

VI. Constructing the National Price and Purity Indices

Once quarterly estimates for the price of a given drug and quantity level in each city are calculated, the national indices for the drug-quantity level are constructed.¹⁹ There are a number of different methods that could be used to generate this index, depending on the assumptions one is comfortable making. The current report uses a methodology employed in previous reports and constructs the index as a weighted average of the prices predicted for each city. However, new weights are used.

The ideal statistic to weight each city-specific average would perhaps be estimates of the number of drug transactions in each city. Unfortunately, the geographical distribution of transactions in the STRIDE data is not representative of the geographical distribution of actual transactions across the country. And there are no existing data that provide accurate estimates of the extent of drug use in the cities of interest for the current analysis. Thus, other alternatives had to be considered.

The weights used in the previous report were based on emergency-room (ER) mentions of the particular drugs in the Drug Abuse Warning Network (DAWN) data. There are several drawbacks to using these data for weights. First, eight of the 29 cities are not represented in DAWN, so their weights must be based on what the analyst considers to be similar cities. For example, Seattle was used as a proxy for Portland, adjusted for population. Second, there are limitations in the DAWN data that raise some doubt about their appropriateness. For example, the number of drug-specific ER mentions within particular cities is very volatile, causing the implied weight constructed from them to be volatile. For cocaine, the weights for New York range from 6.8 percent to 19.2 percent over the 1981–2000 time period. For the “other” category (i.e., those locations not falling in one of the 29 cities), the weight ranged from 2.9 to 47.0 percent. There are similar examples of extraordinary volatility for the other drugs as well. Finally, it is not entirely clear that cross-sectional differences in the number of hospital-related drug episodes are truly reflective of differences in drug use across cities. For example, based on only DAWN data, cocaine use in St. Louis in 2002 would appear to be greater than that in Washington DC. State-level estimates of cocaine use from the 2001 National Household Survey on Drug Abuse suggest that this is unlikely to be the case.

This analysis uses the population of cities and regions as the weights because population data are readily available for all cities, not just those in which DAWN is estimated, and because population estimates are reliable and universally understood. The estimates come from the Census estimates for 1980, 1990, and 2000 reported in *Statistical Abstract*.²⁰ *Statistical Abstract* population estimates for small cities (e.g., Buffalo) tend to include only the MSA, while larger cities (e.g., Boston) usually include the PMSA and/or the CMSA. Because the CMSAs

¹⁹ We recognize that developing a national average from these relatively sparse and unrepresentative data is not advisable for a number of reasons. We do not, therefore, interpret the numbers calculated as true national averages. Instead, we view them as potential indices, the validity of which will be tested by examining their correlation to external drug-related data obtained through more reliable sources.

²⁰ Department of Treasury (2002), *The Statistical Abstract of the United States*, Government Printing Office, Washington D.C.

incorporate several PMSAs, this analysis uses only estimates reported for the PMSA. The one exception is San Francisco, where the population estimates include both the San Francisco and Oakland PMSAs because we group drug transactions from these PMSAs together in the San Francisco city group. Populations for the Census divisions are calculated by summing the populations for the states in each division and subtracting the populations of the MSAs in a Census division that are already included in the model. This will not be a perfect measure because some MSAs include counties in other states. However, it does represent a reasonable approximation, given that the error is consistently applied in each year.

Using data from the decennial population estimates for each city (and Census division), a linear (constant) growth trend is presumed for all years between 1980 and 1990 and then again for 1990 through 2000 so that city (or region) population estimates can be obtained for every year. The constant growth rate identified for 1990–2000 is also applied to 2001–2003 to generate population estimates for these years. Weights for the cities are constructed by determining in each year the fraction each city’s population constitutes of the total 29 cities’ population. The 29 cities sum up to 29/30 of the total weight in the analysis. The other 1/30 is partitioned across the 9 Census divisions, whose weights were determined by their share of the total population across regions in a given year. More weight is given to the cities than to the regions for several reasons. First, this weighting is consistent with what was done in previous reports and is based on the assumption that drug transactions are more concentrated in larger cities than in small cities or rural areas. Second, it seems reasonable to presume that more useful information can be obtained from multiple observations within the same small geographic location (a city) than from observations obtained from a geographically large location (e.g., a Census region). Actual weights used to weight city-specific averages are given in Table E.1 in Appendix E.