

2018 Annual Synar Report

**Division of Health
Informatics**

August 2019



pennsylvania
DEPARTMENT OF HEALTH

Table of Contents

Introduction.....	3
Statewide Results.....	3
Historical Results.....	3
Results by Region.....	4
Results by Youth Sex.....	5
Results by Youth Age.....	6
Results by Youth Race.....	6
Results by Youth Ethnicity.....	6
Results by Clerk Sex.....	7
Distribution of Outlet Types.....	7
Results by Outlet Type.....	8
Conclusions.....	9
Technical Notes.....	10
Contacts.....	17

Introduction

The Synar survey was established in July 1992 when Congress enacted the Alcohol, Drug Abuse and Mental Health Administration (ADAMHA) Reorganization Act (P.L. 102-321), which established the Substance Abuse and Mental Health Services Administration (SAMHSA). With this act, the Synar Amendment (named for its sponsor, former congressman Mike Synar of Oklahoma) was incorporated to address access to tobacco by youth under 18 years of age. One of the core requirements of this amendment is to conduct random, unannounced inspections of outlets that sell tobacco. Pennsylvania's annual Synar survey is designed to satisfy this federally regulated requirement and is intended to estimate the rate at which outlets sell cigarettes to minors (known as the retail violation rate or RVR).

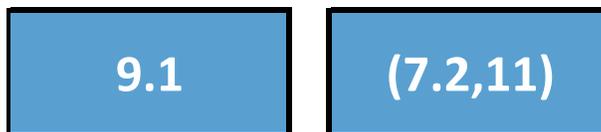
The survey is conducted by youth buyers who attempt to purchase cigarettes from a sample of Pennsylvania cigarette retailers. The outcome of each attempt is recorded and a rate is calculated from the eligible outlets attempted. The 2018 survey was conducted during the summer of 2018.

Statewide Results

In 2018, an estimated 7.2 to 11.0 percent of Pennsylvania cigarette retailers sold cigarettes to minors. The estimate was calculated from the results of the 2018 Synar survey using a 95 percent confidence interval (95% CI) with weighted violation rate (9.1 percent) and standard error (1.0).

Violation rate

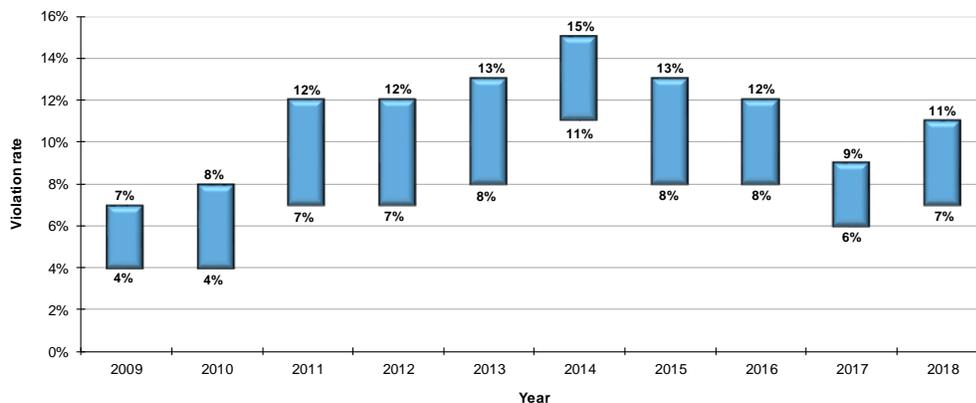
Confidence interval



Historical Results

Pennsylvania has annually conducted the Synar survey since 1996. Significance tests (Rao-Scott Chi-Square) were used to examine the relationships between the 2018 statewide violation rates over the past 10 years (2009-2018). The tests demonstrated that the 2018 violation rate is not statistically different from the 2011, 2012, 2013, 2015, 2016 and 2017 violation rates but is statistically different from the 2009, 2010 and 2014 violation rates.

Graph 1 – Pennsylvania Retail Violation Rates of the Last 10 Years (2009-2018)



Results by Region

Prior to sampling, every eligible outlet on the sample list is placed into one of 10 mutually exclusive and exhaustive geographical regions (**Figure 1**). The geographical regions are the Northcentral (NC), Northeast (NE), Northwest (NW), Southcentral (SC), Southeast (SE) and Southwest (SW) of Pennsylvania, plus the four individual counties: Allegheny (AL), Delaware (DE), Erie (ER) and Philadelphia (PH). Pennsylvania's sampling methodology allows for valid estimates from each region (**Table 1**).

Figure 1 – 2018 Synar Region Map

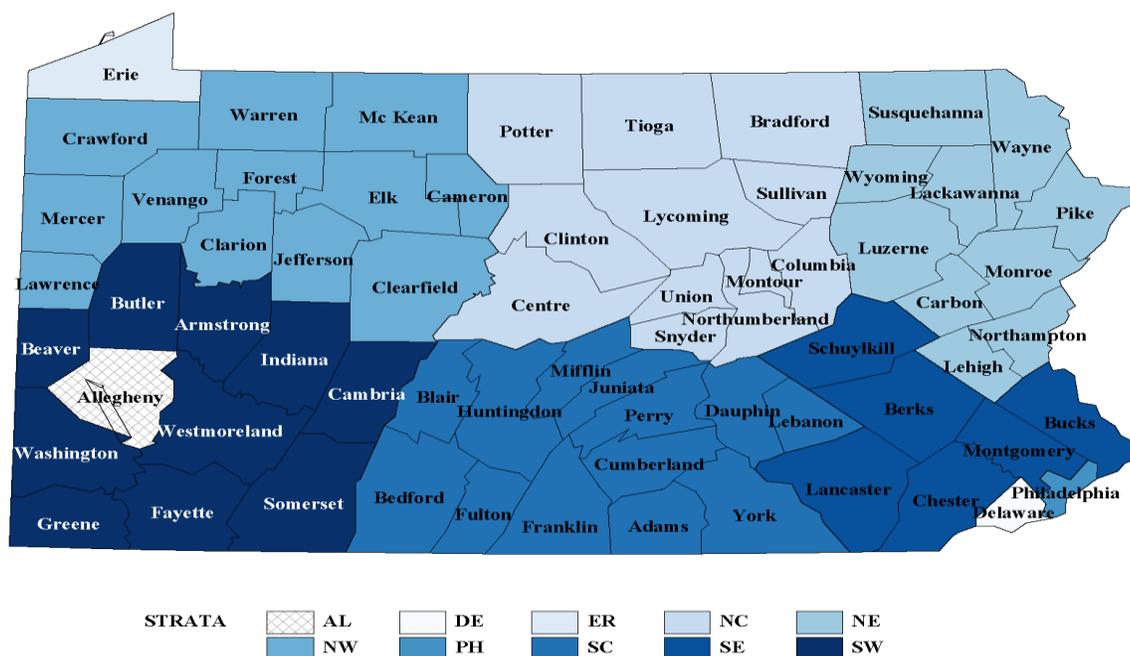


Table 1 – 2018 Synar Region Results

Region	Abbr.	Outlets selected	Outlets completed	Total violations	Weighted rate	Standard error	Lower limit	Upper limit
Statewide	State	1790	1194	89	9.1	1.0	7.2	11.0
Northcentral	NC	119	81	5	6.1	2.9	0.4	11.8
Northeast	NE	236	158	11	7.0	1.8	3.5	10.4
Northwest	NW	116	80	0	0.0	0.0	0.0	3.8
Southcentral	SC	187	137	8	6.1	2.2	1.8	10.4
Southeast	SE	305	204	12	5.9	1.8	2.4	9.4
Southwest	SW	222	144	7	4.9	3.0	0.0	10.8
Allegheny	AL	130	74	0	0.0	0.0	0.0	4.1
Delaware	DE	110	78	3	3.8	2.2	0.0	8.2
Erie	ER	100	68	4	5.9	2.9	0.2	11.5
Philadelphia	PH	265	170	39	22.9	3.2	16.6	29.3

Note 1: Confidence limits ("lower limit," "upper limit") were calculated using the t-distribution with the degrees of freedom (df) determined by subtracting the total strata from the total clusters.

Note 2: The weighted rate takes into account unequal probabilities of selection and non-completions.

Note 3: When "total violations" = 0, the "rule of three (3/n)" is used to calculate upper limit.

Note 4: The rate was marked n/a if "outlets completed" was below 40.

Significance tests (Rao-Scott Chi-Square) and odds ratio calculations were used to examine the relationships between regions with more than 40 completed visits. When total violations is equal to zero, Chi-Square tests cannot be used to test for significance. In this situation, confidence intervals are constructed using the rule of three (3/n), where n is equal to the number of completed outlets. The confidence intervals are used to test for significance by checking for the percentage of overlap.

- ❖ The **Northcentral (NC)** retail violation rate is statistically different from the Philadelphia (PH) rate. Youth buyers are approximately 4.6 times more likely to be sold cigarettes in PH than in the NC.
- ❖ The **Northeast (NE)** retail violation rate is statistically different from the PH rate. Youth buyers are approximately 4.0 times more likely to be sold cigarettes in PH than in the NE.
- ❖ The **Northwest (NW)** retail violation rate is statistically different from the PH rate. An odds ratio could not be calculated due to zero violations in the NW.
- ❖ The **Southcentral (SC)** retail violation rate is statistically different from the PH rate. Youth buyers are approximately 4.6 times more likely to be sold cigarettes in PH than in the SC.
- ❖ The **Southeast (SE)** retail violation rate is statistically different from the PH rate. Youth buyers are approximately 4.8 times more likely to be sold cigarettes in PH than in the SE.
- ❖ The **Southwest (SW)** retail violation rate is statistically different from the PH rate. Youth buyers are approximately 5.8 times more likely to be sold cigarettes in PH than in the SW.
- ❖ The **Allegheny (AL)** retail violation rate is statistically different from the PH rate. An odds ratio could not be calculated due to zero violations in AL.
- ❖ The **Delaware (DE)** retail violation rate is statistically different from the PH rate. Youth buyers are approximately 7.4 times more likely to be sold cigarettes in PH than in DE.
- ❖ The **Erie (ER)** retail violation rate is statistically different from the PH rate. Youth buyers are approximately 4.8 times more likely to be sold cigarettes in PH than in ER.

Results by Youth Sex

Male youth attempted to purchase cigarettes in 645 different outlets, while female youth attempted this in 549 (**Table 2**). Significance tests showed that female youth are approximately 3.7 times more likely to be sold cigarettes than male youth.

Table 2 – Results by Youth Sex

Sex	Total visited	Violations	Weighted rate	Lower limit	Upper limit
Male	645	28	4.2	2.5	5.8
Female	549	61	13.8	10.4	17.1

Note 1: The weighted rate takes into account unequal probabilities of selection and non-completions. It is different than the unweighted rate, which is calculated by dividing the "violations" by the "total visited." The weighted rate should be used at all times.

Note 2: The rate was marked n/a if "total visited" was below 40.

Results by Youth Age

The age of the youth surveyors ranged from 15 to 17 years old (**Table 3**). There was a significant difference between the rate at which outlets sold cigarettes to 15-year-olds and 16-year-olds. A 16-year-old is approximately 5.3 times more likely to be sold cigarettes than a 15-year-old.

Table 3 – Results by Youth Age

Age	Total visited	Violations	Weighted rate	Lower limit	Upper limit
15	280	7	2.3	0.0	4.8
16	908	82	11.2	8.9	13.4
17	6	0	n/a	n/a	n/a

Note 1: The weighted rate takes into account unequal probabilities of selection and non-completions. It is different than the unweighted rate, which is calculated by dividing the "violations" by the "total visited." The weighted rate should be used at all times.

Note 2: The rate was marked n/a if "total visited" was below 40.

Results by Youth Race

There was a significant difference between the rates at which outlets sold cigarettes to white youth and African American/black youth and youth who listed their race as other and African American/black youth. African American/Black youth are approximately 3.0 and 6.6 times more likely to be sold cigarettes than white youth and youth who listed their race as other, respectively.

Table 4 – Results by Youth Race

Race	Total visited	Violations	Weighted rate	Lower limit	Upper limit
White	706	36	5.2	3.6	6.8
Black	365	41	14.2	9.8	18.7
Asian	37	10	n/a	n/a	n/a
Other	86	2	2.5	0.0	6.9

Note 1: The weighted rate takes into account unequal probabilities of selection and non-completions. It is different than the unweighted rate, which is calculated by dividing the "violations" by the "total visited." The weighted rate should be used at all times.

Note 2: The rate was marked n/a if "total visited" was below 40.

Results by Youth Ethnicity

Significance tests showed that non-Hispanic youth are approximately 2.3 times more likely to be sold cigarettes than Hispanic youth.

Table 5 – Results by Youth Ethnicity

Ethnicity	Total visited	Violations	Weighted rate	Lower limit	Upper limit
Hispanic	134	6	4.5	1.2	7.7
Non-Hispanic	1060	83	9.7	7.6	11.8

Note 1: The weighted rate takes into account unequal probabilities of selection and non-completions. It is different than the unweighted rate, which is calculated by dividing the "violations" by the "total visited." The weighted rate should be used at all times.

Note 2: The rate was marked n/a if "total visited" was below 40.

Results by Clerk Sex

Significance tests showed that there was no statistical difference between attempts with a male clerk and a female clerk.

Table 6 – Results by Clerk Gender

Sex	Total visited	Violations	Weighted rate	Lower limit	Upper limit
Male	583	48	10.1	7.3	12.9
Female	608	41	8.2	5.7	10.7

Missing=3

Note: The weighted rate takes into account unequal probabilities of selection and non-completions. It is different than the unweighted rate, which is calculated by dividing the "violations" by the "total visited." The weighted rate should be used at all times.

Distribution of Outlet Types

Cigarettes are sold by a variety of outlets in Pennsylvania. Based on the surveyor's description, each outlet was categorized. The 2018 outlet definitions can be found in the technical notes. The distribution of outlets is shown in **Table 7**. Over 50 percent of the surveyed outlets belong to the convenience/gas and convenience/grocery/no gas categories.

Table 7 – Distribution of Sampled Outlets

Outlet Type	Visited	Percent
Bar/tavern	21	1.8%
Beer distributor	79	6.6%
Convenience/gas	406	34.0%
Convenience/grocery/no gas	278	23.3%
Dollar store	84	7.0%
Pharmacy/drug store	73	6.1%
News outlet	10	0.8%
Restaurant/deli	51	4.3%
Supermarket	109	9.1%
Tobacco	61	5.1%
Other	22	1.8%
TOTAL	1194	100%

Results by Outlet Type

Significance tests were performed among each pair of categorized outlets with at least 40 visits to each outlet type.

- ❖ The retail violation rate for **beer distributor outlets** was statistically different from restaurant/deli outlets. Youth buyers are approximately 5.1 times more likely to be sold cigarettes in restaurant/deli outlets than in beer distributor outlets.
- ❖ The retail violation rate for **convenience/gas outlets** was statistically different from pharmacy/drug store outlets and restaurant/deli outlets. An odds ratio could not be calculated due to zero violations at pharmacy/drug store outlets. However, youth buyers are approximately 3.3 times more likely to be sold cigarettes in restaurant/deli outlets than in convenience/gas outlets.
- ❖ The retail violation rate for **convenience/grocery/no gas outlets** was statistically different from pharmacy/drug store outlets and restaurant/deli outlets. An odds ratio could not be calculated due to zero violations at pharmacy/drug store outlets. However, youth buyers are approximately 2.7 times more likely to be sold cigarettes in restaurant/deli outlets than in convenience/grocery/no gas outlets.
- ❖ The retail violation rate for **dollar store outlets** was statistically different from restaurant/deli outlets. Youth buyers are approximately 6.2 times more likely to be sold cigarettes in restaurant/deli outlets than in dollar store outlets.
- ❖ The retail violation rate for **restaurant/deli outlets** was statistically different from pharmacy/drug store outlets. An odds ratio could not be calculated due to zero violations at pharmacy/drug store outlets.
- ❖ The retail violation rate for **supermarket outlets** was statistically different from convenience/gas outlets, convenience/grocery/no gas outlets, restaurant/deli outlets and tobacco outlets. Youth buyers are approximately 10.4, 12.8, 34.7 and 7.7 times more likely to be sold cigarettes in convenience/gas outlets, convenience/grocery/no gas outlets, restaurant/deli outlets and tobacco outlets, respectively, than in supermarket outlets.
- ❖ The retail violation rate for **tobacco outlets** was statistically different from restaurant/deli outlets. Youth buyers are approximately 4.5 times more likely to be sold cigarettes in restaurant/deli outlets than in tobacco outlets.

Table 8 – Results by Outlet Type

Outlet type	Total visited	Violations	Weighted rate	Lower limit	Upper limit
Bar/tavern	21	1	n/a	n/a	n/a
Beer distributor	79	5	6.3	0.4	12.1
Convenience/gas	406	34	9.4	6.2	12.5
Convenience/grocery/no gas	278	25	11.3	7.0	15.5
Dollar store	84	3	5.2	0.0	11.2
Pharmacy/drug store	73	0	0.0	0.0	4.1
News outlet	10	2	n/a	n/a	n/a
Restaurant/deli	51	12	25.5	12.6	38.5
Supermarket	109	1	1.0	0.0	2.9
Tobacco	61	3	7.1	0.0	15.0
Other	22	3	n/a	n/a	n/a

Note 1: The weighted rate takes into account unequal probabilities of selection and non-completions. It is different than the unweighted rate, which is calculated by dividing the "violations" by the "total visited." The weighted rate should be used at all times.

Note 2: The rate was marked n/a if "total visited" was below 40.

Conclusions

Table 9 – SAMHSA Target Rates vs. Actual Survey Rates

Year	Original target rate	Revised * target rate	Actual survey rate	Actual survey error
1996	Baseline	n/a	50%	± 12%
1997	42%	n/a	30%	± 6%
1998	31%	n/a	31%	± 5%
1999	25%	30%	41%	± 5%
2000	20%	29%	27%	± 4%
2001	20%	25%	28%	3%
2002	20%	20%	15%	3%
2003	20%	20%	11%	2%
2004	20%	20%	7%	2%
2005	20%	20%	8%	2%
2006	20%	20%	7%	2%
2007	20%	20%	7%	2%
2008	20%	20%	5%	2%
2009	20%	20%	6%	2%
2010	20%	20%	6%	2%
2011	20%	20%	10%	2%
2012	20%	20%	10%	2%
2013	20%	20%	11%	2%
2014	20%	20%	13%	2%
2015	20%	20%	11%	2%
2016	20%	20%	10%	2%
2017	20%	20%	7%	2%
2018	20%	20%	9%	2%

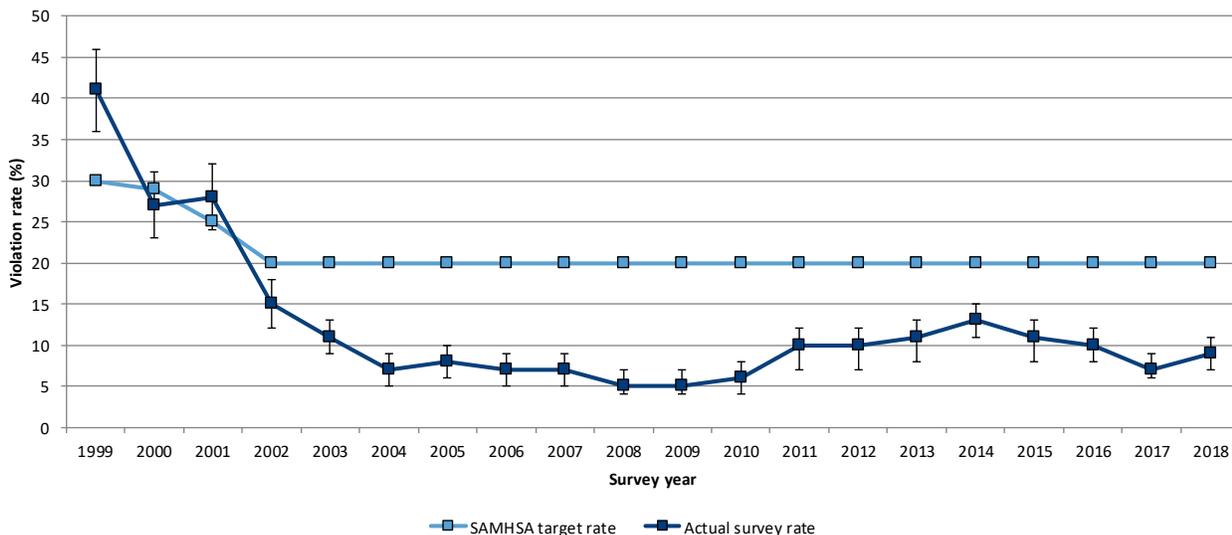
The federal government sets maximum target violation rates for each state, including Pennsylvania. Pennsylvania is expected to be at or below the target rates shown in **Table 9**. The Synar survey only measures the violation rates; it cannot lower them. Violation rates must be lowered through enforcement or other methods. In 1999, Pennsylvania was penalized for exceeding the maximum allowable rate set by SAMHSA. The penalty prompted a massive statewide campaign of enforcement, public awareness and education that continues today. According to the data obtained from the Synar survey, Pennsylvania’s prevention and enforcement efforts appear to be successful. The estimated violation rate dropped significantly after the first few years of the campaign and eventually leveled off well below federal targets. Since 2002, the violation rate has been significantly lower than the federal target rate of 20 percent (**Graph 2**). In 2018, the Philadelphia region sold at a significantly higher rate than the other regions. Restaurant/deli outlets sold at a significantly higher rate than the other outlet types. Allegheny had a significant decrease in their violation rate this year. This should be investigated to see if any techniques could possibly be used in other areas.

*Revised 3/8/00

Note 1: Since 2001, SAMHSA has required a one-sided 95% CI

Note 2: All rates and errors are rounded to the nearest %.

Graph 2 – SAMHSA Target Rates vs. Actual Survey Rates



Technical Notes

Background

The U.S. Department of Health and Human Services clarified the Synar Amendment by issuing the Synar Regulation in January 1996. SAMHSA, an agency of the U.S. Department of Health and Human Services, was chosen to implement the regulation. The Center for Substance Abuse Prevention (CSAP) is an agency of SAMHSA in charge of this regulation. According to the “Synar Regulation Implementation Report,” each state must:

- Have in effect a law prohibiting any manufacturer, retailer or distributor of tobacco products from selling or distributing such products to any individual under the age of 18;
- Enforce such laws in a manner that can reasonably be expected to reduce the illegal sales of tobacco products to individuals under the age of 18;
- Conduct annual, random, unannounced inspections to ensure compliance with the law to be conducted in such a way as to provide a valid probability sample of outlets accessible to youth;
- Develop a strategy and negotiate with SAMHSA a timeframe for achieving an inspection failure rate of less than 20 percent of outlets accessible to youth; and
- Submit an annual report describing in detail the state's survey methodology and activities to enforce its law.

Failure to meet the requirements of the Synar Regulation could result in graduated penalties against a state's Substance Abuse Prevention and Treatment (SAPT) Block Grant, as specified in the statute. The intent of the survey is to assess the effectiveness of Pennsylvania's enforcement programs by measuring the rate at which outlets sell cigarettes to minors. If the rate is high, then enforcement is failing. The survey does not lower the rate; it only measures it.

Pennsylvania Checklist for Meeting SAMHSA Survey Requirements

SAMHSA clarified the Synar Regulation and provided specific survey requirements in the “Synar Regulation: Sample Design Guidance (May 2003).” Below is a list of these requirements and how Pennsylvania fared in 2018.



Obtain approval from SAMHSA in writing for any changes in sampling methodology prior to implementation of the Synar survey.

There were no methodology changes for the 2018 survey.



Develop a sampling frame that includes both over-the-counter and vending machine locations accessible to youth.

Pennsylvania only samples over-the-counter locations. Act 2002-112 restricted the placement of vending machines to locations inaccessible to minors. Since vending machines could only be located in areas inaccessible to minors, SAMHSA approved the exclusion of vending machines from the Pennsylvania survey.



Develop a sampling frame that includes, at a minimum, 80 percent of the tobacco outlets in the state.

Pennsylvania develops the sampling frame from the Department of Revenue's Cigarette License File (CLF). Pennsylvania requires a license to sell cigarettes, and the CLF contains a complete list of all locations licensed to sell cigarettes. The latest coverage

survey, conducted in 2018, estimated that Pennsylvania's sampling frame included 98.4 percent of the tobacco outlets in the state.



Select a sample of outlets to inspect that is representative of the geographic distribution of all tobacco outlets accessible to youth in the state. Decide whether to use clustering or stratification or both.

Pennsylvania used both clustering and stratification, ensuring that the methodology was truly representative of the geographic distribution. The sample size for each stratum roughly mirrored the population distribution of outlets.



Design a sampling methodology and implementation plan that are based on sound survey sampling methodology. Choose a sample design and decide on a random sampling method for each stage of sampling.

Pennsylvania exacted a methodology and implementation plan with a valid probability sample for which the probability of selection for each outlet was non-zero. A two-stage sampling design was used, and the first stage selected primary sampling units (PSU) from within each stratum using probability proportionate to size (PPS) sampling technique. Stage two involved randomly selecting a pre-determined number of outlets from each of the sampled PSUs.



Estimate the original sample size before implementing the Synar survey. Base the estimate of the original sample size on the results of calculations of the minimum sample size needed to meet SAMHSA's precision requirement, plus extra sample needed to account for the expected completion rate and the expected accuracy rate.

Pennsylvania calculated the sample size by first calculating the effective sample size. According to CSAP requirements, the width (w) of the upper limit of the confidence interval (c.i.) must be less than or equal to 3 percent.

Using the equation for the upper limit of a 95% CI of the sample mean \bar{x} gives

$$(S1) \\ \bar{x} + w$$

applying the CSAP requirement for w gives

$$(S2) \\ w \leq 3$$

Where w is defined as

$$(S3) \\ w = z(s.e.)$$

Substituting S3 into S2

$$(S4) \\ z(s.e.) \leq 3$$

Where z is the critical value of the standard normal distribution for a one-sided 95% CI and $s.e.$ is the standard error or standard deviation estimated from the sample data. Substituting 1.645 for z and solving equation S4 for $s.e.$ gives

$$s.e. \leq \frac{3}{1.645} \leq 1.82$$

Therefore, the $s.e.$ must be less than or equal to 1.82 to maintain a width of 3 percent or less for a right-sided 95% CI.

Ignoring the finite population correction, the $s.e.$ is defined as,

$$(S5) \\ s.e. = \frac{\sqrt{p(1-p)}}{\sqrt{n_e}}$$

Substituting S5 into S3 gives

$$w = z \left(\frac{\sqrt{p(1-p)}}{\sqrt{n_e}} \right)$$

Solving for n_e gives the equation for the effective sample size

$$n_e = \left(\frac{z}{w} \right)^2 p(1-p)$$

Where $z = 1.645$, $w = 0.03$ (both z and w are based on 95% CI with tolerance of 3 percent) and $p = 3$ percent over the target rate (20 percent + 3 percent = 23 percent).

Next, the target sample size was calculated using the equation:

$$n_t = \text{Deff}_h \times n_e$$

Where Deff_h is the highest design effect from historical Synar surveys of a similar design.

Finally, the original sample size is calculated using the combined equation:

$$n_o = \frac{n_t}{r_l r_c} + n_A + n_S$$

r_l = lowest eligibility rate of historical Synar surveys of similar design.

r_c = lowest completion rate of historical Synar surveys of similar design or 80 percent (whichever is lower).

n_A = sample added or subtracted needed to fit the clustered sample design.

n_S = supplemental sample.

n_A is the number of sample added or subtracted to guarantee that our precision goals are met and the sample size fits the design. The size of n_A is estimated after reviewing output created by a Statistical Analysis System (SAS) program designed to simulate survey outcomes with varying designs. n_S is the number of supplemental sample allocated to the clustered areas due to sample attrition. Supplemental sample is issued if a cluster does not obtain the minimum number of completions allowed per cluster.



SAMHSA requires the results to be reported with a right-sided 95% CI. The precision requirement for the estimate of the violation rate must have the right-side limit within 0.03 or 3 percentage points from the violation rate estimate. Using the normal distribution, the requirement can be translated into the statement that 1.645 times the (s.e.) of the estimate be within 0.03. That is, $1.645 \times (s.e.) \leq 0.03$ or $(s.e.) \leq \frac{0.03}{1.645} = 0.0182$

Pennsylvania is required to report the results of the survey within the Annual Synar Report (ASR). The c.i. reported in the ASR is different from what is reported in this document because of rounding error and the different methods of calculation. The ASR requires a one-sided c.i. that assumes a normal distribution. This document employs a two-sided c.i. assuming a t-distribution. Confidence intervals may be either one-sided or two-sided, although a two-sided c.i. is most commonly used. In the case of the ASR, where the objective is to determine whether the retailer violation rate is equal to or less than the state target rate (20 percent), the right-sided c.i. is used by the federal government, rather than the two-sided interval.

The right-sided 95% CI is always bounded by zero on the left. The right-side limit is given by (violation rate estimate) + (critical value for a normal one-sided 95% CI) \times (standard error of the estimate). The two-sided 95% CI used in this report and most publications is calculated by (violation rate estimate) + (critical value for a t-distribution two-sided 95% CI) \times (standard error of the estimate). The critical value for a normal one-sided 95% CI is always 1.645, and critical value for a t-distribution two-sided 95% CI approaches 1.96 as the sample size increases. Results might differ slightly from results in other portions of the report due to rounding differences – in particular, differences in the number of significant figures used in the calculations. For example, if the violation rate = 7.3, standard error = 0.9 and there is a sufficiently large sample size, then the confidence intervals for the two methods are calculated as follows:

95% one-sided CI (Normal-dist)

$$9.1 + (1.645 \times 1.0) = [0, 10.7]$$

95% two-sided CI (t-dist)

$$9.1 \pm (1.96 \times 1.0) = [7.0, 11.1]$$

The precision level was achieved for the 2018 survey. The survey had a standard error of 0.0096, which is less than the required 0.0182.



Determine a method of selecting additional outlets to inspect should it not be possible to reach the required minimum number of completed inspections due to sample attrition.

Pennsylvania uses an approved supplemental sample system; additional outlets are issued when a minimum number of outlets aren't completed per cluster and random area. The outlets are randomly selected from the remaining outlets in each survey area. There were 168 supplemental sample outlets issued in 2018.



Obtain a completion rate of 90 percent or better.

Pennsylvania had a 99.6 percent completion rate in 2018.



Record the actual steps of the survey process in the field and keep records of all sources of sample attrition in the field.

Pennsylvania reported the actual steps of the survey process in the Annual Synar Report and kept all records.



Incorporate the complexity of the sample design as a factor when analyzing the survey results.

Pennsylvania used the Taylor series (linearization) method to estimate sampling errors of estimators based on complex sample designs. This method takes into account the variances among PSUs.



Weight the results of the Synar survey to account for unequal probabilities of selection, differences in percentages of eligible outlets between strata or clusters, and other deviations from the intended design.

A base weight is calculated for each outlet using the inverse of the probability of selection for each outlet divided by the total eligible outlets in the stratum (ELIGN). The base weight gives each sampled outlet a weight such that it sums to the number of eligible outlets in the state.

First, the probability of selection was calculated. In a complex design, such as this, the overall probability of selecting an outlet is the product of each stage's probability of selection.

PROBCL = Probability of selecting a cluster

PROBOUT = Probability of selecting an outlet within the cluster

PROBST = Probability of selection for each outlet in the stratum

NCLUST = Number of clusters in the stratum

CPS = Cluster population size

ELIGN = Eligible stratum population size

SAMPSIZE = Sample size of the cluster

SAMPOBS = The number of completed and eligible sample per cluster

$$PROBCL = (NCLUST) \times \left(\frac{CPS}{ELIGN} \right)$$

$$PROBOUT = \left(\frac{SAMPsize}{CPS} \right)$$

$$PROBST = (PROBCL) \times (PROBOUT)$$

$$PROBST = (NCLUST) \times \left(\frac{CPS}{ELIGN} \right) \times \left(\frac{SAMPsize}{CPS} \right)$$

$$PROBST = (NCLUST) \times \left(\frac{SAMPsize}{ELIGN} \right)$$

The base weight (*BASEWGT*) is the inverse of the probability of selection.

$$BASEWGT = \frac{1}{\left(\frac{(NCLUST) \times (SAMPsize)}{ELIGN} \right)}$$

$$BASEWGT = \frac{ELIGN}{(NCLUST) \times (SAMPsize)}$$

The final weight adjusts the base weight for non-completion. The final weight gives each completed eligible outlet a weight such that it sums to the number of eligible outlets in the state. The final weight will always be greater than the base weight, unless all sampled outlets are completed and eligible.

$$FINALWGT = (BASEWGT) \times \left(\frac{SAMPsize}{SAMPOBS} \right)$$



Meet Synar regulation reporting requirements for the survey sampling methodology when completing the Annual Synar Report.

Pennsylvania met all methodology reporting requirements.

Survey Design

The population is defined as Pennsylvania outlets that sell cigarettes and are accessible to minors. The survey uses a sampling frame created from the Department of Revenue’s Cigarette License File, which contains the name and address of every outlet that purchased a license to sell cigarettes in the state.

The survey employs a stratified and clustered design (**Figure 1**) where every eligible outlet location on the sampling frame is grouped into 10 mutually exclusive and exhaustive geographical strata consisting of the Northcentral Health District (NC), Northeast Health District (NE), Northwest Health District (NW), Southcentral Health District (SC), Southeast Health District (SE), Southwest Health District (SW), Allegheny (AL), Delaware (DE), Erie (ER) and Philadelphia (PH) counties.

The outlets within the six “District” strata (NC, NE, NW, SC, SE and SW) are grouped into geographic clusters of adjacent zip codes. Clusters are selected using probability proportional to size sampling, and a predetermined number of outlets within the cluster are selected. The outlets within the four “random” strata (AL, DE, ER and PH) are not clustered but are selected using a simple random selection process.

Survey Procedures

The survey is the result of the combined efforts of the Bureau of Informatics and Information Technology (BIIT), the Bureau of Health Promotion and Risk Reduction (BHPRR), the Department of Drug and Alcohol Programs (DDAP), private contractors, and youth from across the state. Survey teams consisting of adult supervisors and youth between the ages of 15 and 16 are provided with a list of sampled outlets to visit. The youth enter the outlets, attempt to purchase cigarettes and record the outcome of the attempts. The survey was conducted from July 1, 2018, to Aug. 23, 2018.

Outlet Definitions (2018)

Bar/tavern – The primary purpose of a bar or tavern is to sell alcoholic beverages for onsite consumption. Some bars or taverns provide snacks or entire meals, and some don't.

Beer distributor – A beer distributor sells beer by the case. It may provide either walk-in or drive-thru service or both. It does not allow onsite consumption, and it may also sell other items such as soda or snacks.

Convenience/gas – This is a store selling a limited variety of food and an assortment of convenience items for the house and vehicle. It is part of a regional or national chain of stores and has multiple outlets in Pennsylvania. The store is usually open long hours for the convenience of customers. Some stores have a self-service microwave oven for heating purchased food. It sells gasoline and over-the-counter drugs or provides take-out foods, but its major sales items are food. Here is a list of the more popular outlets that should be placed in this category: 7-Eleven, AmPm, A-Plus, Circle K, Cogo's, Convenient Food Marts, Crossroads, E-Z mart, GetGo, Git N Go, Go-Mart, Kwik Fill, QuickStop, Rutters, Sheetz, Stop-N-Go, Stuckey's, Town and Country Food Stores, Turkey Hill, Uni-Mart and Wawa.

Convenience/grocery/no gas – This is a store that sells a limited variety of food and an assortment of convenience items for the house and vehicle but is independently owned; it does not belong to a regional or national chain. These outlets are sometimes referred to as country stores, corner stores, general stores, local markets, mini markets, convenience stores, grocery stores or "Mom and Pop" establishments. These outlets may or may not be open for long hours. Outdoor produce markets are included in this category.

Pharmacy/drug store – Drug stores sell prescription and over-the-counter medications. They may be part of a national or regional chain of outlets or owned and operated by an independent pharmacist. They may sell other items, but their major image is as a pharmacy.

News outlet – News outlets sell newspapers and magazines. They usually sell other items like candy, but their main purpose is selling newspapers and magazines. Include outdoor newsstands in this category.

Restaurant/deli – The primary purpose of an eat-in restaurant is the preparation and service of food for onsite consumption. It may offer alcoholic beverages and meals for take-out, but its major focus is food service for onsite consumption. Diners are included in this category. This kind of shop sells cooked or prepared foods ready for consumption, such as cheeses, cold cooked meats, sandwiches and salads. Most delicatessens have a sandwich menu, most of which are made to order behind the counter at the time of sale. In addition to made-to-order sandwiches, many delicatessens offer made-to-order green salads. Equally common is a selection of pre-made pasta, potato, chicken, tuna, shrimp or other variety of salads. Delicatessens also offer a variety of beverages, chips and snacks. Take-out establishments offer prepared foods primarily for consumption off the premises. Some may

not offer entire meals. Examples include pizza/sub shops, Chinese take-out, bagel shops and donut/coffee shops.

Supermarket – Supermarkets sell food and household items in a large facility. It is a departmentalized self-service store offering a wide variety of food and household merchandise. It is larger in size and has a wider selection than a traditional grocery store. The supermarket typically has meat, produce, dairy and baked goods departments. Along with the items for sale in the various departments, additional items for sale may include canned and packaged goods, as well as various nonfood items, such as household cleaners, pharmacy products and pet supplies. This category will include the major chains such as ACME, Food Lion, Giant, Giant Eagle, Karns, Save-A-Lot, Shop 'n Save and Weis.

Tobacco – The tobacco category covers all tobacco outlets, cigarette outlets and cigar shops. These outlets sell tobacco (cigarettes, cigars and/or smokeless tobacco) as their main product.

Other – This is a last-resort category. Use this category for locations that do not fit in any of the above categories. Describe the type of outlet in the space provided next to the “other” category. Category examples include check cashing outlets, laundromats, hotels, motels, record outlets, clothing outlets, book stores, hardware stores, video stores, campgrounds, prisons, bowling lanes (not in the bar), fire halls, The Gateway Clipper (boat), train stations, auto auctions, auto repair or service stations, bait shops, car dealerships, etc.

Contacts

Stephen Muccioli (statistician supervisor)
Nathanael Tinik (statistician)
Pennsylvania Department of Health
Bureau of Informatics and Information Technology
Division of Health Informatics
Statistical Services and Survey Analysis Section
Email (Stephen Muccioli): smuccioli@pa.gov
Email (Nathanael Tinik): ntinik@pa.gov