

ECONOMIC SCENE

Until a more thorough study is done, it is impossible to know whether private or federal airport screeners perform better.

By ALAN B. KRUEGER

Published: June 24, 2004

IN remarkably short order, the federal government hired 45,000 workers and took over passenger screening at virtually all airports in response to Sept. 11. Would the private sector do a better job?

This question is not just academic: on Nov. 19, the Transportation Security Administration will permit airports to apply to opt out of using federal screeners in lieu of private screeners operating under federal supervision. Initial guidelines for the application and selection process are due this week.

The Aviation and Transportation Security Act, which set up the system, required the new agency to select up to five airports representing different security risks to use private screeners. This pilot program, known as PP5, was intended to provide evidence on whether privately contracted screeners could achieve the efficacy and cost-effectiveness of federal screeners.

In truth, the private screeners are part of a hybrid system because they are overseen by federal employees and highly regulated. For example, private screeners are required to be paid at least as much as federal ones, and to undergo the same training. Still, the PP5 program is the only available evidence on how well private screeners perform compared with federal ones in the current system.

Results of three studies -- by the General Accounting Office, by the inspector general of the Homeland Security Department, and by BearingPoint, a consulting firm hired by the Transportation Security Administration -- comparing performance at privately and federally screened airports were announced at a Congressional hearing in April. Although the detailed results are classified, it is clear that the studies reached the same conclusion: there was no statistical evidence of different performance at federally and privately screened airports.

For example, in its summary report, BearingPoint describes four separate security checks, including covert tests in which inspectors tried to sneak banned items past checkpoints, and computerized Threat Image Projection tests, which superimposed images of threatening objects on X-rays of passengers' luggage. Of the five privately screened airports, only Kansas City performed better than federally screened ones in its class -- and that difference could easily have occurred by chance.

The cost of operating the privately screened airports was about \$1.3 million less, on average, than what the cost was estimated to have been under federal screeners based on traffic flows and other factors, but this figure does not do justice to the uncertainty underlying the calculation. We can be 95 percent confident that the true difference from using private screeners falls between it being \$3 million more costly to \$5.6 million cheaper, a huge range.

There is simply not enough evidence to conclude with any reasonable degree of certainty whether, as a group, private or federal screeners performed better or cost less.

Each study suffered from two unavoidable flaws because of the design of the Pilot Program.

The first flaw is what the econometrician Arthur Goldberger calls "micronumerosity," meaning the sample size is very small. Five privately screened airports are not enough to yield results worth placing much confidence in. If you flip a coin five times, it will be hard, from the results, to tell whether the coin is balanced. More than five airports are required to distinguish consequential differences between federal and private screeners from random patterns in the data.

This problem was compounded because BearingPoint examined performance of each airport against the others in its class, but did not average the results across the five airports. Averaging would have yielded a more precise estimate of the private-public difference.

The second flaw is equally serious. The five privately screened airports were not selected at random. Instead, the agency used 11 criteria -- including the availability of law enforcement officials to the airport and the screening company's resources -- to choose them from among the 19 airports that applied to use private screeners. It is likely that this process yielded airports that gave the private screeners an advantage over the remaining 424 federally screened airports for reasons having nothing to do with whether the screeners were federal or private. This problem is known as selection bias, as the selection and application process stacked the comparison in a particular direction.

Indeed, micronumerosity is probably the only reason that selection bias did not cause the studies to find that private screeners outperformed federal ones, irrespective of the true state of affairs.

Random assignment of some airports to use private screeners and others to a control group with federal screeners would solve the selection problem.

A recent memorandum by the Office of Management and Budget explains the benefit of random assignment: it "ensures, to a high degree of confidence, that there are no systematic differences between the groups in any characteristics (observed and unobserved) except one -- namely, the intervention," which in this case means the intervention to allow airports to opt out of federal screening.

In a noteworthy break from past practices, the budget office advised agencies in its revised Program Assessment Rating Tool, "The most definitive data supporting a program's overall effectiveness would be from a randomized controlled trial, when appropriate and feasible."

The acting administrator of the Transportation Security Administration, David M. Stone, testified that he planned to continue evaluating the effectiveness of private and federal screeners after the opt-out program took effect. So here is a proposal. Suppose 70 airports apply to opt out of federal screening and are deemed eligible. Rather than grant all 70 authority to use private screeners at once, randomly select half of them to opt out and half to continue with federal screeners. This sample is large enough to estimate the cost difference to within \$1 million or so.

Allowing all the applicants to opt out at once could strain the agency's capacity to supervise and monitor the program. Random assignment is not only feasible, it is the fairest way to select the eligible applicants. Airports that are not selected could be phased in later, and serve as a control group in the meantime.

With random assignment of enough airports, there would be no reason to suspect that differences in outcomes between federally and privately screened airports resulted from anything other than whether screening was done by the government or the private sector. Without it, we will be right back where Clark Kent Ervin, the inspector general of Homeland Security, said we are now: without "a sufficient basis at this time to determine conclusively whether the pilot airport screeners performed at a level equal to or greater than that of the federal screeners."