

SUPPORTING MATERIAL

We estimate the similarity score PDFs $F(x)$ and $G(y)$ in §1, the Israeli casings parameter values in §2, and the U.S. parameter values in §3. Justification for using a threshold-based approach is presented in §4 and a comparison of the false negative constraint and the lot size constraint is performed in §5. Additional numerical results discussed in the main text appear in Fig. 11 and Tables 3-4.

1 Estimating the Similarity Score PDFs

For cartridge casings, we estimate the lognormal parameters $(\mu_f, \sigma_f, \mu_g, \sigma_g)$ using the lower left graph in Fig. 12 of [1], which plots the probability that a true match is ranked among the top 10 scores when the database contains one true match and N nonmatches, as N varies from 0 to 10^6 . If the intra-gun similarity score PDF is $f(x)$, then this probability is

$$\sum_{M=0}^9 \frac{N!}{M!(N-M)!} \int_0^\infty (1-G(t))^M G(t)^{N-M} f(t) dt. \quad (1)$$

Using seven points along the performance curve in Fig. 12 of [1], we derive the least-squares estimates $\mu_f = 6.7912$, $\sigma_f = 0.5927$, $\mu_g = 4.5207$, $\sigma_g = 0.5016$. Fig. 2 compares the seven data points to the performance curve predicted (via (1)) by the lognormal distributions, and Fig. 3 shows the resulting PDFs.

For bullets, the inter-gun similarity score parameter values $\mu_g = 5.6817$ and $\sigma_g = 0.2014$ are estimated using maximum likelihood estimation directly from the PDF in Fig. 2 in [2]. The resulting fit appears in Fig. 4. The intra-gun similarity score parameters are derived from lead bullet matching results in [3], where in a database of 475 bullets that contains one true match, seven out of 10 arrivals generated an intra-gun score that ranked first, two arrivals came in second, and one arrival came in third place. We use the CDF $G(y)$ estimated above and a least-squares approach to match the probabilities that an intra-gun similarity

score in a database of size 475 comes in first, second and third positions with probabilities 0.7, 0.2, and 0.1, respectively. The final result is $\mu_f = 6.3369$, $\sigma_f = 0.0718$, which gives the three probabilities 0.690, 0.199, 0.065. The two lognormal PDFs for bullets are displayed in Fig. 5.

2 Estimating Israeli Casings Parameter Values

We estimate the probabilities q_i , r_j and p_{ij} in §2.1, the PMF $g_i(n)$ and the historical threshold t_h in §2.2, and the age PDFs $h(a)$ and $h_m(a)$ in §2.3.

2.1 Estimating the Probabilities q_i , r_j and p_{ij}

The evidence database contains 14,979 entries covering the entire country during 1980-2000. The arrivals data consist of all matches detected between January 1, 2006 and December 31, 2008. There were 7138 arrivals during this time period, and 697 of them generated matches. We start by estimating the three probabilities, q_i , r_j and p_{ij} . Of the 7138 arrivals, 3598 were nonevidence and 3540 were evidence, yielding $q_0 = 0.504$, $q_1 = 0.496$.

We know which of the 89 Israeli police precincts collected each arrival and each database entry. Because many pairs of precincts generated no matches during 2006-2008, we use only two categories (e.g., $J = 1$): intra-location and inter-location. However, to fully exploit the spatial information, it is desirable to merge two locations if it results in a more favorable ratio of intra-location matches to inter-location matches. Each of the 89 Israeli police precincts has an 89-dimensional vector stating the number of matches during 2006-2008 that it has with each precinct. In our analysis, two locations are good candidates for merging if the dot product of their vectors is large; we refer to this dot product as the two locations' similarity, and solve the graph partitioning problem that maximizes the average similarity within each group. We solve this problem using the k -means algorithm [4] with $k = 20$, which merges

the 89 precincts into 20 groups so as to maximize the average similarity within each group. The k -means algorithm is random, and solutions can vary from run to run. We solve the problem 1000 times, and group precincts into stations if they are classified in the same group in $> 90\%$ of the solutions. We left all other precincts as isolated to prevent overfitting. This procedure results in 18 merged groups containing 55 precincts, and 34 isolated precincts, for a total of 52 “new” stations. This merging process increases the intra-location matching proportion from 0.551 to 0.832 for nonevidence, and from 0.595 to 0.875 for evidence.

If we let a_k and d_k denote the number of arrivals from station k and the number of database records from station k , then $r_0 = \sum_{k=1}^{55} \frac{a_k d_k}{(\sum_{l=1}^{55} a_l)(\sum_{l=1}^{55} d_l)} = 0.101$ and $r_1 = 0.899$, where the subscript $j = 0$ corresponds to intra-station and $j = 1$ corresponds to inter-station. Of the 697 matches, 107 were from nonevidence arrivals and 590 were from evidence arrivals. Of the 107 nonevidence arrivals generating matches, 89 were intra-station, giving $p_{00} = 89/107 = 0.832$ and $p_{01} = 0.168$. Similarly, we have $p_{10} = 516/590 = 0.875$ and $p_{11} = 0.125$.

2.2 Estimating the Historical Threshold t_h and the PMF $g_i(n)$

The observed PMF $g_i^*(m)$ (Fig. 6) allows us to estimate the historical threshold t_h for a constant threshold policy that generates an average candidate list size of 10, by assuming that the mean number of true positives is much less than the database size (i.e., $\sum_{i=0}^1 q_i \sum_{n=1}^{\infty} n g_i^*(n) \ll N$), yielding

$$\sum_{i=0}^1 q_i \sum_{n=1}^{\infty} n g_i^*(n) + N[1 - G(t_h)] = 10. \quad (2)$$

Setting $N = 11,350$, which is the average value of the database size during 2006-2008, in (2) gives $t_h = 442.6$.

In the remainder of this subsection, we estimate $g_i(n)$, which is the PMF of true matches, from three quantities: $g_i^*(m)$, which is the observed PMF of detected matches, the

historical threshold t_h under a constant threshold system, and the intra-gun similarity score CDF $F(x)$. Let P_{mn} be the probability that m matches are detected from an arrival in the Israeli database, given that n matches to this arrival exist and the arrival has evidence status i (the argument i in P_{mn} is suppressed for ease of presentation). It follows that $g_i^*(m) = \sum_{n=m}^{\infty} P_{mn} g_i(n)$. However, for practical purposes, we truncate this system of equations at $m = n = 14$ (i.e., we set $g_i(n) = 0$ for $n \geq 15$) because $g_i^*(m)$ is only nonzero for $m \leq 14$ in our data set. This truncated system of equations can be expressed as

$$\begin{pmatrix} P_{11} & P_{12} & \cdots & P_{1,14} \\ 0 & P_{22} & \cdots & P_{2,14} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & P_{14,14} \end{pmatrix} \begin{pmatrix} g_i(1) \\ g_i(2) \\ \vdots \\ g_i(14) \end{pmatrix} = \begin{pmatrix} g_i^*(1) \\ g_i^*(2) \\ \vdots \\ g_i^*(14) \end{pmatrix}. \quad (3)$$

Because $g_i^*(m)$ is known and the system of equations in (3) is invertible, our estimation problem for $g_i(n)$ reduces to determining the probabilities P_{mn} , which are a function of m , n and the known false negative probability $F(t_h)$ under the constant threshold system, which is denoted by p .

Conditioned on n matches existing in the database, the probability P_{mn} that a new arrival detects m of them depends on two things. First, it depends on the current knowledge about how these n entries are related, which can range from not realizing that any of them are matched to each other (i.e., there are n singletons) to realizing that they are all matched to each other (i.e., they are in a single group of n matches). Second, it depends on which matching groups, if any, the new arrival is detected to belong. Hence, to derive P_{mn} , we need to construct a detailed dynamic model that tracks the evolution of the n true matches as they sequentially arrive to the system and undergo a matching process (with false negative probability $p = F(t_h)$) with the prior arrivals.

This model is a Hidden Markov Model (HMM) because we cannot observe the state (or transition probabilities among states) [5]. More generally, the state of the HMM is defined by the current matched groupings of the true matches that have already arrived, where

the groupings are given in the ascending order of the size of the groups; e.g., state 1,1,2 means that there are a total of four true matches currently in the database, two of them are connected (i.e., have been correctly detected to be a match) and the remaining two are isolated (i.e., have not been correctly matched with any of the other three true matches). We use the notation $P(A|B)$ to represent the probability that, conditioned on the current matchings being in state B , after a new arrival undergoes the matching process with the prior arrivals, the new state is A (where, by construction, the sum of the numbers in A is always one greater than the sum of the numbers in B).

For illustrative purposes, we show how to use the HMM to compute P_{m3} for $m = 0, 1, 2, 3$, and then we provide a broad description of a general algorithm for any value of n . The initial state of the HMM is 1, which refers to the first of the n true matches having already entered the database. For this examples where $n = 3$, we need to track the HMM through n more arrivals (i.e., until after the fourth arrival) because the fourth arrival has $n = 3$ true matches in the database. These dynamics are described by the following transitions.

Transitions caused by the second arrival: $P(1, 1|1) = p$, $P(2|1) = 1 - p$.

Transitions caused by the third arrival: $P(1, 1, 1|1, 1) = p^2$, $P(1, 2|1, 1) = 2p(1 - p)$, $P(3|1, 1) = (1 - p)^2$, $P(1, 2|2) = p^2$, $P(3|2) = 1 - p^2$.

Transitions caused by the fourth arrival: $P(1, 1, 1, 1|1, 1, 1) = p^3$, $P(1, 1, 2|1, 1, 1) = 3p^2(1 - p)$, $P(1, 3|1, 1, 1) = 3(1 - p)^2p$, $P(4|1, 1, 1) = (1 - p)^3$, $P(1, 1, 2|1, 2) = p^3$, $P(1, 3|1, 2) = p(1 - p^2)$, $P(2, 2|1, 2) = p^2(1 - p)$, $P(4|1, 2) = (1 - p)(1 - p^2)$, $P(1, 3|3) = p^3$, $P(4|3) = 1 - p^3$.

Hence, the hidden states after the second arrival are (1, 1) and (2), the hidden states after the third arrival are (1, 1, 1), (1, 2) and (3), and the hidden states after the fourth arrival are (1, 1, 1, 1), (1, 1, 2), (1, 3), (2, 2) and (4). The HMM transition probabilities derived above

can be written as the following stochastic matrices, denoted by M_k ,

$$M_1 = \begin{pmatrix} p & 1-p \end{pmatrix},$$

$$M_2 = \begin{pmatrix} p^2 & 2p(1-p) & (1-p)^2 \\ 0 & p^2 & 1-p^2 \end{pmatrix},$$

$$M_3 = \begin{pmatrix} p^3 & 3p^2(1-p) & 3(1-p)^2p & 0 & (1-p)^3 \\ 0 & p^3 & p(1-p^2) & p^2(1-p) & (1-p^2)(1-p) \\ 0 & 0 & p^3 & 0 & 1-p^3 \end{pmatrix}.$$

Note that the product M_1M_2 equals the probability of arriving at the various states that the fourth true match will see upon arrival: $P(1,1,1) = p^3$, $P(1,2) = 3p^2(1-p)$, $P(3) = (1-p)^2(1+2p)$.

Finally, we group all the transitions from the fourth arrival according to how many matches are actually found. For example, the transition from $(1,1,1)$ to $(1,3)$ means that two matches are found. Using the law of total probability yields our final result:

$$P_{03} = P(1,1,1,1|1,1,1)P(1,1,1) + P(1,1,2|1,2)P(1,2) + P(1,3|3)P(3) = p^3,$$

$$P_{13} = P(1,1,2|1,1,1)P(1,1,1) + P(2,2|1,2)P(1,2) = 3p^4(1-p),$$

$$P_{23} = P(1,3|1,1,1)P(1,1,1) + P(1,3|1,2)P(1,2) = 3p^3(1-p)^2(1+2p),$$

$$P_{33} = P(4|1,1,1)P(1,1,1) + P(4|1,2)P(1,2) + P(4|3)P(3) = (1-p)^3(6p^3 + 6p^2 + 3p + 1).$$

For a general value of n , the calculation of P_{mn} for $m = 0, 1, \dots, n$ can be carried out by the following algorithm.

1 - Construct the HMM through the $n + 1^{\text{st}}$ arrival, and derive the transition matrices M_k for $k = 1, \dots, n$.

2 - Compute $\prod_{k=1}^{n-1} M_k$, which gives the probability distribution for the various hidden states that the $n + 1^{\text{st}}$ true match sees upon arrival.

3 - In the transition matrix M_n , find the number of matches m detected for each possible transition $A \rightarrow B$. Using the law of total probability, multiply the transition probability

from A to B in M_n by the probability of observing state A , which is $\prod_{k=1}^{n-1} M_k$, and add these products over all possible transitions to get P_{mn} .

Running this algorithm for $n = 1, \dots, 14$ results in the $g_i(n)$ PMF plotted in Fig. 6.

2.3 Estimating the Age Distributions

Of the 1364 matching pairs during 2006-2008, 1227 (or 90.0%) of them have database entries that occurred during 2006-2008, in which case we know the exact age in days. For the remaining 137 (or 10.0%) matching pairs, we only know the year when the database entry was acquired, and so the age is in the interval $[t_j, t_j + 365]$ for $j = 1, \dots, 137$. To estimate the PDF $h_m(a)$, we solve a maximum likelihood estimation problem with uncertain data. Let the known ages be denoted by the 1227-dimensional vector \mathbf{X} , and the age of records with uncertain age be given by the 137-dimensional vector \mathbf{Z} . Given lognormal parameters (μ_a, σ_a) , the likelihood function is $\prod_{i=1}^{1227} f(X_i) \prod_{j=1}^{137} \int_{t_j}^{t_j+365} f(Z_j) dZ_j$. Choosing (μ_a, σ_a) to maximize the log-likelihood function, which is $\sum_{i=1}^{1227} \log[f(X_i)] + \sum_{j=1}^{137} \log[F(t_j+365) - F(t_j)]$, yields $\mu_a = 4.904$ and $\sigma_a = 1.687$. The frequency distribution of these 1364 ages and the resulting lognormal are plotted in Fig. 7.

To estimate the age PDF $h(a)$ for all database entries, we assume that the age of an entry is December 31, 2010 minus the acquisition date of the entry. However, we only know the year in which each entry in the Israeli database was acquired. We estimate $h(a)$ by fitting a piecewise cubic hermite interpolating polynomial [6] on the yearly aggregates to estimate an increasing, smooth CDF, and then numerically differentiating it to yield a PDF (Fig. 8).

3 Estimating U.S. Parameter Values

In this section, we review the parameter estimation procedure for the three U.S. cartridge casing scenarios. The CDF's $F(x)$ and $G(y)$, which characterize the accuracy of ballistic

imaging, are assumed to be the same as in the Israeli scenario. Of U.S. casings arrivals, 28% are evidence [7], and hence $q_0 = 0.72$, $q_1 = 0.28$. The r_j probabilities depend on the geographical approach. For the partition approach, $J = 0$ and $r_0 = 1$. For lack of data, we assume that each of the 47 partitions has the same arrival rate and the same number of database entries, and we make the same assumption for each of the 12 regions. Consequently, for the regional approach, $J = 1$, $r_0 = 1/12$ and $r_1 = 11/12$. For the national approach, $J = 2$ (where $j = 0, 1, 2$ corresponds to intra-partition, inter-partition and intra-region, and inter-region, respectively), and $r_0 = 1/47$, $r_1 = (1/12) - (1/47)$ and $r_2 = 11/12$. Turning to p_{ij} , $p_{00} = p_{10} = 1$ for the partition approach. For lack of data, because the number of partitions in the U.S. is similar to the number of stations in Israel (47 vs. 52), we simply use the Israeli values in the U.S. regional approach, so that $p_{00} = 0.832$, $p_{01} = 0.168$, $p_{10} = 0.875$, $p_{11} = 0.125$. In the U.S. national approach, because we have no data, we arbitrarily assume that half of inter-partition matches are intra-region and half are inter-region, which implies that $p_{00} = 0.832$, $p_{01} = p_{02} = 0.084$, $p_{10} = 0.815$, $p_{11} = p_{12} = 0.0625$.

The next step is to derive the PMF $g_i(n)$ and t_h for the three U.S. geographical approaches. Equation (2) solves for t_h in terms of the evidence database size N . The NIBIN database for cartridge casings has 554,578 records, which includes 28% evidence and 72% nonevidence [7]. Hence, we set $N = 0.28(554,578)$ in the national approach, $N = 0.28(554,578)/12$ in the regional approach, and $N = 0.28(554,578)/47$ in the partition approach. Note that in the Israeli scenario, we knew $g_i^*(m)$, which allowed us to solve for the historical threshold t_h via equation (2), and then solve for $g_i(n)$ in equation (3) because P_{mn} is a function of $F(t_h)$. We do not know $g_i^*(m)$ in the U.S. scenarios, and approach the problem in reverse by first estimating $g_i(n)$. The easiest case is the U.S. national approach, where we assume that the U.S. $g_i(n)$ under the national approach is the same as the Israeli $g_i(n)$. Then we solve equations (2)-(3) simultaneously for t_h and $g_i^*(m)$. For the U.S. partition and

regional approaches, when estimating $g_i(n)$ we need to first account for the fact that some of these true matches will occur outside of the area being searched. More specifically, let \hat{p}_i be the probability that a true match involving an arrival of evidence status i occurs in the area being searched (i.e., in the partition under the partition approach, and in the region under the regional approach). Then $\hat{p}_0 = 0.832$ $\hat{p}_1 = 0.875$ for the partition approach, and $\hat{p}_0 = 0.832 + 0.084 = 0.916$, $\hat{p}_1 = 0.875 + 0.0625 = 0.9375$ for the regional approach. If we let $\hat{g}_i(n)$ denote the probability of n true matches in the area being searched (i.e., partition or region) for an arrival of evidence status i , then it follows that

$$\sum_{n=m}^{\infty} \binom{n}{m} \hat{p}_i^m (1 - \hat{p}_i)^{n-m} g_i(n) = \hat{g}_i(m) \quad \text{for } i = 0, 1, \quad (4)$$

where $g_i(n)$ is the U.S. national $g_i(n)$. For the partition approach and the regional approach, we solve (4) for $\hat{g}_i(n)$, plug this solution in for $g_i(n)$ in equation (3), and jointly solve equations (2)-(3) to get $g_i^*(m)$ and t_h . Fig 9 shows $g_i(n)$ and $g_i^*(m)$ for the partition, regional and national approaches.

For the U.S. bullets scenarios, the NIBIN database for bullets has 226,830 records, which includes 19% evidence and 81% nonevidence [7]. Hence, we set $N = 0.19(226, 830)$ in the national approach, $N = 0.19(226, 830)/12$ in the regional approach, and $N = 0.19(226, 830)/47$ in the partition approach. Substituting these values into equation (2) gives $t_h = 466, 513$ and 594 for the partition, regional and national approaches, respectively.

4 Comparison of a Constant Threshold System and a Rank-based System

In this section, we argue that for a simplified setting in which there is never more than one match in the database for an arrival, the rank-based system and the constant threshold system perform very similarly. Equation (1) gives the detection probability for a rank-based

system that generates a candidate list of size 10 when the database has N entries. In approximating the detection probability for the constant threshold system that generates an average candidate list size of 10, we make use of the fact that the average number of true positives in the Israeli data is 0.214: 90.24% of the candidate lists have no true matches, and the average number of true matches in the other 9.76% of lists is 2.19. Substituting 0.214 for the first term in equation (2) and solving for the threshold gives $t_h \approx G^{-1}\left(1 - \frac{9.786}{N}\right)$. The detection probability for the constant threshold system in this simplified setting is $1 - F(t_h)$, which can be approximated by

$$1 - F\left(G^{-1}\left(1 - \frac{9.786}{N}\right)\right). \quad (5)$$

Using the cartridge casings lognormal distributions derived in §1 and plotting the detection probabilities in equations (1) and (5) as functions of the database size N reveal that the performance curves are almost identical (Fig. 10).

5 The False Negative Constraint vs. the List Size Constraint

Here we explore the difference between constraining the expected number of false negatives, as in equation (3) in the main text, and constraining the expected list size, which is much more difficult analytically. Our concern is that our choice of the false negative constraint may favor the optimal policy because the solution to (1)-(3) in the main text may raise the average number of true positives in the candidate list. In this case, restricting only the average number of false positives would increase the average list size beyond 10. Calculations in this section suggest that the effect of our choice of constraint on the results is extremely small.

We focus on Israeli cartridge casings first. We know that the probability that there

is at least one true positive in the candidate list (taking into account arrivals that have no true matches in the database) is 0.0976 for the constant threshold policy and 0.1035 for the optimal policy. Although we know that the mean number of true positives in the candidate lists with at least one true positive is 2.19 for the constant threshold policy, we do not know this quantity for the optimal policy. By assuming that this value is 5.0, which should be a significant overestimate, the mean number of true positives in the candidate list increases to $5 \times 0.135 = 0.5175$. To approximate the new detection probability under a list size constraint, we should resolve (1)-(3) in the main text with the right side of the constraint reduced by 0.5175-0.214. To be conservative, we reduce the right side of the constraint by 0.5175. The new detection probability is 0.9870, compared to 0.9872 under the original constraint. Similarly, when the same procedure is carried out for U.S. casings under the national approach, the detection probability is reduced from 0.9699 to 0.9694.

References

- [1] Beauchamp A, Roberge D (2005) Model of the behavior of the IBIS correlation scores in a large database of cartridge scores. Unpublished manuscript. Accessed at www.forensictechnology.com/Default.aspx?app=LeadgenDownload&shortpath=docs/LargeDatabaseFinal.pdf on September 2, 2011.
- [2] Roberge D, Beauchamp A (2006) The use of BulletTRAX-3D in a study of consecutively manufactured barrels. *AFTE Journal* 38:166-172.
- [3] Binck TB (2008) Comparing the performance of IBIS and BulletTRAX-3D technology using bullets fired through 10 consecutively rifled barrels. *J. Forensic Science* 53:677-682.
- [4] Hespanha J (2004) grPartition - a MATLAB function for graph partitioning. Accessed at www.ece.ecsb.edu/hespanha/software/grPartition.html on September 8, 2011.

- [5] Durbin R, Eddy S, Krogh A, Mitchison G (1998) Biological sequence analysis: probabilistic models of proteins and nucleic acids. Cambridge U. Press, Cambridge, UK.
- [6] Fritsch FN, Carlson RE (1980) Monotone piecewise cubic interpolation. *SIAM J. Numerical Analysis* 17:238-246.
- [7] Office of the Inspector General, U.S. Department of Justice (2005) *The Bureau of Alcohol, Tobacco, Firearms and Explosives' National Integrated Ballistic Information Network program*. Audit Report 05-30, Washington, D.C.

Optimized Information	Thresholds	Detection Probability
None	$t_h = 443$	0.931
Evidence	$t_0 = 486, t_1 = 416$	0.936
Spatial	$t_0 = 321, t_1 = 528$	0.965
Age	$t_a = -319.1e^{-0.00152a} + 537.4$	0.972
Evidence, Spatial	$t_{00} = 354, t_{01} = 562, t_{10} = 301, t_{11} = 502$	0.967
Evidence, Age	$t_{0a} = -302.7e^{-0.00149a} + 507.9, t_{1a} = -345.9e^{-0.00153a} + 587.3$	0.974
Spatial, Age	$t_{0a} = -283.3e^{-0.00100a} + 435.7, t_{1a} = -420.4e^{-0.00128a} + 674.0$	0.986
Evidence, Spatial, Age	$t_{00a} = -292.6e^{-0.00121a} + 459.0, t_{01a} = -408.5e^{-0.00136a} + 691.6$ $t_{10a} = -223.1e^{-0.00142a} + 364.9, t_{11a} = -339.2e^{-0.00158a} + 587.6$	0.987

Table 3: Detection probability for all combinations of optimizing evidence status, spatial category and age information. The subscripts of t_{ija} are suppressed if they are not being optimized.

Policy	Parameter Values			U.S. Casings Scenarios		
	p_{00}	p_{10}	$\frac{p_{i2}}{1-p_{i0}}$	Partition	Regional	National
Constant Threshold	0.832	0.875	0.5	0.873	0.881	0.817
Optimal Thresholds	0.832	0.875	0.5	0.899	0.946	0.970
Constant Threshold	0.832	0.875	0.2	0.873	0.906	0.817
Optimal Thresholds	0.832	0.875	0.2	0.899	0.962	0.973
Constant Threshold	0.888	0.917	0.5	0.903	0.895	0.817
Optimal Thresholds	0.888	0.917	0.5	0.929	0.953	0.973
Constant Threshold	0.888	0.917	0.2	0.903	0.911	0.817
Optimal Thresholds	0.888	0.917	0.2	0.929	0.969	0.975
Constant Threshold	0.944	0.958	0.5	0.932	0.909	0.817
Optimal Thresholds	0.944	0.958	0.5	0.958	0.969	0.977
Constant Threshold	0.944	0.958	0.2	0.932	0.917	0.817
Optimal Thresholds	0.944	0.958	0.2	0.958	0.969	0.978

Table 4: Sensitivity analysis results for the three U.S. casings scenarios under both policies, where we vary the probabilities (p_{00} and p_{10}) of intra-partition matches and the probability ($\frac{p_{i2}}{1-p_{i0}}$) that an inter-partition match is inter-region. The fraction of true matches that are intra-partition is $0.72p_{00} + 0.28p_{10}$, which is 0.844 in the top four rows, 0.896 in the middle four rows, and 0.948 in the bottom four rows. The performance of the partition approach is unaffected by $\frac{p_{i2}}{1-p_{i0}}$. Similarly, the constant threshold policy under the national approach is unaffected by any of these parameter values.

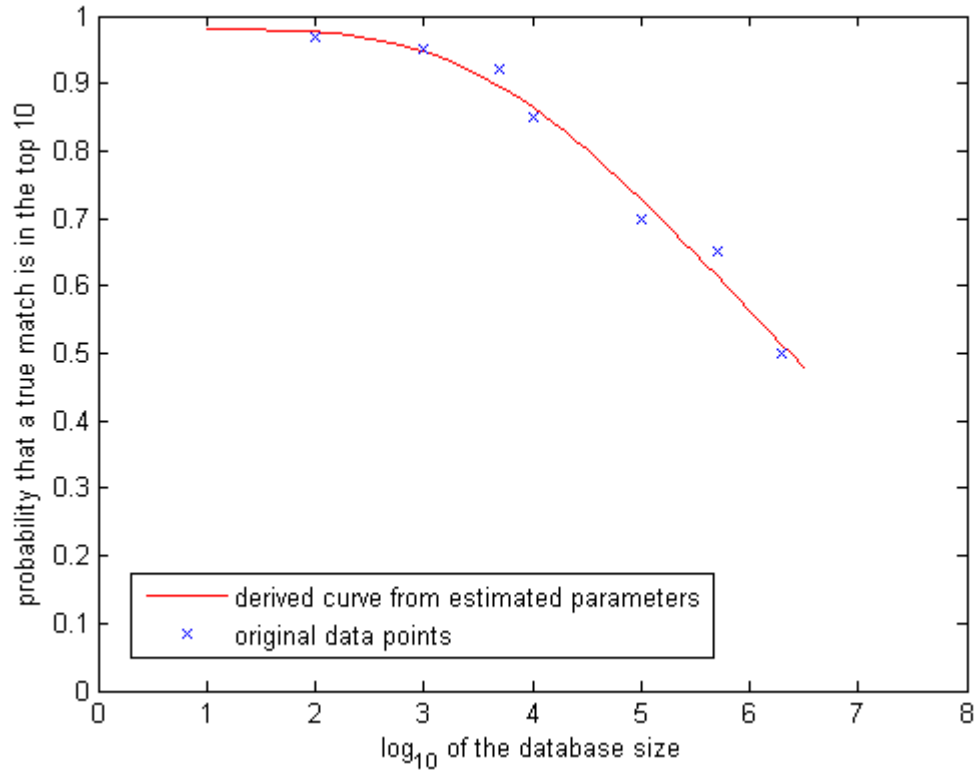


Figure 2: Actual points (x) on the performance curve for casings from data in [1] vs. the performance curve generated by the best-fit lognormal distribution.

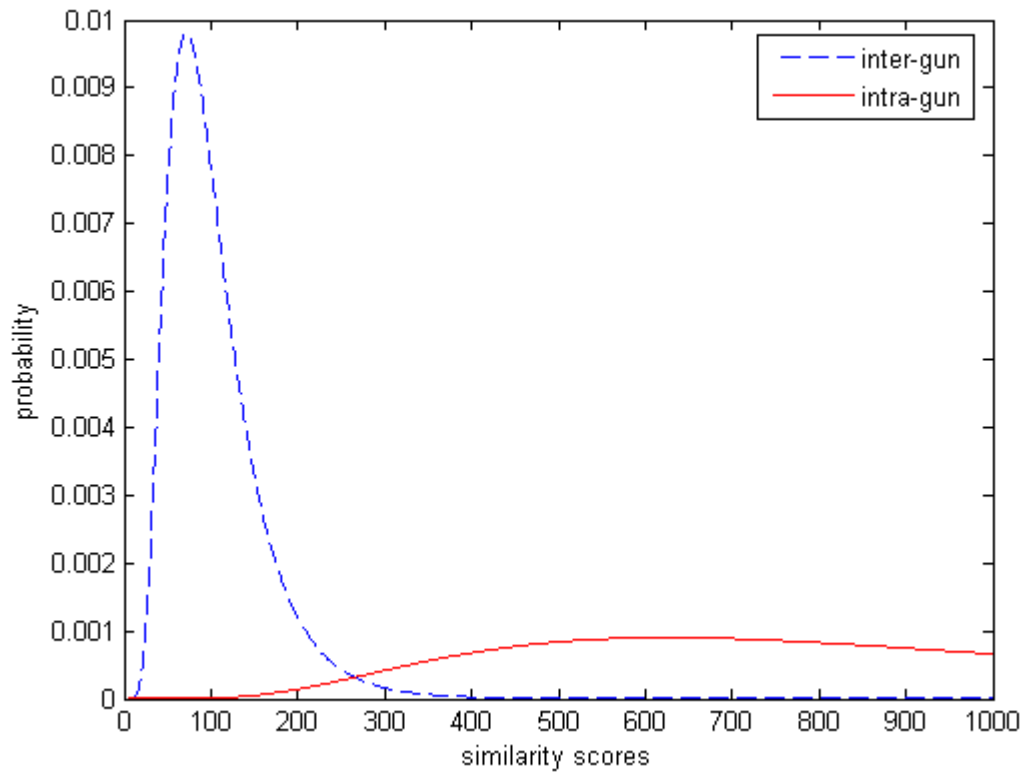


Figure 3: The lognormal similarity score PDFs, intra-gun $f(x)$ and inter-gun $g(y)$, for casings.

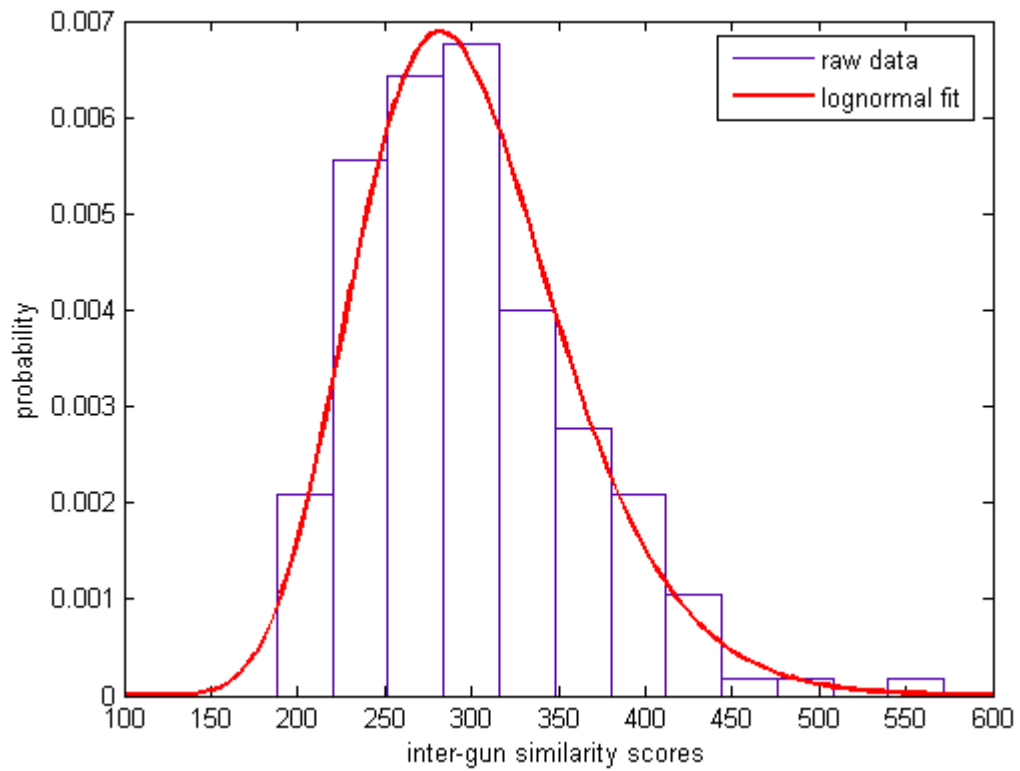


Figure 4: Actual (based on BulletTrax-3D data in [2]) vs. lognormal inter-gun similarity score PDF for bullets.

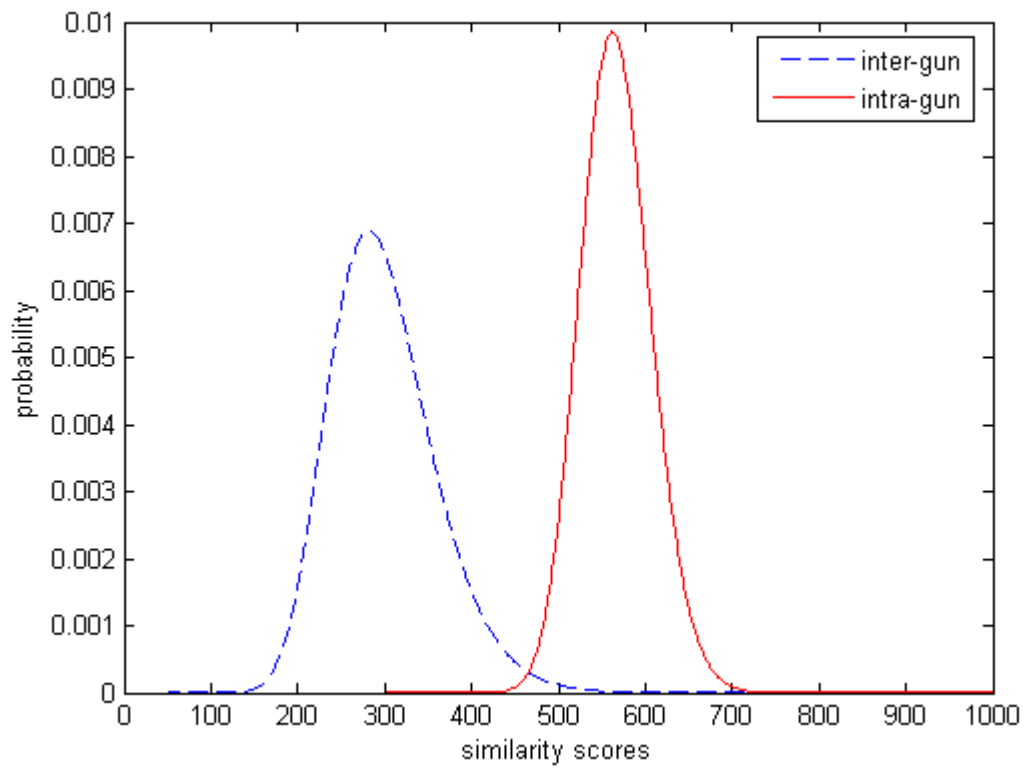


Figure 5: The lognormal similarity score PDFs, intra-gun $f(x)$ and inter-gun $g(y)$, for bullets.

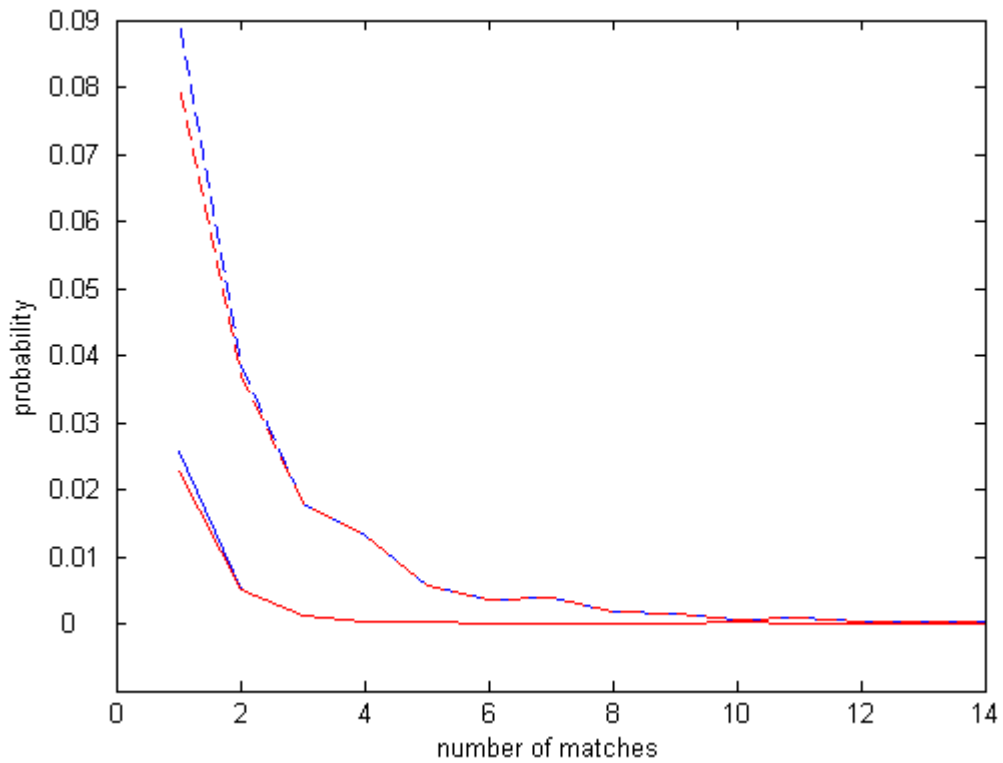


Figure 6: The PMFs for the true ($g_i(n)$ in blue) and detected ($g_i^*(m)$ in red) matches for arrivals of evidence status i ($i = 0$ is nonevidence (—) and $i = 1$ is evidence (- - -)).

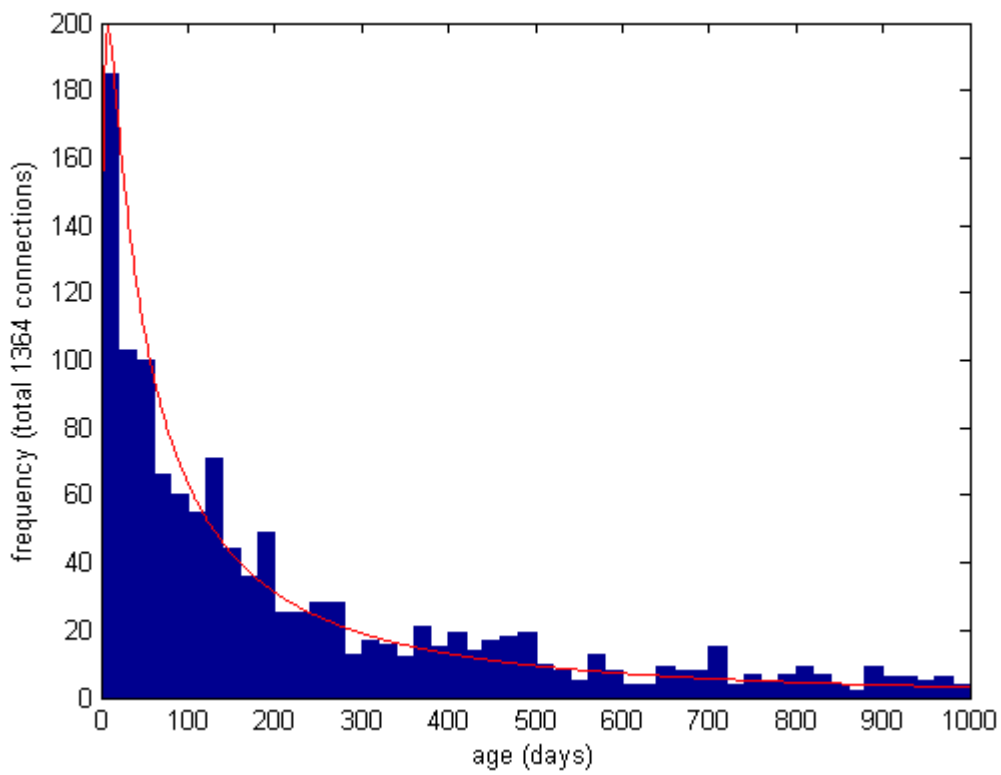


Figure 7: Frequency distribution of ages of the 1364 matching pairs in the Israeli database, containing 1227 exact ages and 137 ages that are randomly sampled from the correct year, along with the best-fit lognormal PDF, $h_m(a)$.

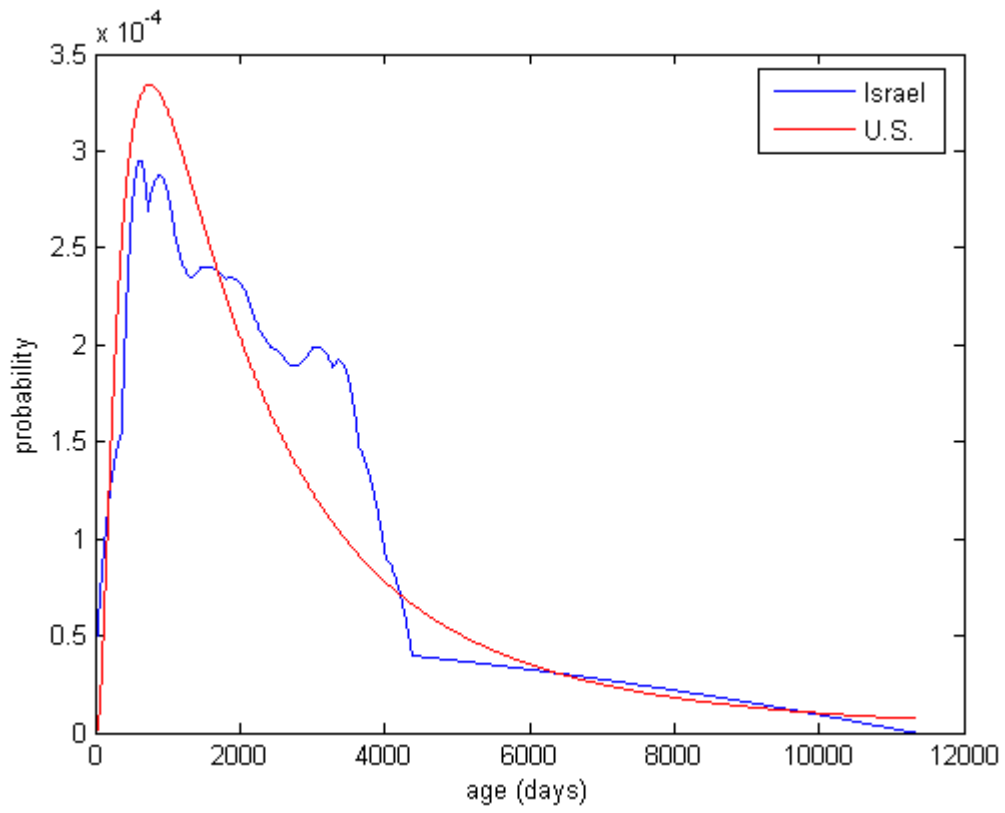


Figure 8: Age PDF $h(a)$ for all database records.

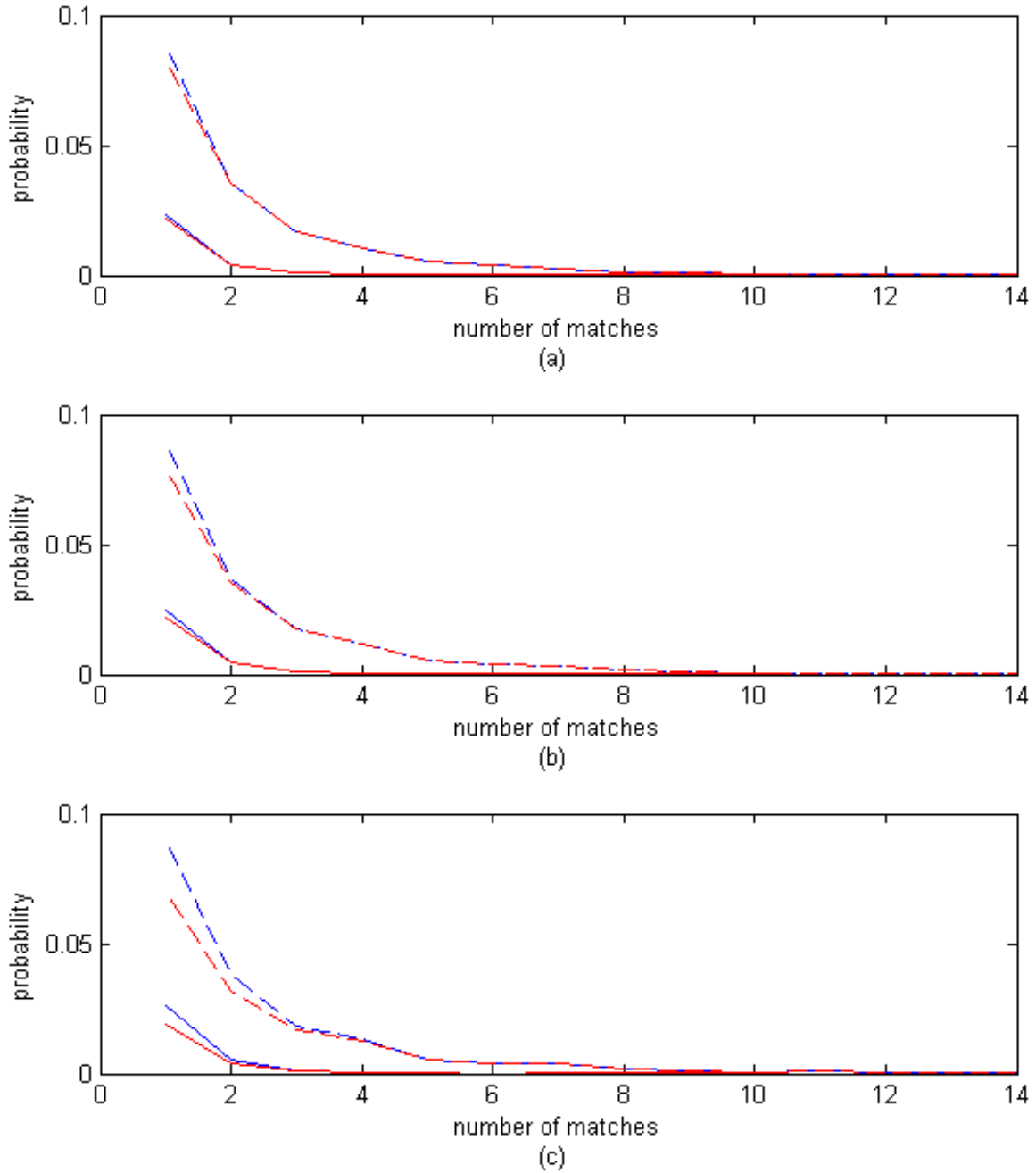


Figure 9: The PMFs for the true ($g_i(n)$ in blue) and detected ($g_i^*(m)$ in red) matches for arrivals of evidence status i ($i = 0$ is nonevidence (—) and $i = 1$ is evidence (- - -)) for **(a)** U.S. national approach, **(b)** U.S. regional approach, **(c)** U.S. partition approach.

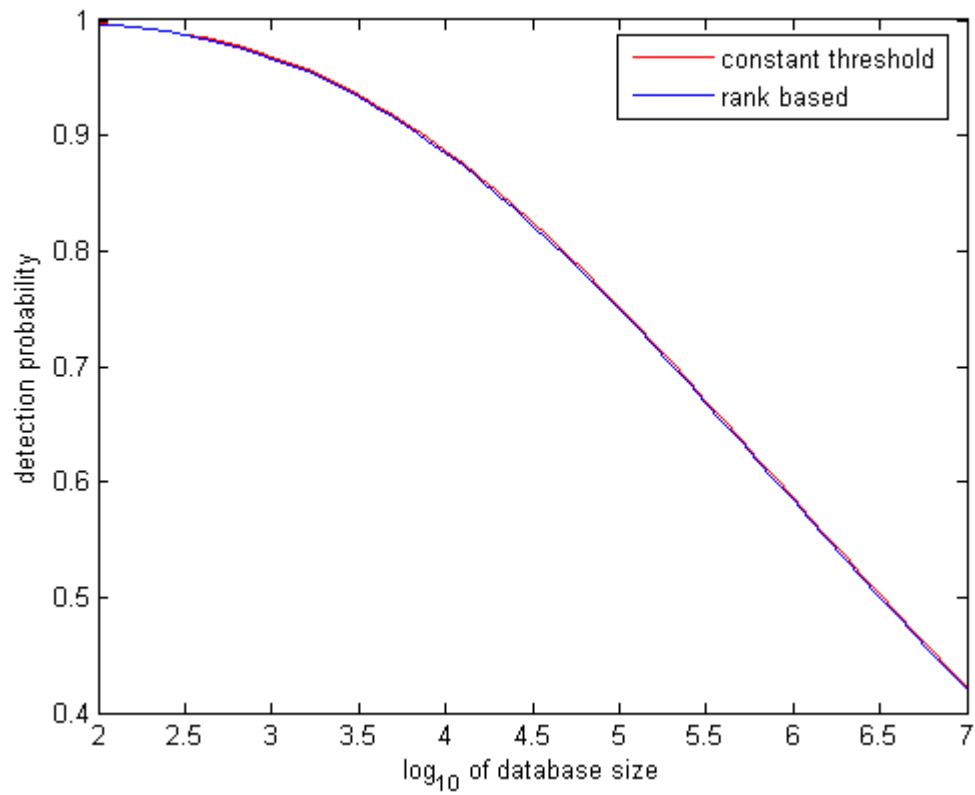


Figure 10: Comparison of the performance of the rank-based system and the constant threshold system.

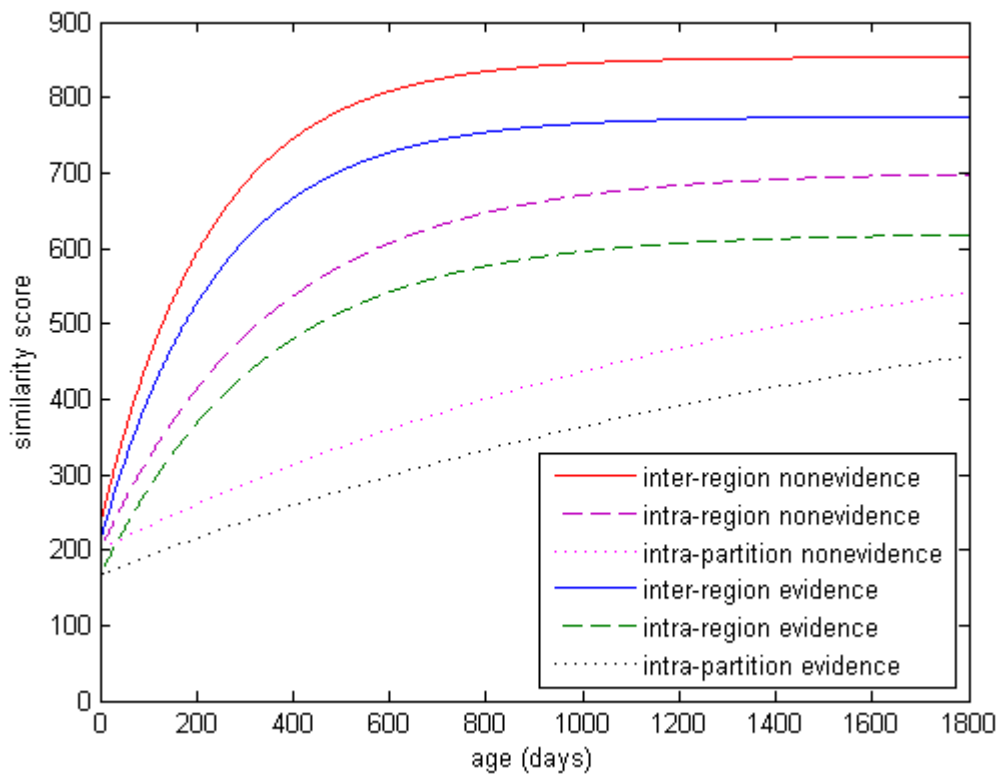


Figure 11: Optimal thresholds for U.S. cartridge casings under the national approach.