Which is worse: treating a Luke Skywalker like a Darth Vader or treating a Darth Vader like a Luke Skywalker? And then, how much worse? For example, it is much worse to treat a Luke Skywalker like a Darth Vader, one wants to be very sure that a case is indeed a Darth Vader. Stated more formally, If false positives are much more costly than false negatives, one should quite sure before a case is forecasted as Darth Vader. In contrast, if it is much worse to treat a Darth Vader like a Luke Skywalker, one wants to be very sure that a case is indeed a Luke Skywalker. Stated more formally, If false negatives are much more costly than false positives, one should quite sure before a case is forecasted as Luke Skywalker. These ideas can be made quite precise.

Consider the following example. Suppose one has developed a risk scale for violent crime. That scale ranges from a probability of 1.0 that a violent crime will be committed to a probability of 0.0 that a violent crime will be committed. At the extremes, one is certain that a violent crime will be committed or that a violent crime will not be committed. How this scale might be used is illustrated in Figure 1.

¹In economics, the cost of a product is not necessarily the same as the price of a product. The former speaks to the value used up when the product is made. The latter speaks to the value at which that product is sold. But parole decisions are not undertaken in a market, so the term cost (or loss) will be used.

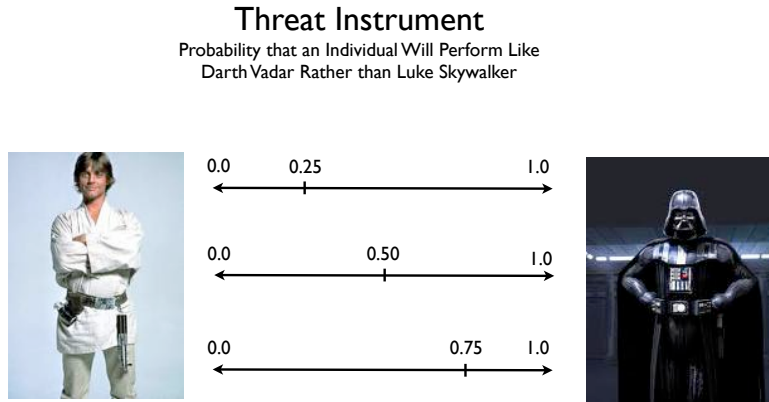


Figure 1: Changing the Forecasting Threshold

If the costs of false negatives and false positives are the same, an individual with violent crime risk probability greater than .50 should be treated as a potential Darth Vader. An individual with a violent crime risk probability of .50 or less should be treated as a potential Luke Skywalker.² This is represented by the middle two-headed arrow in Figure 1. The threshold of .50 splits the arrow into two segments of equal length, implying that false negatives and false positives have equal costs³

Suppose now that the costs of a false negative are taken to be three times the costs of a false positive. It follows that the violent crime risk threshold should be dropped and in particular, the probability threshold should be set at .25 ($.75/.25 = 3.0$). This is represented by the top two-headed arrow in Figure 1. The threshold splits the two-headed arrow so that the right

²What is done with individuals whose risk is exactly .50 does not matter in practice. Those individuals can be placed on either side of the .50 threshold. In effect, the two outcomes are equally likely, and one could just as well flip a coin.

³As the probability of being a Darth Vader drops, the probability that an individual is not a Darth Vader necessarily increases. In this example, that is the same as saying the probability that an individual is a Luke Skywalker increases.

segment is 3 times longer than the left segment. Locating the threshold in this fashion captures more potential Darth Vadars and is required by the 3 to 1 cost ratio. Put in other terms, the evidence hurdle that must be exceeded in order to treat an individual as a Darth Vadar has been reduced by a factor of three compared to when the probability threshold was .50. Consequently, false negatives are reduced; fewer Darth Vadars are treated as if they were Luke Skywalker. The tradeoff, therefore, is that there will be more Luke Skywalkers treated as if they were Darth Vadar — more false positives. But this is fully consistent with the 3 to 1 cost ratio.

Conversely, suppose the costs of a false negative are taken to be one third the costs of a false positive. Now, the violent crime risk threshold is raised and in particular, the probability threshold should be set at .75 ($.25/.75 = 1/3$). This is represented by the bottom line in Figure 1. The threshold splits the two-headed arrow so that the right segment is one-third the length of the left segment. The threshold is set in this fashion to capture more Luke Skywalkers as required by the 1 to 3 cost ratio. The evidence hurdle that must be exceeded to treat an individual as a Darth Vadar has been raised by a factor of three compared to when the probability threshold was .50. Consequently, false positives are reduced; fewer Luke Skywalkers are treated as if they are Darth Vadar. The tradeoff is that more Darth Vadars will be treated as if they were Luke Skywalker — more false negatives. This is fully consistent with the 1 to 3 cost ratio.⁴

Looking back at Table 1, lowering the threshold has the effect of moving cases from the right column to the left column. Conversely, raising the threshold value has the effect of moving cases from the left column to the right column.

In summary, when parole forecasts are being made, one is in practice always trading false negatives against false positives. Even when these tradeoffs are not recognized or acknowledged, they are in play. And in those tradeoffs are necessarily their relative costs. The evidence threshold can be moved up or down to take those relative costs into account. The usual assumption implicitly made is that the costs are identical in which case the evidence threshold is located exactly in the middle of its range favoring neither false positives or false negatives. If this balance is not consistent with parole board preferences, the wrong evidence threshold is being used and

⁴Any cost ratio can be represented in this fashion. The 1 to 1, 3 to 1, and 1 to 3 ratios are just illustrations.

parole board preferences are being overridden. The good news is that the evidence threshold can be changed to be more consistent with parole board preferences.

3 Determining Relative Costs

As already noted, it can be extremely difficult to accurately quantify the costs of forecasting errors. For example, how would one assign a dollar value to a heinous crime committed while on parole or to the damage two extra years in prison can do to an inmate's employability. We have also seen that, fortunately, it is not necessary to quantify such costs. All that is required is relative costs. But, how can these relative costs be determined?

One approach is to ask explicitly how many false positives one can accept for every false negative or how many false negatives one can accept for every false positive?⁵ This seems like a manageable task. Nevertheless, arriving at such ratios can be challenging.

To begin, one must determine whose ratios matter? Who should be the relevant stakeholders? Presumably parole officials are in the mix. But what about former victims, the offender, or the offender's family? What about corrections officials, state legislators or members of the state's executive branch? To the best of my knowledge, there has been no systematic discussion of such matters.

Another task is to elicit the cost ratios. Perhaps surprisingly, I have found that experienced criminal justice officials have little trouble when asked. Once reminded of the consequences that can follow from false positives and false negatives, they can rapidly decide "how much worse" one is than the other. For the jurisdictions I have worked in, parole officials treat false negatives as far more costly than false positives. Costs ratios of 10 to 1 or even 20 to 1 are common. However, I have no experience with other kinds of stakeholders.

Even if the set of relevant stakeholders is relatively homogeneous (e.g., the state parole board), there will often be different views about what the ratio of false positives to false negatives should be. Then, the problem is

⁵Alternatively and equivalently, one can ask how many false positives one can accept for every true positive or how many false negatives one can accept for every true negative? Sometimes this is an easier method to implement, but it is a detail that will not be discussed here.

how to work with the different cost ratios. One option is to average them. But this can be unsatisfactory because some stakeholder's views should have more weight than others. For example, the view held by a parole board chair might be more important than the view held by any given parole board member. Unfortunately, weighting the ratios as part of the averaging introduces another messy complication. Precise weights are required and it is not clear how such weights would be determined.

An alternative solution is to provide different forecasting procedures depending on the ratio of false positives to false negatives. This can help the process by which the implications of the forecasts can be addressed. At the very least, playing through the results of several different forecasting procedures derived from several different cost ratios could be a useful exercise undertaken every few years.

The best solution is for the stakeholders to arrive at a single consensus cost ratio. In my experience, this has not been difficult. The precise cost ratios usually articulated by individual stakeholders are not held with great conviction. The difference, for example, between a cost ratio of 10 to 1 and 5 to 1 may not represent a principled disparity. In effect, the difference is noise. As a result, most stakeholders can readily compromise. Matters would no doubt be more contentious should the stakeholders be more heterogeneous.

3.1 Some Further Complications

In practice, there are several of additional complications. First, whether an individual turns out to be a true positive or true negative depends heavily on how behavior on parole is measured. Arrests, for instance, will miss crimes for which no arrest is made. Moreover, the arresting charge may not correspond well to the criminal behavior actually undertaken. Then, there is usually a translation into the categories used to characterize violations of parole. The hope is that those categories are useful and yet broad enough to sufficiently accurate. Thus, for parole violation purposes it may not matter which kind of murder was committed. The individual has failed because of a violent crime.

Second, there may be a desire to have different cost ratios for different individuals. For example, the costs to an individual from additional time in prison can vary over time and over individuals. However, it will usually not be practical to arrive at a variety of cost ratios depending on the circumstances surrounding a given parole review. In addition to deciding about what to do

with the individual, the cost ratio would have to be determined. The time needed to make a parole decisions could increase substantially. More practical would be a few different cost ratios determined in advance and applied to broad categories of offenders. For example, the costs of incarceration are perhaps higher for individuals who have a wife and children.

Third, there are systemic issues that go beyond cost ratios surrounding particular individuals. A common problem is that if the cost ratio implies a large number of false positives, prison capacity may at some point be sorely taxed. That is, there is one cost ratio when there are plenty prison beds and a very different one when there are not. If changing circumstances such as these can be anticipated, they can be built into the forecasts. When they cannot, there needs to be the capacity to rebuild the forecasting procedures taking the new circumstances into account.

Fourth, there must be clarity among decision makers about the consequences of overriding the forecasts. Overriding the forecasts can be useful when there is *new* information available not used when the forecasting procedures were developed. For example, the forecasting procedures may not have taken into account health conditions that could not be properly treated in prison hospitals or that may imply the need for a parole on humanitarian grounds. In effect, the costs of a false positive are increased as new facts are brought to bear. In contrast, overriding a forecast because of a view that the *existing* information has not be properly evaluated, will likely degrade the quality of the forecasts in an *ad hoc* manner. At least since the middle 1970s, it has become clear that subjective judgments about risk and uncertainty are subject to a wide variety of biases and that in general, statistical procedures can do a better job (Tversky and Kahneman, 1974; Kahneman, Slovic, and Tversky, 1982). Subjective judgements informed by clinical expertise are compromised by the same sorts of problems (Dawes, 1996; Kreuger, 2008). Finally, the override may actually be an implicit effort to alter the cost ratio. If so, that should be addressed directly.

Fifth, there may be concerns that the cost ratios imposed have little empirical justification. That is, a complete cost analysis might produce very different results. On the one hand, that may not matter. What matters is the *preferences* of stakeholders, and a key feature of these preferences can be captured in a elicited cost ratios. On the other hand, providing more thorough information about the relative costs of false negatives and false positives can usefully inform stakeholder preferences. In other words, the forecasts should respond to stakeholder preferences about the costs of fore-

casting errors, but well-informed preferences are surely desirable. If nothing else, forecast-informed parole decisions may be less subject to challenge.

Finally, there is only one way to reduce both false negatives and false positives: improve the performance of the forecasting procedures. This is not the place to consider the issues, but over the past several years there has been some real success using new statistical procedures coupled with the power of modern computers. Rather than trying to develop theoretically informed models based on factors that assumed to affect behavior on parole, the new approaches are crassly empirical. They “mine” very large data bases containing hundreds of variables to find patterns that can help in forecasting (Berk, 2008b). At least so far, these methods seem to perform far better than conventional modeling approaches (Berk et al., 2009).

4 An Example

It should be clear that if the cost ratio of false negatives to false positives (or equivalently vice versa) is properly taken into account, forecasts can be affected in a useful manner. To fix this idea, we turn to an example. The data are comprised of a little over 10,000 individuals for whom parole forecasts were needed. The outcome is an arrest for any violent crime committed within 18 months of release on parole. For this sample of offenders, around 10% are arrested for a violent crime within 18 months of release on parole.

From a statistical perspective, a violent crime is for this population a rare event. As such, it is difficult to find predictors to improve overall forecasting accuracy. In this instance, if one simply forecasted that no violent crime would be committed within 18 months of release on parole, that forecast would be correct about 90% of the time. However, this is probably unsatisfying because all of the true positives would be false negatives — all of the Darth Vadars would be forecasted as Luke Skywalkers. In effect, one is assuming that false positives are far more costly than false negatives. In practice, it is likely that the reverse would be true. The simple forecast is not responsive to the likely cost ratio. It is important, therefore, to try to do better.

For this example, most of the predictors come from the usual sources that parole officials consult when they make parole decisions. Therefore, the usual predictors are included (e.g., age, gender, marital status, prior record). There are also some predictors that are often not available (e.g., behavior in prison, juvenile record). A machine learning procedure called Random

Forests (Breiman, 2001), implemented in the programming language R, was applied. Random forests have been effective in other related work (Berk et al. 2009, Berk, 2009).

	Predict Not Violent	Predict Violent	Percent Correct
Not Violent	7607	1877	80%
Violent	561	376	40%

Table 2: Forecasting Skill Using 3 to 1 Cost Ratio of False Negatives to False Positive

Table 2 show the result when the cost ratio of false negatives to false positives is set to about 3 to 1. The numbers in the middle two columns are counts of the number of cases that fall in each cell. The two numbers in the far right column are the percentage forecasted correctly in each row. The top row is for individuals who are not arrested for violent crimes on parole. The bottom row is for individuals who are arrested for violent crimes on parole. All of the numbers in the table are the result of real forecasts. They are derived from data not used to build the forecasting algorithm.

One can see that about 80% of the negatives (i.e., no failure) are correctly forecasted; given that an individual did not commit a violent crime, the statistical procedures correctly forecasts that outcome 80 times out of 100. About 40% of the positives (i.e., failure) are correctly forecasted; given that an individual committed a violent crime, the statistical procedure correctly forecasts the outcome 40 times out of 100.

Because for these data, approximately 90% are not arrested for a violent crime within 18 months of release, it is not surprising that such cases are identified with substantial accuracy. The performance for the positives is disappointing by comparison. However, one must not forget that arrests for violent crimes are relatively uncommon. Then, the improvement from a base rate of 10% to 40% looks better. That is, in the overall sample, the failure rate is about 10%. In the subset identified by the forecasts, the failure rate is about 40%, a four-fold improvement.

Table 3 show the results when the cost ratio of false negatives to false positives is set to 30 to 1. False negatives are taken to be far more costly than false positives and are given more weight as the forecasts are constructed. In effect, the threshold to identify a positive as been lowered substantially. One result is that there should be more false positives.

	Predict Not Violent	Predict Violent	Percent Correct
Not Violent	3998	5486	42%
Violent	156	781	84%

Table 3: Forecasting Skill Using 30 to 1 Cost Ratio of False Negatives to False Positive

In Table 3, the negatives are forecasted with only about a 42% accuracy, and the positives are forecasted with about an 84% accuracy. Performance for the negatives has been cut approximately in half. But, within the subset of offenders who were actually arrested for violent crimes, well over 80% can be forecasted correctly. At the same time, there are costs. There are about 7 false positives for every true positive (5486/781). When the cost ratio was set to 3 to 1, there were about 5 false positives for every true positive. Just as expected, raising the relative cost of false negatives results in more false positives.

The two cost ratios were set to provide a likely lower and upper bound for those ratios. In practice, something in between is a likely choice. It is also important to stress that both sets of results are formally correct. The underlying data are identical, and no technical errors have been made. Table 2 differs substantially from Table 3 because the imposed cost ratios differ dramatically. Which results (or some others) would be more responsive to the needs of stakeholders depends on the cost ratio they favor. And that is a policy decision to be made before the data are analyzed. It is a constraint imposed on the forecasts.

Finally, there is extensive and informative output from random forests that would ordinarily be examined. In particular, there are measures of each predictor's unique contribution to forecasting accuracy. There are also plots of the functional form though by each predictor operates on the forecasts. Although both kinds of output can be very instructive, they are a diversion for this paper. Nevertheless, it is useful to add that one strong feature of random forests is that all of its output takes the cost ratio of forecasting errors into account. This is not true of most other methods.

5 Conclusions

Forecasting errors are part of the forecasting process, and all forecasting errors have costs. To ignore these costs is to accept a default that the costs of false negatives are the same as the costs of false positives. Forecasts build in this equivalence. When equal costs are inconsistent with the preferences of stakeholders, forecasts assuming equal costs can be very misleading. The solution is to elicit stakeholders' preferences for the relative cost of false negatives and false positives and allow the cost ratio to affect the forecasts. In some jurisdiction, this is being done.

References

- Berk, R.A., (2008b) *Statistical Learning from a Regression Perspective*. New York: Springer.
- Berk, R.A. (2008a) Forecasting Methods in Crime and Justice.” *Annual Review of Law and Social Science*, Vol 4: 173-192, J. Hagan, K.L. Schepple, and T.R. Tyler (eds.), Palo Alto: Annual reviews.
- Berk, R.A., (2009) The Role of Race in Forecasts of Violent Crime.” *Race and Social Problems* 1: 231–242.
- Berk, R.A. (2010) Asymmetric Loss Functions for Forecasting in Criminal Justice Settings. *Journal of Quantitative Criminology*, forthcoming.
- Berk R.A. and Cooley TF. 1987. Errors in Forecasting Social Phenomena. *Climatic Change* 11: 247-265.
- Berk, R.A., Sherman, L., Barnes, G., Kurtz, E., and Ahlman, L. (2009) Forecasting Murder within a Population of Probationers and Parolees: a High Stakes Application of Statistical Learning. *Journal of the Royal Statistical Society* (Series A) 172, part 1: 191–211.
- Borden, H.G. (1928) Factors Predicting Parole Success. *Journal of the American Institute of Criminal Law and Criminology* 19: 328-336.
- Breiman, L. (2001) Random forests. *Machine Learning*, **45**, 5-32.
- Burgess, E.W. (1928) Factors Determining Success or Failure on Parole. In Bruce, A.A., Harno, A.J., Burgess, E.W., and Landesco, J. (eds.) *The Working of the Indeterminate Sentence Law and the Parole System in Illinois*. Springfield, Illinois, State Board of Parole: 205–249.
- Dawes, Robin (1996) *House of Cards*. New York: Free Press.
- Dean, C.W., and Dugan, T.J. (1968) Problems in Parole Prediction: An Historical Analysis. *Social Problems* 15: 450-459.
- Farrington, D.P. (1987) Predicting Individual Crime Rates. In D. M. Gottfredson and M. Tonry (eds.), *Prediction and Classification*. Chicago: University of Chicago Press.

- Gottfredson, S.D., and Moriarty, L.J. (2006) Statistical Risk Assessment: Old Problems and New Applications. *Crime & Delinquency* 52(1): 178–200.
- Kahneman, D., Slovic, P., and Tversky, A. (1982) *Judgment under Uncertainty: Heuristics and Biases*. Cambridge: Cambridge University Press.
- Kreuger, J.I. (2008) *Rationality and Social Responsibility*. New York, Free Press.
- Tversky, A. and Kahneman, D. (1974) Judgement Under Uncertainty: Heuristics and Biases. *Science* 185 (4157): 1124-1131.