Routledge
Taylor & Francis Group

Check for updates

Estimating the Amount of Nuclear Weapon-usable Material Outside Government Control Using Data on Reported Seizures

Valentin Stanev[a] and Steve Fetter[b]

[a]Department of Materials Science and Engineering, University of Maryland, College Park, USA; [b]School of Public Policy, University of Maryland, College Park, MD, USA

ABSTRACT

Terrorists could acquire nuclear weapons by using weapon-usable nuclear material that was stolen or otherwise diverted from legitimate authorities. Multiple well-documented seizures suggest the existence of a black market that draws on an unknown stock of weapon-usable nuclear material that is not under the control of authorities. We estimate the total amount of uncontrolled material based on publicly reported seizures and several different statistical methods and models. We estimate that 90 to 250 kilograms—sufficient for up to ten nuclear weapons—remain outside the control of legitimate authorities. While this estimate is subject to large uncertainties and potential bias, governments may have additional information about nuclear material seizures that could be used to improve estimates.

Introduction

It is widely accepted that the largest barrier to the acquisition of nuclear weapons is access to weapon-usable nuclear material (WUNM)—highly enriched uranium (HEU) and plutonium. Production of these materials is difficult for nation states and beyond the capacity of even the best-equipped terrorist group. A central element in initiatives to reduce the risk of nuclear terrorism has been an effort to reduce the production and use of WUMN, to remove WUNM where possible, to increase the security of remaining stockpiles, and to detect and intercept WUNM in transit or in black markets.[1]

Despite these efforts, it is likely that some unknown but substantial quantity of WUNM remains outside the control of any legitimate authority. As shown in Table 1, 21 seizures involving a total of 19.72 kg of WUNM (19.35 kg of HEU and 0.37 kg of plutonium) were publicly reported between 1992 and 2015. These seizures resulted from inspections at international borders, sting operations, and other police actions. Additional seizures may have occurred that have not been publicly reported. It is

CONTACT Valentin Stanev ✉ vstanev@umd.edu 📧 Department of Materials Science and Engineering, University of Maryland, College Park, MD 20742-4111, USA.

Table 1. Documented cases of seizures of weapon-usable nuclear material. The amount is given in grams.

Date	Location	Material	Amount, g
6 Nov 1992	Podolsk, Russia	HEU (90%)	1,500
29 Jul 1993	Andreeva Guba, Russia	HEU (36%)	1,800
28 Nov 1993	Murmansk, Russia	HEU (20%)	4,500
24 May 1993	Vilnius, Lithuania	HEU (50%)	150
March 1994	St. Petersburg, Russia	HEU (90%)	2,972
10 May 1994	Tengen-Wiechs, Germany	Pu	6.2
13 Jun 1994	Landshut, Germany	HEU (87.7%)	0.795
25 Jul 1994	Munich, Germany	Pu	0.24
8 Aug 1994	Munich Airport, Germany	Pu	363.4
14 Dec 1994	Prague, Czech Republic	HEU (87.7%)	2,730
Jun 1995	Moscow, Russia	HEU (21%)	1,700
6 Jun 1995	Prague, Czech Republic	HEU (87.7%)	0.415
8 Jun 1995	Ceske Budejovice, Czech Rep.	HEU (87.7%)	16.9
29 May 1999	Rousse, Bulgaria	HEU (72.6%)	10
May 2000	Electrostal Russia	HEU (20%)	3,700
Dec 2000	Karlsruhe, Germany	Pu	0.001
16 Jul 2001	Paris, France	HEU (72.6%)	0.5
26 Jun 2003	Sadahlo, Georgia	HEU (89%)	170
1 Feb 2006	Tbilisi, Georgia	HEU (89%)	79.5
5 Oct 2009	Chisinau, Moldova	HEU (72.6%)	4
3 Nov 2011	Tbilisi, Georgia	HEU (89%)	18
Total			19,722

reasonable to assume that seized materials are part of a larger stock of WUMN that exists outside any government and regulatory control, and that there remains some quantity of WUNM that is potentially available to buyers on black markets. Reliable estimates of this quantity of material would be useful for evaluating the risks of nuclear terrorism and the level of effort that should be devoted to reducing these risks.

Although individual states maintain nuclear material accounting systems, uncertainties and errors in measurement and accounting limit their use to determine how much—if any—material may have been lost or stolen. For example, the United States estimates that it produced or otherwise acquired a total of 111,700 kg of plutonium, but current inventories and known removals account for only 109,300 kg, leaving about 2,400 kg unaccounted for.[2] There is no evidence that any of this material was stolen; most or all the 2,400 kg shortfall is probably the result of overestimates of plutonium production and/or underestimates of plutonium disposed with wastes or remaining in processing facilities. The Russian government does not provide a public accounting of its plutonium balance, but it is generally believed that accounting uncertainties are even larger than those in the United States. Thus, we cannot rely on official material accounting systems to provide a plausible upper bound on the amount of WUNM that may be outside of state control.[3]

In this article, we use information about known seizures to estimate the total amount of WUNM that remains outside government control. Our approach is statistical.[4] We begin with the simple assumption that seizures of WUMN are random events: in any black market transaction or smuggling operation there is some probability that the material will be seized by authorities, with the outcome determined by random factors. With additional—and more questionable—assumptions, we can

build simple models to describe the seizure or recovery process. We consider two such models: a binomial model and a capture-recapture model. These models represent complementary approaches: the former assumes a constant flow of material from which some unknown fraction is recovered; the second assumes a fixed stock of material from which there is an occasional recovery. We also consider the potential bias of the mass distribution of intercepted items and introduce a corrected distribution that includes the possibility of unobserved items of larger mass.

We acknowledge that our estimates have very large uncertainties and that they may systematically under-estimate or over-estimate the quantity of uncontrolled WUNM. The estimates obtained here should be treated as a preliminary attempt to address this important question.

Publicly known seizures of WUNM

There are three well-known databases on incidents involving WUNM. The International Atomic Energy Agency (IAEA) supports the Incident and Trafficking Database, which is protected by confidentiality agreements between IAEA and the member states and is not publicly accessible (although the agency publishes annual summary reports).[5] A second database, maintained by the Center for Nonproliferation Studies, is open-source but includes data only for recent years and the quality and accuracy of the information varies considerably.[6] Here we use the Database on Nuclear Smuggling, Theft and Orphan Radiation Sources (DSTO), maintained by University of Salzburg. Although not available to the public, several detailed reports on the information contained in DSTO have been published.[7] Unless stated otherwise, the information used here comes from these two reports. Table 1 summarizes the well-documented seizures of WUNM.

Estimates based on the recovery probability distribution

We begin with a simple method for estimating the quantity of uncontrolled WUNM. We aggregate all seized WUNM, neglecting differences in chemical and isotopic composition. The total quantity of WUMN seized between 1992 and 2015 is about 19.72 kilograms, which for simplicity we round to 20 kg. The fraction of WUMN that has been recovered is given by

$$t = \frac{r}{r + x} \tag{1}$$

where $r = 20$ kg and x is the amount of WUNM that remains outside control. We assume that the seizure of WUNM is an intrinsically random process with a fixed probability. The probability that x is less than a given amount X is given by

$$P(x < X|r) = \int_{\frac{r}{X+r}}^{1} f(t)dt = 1 - CDF_f\left(\frac{r}{X + r}\right) \tag{2}$$

where $f(t)$ is the probability distribution and CDF_f is the cumulative probability distribution of the fraction of material that has been recovered. We represent $f(t)$

with the Beta distribution, which is a family of continuous distributions used extensively for Bayesian estimations:

$$f(t) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \, t^{\alpha-1}(1-t)^{\beta-1}, \tag{3}$$

where α and β are shape parameters. Using the values $\alpha = \beta = 1$ produces uniform distribution $f(t) = 1$, which is equivalent to complete ignorance about the fraction of material that has been seized. In that case, Equation (2) reduces to

$$P(x < X|r) = 1 - \frac{r}{X + r} = \frac{X}{X + r}. \tag{4}$$

The uniform distribution gives a 50% probability that less than 20 kg remains unrecovered, 75% probability of less than 60 kg; 90% probability of less than 120 kg; and 96% probability of less than 200 kg. This is an unbiased estimate because it does not require any assumptions apart from the randomness of the recovery process. This provides an important baseline for comparison with other estimation methods that rely on additional assumptions.[8]

But we need not assume complete ignorance about the recovery probability. Black markets exist in a wide range of contraband materials, including illegal drugs, stolen and counterfeit goods, currency, weapons, humans, and wildlife. In some cases, it is possible to estimate the total size of the market and therefore the recovery probability (the percent of contraband that is seized). Studies of these black markets generally indicate that seizures are a small fraction of the total amount of contraband in circulation. In the case of illegal drugs, consumption is estimated with surveys, drug treatment, and arrest data. If we divide reported seizures by the sum of seizures and estimated consumption in the United States, we find that the average recovery fraction during the period 2000–2010 was 6–16% for heroin, 8–21% for methamphetamine, 21–35% for marijuana, and 24–43% for cocaine.[9] In the case of elephant ivory, the size of the market generated by illegal poaching can be estimated by modeling and tracking elephant populations. Dividing reported seizures of African elephant ivory by estimates of the total amount of ivory derived from illegal poaching during the period 2010–2012 yields a recovery probability of 4–19%.[10] Based on reported seizures and estimates of the number of small arms purchased with the intention of trafficking them, one study estimates that U.S. and Mexico authorities seized 9–35% of total arms bought in 2010–2012 with the intention of trafficking them.[11] Estimated recovery rates for stolen fine art range from 6% to 20%.[12]

There are important differences between the black market in WUNM and other black markets. The largest black market, illegal drugs, has well-developed and well-financed networks of producers and distributors with decades of experience in smuggling, but there also exist dedicated intelligence and police efforts to counter the drug trade. The black market in ivory and other wildlife is much smaller and smugglers are less organized and professional, but police and customs officials in the supplier countries are also less effective. Gun smuggling is strongly linked to other criminal activities, and the same groups smuggling guns also smuggle illegal drugs
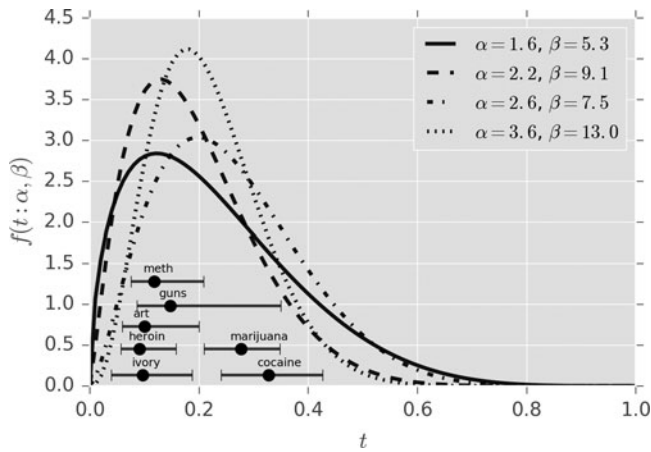
Figure 1. Notional probability distributions for the WUMN recovery probability, based on recovery probabilities estimated for other black markets.

and other contraband. Despite these differences, Foss concludes that "characteristics associated with smuggling and trafficking of nuclear material are no different than the characteristics associated with smuggling and trafficking other illicit commodities."[13] Foss notes in particular similarities between the black markets for stolen art and nuclear material: both involve small, portable objects that are typically subject to strong protections from theft, and both have niche markets in which legal purchases occur only in limited and well-defined circumstances. In both markets, sellers have a difficult time finding buyers, and hoaxes and scams are common. Thus, we judge that it is not unreasonable to assume that the recovery probability for WUMN is likely to be in the same range as that for other black markets, for which recovery probability estimates range from a low of 6–10% to a high of 35–45%.

We use Equation (3) to create a set of distributions for the WUMN recovery probability with a lower 10% confidence level in the range of 6–10% and an upper 90% confidence level in the range of 35–45%. Figure 1 shows four representative distributions that satisfy these conditions with shape parameters (α, β) of (1.6, 5.3), (2.2, 9.2), (2.6, 7.5), and (3.6, 13). These are reasonable prior probability distributions based on estimated recovery probabilities in other black markets, but they allow for the possibility that the recovery probability for WUMN could be significantly higher or lower.

Integrating these functions, we obtain estimates for the probability $P(x\langle X|r)$. As shown in Figure 2, there is a 50% probability that the amount of uncontrolled WUMN is less than 65–95 kg; a 75% probability of less than 105–165 kg; and a 90% probability that less than 180–320 kg remains outside control. These estimates are significantly higher than those obtained using the uniform probability distribution $f(t) = 1$ because the prior probability distributions are skewed to left.

Binomial approximation

In this section we use the variation in the number of seizures with time to estimate the recovery probability. Although this method relies on the questionable
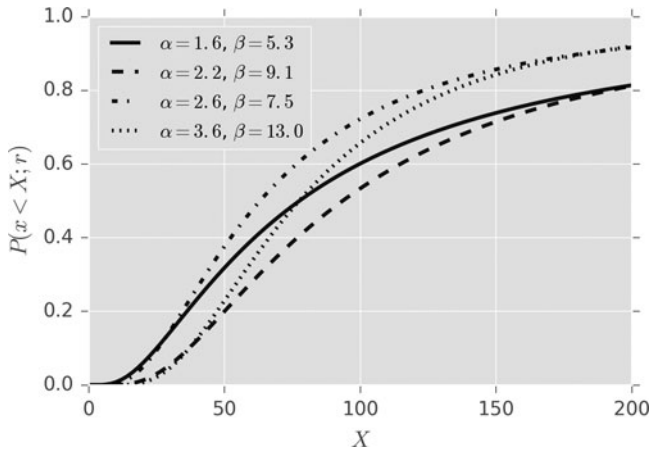
Figure 2. Plot of the probability $P(x < X \vee r)$ that the amount of WUNM that has not been recovered is below X.

assumption that the recovery probability is constant with time, it provides a natural extension of the estimates obtained in the previous section.

A severe problem in developing a statistical model for the recovery process is the very small number of seizures of WUNM. To compensate, we expand the dataset by considering seizures of all radioactive materials and assume that the recovery probability for WUNM is roughly equal to that for all radioactive materials. We consider incidents involving radioactive materials where there was clear criminal intent, excluding incidents unlikely to be connected to potential black market transactions. We restrict our analysis to transactions in the Black Sea region.[14] Although this choice is driven by the availability of data, the countries of this region account for more than half of the WUNM seizures and more than 85% of the total amount of WUMN seized.

Figure 3 shows the number of seizures of radioactive materials per year in the Black Sea region from 1991 to 2012.[15] There is a significant variation in the number
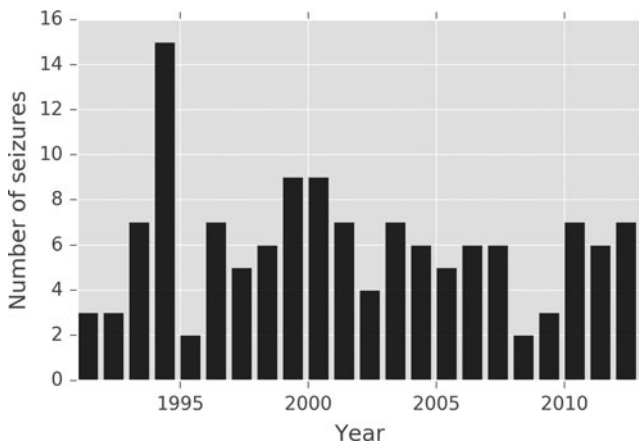

Figure 3. Annual number of seizures of radioactive materials in the Black Sea region, 1991–2012.

of seizures per year. Some of this variability may be due to changes in political, economic, and security conditions. For example, the collapse of the Soviet Union and the accompanying widespread political and economic crisis might explain the sharp rise in the number of seizures from 1991 to 1994, and the subsequent improvements in the economic situation in the countries of the former Soviet Union and Eastern Europe might help explain the gradual decrease in the number of seizures in from 2000 to 2010. Although political and economic events are likely to be important determinants of nuclear smuggling attempts and the recovery probability, we ignore these factors and assume a constant rate of smuggling and a constant recovery probability, to create a simple model for the relationship between the observed variability of seizures and the recovery probability. In other words, we assume that every year there is some fixed number of attempts to smuggle or sell radioactive material and that there is a constant probability q that each such attempt will be detected and the material will be seized.

There is, of course, no reason to expect the same number of smuggling attempts each year. On the contrary, we expect this number to vary in response to changes in economic and security conditions, as well as other factors. We can, however, mitigate the effects of this extrinsic variability by combining several non-consecutive years together, to average away the deterministic time dependence of N_{tot}.

Before we do this, however, let us first attempt to directly fit the data shown in Figure 3. We use a standard binomial model, with each year providing an independent sample of a process with N_{tot} trials, each of which may result in a recovery with probability q, and leads to average number of seizures per year $n = qN_{tot}$.[16] From the data we can estimate n, but not directly q or N_{tot}. However, we can use the observed mean *and* variance of the annual number of seizures to estimate simultaneously both parameters of the model. Because the variance of the binomial distribution is given by $\sigma^2 = N_{tot}q(1 - q) = n(1 - q)$,

$$q = 1 - \frac{\sigma^2}{n}$$

or, equivalently,

$$\sigma(N_{tot}) = \sqrt{n\left(1 - \frac{n}{N_{tot}}\right)}.$$

For the data shown in Figure 3, the mean and variance in the sample are 6.0 and 8.1, respectively, which is not compatible with the binomial model (the variance is too large, resulting in a negative value for q). Note, however, that the number of seizures in 1994 (15) is more than four standard deviations above the mean for all years except 1994, and thus is clearly an outlier. If we omit this year, then $n = 5.6$ and $\sigma^2 = 4.3$, which leads to $q = 0.24$, suggesting that about one-quarter of all smuggled material was seized.

Although removing one year from the data may be justified by the need to remove an outlier, it also has the unfortunate consequence of ignoring the year with highest
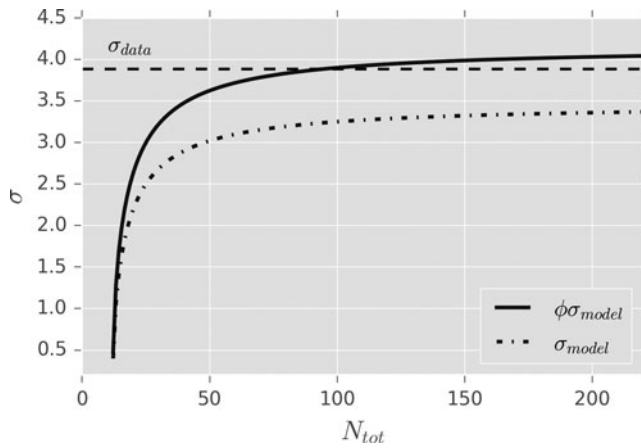
Figure 4. The dependence of the variance on N_{tot}, directly as given by the binomial model (dashed-dotted curve), and corrected for overdispersion of the data with ϕ (solid curve). The intersection of the σ_{data} (dashed line) with corrected σ_{model} gives an estimate of N_{tot} of around 100 (the recovery probability of about 15%). Note the change of scales caused by combining years in groups of two: n = 12.

number of incidents. This artificially reduces the variance, which might result in a significant underestimation of N_{tot} and, accordingly, an overestimation of q. Because overdispersion (i.e., variance significantly larger than what a simple binomial model would predict) is a common occurrence in real data, various approaches for tackling it have been suggested.[17] Here we will follow a simple procedure. First, we combine all years in random groups of two to reduce the effects of the historical trends in the data.[18] We then estimate the standard deviation of the binned data and assume it has the form $\sigma_{data} = \phi\sigma_{model}$, where ϕ is the variance inflation parameter. This is one of the simplest ways to introduce overdispersion in the model, and has been extensively used in the literature.[19] It has the advantage that ϕ can be estimated directly, without assuming prior knowledge of q or N_{tot}. Using this parameter, we can correct the theoretical curve $\sigma(N_{tot})$, and find its intersection with the standard deviation of the data. We estimate ϕ using the Pearson's chi-square statistic:

$$\phi = \left(\sum \frac{(n_i - n_{mean})^2}{n_{mean} d_f} \right)^{\frac{1}{2}}$$

where n_i is the data for ith two-year bin, and $d_f = N_{bins} - 1$ is the number of degrees of freedom (N_{bins} is the total number of two-year intervals).

In Figure 4 we show the corrected curve, which now intersects with the data and provides an estimate of N_{tot} slightly below 100. Because the average for the binned data is $n = 12$, this suggests recovery probability of about 15%. As expected, this is significantly less than the previous estimate of 24%, which ignored data from 1994.

To get a rough idea of the accuracy of this estimate (at least within the confines of this highly simplified model), we run simulations by repeatedly drawing binomially distributed samples for fixed n and N_{tot}, and calculate their mean and variance. We
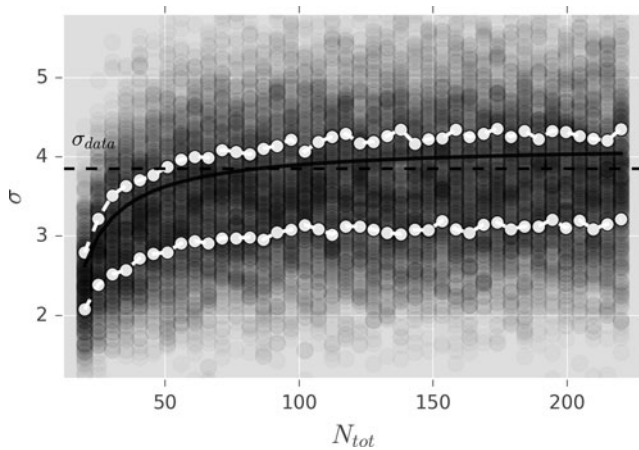
Figure 5. Plot of simulated binomial processes, overlaid on the observed and theoretical standard deviations (the latter is corrected for overdispersion). The white points show the boundaries of the first and fourth quartiles for each simulation, and the white lines were obtained from these points with third-degree spline fitting. The distribution of simulated points is noticeably asymmetric with respect to the theoretical curve.

multiply the standard deviation by the calculated variance inflation factor and plot the lines which contain 50% of the simulations. This procedure provides us with a measure of the confidence interval for the estimate of N_{tot}. The results of the simulation are shown in Figure 5. We can say with 75% confidence that N_{tot} is greater than 56, and thus the recovery probability is less than $12/56 \approx 20\%$. Unfortunately, this method does not provide an upper confidence limit for N_{tot} (or a lower confidence limit for the recovery probability).

Note that by including the variance inflation parameter we are correcting for some of the most egregious simplifications of the model (the assumptions of constant q and N_{tot}). Although many other treatments of overdispersion are possible, these estimates could be improved if a subset of the data was found where more underlying assumptions are satisfied

This analysis is based on data for seizures of radioactive materials. Most seizures result from the activities of police or border agents, who do not have knowledge of the physical or chemical characteristics of the materials involved, which suggests that the recovery probability for WUMN may be like that for all radioactive materials.[20] If the recovery probability for WUNM is equal to the estimates given above for radioactive materials, there would be 75% confidence that the amount of uncontrolled WUMN is more than 80 kg. On the other hand, knowledgeable smugglers should consider WUNM much more valuable than other radioactive materials and should take greater care to prevent detection by police and border agents. This suggests that the seizure probability for WUNM could be significantly lower than for all radioactive materials and that the amount of uncontrolled material is significantly greater than 80 kg. We can compare this with the estimate obtained in the previous section (75% probability of less than 105–165 kg).

What can we learn from capture-recapture methods?

The methods described in the preceding sections are based on the perspective of a constant flow of uncontrolled WUNM, from which some unknown fraction is recovered year after year. A complementary approach is to consider a fixed combined stock of such material, from which small amounts are recovered. This bears some resemblance to the problem in biostatistics of estimating an unknown population size from a limited sample, which is commonly addressed using capture-recapture methods. An initial random sample of animals is captured, tagged, and released, and then another random sample of animals is captured some time later. The total size of the population can be estimated by dividing the tagged in the first sample by the fraction of tagged animals in the second sample.

Capture-recapture methods cannot be directly applied to the problem of uncontrolled WUNM, but we can devise a surrogate measure for recapture events. If WUMN intercepted at different times and locations originated from the same source, we can treat the cases connected to a single source as recapture events from the sample of all cases involving WUNM, which in turn is drawn from the entire WUNM population.[21] The capture-recapture model developed here assumes that all smuggled WUNM comes from a relatively small number of well-defined sources; that each such source has contributed to at least two separate smuggling instances so there could be "capture" and "recapture" events (i.e., multiple seizures from the same source), that the probability of recapture is the same for all sources, and that captured items coming from the same source can be reliably identified and connected (e.g., by using nuclear or other forensic analysis). These assumptions are highly questionable, but this method provides a fundamentally different approach than the binominal model.

We have only a few possible instances of recapture, so the sparse events formulation of the problem is used.[22] It relies only on the number of cases without recapture (denoted by f_1) and number of single recaptures (known as f_2). In this case we have:

$$N_{tot} = S + \frac{\left(f_1^2 - \sum_i Z_i^2\right)}{2f_2},$$

where S is the total number of captures and Z_i is the number of cases captured only in a single year i (we consider each year as a separate "trap" and the sum is over all years). From these definitions it follows that $S = f_1 + f_2$ and $\sum Z_i = f_1$ (where the sum is again over all years). If we take 21 seizures of WUMN and with the assumption of one recapture event ($S = 21$ and $f_2 = 1$) we arrive at $N_{tot} \approx 105$, corresponding to a recovery probability of about 20%. Note that if there were no recapture events ($f_2 = 0$) the general formula gives infinity for N_{tot}, as it should: the observed probability for recapture is zero, implying infinitely large population.

We can obtain the standard deviation of this estimate. The general formula for the variance of N_{tot} is:

$$\sigma^2(N_{tot}) = (N_{tot} - S) + (N_{tot} - S)^2 \left(\frac{1}{f_2 + 1} + \frac{4}{N_{tot}} \right)$$
$$+ \left(\frac{\sum_i (f_1 - Z_i)^2 Z_i - [\sum_i (f_1 - Z_i)Z_i]^2 / N_{tot}}{(f_2 + 1)^2} \right).$$

With this we get $\sigma(N_{tot}) \approx 70$, which is very large. If we assume that the interval $N_{tot} \pm \sigma(N_{tot})$ contains roughly 68% of the possible N_{tot} calculated from similar observations, we can construct a confidence interval for the number of the uncontrolled WUNM items: with 75% confidence there are fewer than 130 items not captured, or less than 120 kg of uncontrolled WUMN.[23] This interval is similar to the one obtained in the section "Estimates Based on the Recovery Probability Distribution" for comparable recovery probabilities. Note that if either of the key assumptions of the capture-recapture model is incorrect (that there is always another member of the same group and that we can reliably link them), it would overestimate the amount of missing material.

Good-Turing formula

In the discussion above we estimated the total number of uncontrolled WUMN items based on the number of known seizures, and converted the derived total number of items to a total mass of uncontrolled material using the average mass per seizure. This assumes that the mass distribution of all uncontrolled material is the same as the mass distribution of the seized material. This assumption merits further examination because it can significantly bias the resulting estimates.

To demonstrate the possible pitfalls of using the observed distribution, we divide the mass of intercepted materials into 1-kilogram bins, as shown in Figure 6. The distribution seems reasonably smooth, but this is a consequence of having relatively large bins; splitting the data in smaller bins leads to a much more irregular and jagged distribution (shown in the inset of Figure 6). Thus, collecting the data in
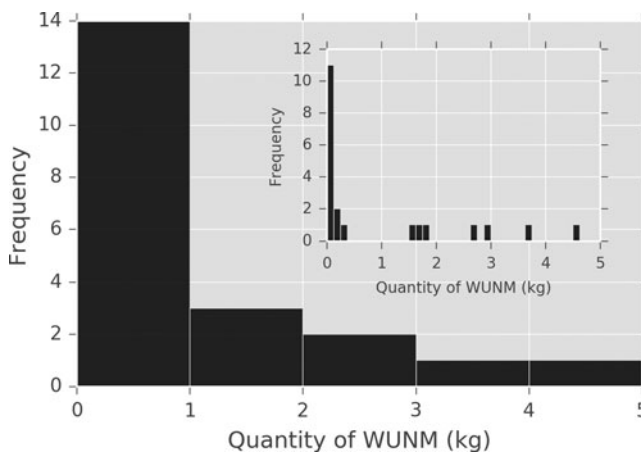

Figure 6. The distribution of intercepted weapon-usable material by mass using 1-kilogram bins. In the inset the distribution using 0.125-kilogram bins is shown.
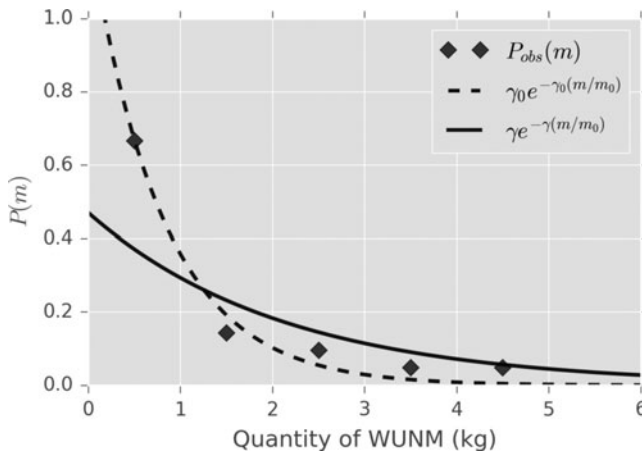
Figure 7. Proposed probability density function of intercepted weapon-usable material by mass. The diamonds represent the observed relative frequencies. The dashed line is a simple exponential fit ($\gamma_0 = 1.25$), and the solid line is modified exponent with $\gamma = 0.47$ adjusted for the possibility of heavier shipments calculated by Good-Turing formula.

1-kilogram bins is a simple but effective, and, for the following discussion, necessary smoothing procedure. The larger problem is that by using the distribution of observed cases to estimate the mass distribution of the entire unobserved pool of uncontrolled material we are implicitly assuming that the probability than any single item in the unobserved pool has a mass of more than 5 kilograms is negligible. This is a result of the fact that we have not observed seizures of more than 5 kilograms of WUNM.

One solution is to adjust the observed distribution to account for the possibility of larger but unobserved smuggling events. This can be done by fitting the observed distribution with a curve and extrapolating to larger masses. (Criticality concerns limits the mass of a single item, but we ignore this here.) A simple exponential, $P(m) = \gamma_0 e^{-\gamma_0 (m/m_0)}$, with m_0 as the size of the bin, in our case, 1 kilogram, and a single adjustable parameter γ_0 provides a reasonable fit to the observed points (see Figure 7). The fitting gives $\gamma_0 = 1.25$, which leads to a small probability of items greater than 5 kg: $P(m > 5\,\text{kg}) \approx 0.002$.

Another approach is to use a method developed by Alan Turing and Irving Good, which provides a way to modify the simple estimate of population frequency based on observed sample frequencies to include the possibility of unobserved events.[24] Turing and Good suggested replacing the naive estimate of population probability for a particular item, given by $P_n = n/N$, where n is the number of times the item has been observed out of the total sample size N, by a modified probability $P_n^* = n_n^* /N$. There are many ways to construct this corrected count n_n^*, but there are two general properties it must have to be useful: $n_0^* > 0$ (which accounts for the unobserved items) and $n_n^* < n$ (so that the total probability can be properly normalized). Turing proposed a simple way to estimate the frequency of unobserved items with the formula $n_0^* = \frac{\nu_1}{N}$, where ν_1 is the number of species of items that have been observed only once. This is intuitively appealing because the frequency of

Table 2. The recovery probability and corresponding amounts of uncontrolled WUMN for the various methods discussed, for confidence levels of 25%, 50%, and 75%. The actual quantities represent the upper bounds estimated from the utilized methods for each confidence level.

	Recovery Probability (%)			Uncontrolled WUMN (kg)		
	25% CL	50% CL	75% CL	25% CL	50% CL	75% CL
Probability distribution	11–16	18–24	26–34	40–60	60–90	110–170
Binomial	—	24	—	—	60	—
Corrected binomial	—	15	20	80	110	—
Capture-recapture	14	20	—	—	80	120
Summary		∼20 (11–34)			∼80 (40–110)	
Good-Turing					∼180 (90–250)	

items we have seen only once should give us a good idea about the number of things we have not seen at all. From this estimate we obtain $P(m > 5\text{kg}) \approx 0.095$, which is more than an order of magnitude greater than that given by a simple exponential fit. Guided by this result, we modify the parameter of the exponential fit, where the value of γ is adjusted to incorporate a estimate for the probability $P(m > 5\text{kg})$. Using $P(m > 5\text{kg})$ from the Good-Turing formula, we obtain $\gamma = 0.47$ as our best estimate for the correction factor. As shown in Figure 7, this $P(m)$ has a much fatter tail. Note that the corrected distribution function is not a good fit for the observed probabilities, which is to be expected because we are accounting for unobserved events rather than simply fitting the observed data.

One objection to the use of the Good-Turing method in this context is that it is typically used to deal with categorical data, rather than continuous variables like mass. Applying it to the WUNM dataset requires discretizing the mass of the items by binning it, as we have done above. The choice of bin size is arbitrary and the results may depend on that arbitrary choice. In this case, however, the choice of bin size was constrained by the requirement to obtain a smooth distribution. In the binning procedure we have used above, these two typically separate steps of the method are combined and are mutually constraining one another. Indeed, even in the case of *bona fide* categorical variables, the necessary smoothing redistributes the original observations, thus effectively making the definition of a category somewhat vague.[25]

The average mass per seizure in the 21 observed seizures was 0.95 kg. The corrected exponential distribution leads to a much higher expectation value per smuggling event ($w' = 2.13$ kg). In other words, the absence of (rare) heavier shipments in observed seizures could significantly bias our estimates of the weight of the missing material, by a factor of about $2.13/0.95 = 2.24$. Thus, total amount of uncontrolled material could be more than twice as large as indicated above.

Conclusion

Weapon-usable nuclear material outside the control of authorities can be used for nuclear terrorism or development of clandestine nuclear-weapons program. We have not found in the open literature an attempt to estimate the total quantity of

uncontrolled WUNM. We provide estimates using several different statistical methods, based on publicly known seizures of WUMN (total of 20 kg in 21 separate seizures) and the assumption that the seizure of WUNM is a random process. The results are summarized in Table 2.

The probability distribution, binomial, corrected binomial, and capture-recapture methods give estimates ranging from 40 to 110 kg for the total amount of uncontrolled WUMN, with a best estimate of about 80 kg; applying the Good-Turing correction would roughly double the total mass, to 90 to 250 kg, with a best estimate of 180 kg. Note that 25 kilograms of HEU is considered a "significant quantity" by the International Atomic Energy Agency—enough for a first nuclear weapon. Thus, enough HEU to build up to ten nuclear weapons may exist outside the control of authorities. Although these methods have serious shortcomings and are based on highly questionable assumptions, they provide a starting point for a further discussion.

Appendix A

The data presented in Table 1 summarize well-documented cases involving seizure of WUNM materials. There are many other incidents involving WUNM reported in open sources, but not definitively confirmed and which we have not included in the analysis. There is one instance, however, which is of special interest: in 1998, Russian security services reported that employees of a nuclear facility in Chelyabinsk were caught attempting to divert 18.5 kg of HEU (the enrichment level was not specified). Adding this quantity to the cases in Table 1 would almost double the total amount of seized WUNM and would significantly change the estimates of the amount of uncontrolled WUNM summarized in Table 2. In this Appendix we repeat the relevant calculations with this case included.

The estimates of recovery probability from the section "Estimates Based on The Recovery Probability Distribution" remain unchanged, but $P(x\langle X|r)$ is modified by the new total. In Figure A1 (an analogue of Figure 2) we show the new probability
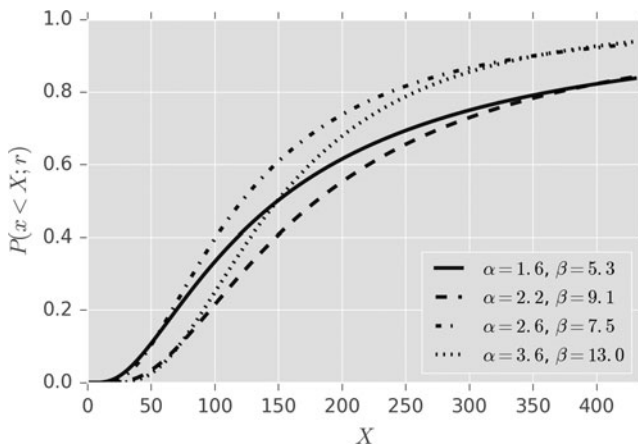

Figure A1. Plot of the probability $P(x\langle X|r)$ that the amount of WUNM that has not been recovered is below X, with the additional 18.5 kg of HEU included in the total amount.

Table A1. The recovery probability and corresponding amounts of uncontrolled WUMN for the various methods discussed, for confidence levels of 25%, 50%, and 75%.

	Recovery Probability (%)			Uncontrolled WUMN (kg)		
	25% CL	50% CL	75% CL	25% CL	50% CL	75% CL
Probability distribution	11–16	18–24	26–34	70–110	120–160	200–300
Binomial	—	24	—	—	160	—
Corrected binomial	—	15	20	150	220	—
Capture-recapture	14	20	—	—	160	220
Summary		~20 (11–34)			~160 (120–220)	
Good-Turing					~230 (100–430)	

$P(x\langle X|r)$, with the same recovery probability distributions as in Figure 1. As could have been anticipated, the likely amount of unrecovered WUNM has increased significantly (see Table A1).

The uncontrolled WUNM amount obtained using the binomial method (see section "Binomial Approximation") are also based on first finding reasonable estimates for the recovery probability, which are weakly dependent on adding a single new instance. In fact, the mean and the variance of the data split by years increase with less than a percentage point, from 6 and 8.1 to 6.04 and 8.14 respectively. In view of this, we do not repeat all the steps in Section "Binomial Approximation," but directly use the recovery probability and its confidence levels obtained there. With these we get for our best estimate 220 kg of uncontrolled WUNM, with 75% confidence it is above 150 kg.

The recovery probability estimated by the capture-recapture method described in the section on capture-recapture methods is also relatively insensitive to the addition of a new case for 1998. The N_{tot} is now around 115, corresponding to a recovery probability of about 19% (compared to 20% in section on capture-recapture methods). This leads to 75% confidence that there are fewer than 130 items not captured, or less than 220 kg of uncontrolled WUNM.

The large amount of WUNM from the Chelyabinsk incident (more than triple that of any other single item in Table 1) underlines the need for correcting the mass estimates by considering the possibility for large but so far unobserved cases. We repeat the analysis from the section "Good-Turing Formula" with this case added. Both simple exponential fit and exponential fit corrected for unobserved masses are shown in Figure A2. The fitting yields $\gamma_0 = 1.16$, which leads to $P(m > 5kg) \approx 0.0022$, while the Good-Turing gives $P(m > 5kg) \approx 0.136$ (the unobserved bins should exclude the 18–19 kg bin, but this is a rather small effect, so we neglect it). The corrected parameter is $\gamma_0 = 0.40$. The observed average mass per item is 38.5/22 = 1.75 kg, while the exponential fit now gives 2.5 kg. Thus, the correction factor is 2.5/1.75 = 1.43, which is smaller than in the main text (2.24).

To summarize, including the 1998 incident involving 18.5 kg of HEU does not significantly change the estimates of the recovery probability, but it significantly increases estimates of the amount of uncontrolled WUNM. Estimates for the total amount of uncontrolled WUMN range from 70 to 300 kg, with a central estimate of about 160 kg; applying the Good-Turing correction further increases those to 100 to 430 kg, with a central estimate of 230 kg (Table A1).
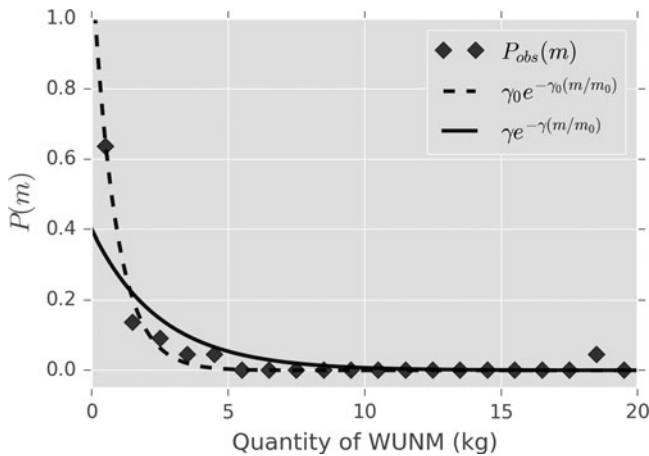
Figure A2. Proposed probability density function of intercepted weapon-usable material by mass. The diamonds are the observed relative frequencies, which include the 1998 incident involving 18.5 kg WUNM. The dashed line is a simple exponential fit ($\gamma_0 = 1.16$), and the solid line is modified exponential function ($\gamma = 0.40$), corrected for the possibility of heavier shipments calculated by Good-Turing formula.

Notes and References

1. Here we use the term "black market" to refer to activities involving attempted illicit sale and purchase of nuclear or other radioactive materials, without necessarily implying the existence of well-developed markets and distribution networks.
2. U.S. Department of Energy, "The United States Plutonium Balance, 1944–2009: An update of Plutonium: The First 50 Years, DOE/DP-0137, February 1996," 2009, http://fissilematerials.org/library/doe12.pdf.
3. For a possible architecture of a future comprehensive international accounting system, see N. Gallagher, J. Siegel, and J. Steinbruner, Comprehensive Nuclear Material Accounting, (University of Maryland, CISSM Monograph, 2014).
4. This is, of course, not the only possible approach to the problem. One alternate is expert elicitation, in which experts in nuclear material smuggling (or black markets more generally) would be asked to estimate seizure probability or uncontrolled stockpile size based on their personal experience, knowledge, and judgment. Expert elicitation has been used to estimate a variety of difficult-to-measure factors, ranging from the impacts of air pollution and climate change to the proliferation resistance of nuclear fuel cycles. See M. Granger Morgan, "Use (and abuse) of expert elicitation in support of decision making for public policy," *Proceedings of the National Academy of Sciences* 111(2014): 7176–7184.
5. IAEA Incident and Trafficking Database, *Incidents of nuclear and other radioactive material out of regulatory control, 2016 Fact Sheet*, 2017, https://www-ns.iaea.org/downloads/security/itdb-fact-sheet.pdf.
6. CNS Global Incidents and Trafficking Database, *CNS Global Incidents and Trafficking Database, 2016 Annual Report*, 2017, http://www.nti.org/media/documents/global_incidents_and_trafficking_2016.pdf.
7. L. Zaitseva, *Nuclear Trafficking: 20 Years in Review*, Contribution to WFS Meeting, Erice (2010). L. Zaitseva, and F. Steinhäusler, "Nuclear trafficking issues in the Black Sea region." *Non-Proliferation Papers* 39 (EU Non-Proliferation Consortium) 1–24 (2014).
8. The assumption of $f(t) = 1$ is similar to the so-called Copernican principle (Gott, J. Richard. "Implications of the Copernican Principle for Our Future Prospects." *Nature* 363(27 May 1993): 315–19. doi:10.1038/363315a0. Gott noted that if an observer saw a

structure that had been in existence for time *t*, and if it can be assumed that *t* is randomly distributed within the total time interval *T* of the structure's existence, the probability for an observation to have happen between *T/3* and *3T/4* is 0.5, and thus the probability that the remaining time *(T – t)* is between *t/3* and *3t* is also 0.5. Higher confidence levels are achieved by expanding the interval *(T – t)*. Our approach is similar, with the key assumption that there is nothing special about the total amount of material recovered to date. For more technical discussion on the Copernican principle, see Gott, J. Richard. "Future prospects discussed." *Nature* 368(1994): 108–108, and Caves, Carlton M. "Predicting future duration from present age: A critical assessment." *Contemporary Physics* 41(2000): 143–153.

9. For reported seizures see The White House, *National Drug Control Strategy, Data Supplement 2016*, Table 80, (https://obamawhitehouse.archives.gov/sites/default/files/ondcp/policy-and-research/2016_ndcs_data_supplement_20170110.pdf). For consumption estimates see B. Kilmer, S. Everingham, J. Caulkins, G. Midgette, R. Pacula, P. Reuter, R. Burns, B. Han, and R. Lundberg, *What American's Users Spend on Illegal Drugs: 2000–2010*, (RAND Corporation, 2014). Note that the ranges in the text do not correspond to a well-defined confidence level. According to Kilmer et al., "The lower and higher ends of the range are meant to give some sense of the uncertainty, but they have a very specific and nuanced meaning that is vulnerable to misinterpretation. For cocaine, heroin, and meth, they reflect only one source of uncertainty: the 95-percent confidence interval surrounding the share of adult male arrest events involving a positive drug test. For marijuana expenditures and consumption, the lower estimate is based on NSDUH estimates with no adjustment for underreporting, and the higher estimate multiplies this value by two. Since there are many other sources of uncertainty, readers should not consider these as lower or upper bounds or as 95-percent confidence intervals. The range should be considered plausible, but not extreme."

10. For reported seizures of African elephant ivory see Environmental Investigation Agency, *Large Scale Ivory Seizures, version 3*, 2016 (https://eia-international.org/wp-content/uploads/EIA_Large-scale-ivory-seizures_V3-January-2000-to-December-2016.pdf). For estimates of the total amount of ivory derived from illegal poaching see Wittemyer, George, Joseph M. Northrup, Julian Blanc, Iain Douglas-Hamilton, Patrick Omondi, and Kenneth P. Burnham. "Illegal killing for ivory drives global decline in African elephants." *Proceedings of the National Academy of Sciences* 111(2014): 13117–13121. These authors assume an average of 6.7 kg of ivory per elephant; approximate 80% confidence interval given by multiplying the interquartile range by 1.9.

11. See McDougal, Topher L., David A. Shirk, Robert Muggah, and John H. Patterson. "The way of the gun: estimating firearms trafficking across the US–Mexico border." *Journal of Economic Geography* 15(2014): 297–327, Topher McDougal, David A. Shirk, Robert Muggah, and John H. Patterson. "The Way of the Gun: Estimating Firearms Traffic Across the U.S.-Mexico Border." Igarapé Institute and University of San Diego, March 2013. https://justiceinmexico.org/wp-content/uploads/2014/09/2013_WayoftheGun.pdf.

12. Robert E. Spiel, *Art Theft and Forgery Investigation: The Complete Field Guide* (Springfield, IL: Charles C. Thomas, 2000).

13. J. D. Foss, "Plutonium and Picasso: A Typology of Nuclear and Fine Art Smuggling." *Inter-Agency Journal* 8(2017): 34–43.

14. The Black Sea region includes the following 12 countries: Bulgaria, Romania, Ukraine, Moldova, Russia, Georgia, Armenia, Azerbaijan, Greece, Serbia, Albania, and Turkey (L. Zaitseva, and F. Steinhäusler, "Nuclear trafficking issues in the Black Sea region"). Note that Turkey, Greece, Bulgaria, Romania, Macedonia, and Serbia lie along one of the most active routes for smuggling drugs, goods, and people to and from Europe.

15. See Figure 8 in L. Zaitseva, and F. Steinhäusler, "Nuclear trafficking issues in the Black Sea region."

16. A more realistic model would be one in which N_{tot} itself is a Poisson-distributed random variable. Such model would be doubly stochastic (with parameters of the probability distribution being themselves random variables), which could account for the observed overdispersion (see below). Unfortunately, this would also invalidate the simple variance formula we use to estimate N_{tot}.

17. R. W. M. Wedderburn, "Quasi-likelihood functions, generalized linear models, and the Gauss-Newton method." *Biometrika* 61(1974): 439–447. J. Hinde, and C. Demétrio, "Overdispersion: Models and estimation." *Computational Statistics & Data Analysis* 27(1998): 151–170.

18. Grouping the data is not unique, which introduces randomness in the procedure. To minimize its effects, we have simulated 10,000 random groupings, calculated the standard deviation of each one of them, and took the median of this sample ($\sigma_{data}^{median} \approx 3.88$).

19. K. P. Burnham, D. R. Anderson, G. C. White, C. Brownie, and K. H. Pollock, "Design and Analysis Methods for Fish Survival Experiments Based on Release-Recapture." In American Fisheries Society Monograph No. 5, 1987, 252–254 (Bethesda, Maryland, USA).

20. Most seizures of WUNM involved HEU, which is weakly radioactive and emits low-energy gamma rays that are easily shielded. HEU is therefore more difficult to detect with radiation detection equipment than many other radioactive materials. We do not know what fraction of seizures of radioactive materials or WUMN were triggered by instrument detection, but if instrument detection was responsible for a large fraction of seizures of radioactive material, then the seizure rate for radioactive materials may be much higher than that the seizure rate for WUMN. If so, the recovery probability estimates in this section could significantly underestimate the amount of uncontrolled WUNM.

21. The analogy is imperfect: apart from the implicit assumption of "recapture" always being possible, it also ignores the changes in the recovery probability from the first to the second and so on captures (due to the changes of the unobserved source population). We use the original formulation of the recapture problem, but we expect that the method could be refined to address these issues.

22. We consider as a recapture event the 2001 incident in Paris, with initial capture in 1999 in Rousse, Bulgaria. There are other potential candidates: 2003, 2006, and 2009 interceptions in Georgia, and seizures in 1994 and 1995 in the Czech Republic. Since it is difficult to be certain about connection between these cases, we make the most conservative assumption of a single recapture. For sparse formulation of the capture-recapture method see A. Chao, "Estimating Population Size for Sparse Data in Capture-Recapture Experiments." *Biometrics* 45(1989), 427–438.

23. With $f_1 = 19$ and $f_2 = 2$ (two recapture events), the method yields $N_{tot} \approx 71$ and $\sigma(N_{tot}) \approx 38$. This leads to a recovery probability of around 30 percent, and 75 percent confidence there are fewer than 110 items not captured, or less than 105 kg of uncontrolled WUMN.

24. I. J. Good, "The Population Frequencies of Species and the Estimation of Population Parameters." *Biometrika* 40(1953), 237–264.

25. Much research on the Good-Turing method has focused on finding consistent and convenient smoothing techniques that effectively redistribute the observation. See, for example W. Gale, and G. Sampson, "Good-Turing Frequency Estimation Without Tears." *Journal of Quantitative Linguistics* 2(1995), 217–237.