Measuring Illegal Immigration at US Border Stations

by Sampling from a Flow of 500 Million Travelers

Palmer Morrel-Samuels

Employee Motivation and Performance Assessment,

and

University of Michigan Business School

Measuring Illegal Immigration at US Border Stations

by Sampling from a Flow of 500 Million Travelers

ABSTRACT

500 million travelers enter the US yearly at a Port of Entry (POE) after an

Immigration & Naturalization Service (INS) interview. We describe a general method for

sampling from a flow, and summarize results from random reinspections of travelers at 20

POEs. Analyses reveal that 47 in 5614 travelers (0.8% – 0.24%) were erroneously

granted entry. Results suggest INS intercepts 9.3% to 16.0% of travelers attempting

illegal entry at a POE, and that INS mistakenly admits 3.51 to 6.01 million illegal

immigrants at POEs annually. Additional applications of our sampling method (e.g., for

quality control, population studies) are briefly discussed.

Approximately 500 million travelers enter the United States each year after a brief interview with an Inspector of the US Immigration & Naturalization Service (INS). Ever since the US began its concerted effort to control entry into the country in 1882, there has been a lingering controversy about illegal admission and the illegal immigration that it facilitates.[1] When the Nationality Act was passed in 1906 many of the country s current policies and guidelines were codified, and they remain in place today with surprisingly little modification (United States Department of Justice — Immigration and Naturalization Service, 1991). For example, INS still bars entry for the insane, convicts professional beggars travelers likely to become public charges and (as specified in the Immigration Act of 1891) those who have committed crimes of moral turpitude. Research suggests that illegal immigrants constitute a surprisingly large portion of the current US population (e.g., Beck, 1997), a trend with important political, legal, clinical, cultural, and occupational ramifications (e.g., see respectively Alvarez & Butterfield, 2000; Goldman, 1999; Smart & Smart, 1995; Padilla, 1993; and Halcrow, 1987). But, despite considerable interest, no one has a precise measurement of how many illegal immigrants enter the United States annually by evading detection at an airport or traffic checkpoint.

This paper describes the outcome and the methodological details of a new process that INS has adopted to measure the number of illegal immigrants who come into the US

---

[1] Although some writers prefer the term illegal alien this paper follows INS s convention of using the term illegal immigrant to denote non-citizens who enter the US without legitimate grounds to do so, and who — by virtue of that entry — remain in the country for some undetermined period of time.

at its Ports of Entry — the border stations and airports that ring the nation's borders[2] the research does not address illegal entry occurring between border stations (viz., at places on the international boundary where there is no staffed border crossing facility) nor does it examine the residency duration of those who entered the country illegally. In addition to providing an estimate of illegal immigration, our method can serve as a general approach for sampling from a large flow in applied settings where the elements of the sample frame do not reside in a static stock, and where a full list of the population members cannot be compiled prior to sampling. As far as we can determine, the approach appears to be unique.

Previous Attempts to Enumerate Subsets of Travelers: The US Customs Service is currently implementing a program to determine the number of travelers slipping through Customs with contraband. Their approach involves a computer generated selection signal that appears on a Customs Inspector's computer screen as he or she processes license plate numbers (Bolstein & Hill, 1996). The method is promising, but several drawbacks (viz., reliance on sampling every *nth* vehicle, allowing manual over-rides where vehicles are intentionally added to or deleted from the sample by Inspectors, and loss of double-blind procedure) make claims to randomization tenuous and threaten the assumptions that lie at the heart of margin of error computations. The procedure at

---

[2] This description for a Port of Entry is something of an oversimplification, as are virtually all generalizations about INS. There are approximately 300 Ports of Entry processing travelers, but a substantial number are far inside or far outside the nation's boundaries. Seaports have been excluded from this discussion because the sampling program described in this paper has yet to be implemented there.

Customs is apparently being refined on a continuing incremental basis, so it is likely that methodological rigor will be increased as time passes.

An earlier approach to estimating the number of inadmissible travelers coming through US Ports of Entry highlights the progress made during recent years. Just over twenty years ago the INS set about estimating illegal entry at POEs (Ports of Entry) and used the Delphi method — where a small group of informed colleagues simply hypothesize in a structured discussion until consensus is reached (Lesko Associates, 1975). One can hardly imagine a more error-prone approach; and indeed the estimate that grew out of that process (8.2 million undocumented aliens presumed to be living in the US as of 1975) has been disconfirmed by a more definitive study that used quasi-experimental methods (Cook & Campbell, 1979) to conduct interviews with approximately 6000 illegal aliens (US Department of Justice Immigration and Naturalization Service, 1992).

That same empirical study (US Department of Justice Immigration and Naturalization Service, 1992) is the source of a suggestive statistic that bears on the issue at hand: The interviewers found that 0.7% of the undocumented non-citizens living in the US had used an illegal passport, visa, or Border Crossing Card to enter the country. As suggestive as the figure is, its generalizability is unfortunately open to question because the travelers who are most likely to use forged or misappropriated documents would also be unlikely participants in the study, and, if participating, might be somewhat circumspect about volunteering information concerning extensive fraud; that is, the 0.7% might represent non-habitual violators who are quite different from travelers who

routinely turn to fraud, contraband, and illegality for their sustenance. Nevertheless, the outcomes of this well-designed field experiment, rare in the discipline of immigration studies, have been corroborated by the little information available from alternative (albeit less rigorous) methods (e.g., Edmonston, Passel & Bean, 1990; Hill, 1985; J. A. Reyes Associates, 1977) that involve counting or estimating the number of illegal aliens residing in the US.

An interesting approach is found in Kish s work for the New York Port Authority (Kish, Lovejoy & Rackow, 1961), where the goal was to compile a profile of vehicular traffic entering and leaving the city. They used a continuous sampling method (running all hours of the day, seven days a week) that was stratified by volume and rolled from toll booth to toll booth in one hour blocks gathering information about the state of origin, number of passengers, destination, and travel purpose of every 12th vehicle. It is only the incorporation of a sampling routine based on *nth* selections that weakens the method, because of course, vehicles in groups (formal or not) may represent an important atypical part of the traffic flow; specifically, once a vehicle in a caravan with less than 12 cars or trucks has been selected, the remaining members of the caravan have no chance of entering the sample, despite the fact that origins and travel purposes may vary substantially within the group. In some respects the approach is similar to recent work on regulated random schedules (Keohler & Levin, 1998) but lacks the benefits that variable start points and variable phase extent provide.

The Experience Sampling Method: The current work for INS uses a time sampling method that is similar, but not identical, to much of the work in experience sampling. Since the late 70 s a growing body of research has adopted what is often called the experience sampling method; it typically uses a pager to solicit journal entries from subjects at randomly selected times during their normal daily routines (Larson & Csikszentmihalyi, 1978). Experience sampling is based on features common in conventional sampling work (Kish, 1965; Kalton, 1983) and has been applied successfully in a range of disciplines. It has been used in clinical studies to analyze preventive health care (Abramowitz, Obten & Cohen, 1998), eating disorders (Swarr & Richards, 1996), anger management among forensic patients (Hillbrand & Waite, 1992), and post-hospitalization rehabilitation (Filstead, 1988). In field research it has been used with good success in many studies, including research on assistance seeking (Bornstein, 1998), intrinsic motivation (Csikszentmihalyi & Rathunde, 1992; Moneta & Csikszentmihalyi, 1996), hormones and cognition (Dabbs, Strong & Milun, 1997), action and affect (Gauvin, Rejeski & Norris, 1996), television viewing (Kubey, 1986), mood (McConville & Cooper, 1995), personality and cognition (Prescott, Csikszentmihalyi & Graef, 1981), and communication complexity (Rathunde, 1997) to name but a few. The method has good reliability and validity (Csikszentmihalyi & Larson, 1987; Csikszentmihalyi & Larson, 1992), and — of particular importance for work with INS — has held up well in a broad range of cultures that includes subjects from Austria (Brandstaetter & Gaubats, 1997), the Netherlands (Barge-Schaapveld, Nicolson, Van Der Hoop & Devries, 1995), Britain (Clarke & Haworth, 1994), Australia (Hnatiuk, 1991),

and South Korea (Lee, 1995), among others. However, despite cogent arguments in favor

of using experience sampling in applied settings (Alliger & Williams, 1993) and studies of

observable behavior (Csikszentmihalyi & Rathunde, 1992), relatively little has been done

with this methodology in the applied arena. Some recent exceptions are a cross validation

of pain assessments (Lousberg, Schmidt, Groenman, Vendrig et al., 1997), a study on

employee attitude (Alliger & Williams, 1993), an analysis of the relation between

motivation and mathematical computation accuracy (Schiefele & Csikszentmihalyi, 1995),

research on hiring at a large company (Brandstaetter & Gaubats, 1997), and an analysis of

stress in nurses (Shiu, 1998).

Method

To sample travelers from the flow at the nation s Ports of Entry (POEs) several

assumptions were adopted about similarities and differences between POEs, hours of the

day, and specific inspection lanes within POEs in a straightforward manner consistent

with quasi-experimental methods for field settings (Cook & Campbell, 1979). These

assumptions (e.g., that Inspectors are similar but not identical in the rigor with which

they conduct traveler interviews) led to a decision to stratify the 300 POEs into 5

quintiles on the basis of volume, to select travelers from the flow at virtually any hour of

any day throughout all days of the month, and to rotate from lane to lane on a randomized

basis without notifying the lane s Inspector beforehand.

The purpose of the random selection procedure was to re-inspect a representative group of travelers who had already been approved for entry into the US after a brief conventional INS interview, but who were still within the confines of the Federal Inspection Area at the POE, and could therefore still be detained, admitted, or denied entry if upon closer examination the traveler was found to be technically inadmissible. These supplementary inspections (called INTEX exams, for Immigration and Naturalization Traveler Examinations) followed a set procedure and lasted about 20 minutes. The reader should note that this is substantially longer than the minute or two that INS Inspectors typically spend examining a passport, visa, or Border Crossing Card before granting admission to a non-citizen traveler.

Because a recent federal law — The Government Performance and Results Act of 1993, commonly called GPRA — requires all federal agencies to furnish objective empirical evidence of their actual results and effectiveness, INTEX examinations were designed to measure one crucial part of the INS mission: To interdict inadmissible travelers who would otherwise enter the US. Because the vast majority of the 500 million travelers who present themselves at a POE are admitted, the task of evaluating Inspectors accuracy by obtaining a random sample is similar to field research in public health studies where the incidence of a sought disease is very low: Only 0.56 million, or 0.11%, of all travelers were denied entry by INS Inspectors last year. A primary assumption of the work, therefore, was that the effect size for the phenomenon of interest is very small; accordingly, when an INS Inspector mistakenly allows entry for a traveler who is technically inadmissible — a miss, in signal detection theory terms — detection will

necessarily be difficult because the base rate (justified exclusions) is itself exceptionally

low even though Inspectors are free to take as much time as they choose with a traveler,

and can send questionable or complex cases on to a Senior Inspector examining travelers in

a special inspection area. Of course, because INS has never conducted a full empirical test

of Inspector accuracy, we had no way of knowing what the actual miss rate would be

until the random selection process was implemented and results analyzed.

To compile a sample of travelers for re-examination, we constructed a multi-stage

multi-phase proportional stratified random sample. There were 10 airport POEs and 10

land POEs involved,[3] each from either the top, bottom, or middle quintile of traveler

volume. A comprehensive list was constructed of all terminals and all lanes (i.e., the

desks or booths where INS Inspectors interview travelers) at each of these 20 ports.

High- volume ports were sampled 5 times per day, low-volume ports once a day, and

intermediate-volume ports were sampled 3 times a day; every inspection lane at every

minute of every hour when the POE had sufficient staff and a sufficient volume of

travelers[4] was included in the sampling frame. At a randomly selected time (listed on a

---

[3] An *a priori* test of statistical power (Cohen & Cohen, 1983) determined that the minimum requirement
for monthly analyses would be 20 POEs if each was sampled 3 times per day (scheduled n = 1860,
projected n = 1302; r = 0.1, p = .85, alpha = .05 two-tailed). The 20 participating POEs were chosen at
random by two senior INS officials from a list of all POEs; (ports were excluded from the study if they
had less than 3 Inspectors, or if the port's layout raised concerns about participants' safety).
Randomization proceeded as follows: Two lists were compiled, one of airport POEs and one of land-based
POEs; each list was divided into 5 quintiles according to volume; 10 airport POEs and 10 land-based
POEs were selected at random with 3 coming from the top quintile, 3 from the bottom quintile, and 4
from the middle quintile of each list. The 20 selected POEs accounted for approximately 90% of INS
volume during the previous year.
[4] A survey was developed by personnel from INS and EDS to determine the universe of suitable hours and
locations for these INTEX exams. In some POEs, INTEX exams could occur virtually around the clock,
and at a few POEs (where staff or traffic was limited) INTEX exams might occur only during a 2 or 3 hour
window.

limited-distribution memo sent to the POE s director) a specially trained inspector[5]

would approach the numbered lane listed on the random schedule, and would conduct an

INTEX examination on the next traveler who received approval for entry.[6] The process

was double-blind, with neither the traveler nor the initial INS Inspector knowing that an

INTEX exam was forthcoming. (This is an important part of the procedure because

travelers are legally free to withdraw from an INS inspection without penalty; it is an

option that is exercised with varying frequency at the nation s POEs.) The multi-phase

component of the random schedule kept hours and minutes shifting; the multi-stage

component kept airport terminals and lanes shifting, and the stratification insured that

inspections at high-, low-, and intermediate-volume ports would be proportionately

represented.

Construction of the randomized schedule proceeded as follows: For any given

month, each day of the calendar was given one row for each of the 20 ports, and these

appeared either 1 or 3 or 5 times depending on the port s volume quintile. This generated

1800 rows for 30-day months, and 1860 rows for 31-day months; (we will hereafter use a

1860 row matrix for our explanation). Each of the 20 POEs in these 1860 rows also

included the start-time and end-time of the temporal window during which staffing

strength and traveler volume would allow Inspectors to conduct INTEX exams.

Appended to these hour designations was a minute designation, also selected by random

---

[5] Training sessions and training materials were provided by the Government Consulting Group of EDS.
[6] Families, touring clubs, vehicular passengers, and similar groups often present their passports together to an INS Inspector; in such cases all travelers in the group were given an INTEX exam. If the Inspector was unable to conduct the exam because no traveler was emerging from the specified lane at the specified time, then a group of predetermined rules would alter the time and/or location of the INTEX interview.

from the universe of minutes in an hour, so that, within this temporal window, INTEX

exams would be scheduled for any minute of any hour with equal probability across the

20 ports. INTEX exams for this analysis were conducted during four months at these 20

POEs.[7] During that period a total of 6484 INTEX exams were scheduled to occur at

randomly selected times and at randomly selected inspection lanes. Actual randomization

was achieved by using a random shuffle with unique seed values to reorder the rows of

each monthly matrix, and then, in the appropriate column, inserting incrementing values

for hours or minutes or POE sectors (being either for vehicles or foot passengers, and

utilized in a volume-sensitive ratio) or lane numbers. By this procedure, (what we ll call

Time-Location Sampling, for the sake of convenience) times and locations for INTEX

exams were randomly selected in a proportional stratified sample with fully known

probabilities even though the elements of the sample (the travelers) were never

enumerated in a stable static list.

INTEX exams were conducted as follows: Port Directors would receive the Time-

Location Sampling schedule, and communicate it on a daily basis to the designated

INTEX Inspector at the port. The INTEX Inspector would go to the primary inspection

lane specified by the schedule and wait outside the view of the INS Inspector screening

travelers at that location. At the time specified by the schedule, the INTEX Inspector

would intercept the next traveler authorized by an inspection at that primary inspection

lane and give him or her a brief document describing the INTEX re-examination procedure

and its random selection process. If the primary lane was inactive, a set of default rules

---

[7] Data were made available by INS for the period between October 1st 1998 and January 24th 1999.

would direct the INTEX Inspector to specific different locations. By this process, a

randomly selected traveler who had just been approved to enter the US would be directed

to a separate area in the port where a more thorough inspection of documents, databases,

and possessions could be completed. INTEX Inspectors used a data recording sheet that

facilitated both data collection and standardization of the re-inspection procedure. If the

traveler was found to be inadmissible on the grounds of new information gleaned during

the INTEX examination, then he or she was denied entry to the US, and ushered out of

the port s inspection area.

Results

When data collection for this analysis terminated, we had received reports

summarizing INTEX inspections occurring in 4718 (i.e., 73%) of 6484 scheduled time

slots. In total, 5614 travelers received the supplementary inspections (average duration

14.2 minutes). Of these, 2497 (44.5%) were US citizens and the remainder were citizens

in one of 102 other countries.[8]

Excludable Travelers Admitted in Error: The primary statistic furnished by the

INTEX program is derived from the number of travelers who had just been approved for

---

[8] By federal law, US citizens are not required to carry identifying documents; accordingly, in most cases a verbal claim of US citizenship is tantamount to entry. For the 2497 INTEX exams of US citizens, the average duration was 6.6 minutes and the median duration was 5.0 minutes.

entry into the US by an INS Inspector, but who were found to be inadmissible by virtue of information discovered during their INTEX examination. Specifically, the number of travelers found inadmissible during an INTEX exam divided by the total number of travelers re-examined, gives a proportion showing the extent to which — all other things being equal — excludable travelers are being admitted in error.

During the 116 days of this project, 47 inadmissible travelers were found among the 5614 travelers who received an INTEX exam; accordingly, INS Inspectors erroneously authorized entry during 0.8% of their inspections, and granted entry correctly during 99.2% of their inspections, at least in the months and locations covered by the study. Moreover, by applying Kalton s formula for computing margins of error in stratified random samples (Kalton, 1977) it is clear that, with a response rate of 73%, a design effect of 0.8, a q of 99.2, an alpha of .05 and a 95% confidence interval, the margin of error surrounding this proportion is quite small: approximately one quarter of one percent. That is — at least during the months examined at the locations surveyed — our data show that INS mistakenly granted admission to 0.8% – 0.24% of the travelers who requested admission at a Port of Entry.[9]

Evaluation of the Stratification Variable: Despite the fact that there was no way to anticipate exactly how the stratifying variable (size of the POE) would function, the results were encouraging: In the current database of 5614 INTEX exams (where

---

[9] These results report an unweighted ratio, and do not account for the fact that the proportion of air travelers to land travelers varies somewhat from month to month and year to year.

inadmissibility s standard deviation was 0.05 at small POEs, 0.08 at intermediate POEs, and 0.10 at large POEs) an analysis of the design coefficient shows that the cost of conducting INTEX exams in this four-month period was only 80% of what it would have been had stratification not been used in our design. Specifically, without stratification, we would have had to conduct 1402 additional INTEX exams to maintain our current margin of error. Results suggest that stratifying on the basis of volume was suitable for the conditions of the study.

       Reliability, Validity, and Business Utility: It is reasonable to assume that the rigor of these INTEX exams varied, at least to some extent, from location to location. This is natural enough, given the fact that an Inspector s job involves a degree of subjective judgement (e.g., while deciding whether or not a traveler is likely to become a ward of the state). The INTEX program incorporated five features to enhance reliability: 1) INTEX responsibilities were restricted to senior staff; 2) As mentioned above, INTEX Inspectors attended training sessions led by a small team of specialists; 3) A committee of senior executives at INS headquarters reviewed, and in some cases overruled, the outcome of each INTEX inspection prior to data analysis; 4) INTEX inspectors were required to follow a check-off sheet that both outlined the steps of the examination and facilitated data collection, and; 5) Directors at each POE monitored staff adherence to the schedule of random times and locations. It is consistent with these safeguards that the split-half reliability coefficient of the INTEX exams was within the acceptable range ($r = 0.7$, $n = 47$; $p < .0001$). The findings confirm a substantial body of research showing that

employees intrinsic motivation and features of the corporate culture are linked to defect

rates in a number of work settings.

We assessed convergent validity of the INTEX procedure by evaluating the

relation between INTEX results and data from the previous year s employee survey at

INS. The employee survey itself had adequate reliability and validity, (at least as far as

we could determine), making it a reasonable resource for cross-validation.[10] Of the eight

topics addressed in the employee survey, five had a significant and coherent relation to

the INTEX data; the remainder were not significant. In all cases, the more favorable the

rating of the corporate culture, the better the Inspector s ability to catch inadmissible

travelers during the initial (i.e., conventional) interview. That is, we assume that INTEX

exams function as typical quality control assessments, where higher quality during a

primary process (i.e., the conventional brief interview that all travelers receive) is

accompanied by fewer revealed defects during the subsequent secondary process (i.e.,

INTEX Inspector s re-examination of randomly selected travelers). Good convergent

validity is suggested by the fact that fewer revealed defects during the INTEX exam are

predicted by better employee training (n = 3533; df = 1; F = 30.55; p < .0003; Beta = -

.01), greater support for INS policies (n = 3533; df = 1; F = 29.87; p < .001; Beta = -.01),

better customer orientation, (n = 3533; df = 1; F = 9.11; p < .003; Beta = -.01), stronger

---

[10] The employee survey s reliability coefficient was high (Cronbachs s Alpha = .97) despite several
psychometric weak points in the survey s design; evidence of validity is suggested by the correlation
between number of supervisors and rated communication (r = .28, n = 1689), as well as the correlation
between average motivation score in the POE and perceived morale (r = .80, n = 88). The survey was
designed to measure 8 topics, and a principal components factors analysis of the data confirmed that only 8
factors had an eigenvalue equal to or greater than 1; these 8 factors accounted for 69.8% of the variance in
the survey responses.

organizational commitment, (n = 3533; df = 1; F = 11.61; p < .0007; Beta = -.01), and

higher overall employee motivation (n = 3533; df = 1; F = 9.49; p < .002; Beta = -.01).

We assessed internal validity by examining the relation between the Inspector s

thoroughness during the INTEX exam, and the likelihood of his or her discovering grounds

for inadmissibility. Inspectors had the ability to query as many as 30 sources of

information (e.g., inspection of carry-on baggage, inspection of a vehicle s gas tank,

querying of the INS database on felons, etc.) before determining the outcome of the

INTEX interview; some data queries were obligatory, and others were optional. Because

Inspectors had some latitude during INTEX exams, the total number of data sources

queried during the examination could be used as a measure of thoroughness.[11] Good

internal validity of the sampling and re-examination process is suggested by the fact that

the greater the thoroughness (i.e., the higher the number of data sources evaluated by the

Inspector) the greater the ability to detect an inadmissible traveler. The multiple

regression controlled for type of POE (airport vs. land port) and was highly significant (n

= 5614; df = 2; F = 21.87; p < .0001; B = .001).

Suitable discriminant validity of the assessment is suggested by the fact that the

likelihood of being excluded varies with the identity of the POE, presumably, because

economic and societal pressures to emigrate vary by locale. The regression is highly

significant (n = 1698; R^2 = .03; df = 19; F = 2.55; p < .0003) and includes no covariates.

---

[11] We recognize the fact that highly experienced Inspectors may use heuristics which allow them to query fewer data sources without jeopardizing rigor; nevertheless, it seems appropriate to treat this as a measure of thoroughness because all travelers who received an INTEX exam had already passed a conventional primary inspection. Accordingly, it is justifiable to presume that an Inspector s persistence facilitates discovery of grounds for exclusion.

Adequate levels of business utility are demonstrated by the fact (outlined above) that lower defect rates during conventional (non-INTEX) interviews reflect, at least to some extent, predictable features of the corporate culture. This, along with evidence that POEs differ significantly in the defect rate revealed by INTEX exams, suggests the feasibility of remedial interventions designed to enhance both employee motivation and performance at selected POEs.

Sources of Information Leading to Successful Discovery of Inadmissibility: As mentioned above, the INTEX examination procedure included both mandatory and optional searches of 30 sources of information. The overall regression between identity of the queried data source and INTEX outcome was significant (n = 5614; R^2 = .02; df = 29; F = 3.29; p < .0001) showing that detection of inadmissible status could be predicted, at least to some extent, on the basis of information sought by the Inspector. Moreover, examination of individual regression coefficients showed that three sources of information (two databases and one part of the traveler s vehicle) had a significantly positive effect, increasing likelihood of detection by 6.8%, 1.4% and 1.3% respectively (B = .07, .01, .01; t = 3.53, 2.25, 2.00; p = .0004, .024, .045). Conversely, there were also sources of information that, once queried, seemed to lead INTEX Inspectors to conclude that the traveler should be considered legitimately admissible; these three sources of information (a personal belonging, a part of the traveler s vehicle, and an INS database) increased the likelihood of being granted entry by 2.6%, 1.2%, and 1.0% respectively (B = .03, .01, .01; t = 2.04, 1.93, 2.37; p = .041, .054, .018). The results have obvious programmatic value

to INS executives who are concerned about procedures and availability of databases in the field.[12]

Demographics and Predictor Variables Associated with Inadmissibility: Analysis of the data reveals, just as one would hope, that only the duration of the INTEX examination (F = 40.8, df = 1; p < .0001) and the identity of the POE itself (F = 2.0, df = 19; p < .007) have any impact on an Inspector s decision to exclude a traveler. Neither the staffing agency responsible for maintaining the inspection lane (either INS or US Customs Service), nor the number of travelers in the group, nor the traveler s gender, nor the traveler s age, nor the foreign traveler s country of citizenship had any effect on admissibility. (For the overall exploratory regression predicting inadmissibility: n = 2512; R^2 = .08; df = 121; F = 1.72; p < .0001. All non-significant predictors had probabilities above the .05 level.) The finding is important because it suggests that the INTEX procedure, as well as the randomized schedule and the Inspectors who conducted those exams, did not discriminate unfairly against travelers who fit any specific profile based on gender, race, age, or citizenship.

Discussion

---

[12] The reader will understand our decision to refrain from identifying the names of these databases and locations. INS received a full report.

The overall rate of inadmissible travelers missed (0.8% – 0.24%, i.e., 47 out of 5614, with a 95% confidence interval and an alpha of 0.05) suggests that, at least to the extent that we can extrapolate to the entire INS network, a relatively small proportion of the travelers who are granted entry to the US should not be admitted. However, when we map this unweighted proportion onto the actual number of travelers, rather than the proportion, the picture is more sobering: Approximately 500 million travelers were granted entry at a POE last year, and approximately 0.56 million travelers were denied admission during conventional (non-INTEX) screenings. If we are correct to believe, as the INTEX results suggest, that 0.8% of those 500 million, plus or minus 0.24%, should have also been excluded, then INS missed several million inadmissible travelers. Specifically, if we assume that the 0.56 million excluded travelers were correctly denied entry, then the 500.56 million people seeking entry at US POEs last year contained anywhere between 6.01 and 3.51 million people (i.e., 0.56 million plus 5.45 million or 2.95 million) who should not be granted entry according to US federal laws. Current INS procedures enable Inspectors to detect only 0.56 million of these inadmissible travelers annually. Accordingly, our results suggest that INS is, at best, excluding only 16.0% of these inadmissible travelers, and at worst, is excluding only 9.3% of the travelers who should be denied entry.

The results also suggest that the reliability, validity, and utility of Time-Location Sampling are adequate. The fact that POEs are significantly different in their INTEX catch rates suggests that our sampling method can operate even in the presence of some contextual pressures (e.g., presumably to document a high or low catch rate) that are not

uncommon in many organizational settings. Time-Location Sampling should have unique value for researchers who must draw a representative sample from a dynamically changing population where the individual members cannot be identified on a static comprehensive list (e.g., in some quality control programs and in some population studies); the primary benefit of the method is its ability to avoid the central problem with sampling every n^{th} element: namely, the contiguity effects that make neighboring elements similar (or dissimilar) to each other according to unknowable multivariate dynamics. Moreover, the direct business utility of the approach is also demonstrated by the fact that, in this application at least, reinspections driven by Time-Location Sampling successfully allowed INS to discover and exclude a number of illegal immigrants (e.g., a Canadian citizen wanted on an arrest warrant, a British citizen who misrepresented himself as an American to avoid legal authorities, a Honduran woman traveling with a forged passport, a Mexican citizen with previous drug convictions, etc.) who would otherwise have been granted entry into the US.

Over the course of the project INTEX exams revealed failures in vigilance that varied in seriousness. Roughly a third of the INTEX interceptions involved minor technical infractions (e.g., over-staying a previous visa by a few days); a third involved substantive infractions (e.g., making a false claim of US citizenship); and the remaining third involved criminal infractions (e.g., using a forged document).[13]

---

[13] INS executive officers undertook this analysis using criteria that are common within their agency. The distinction (technical vs. substantive vs. criminal) has little impact on the outcome of an INTEX exam because non-citizens can be denied entry for lying during an INS inspection.

It is important to acknowledge that the current approach does entail a number of limitations. For example, we have very little idea about the entry of technically inadmissible travelers at very small POEs (where staffing was insufficient to support the INTEX program), nor about the number of illegal immigrants who subsequently leave the US without attracting attention. It is also not possible to quantify the extent to which INS meets the second part of its federally mandated mission: To facilitate entry for those who should be granted entry. That is, the current work allows us to quantify misses (inadmissible travelers who entered) and hits (admissible travelers who entered), but false alarms (admissible travelers who were excluded) and correct rejections (inadmissible travelers who were excluded) have yet to be analyzed. Historically, false alarms and correct rejections by INS have been less pressing than hits and misses because there are numerous legal and procedural safeguards that prevent overzealous exclusion of travelers.[14]

It is also the case that we cannot be entirely certain about the extent to which we can generalize these results to POEs that were not explicitly included in the program; however, it is reasonable to assume the results have considerable generalizability because the primary process for selecting the 20 POEs relied on randomization, and because the only differences between participating and non-participating POEs had to do with staffing levels and (where relevant) the ability of an Inspector to find a safe but inconspicuous place to intercept vehicles. Currently INTEX has been expanded to run year-round throughout the land-based and airport-based POE system, with results being

---

[14] This may be starting to change now that excluded travelers have less recourse for addressing grievances

posted on the agency s Intranet (J. Walton, Personal Communication, May 1, 2001), an

application of the research that is consistent with recent recommendations (e.g., US

Commission on Immigration Reform, 1994) previous research (e.g., Kish, 1995) and

federal legislation (Government Performance and Results Act, 1993).

Although there is much that the current work cannot tell us, there is a good deal

that it does allow us to specify with a fair degree of precision. Mistakenly authorized

entry of illegal immigrants has been a contentious issue at least since congressional

hearings on the topic just after the turn of the century (US Committee on Immigration and

Naturalization, 1919). One of the uncertainties that has complicated this issue is the fact

that we have had no reliable and practical method for obtaining a clear and valid estimation

of the problem s scope. Now, more than eighty years after that initial debate in

Congress, we finally have data that provide clear information on the entry of inadmissible

travelers at specific locations under specific conditions. Moreover, we also have initial

data suggesting that enhanced vigilance at US POEs is associated with (among other

things) better training, stronger organizational commitment, and higher levels of employee

motivation.

---

under rules governing Expedited Return.

REFERENCES

Abramowitz, S., Obten, N., & Cohen, H. (1998). Measuring HIV/AIDS Case

Management. Social Work in Health Care, 27, 1-28.

Alliger, G., & Williams, K. J. (1993). Using signal-contingent experience sampling

methodology to study work in the field: A discussion and illustration examining task

perceptions and mood. Personnel Psychology: Special Issue: Innovations in research

methods for field settings, 46, 525-549.

Alvarez, R. M., & Butterfield, T. L. (2000). The resurgence of nativism in California?

The case of Proposition 187 and illegal immigration. Social Science Quarterly. 81(1),

167-179.

Barge-Schaapveld, D. Q. C. M., Nicolson, N. A., van der Hoop, R. G., & DeVries, M. W.

(1995). Changes in daily life experience associated with clinical improvement in

depression. Journal of Affective Disorders, 34, 139-154.

Beck, R. (1997). Immigration-fueled U.S. population growth is "spoiler" in economic,

social and environmental efforts. Population & Environment: A Journal of

Interdisciplinary Studies. 18(5), 483-487.

Bolstein, R., & Hill, J. (1996). Estimating the number of occurrences of a rare event: US Customs violations at border crossings. Proceedings of the section on quality and productivity. Chicago, IL: American Statistical Association.

Bornstein, R. F. (1998). Implicit and self-attributed dependency strivings: Differential relationships to laboratory and field measures of help seeking. Journal of Personality & Social Psychology, 75, 778-787.

Brandstaetter, H., & Gaubats, S. (1997). Befindestagebuch am neuen Arbeitsplatz in differentialspychologischer Sicht./ Time sampling daily experiences of newcomers to the workplace: An individual difference perspective. Zeitschrift fuer Arbeits - und Organisationspsychologie, 41, 18-29.

Clarke, S. G., & Haworth, J. T. (1994). Flow experience in the daily lives of sixth-form college students. British Journal of Psychology, 85, 511-523.

Cohen, J. & Cohen, P. (1983). Applied multiple regression/correlational analysis for the behavioral sciences (2nd ed.). Hillsdale: Erlbaum.

Cook, D. C., & Campbell, D. T. (1979). Quasi-experimentation: Design and analysis

issues for field settings. Boston: Houghton Mifflin.

Csikszentmihalyi, M. & Larson, R. (1987). Validity and reliability of the experience-

   sampling method. Journal of Nervous & Mental Disease: Special Issue: Mental

   disorders in their natural settings: The application of time allocation and experience-

   sampling techniques in psychiatry, 175, 526-536.

Csikszentmihalyi, M. & Larson, R. (1992). Validity and reliability of the experience

   sampling method. In M. W. deVries (Ed.), The experience of psychopathology:

   Investigating mental disorders in their natural settings. (17th ed., pp. 43-57).

   Cambridge, England: Cambridge University Press.

Csikszentmihalyi, M. & Rathunde, K. (1992). The measurement of flow in everyday

   life: Toward a theory of emergent motivation. In J. Jacobs (Ed.), Nebraska Symposium

   on Motivation. (8th ed., pp. 57-97).   Lincoln, NE: University of Nebraska Press.

Dabbs, J. M. Jr., Strong, R., & Milun, R. (1997). Exploring the mind of testosterone: A

   beeper study. Journal of Research in Personality, 31, 577-587.

Edmonston, B., Passel, J. S. & Bean, F. D. (1990). Perceptions and estimates of

   undocumented migration to the United States. In B. Edmonston, J. S. Passel, & F. D.

   Bean (Ed.), Undocumented Migration to the United States: IRCA and the Experience

of the 1980s (pp. 11-32). Santa Monica, CA: Rand Corporation.

Filstead, W. J. (1988). Monitoring the process of recovery using electronic pagers as a

treatment intervention. In M. Galanter (Ed.), Recent developments in alcoholism, 6,

181-191. New York, NY: Plenum Press.

Gauvin, L., Rejeski, W. J. & Norris, J. L. (1996). A naturalistic study of the impact of

acute physical activity on feeling states and affect in women. Health Psychology, 15,

391-397.

Goldman, M. (1999). The Violence Against Women Act: Meeting its goals in protecting

battered immigrant women? Family & Conciliation Courts Review. 37(3), 375-392.

Government Performance and Results Act of 1993 [GPRA], Public Law Number 62, 107

Stat. 285.

Halcrow, A. (1987). Should business alone pay for social progress? Personnel Journal.

66(9), 58-73.

Hill, K. (1985). Illegal Aliens: An Assessment. In D. B. Levine, K. Hill 7 R. Warren (Ed.)

Immigration Statistics: A Story of Neglect, 225-50. Washington, D.C.: National

Academy Press.

Hillbrand, M. & Waite, B. (1992). The social context of anger among violent forensic

patients: An analysis via Experience Sampling Method. Journal of Offender

Rehabilitation, 18, 81-89.

Hnatiuk, S. H. (1991). Experience sampling with elderly persons: An exploration of the

method. International Journal of Aging & Human Development, 33, 45-64.

J. A. Reyes Associates (1977). The Survey Design for a Residential Survey of Illegal

Aliens. Report submitted to the US Immigration and Naturalization Service.

Washington, D.C.: J. A. Reyes Associates.

Kalton, G. (1977). Practical methods for estimating survey sampling errors. Bulletin of

the International Statistical Institute, 47, 495-514.

Kalton, G. (1983). Introduction to survey sampling. Beverly Hills, CA: Sage.

Kish, L. (1965). Survey sampling. New York, NY: John Wiley.

Kish, L. (1995). Space/Time variations in survey estimates. Paper presented at the

Annual Meeting of the American Statistical Association.

Kish, L., Lovejoy, W. & Rackow, P. (1961). A multi-stage probability sample for traffic

surveys. Proceedings of the Social Statistics Section, American Statistical Association

(pp. 227-230).

Koehler, M. J. & Levin, J. R., (1998). Regulated Randomization: Potentially Sharper

Analytical Tool for Multiple-Baseline Design. Psychological Methods, 3, 206-217.

Kubey, R. W. (1986). Television use in everyday life: Coping with unstructured time.

Journal of Communication, 36, 108-123.

Larson, R. & Csikszentmihalyi, M. (1978). Experiential correlates of time alone in

adolescence. Journal of Personality, 46, 677-693.

Lee, M. R. (1995). The relationship between high school seniors   experiences of daily

life activities and levels of depression. Korean Journal of Child Studies, 16, 23-36.

Lesko Associates (1975). Final Report: Basic Data and Guidance Required to

Implement a Major Illegal Alien Study during Fiscal Year 1976. Washington, D.C.:
US Immigration and Naturalization Service.

Lousberg, R., Schmidt, A. J. M., Groenman, N. H., Vendrig, L., et al. (1997).
Validating the MPI-DLV using experience sampling data. Journal of Behavioral
Medicine, 20, 195-206.

McConville, C., & Cooper, C. (1995). Is emotional intensity a general construct?
Personality & Individual Differences, 18, 425-427.

Moneta, G. B., Csikszentmihalyi, M. (1996). The effect of perceived challenges and
skills on the quality of subjective experience. Journal of Personality, 64, 275-310.

Padilla, A. M., (1993). La investigacion sobre Migracion desde una Perspectiva
Psicologica: Reflexiones sobre cuatro trabajos./ Investigation about migration from a
psychological perspective: Reflections on four studies. Revista de Psicologia Social y
Personalidad. 9(2), 95-99.

Prescott, S., Csikszentmihalyi, M., & Graef, R. (1981). Environmental effects on
cognitive and affective states: The experiential time sampling approach. Social
Behavior & Personality, 9, 23-32.

Smart, J. F., & Smart, D. W., (1995). Acculturative stress of Hispanics: Loss and

   challenge. Journal of Counseling & Development. 73(4), 390-396.

Rathunde, K. (1997). Parent-adolescent interaction and optimal experience. Journal of

   Youth & Adolescence, 26, 669-689.

Schiefele, U., & Csikszentmihalyi, M. (1995). Motivation and ability as factors in

   mathematics experience and achievement. Journal for Research in Mathematics

   Education, 26, 163-181.

Shiu, A. T. (1998). The significance of sense of coherence for the perceptions of task

   characteristics and stress during interruptions amongst a sample of public health

   nurses in Hong Kong: Implications for nursing management. Public Health Nursing,

   15, 273-280.

Swarr, A. E., & Richards, M. H. (1996). Longitudinal effects of adolescent girls

   pubertal development, perceptions of pubertal timing, and parental relations on eating

   problems. Developmental Psychology, 32, 636-646.

The Committee on Immigration and Naturalization (1919). Percentage Plans for

   Restriction of Immigration. Washington, D.C.:US Government Printing Office.

US Commission on Immigration Reform (1994). US Immigration Policy: Restoring Credibility. Washington, D.C.:US Government Printing Office.

US Department of Justice Immigration and Naturalization Service (1991). An Immigrant Nation: United States Regulation of Immigration 1798-1991. Washington, D.C.:US Government Printing Office.

US Department of Justice Immigration and Naturalization Service (1992). Immigration Reform and Control Act: Report on the Legalized Alien Population. Washington, D.C.:US Government Printing Office.

AUTHOR S NOTE