Published in final edited form as:

J Health Commun. 2011 September ; 16(8): 820–830. doi:10.1080/10810730.2011.561912.

The Effect of Mystery Shopper Reports on Age Verification for Tobacco Purchases

BRAD S. KREVOR,

Responsible Retailing Forum, Inc., Waltham, Massachusetts, USA

WILLIAM R. PONICKI,

Prevention Research Center, Berkeley, California, USA

JOEL W. GRUBE, and

Prevention Research Center, Berkeley, California, USA

WILLIAM DeJONG

Boston University School of Public Health, Boston, Massachusetts, USA

Abstract

Mystery shops (MS) involving attempted tobacco purchases by young buyers have been employed to monitor retail stores' performance in refusing underage sales. Anecdotal evidence suggests that MS visits with immediate feedback to store personnel can improve age verification. This study investigated the impact of monthly and twice-monthly MS reports on age verification. Forty-five Walgreens stores were each visited 20 times by mystery shoppers. The stores were randomly assigned to one of three conditions. Control group stores received no feedback, whereas two treatment groups received feedback communications every visit (twice monthly) or every second visit (monthly) after baseline. Logit regression models tested whether each treatment group improved verification rates relative to the control group. Post-baseline verification rates were higher in both treatment groups than in the control group, but only the stores receiving monthly communications had a significantly greater improvement than control group stores. Verification rates increased significantly during the study period for all three groups, with delayed improvement among control group stores. Communication between managers regarding the MS program may account for the delayed age-verification improvements observed in the control group stores. Encouraging inter-store communication might extend the benefits of MS programs beyond those stores that receive this intervention.

Preventing the illegal sale of tobacco products to minors is a central objective of tobacco control policy (DiFranza & Dussault, 2005; Edwards, Brown, Hodgson, Kyle, Reed, & Wallace, 1999; Institute of Medicine, 1994; US Department of Health and Human Services, 2000). The principal strategy for achieving this objective has been enforcement. The federal Synar Amendment—passed in 1992 (Cummings, Hyland, Perla, & Giovino, 2003; Public Health Service Act, 1992), with final regulations issued in 1996 (Federal Register, 1996)—requires states and territories to pass and enforce tobacco sales-to-minors laws. As a result of enforcement activities mandated by the Synar regulations, tobacco retailer compliance rates rose from 59.9% in FFY 1997 to 89.5% in FFY 2007 (US Department of Health and Human Services, 2008). *Best Practices for Comprehensive Tobacco Control*, a blueprint developed by the Centers for Disease Control and Prevention (CDC), states that enforcement of

tobacco sales-to-minors laws and regulations is an essential component of comprehensive tobacco control programs (US Department of Health and Human Services, 1999).

The tobacco control community has debated whether enforcement reduces underage tobacco use (Cummings et al., 2003; Rigotti, DiFranza, Chang, Tisdale, Kemp, & Singer, 1997) or merely causes underage tobacco users to switch to social sources (Jones, Sharp, Husten, & Crossett, 2002). Some researchers and advocates have argued that enforcement of sales-to-minor laws diverts scarce tobacco control resources from other strategies (e.g., excise tax increases and smoking bans) whose effectiveness have been demonstrated (Ling, Landman, & Glantz, 2002). In fact, reductions in underage tobacco use have been detected in some communities that achieved high levels of compliance with tobacco sales-to-minors laws (DiFranza, Carlson, & Caisse, 1992; Forster, Murray, Wolfson, Blaine, Wagenaar, & Hennrikus, 1998; Jason, Billows, Schnopp-Wyatt, & King, 1996; Jason, Ji, Anes, Brown, & Birkhead, 1991; Jason, Pokorny, & Schoeny, 2003). Moreover, recent literature reviews indicate that strong and comprehensive tobacco control policies, including compliance checks, are associated with decreased youth smoking (Botello-Harbaum, Haynie, Iannotti, Wang, Gase, & Simons-Morton, 2009; Richardson et al., 2009). Efforts to reduce youth access to tobacco are perceived by Americans as among the most important issues facing tobacco control.

The application of consumer protection statutes is an alternative method for boosting compliance with underage sales laws. In all, 43 state attorneys general have executed Assurances of Voluntary Compliance (“AVC”) in which national retail chains commit to making changes in hiring, training, point-of-sales practices, and supervision to improve staff performance in checking IDs and refusing underage sales of tobacco products (Krevor, Lieberman, & Gerlach, 2002). Mystery shops (MS) involving attempted tobacco purchases by young, legal-age customers are often employed to assess progress in age-verification performance at the store- and chain-level.

There is growing recognition that MS can also be used as a point of intervention. One AVC signatory chain, ExxonMobil, has employed MS results as the basis for assigning additional staff training and more intensive monitoring. Another AVC signatory chain, 7-Eleven, also recently adopted the practice of having mystery shoppers provide immediate, real-time feedback to clerks and managers. Similar community-based programs have been implemented to reduce tobacco sales to minors. For example, “Reward and Reminder” provides immediate feedback to clerks, giving either a reward for refusing to sell cigarettes without proper ID or a reminder of the law to those who sell (Biglan, Ary, Smolkowski, Duncan, & Black, 2000).

The immediacy of the communicated feedback is one of the key elements of the mystery shop procedure. All types of retailers use mystery shoppers (sometimes called “secret shoppers”) to test how well staff interact with customers. Detailed reports are typically submitted to store managers at a later time (Wilson, 1998). In contrast, with a focus on preventing illegal sales to minors—where the underlying concern is not customer service or maximizing sales, but compliance with the law—mystery shop vendors provide immediate, real-time feedback and then send a follow-up report. Expressed in terms of McGuire’s (1989) communication/persuasion matrix, providing immediate MS reports increases the likelihood that managers will attend to, comprehend, accept, and remember a mystery shop communication and thus act to improve clerk performance.

Although anecdotal evidence from mystery shop vendors suggests that stores providing MS feedback on a regular basis show improved age-verification rates, no controlled studies have investigated whether MS reports increase the likelihood that clerks will check ID for tobacco

sales. Moreover, no studies have investigated the effect of MS communications on age-verification conduct when provided at various frequencies. The current study was conducted with the support of Walgreens, the first AVC signatory chain. The study aims were (1) to test whether stores receiving MS reports exhibit improved age-verification performance relative to control stores not receiving feedback, and (2) to investigate whether more frequent MS reports lead to greater improvements in age-verification conduct than do less frequent reports.

Method

Procedures

The study involved 45 Walgreens stores, with 15 randomly selected from each of 3 U.S. cities: Houston, Milwaukee, and St. Louis. Five stores from each community were assigned randomly to each of three study conditions: control, monthly MS feedback, or twice monthly MS feedback.

Mystery shoppers were assigned to visit each store twice per month between June 1, 2007 and March 31, 2008. The first five MS attempts served as a baseline to establish pre-treatment sales rates. Because the first mystery shopper report given to both treatment groups occurred *immediately following* the fifth purchase attempt, that attempt is evaluated as a part of the baseline. Thereafter, stores assigned to the first treatment group (N = 15) received real-time MS reports at the end of every attempted tobacco purchase. Stores in the second treatment group (N = 15) received real-time feedback at the conclusion of every second MS purchase attempt, in effect providing a less intense MS reporting schedule. Stores in the control group (N = 15) did not receive feedback after any MS visits.

Mystery shoppers

The project employed the BARS Program of Lakewood, CO to identify and train the mystery shoppers, all men and women age 20 to 22. Underage shoppers were not used to ensure that attempted purchases posed no legal risk to Walgreens, its clerks, or the mystery shoppers.

Visits were scheduled at varying times of day and days of the week. With each visit, the mystery shopper sought to buy a randomly chosen brand of cigarettes from a list of 16 major brands. If the clerk asked the shopper's age, the shopper answered truthfully. If asked for an ID, the shopper claimed not to have one with him or her at that time. Visits in which the clerk refused to sell tobacco without presentation of valid ID were recorded as "Pass," while visits in which the clerk was willing to sell were coded as "Fail."

Mystery shopper reports

Some MS visits were designed for observational purposes only, and these ended immediately upon either a refusal to sell without ID or completion of the transaction; no MS report was issued for these visits. Other MS visits were designed to test the effects of MS reports upon subsequent age verification. For these visits, immediately after the clerk refused or indicated a willingness to sell without ID, the mystery shopper identified himself or herself and asked that the manager on duty join him or her and the clerk. The mystery shopper then provided a Green Card, indicating the clerk's refusal to sell without ID, or a Red Card, indicating that the clerk did not ask for ID or was willing to make the sale without ID being provided. Mystery shoppers were instructed to terminate the visit if they recognized any of the on-duty clerks from a prior visit to the store; such visits were undertaken by another mystery shopper at a later time.

Data

The data collection plan called for 20 MS visits to each of the 45 stores, providing 900 observations. In practice, some store observations had to be excluded from analysis. For example, in some cases the recorded name of the sales clerk could not be matched later to any of that store's employees; these visits were discarded due to the possibility that the wrong store was visited. Such exclusions resulted in a final sample size of 859 store visits, with each store having at least 16 analyzable purchase attempts.

Analyses

The primary analyses were a series of logit regressions. The outcome measure in each model was coded as a dichotomous variable for each purchase attempt (pass = 1; fail = 0).

Treatment condition was coded two ways. In order to test for overall effects of MS reports, Models 1 and 3 used a single dummy variable to indicate post-baseline visits to stores in the two treatment groups, whether monthly or twice-monthly (post-baseline treatment visits = 1, otherwise = 0). The coefficient for this dummy variable indicates the estimated change in age verification resulting from stores receiving MS reports, regardless of frequency, relative to the control group stores receiving no MS reports. Models 2 and 4 addressed differential effects of frequency of MS communications. To this end, two dummy variables, one for monthly MS reports and the other for twice-monthly MS reports were created.

Control variables included the mystery shopper's age and gender. Controls for time of day and cigarette brand requested were evaluated in preliminary analyses but subsequently dropped because they were not significantly associated with the outcome measure.

In each case, we tested two model specifications. Models 1 and 2 included dummy variables for stores in Houston and in St. Louis to control difference in age-verification propensity between the respective city and Milwaukee (the reference category). Models 3 and 4 employed a fixed-effects approach, replacing the city dummy variables with 45 store-specific intercepts. The fixed-effects model is more conservative, as it effectively controls for all unmeasured factors that might affect age verification at a given store (e.g., characteristics of each store's employees and management team, city and state laws and enforcement activities to reduce tobacco sales to minors, and other city and state tobacco control activities). Failure to account for such differences could bias regression coefficients and standard-error estimates (Greene, 2003). In effect, the fixed-effects models ensure that the treatment coefficients measure the increase in age verification for each store relative to its own baseline level. Alternative models treating these underlying store differences as random effects provided similar results to models with store fixed effects.

All logit models also included a linear time variable to control for any underlying trend in age-verification rates operating across all stores in the sample. The trend variable was calculated in days, ranging from 1 for the earliest mystery shopper visit (June 5, 2007) to 298 for the last store observation (March 28, 2008). Similar results were obtained in alternative analyses employing either a quadratic time trend or month-specific dummy variables.

Results

Table 1 presents descriptive statistics for all variables used in the analysis. Store personnel appropriately requested ID and refused a sale in 75.9% of the 859 store visits made during the study. Age-verification rates by experimental group were as follows: control: 68.9%; monthly MS reports: 78.7%; and twice-monthly MS reports: 80.1%. As expected under the study design, approximately one-third of all store observations took place in each city.

Table 2 presents the results for the four logit regression analyses. Models 1 and 2 were the analyses with city dummy variables, but no store-specific fixed effects. Model 1 assumed that MS reports, whether communicated monthly or twice monthly, have identical effects. The results of this analysis indicated that, during the post-baseline period, stores receiving feedback either once or twice monthly were significantly more likely to request age identification than were control group stores. Model 2 estimated separately the effect of each MS report frequency. This analysis indicated that both the monthly and twice-monthly treatment group stores significantly raised their age-verification rates relative to those for the control group stores. The stores receiving feedback once per month exhibited a slightly larger effect than did the stores receiving twice-monthly MS reports, but this difference was not significant.

Models 3 and 4 took a more conservative approach and included store-specific fixed effects. Model 3 suggested that receiving MS reports at either frequency level had a positive, but not quite significant impact on age-verification rates ($p < .09$). Model 4 indicated that stores receiving monthly MS reports significantly improved age-verification rates, but did not show a significant effect for receiving twice-monthly reports.

All four regression models indicated that age verification was significantly less likely with older mystery shoppers, but that verification rates were not significantly associated with the mystery shopper's gender. Significant time trends in all four models suggest that, even after accounting for the impact of the MS reports, stores overall became increasingly more likely to verify the mystery shopper's age over the course of the study. Finally, the city dummy variables in Models 1 and 2 suggested that Houston had significantly lower age-verification performance than did Milwaukee, even after controlling for other predictors.

Figure 1 provides a graphical presentation of changes in average age-verification rates for each experimental group over the course of the study. Whereas the logit regressions were based on separate observations for all 859 store visits, the age-verification rates shown in this figure are averaged by quarter in order to better show long-term patterns while minimizing the effects of short-term random variation. Quarter 1 corresponds to the baseline period. All observations during this quarter occurred before the treatment stores received their first MS reports. The other three quarters represent the post-baseline period after stores in the two treatment groups began receiving feedback.

There are four things to note in Figure 1. First, baseline age-verification rates were higher among stores assigned to the twice-monthly MS report group (66.7%) than among stores assigned to receive monthly MS reports (56.3%) or no reports (54.9%). That is, despite random assignment of stores to experimental conditions, the stores in the twice-monthly treatment group performed much better at baseline than did the other two groups.

Second, both MS report treatment groups exhibited increases in post-baseline age verification. On average, during the second and third quarters of the study, both treatment groups verified mystery shopper age at rates 14 to 20 percentage points higher than for the control group.

Third, providing MS reports had large immediate effects, sharply increasing post-baseline age-verification rates by the second quarter, especially in the monthly report group. The effects of MS reports on age verification became less pronounced thereafter, leveling out between the third and fourth quarters in the monthly MS report group and actually declining by 5 percentage points in the twice-monthly MS report group. In contrast, the control group stores saw no immediate increases in age-verification rates, but did show gradual improvement over time.

Fourth, age-verification rates for stores in all three groups increased significantly over time. Overall, age-verification rates improved by an average of 2.2 percentage points per month during the study period. Moreover, by the final study quarter, the age-verification rates of the three experimental groups had converged. This resulted from the slow but steady improvement among the control group stores, whereas the age-verification rates for the monthly MS report group flattened in the fourth quarter and those for the twice-monthly MS report group decreased slightly.

Discussion

The aims of this study were to ascertain (1) whether immediate mystery shopper feedback has significant effects on age verification for tobacco purchases, and (2) whether the size of this effect varies with the frequency of those communications. Age-verification rates for both treatment groups were significantly higher after the initiation of MS reports than during the baseline period. This simple effect is confounded, however, by evidence that control group stores eventually experienced significant post-baseline increases in age verification despite never receiving any MS reports.

The first study aim was investigated using logit regression, as presented in Models 1 and 3 in Table 2. These analyses established whether receiving MS reports, regardless of frequency, led to higher age-verification rates. With Model 1, without store-specific fixed effects, the estimated effect for receiving MS reports was positive and statistically significant. With Model 3, a more conservative approach with store-specific fixed effects, this effect was only marginally significant ($p < .09$).

Models 2 and 4 addressed the second study aim by testing whether monthly versus twice-monthly MS feedback had differing impacts on age-verification performance. Model 2, without controls for store-specific effects, indicated that post-baseline age-verification rates were higher for both treatment groups compared with the controls. No difference was found between the two treatment groups. Model 4, the more-conservative fixed-effects analysis, indicated a significant improvement in age verification only among stores receiving monthly MS reports, which improved their post-baseline rates by about 8 percentage points more than did the control group stores. In contrast, stores receiving the twice-monthly MS reports did not improve their overall age-verification rates significantly more than did the control group stores. In sum, neither of these models suggested that the MS-report effect increases with the frequency of communicated feedback.

As would be expected, age-verification rates had a significant negative association with customer age: a 22-year-old mystery shopper was almost 10% less likely to be asked for proof of age than was a 20-year-old despite the fact that Walgreen's policy requires verification for any buyer appearing to be younger than 40. Shopper's gender did not affect age-verification rates. The models without store-specific fixed effects also indicated some differences in age-verification between the three study cities, with Houston having significantly lower rates than Milwaukee.

There are two unexpected findings that deserve special consideration. First, Models 1 and 2, without store-specific fixed effects, found significant improvements for both MS report frequencies, whereas Models 3 and 4, with store-specific fixed effects, found that only the monthly treatment group showed significant improvements relative to the control group. This discrepancy between the two sets of models likely stems from the fact that the stores receiving twice-monthly MS reports had higher age-verification rates than the other experimental groups during the baseline period

The significant treatment effects for these two models result from the higher post-baseline rates shown in Figure 1 for stores receiving either frequency of MS reports, especially during the middle two quarters of the study. The fixed-effects approach used in Models 3 and 4 is more conservative. Model 4 failed to find any advantage for the twice-monthly report group over the controls, probably because of the higher baseline rates of age verification in the former group. This can be seen in Figure 1, where the twice-monthly treatment group's line remains above and largely parallel to that of the control group during the first three quarters of the study, with the gap closing somewhat in the last quarter.

Fixed-effects estimates are preferable only to the extent that the higher average baseline rates for the twice-monthly feedback stores reflects real differences in their underlying age-verification propensity. This may not be the case. Figure 1 shows that post-baseline age-verification rates were quite similar for the two treatment groups: both had rates averaging roughly 16 percentage points higher than the control group's rate during the second and third study quarters, before having their advantage disappear in the final months of the study. This similarity between the two treatment groups immediately after the start of the MS reports suggests that the higher baseline rates of the twice-monthly group were possibly due to chance variation over a relatively small number of baseline observations. In retrospect, this problem might have been reduced by assigning stores to treatment groups after collecting baseline data, making it possible to match stores more evenly across the study arms. A longer baseline period might also have ameliorated this difficulty.

The second surprising finding is that age-verification rates rose steadily even among stores that never received any MS reports. Figure 1 shows that stores receiving MS reports, especially the monthly group, showed immediate improvements in age-verification rates following the start of the MS reports, whereas the control group's gains accumulated far more slowly over the post-baseline period. A possible explanation for the control group's improvement is that awareness of the mystery shop program gradually spread through interpersonal communications between managers of treatment group stores and managers of control group stores. A brief follow-up interview of managers at 27 of the 45 stores found that some control group store managers remembered hearing about MS programs through corporate memos, newsletters, or trainings. The managers had difficulty differentiating the program from other ongoing activities, however, making it difficult to draw firm conclusions. From an evaluation perspective, it is problematic that information about the MS visits may have eventually spread from treatment to control group stores, for such cross-group contamination would bias against finding positive treatment effects. Future research designs might reduce the possibility of such contamination by choosing stores that are more distant from each other in terms of both geography and corporate hierarchy.

From a policy perspective, however, the potential effects of inter-store communication concerning MS reports may represent an opportunity rather than a problem. Periodic MS reports produce a demonstrable improvement in age-verification conduct, but they entail significant costs. The eventual improvement in age verification by the control group stores suggests that retail chains might achieve high age-verification rates by visiting only a subset of stores on a monthly basis, communicating those results to all stores, and then encouraging interpersonal communication among the store managers. Such an approach would offer significant cost efficiencies, and this possibility should be further explored.

Acknowledgments

Data collection costs for this research were supported by a grant from Walgreens, Inc. to the Responsible Retailing Forum, Inc. The views expressed in this paper reflect those of the authors and not Walgreens. Analysis was also supported by National Institute on Alcohol Abuse and Alcoholism Research Center Grant P60-AA006282-28.

References

- Biglan A, Ary DV, Smolkowski K, Duncan T, Black C. A randomised controlled trial of a community intervention to prevent adolescent tobacco use. *Tobacco Control*. 2000; 9:24–32. [PubMed: 10691755]
- Botello-Harbaum MT, Haynie DL, Iannotti RJ, Wang J, Gase L, Simons-Morton B. Tobacco control policy and adolescent cigarette smoking status in the United States. *Nicotine & Tobacco Research*. 2009; 11(7):875–885. [PubMed: 19443786]
- Cummings KM, Hyland A, Perla J, Giovino G. Is the prevalence of youth smoking affected by efforts to increase retailer compliance with a minors' access law? *Nicotine & Tobacco Research*. 2003; 5:465–471. [PubMed: 12959784]
- DiFranza JR, Carlson RR, Caisse RE. Reducing youth access to tobacco. *Tobacco Control*. 1992; 1:58.
- DiFranza JR, Dussault GF. The federal initiative to halt the sale of tobacco to children—the Synar Amendment 1992–2000. *Tobacco Control*. 2005; 14:93–98. [PubMed: 15791018]
- Edwards R, Brown J, Hodgson P, Kyle D, Reed D, Wallace B. An action plan for tobacco control at regional level. *Public Health*. 1999; 113:165–170. [PubMed: 10483077]
- Federal Register. Block Grants (45 CFR Part 96) 61 Fed. Reg. 1492. Department of Health and Human Services, SAMHSA; 1996. Tobacco Regulation for Substance Abuse Prevention and Treatment.
- Forster JL, Murray DM, Wolfson M, Blaine TM, Wagenaar AC, Hennrikus DJ. The effects of community policies to reduce youth access to tobacco. *American Journal of Public Health*. 1998; 88:1193–1198. [PubMed: 9702146]
- Greene, WH. *Econometric analysis*. 5. Upper Saddle River, NJ: Prentice Hall; 2003.
- Institute of Medicine. *Growing up tobacco free: Preventing nicotine addiction in children and youths*. Washington, DC: National Academy Press; 1994.
- Jason LA, Billows WD, Schnopp-Wyatt DL, King C. Long-term findings from Woodbridge in reducing illegal cigarette sales to older minors. *Evaluations and the Health Professions*. 1996; 1996(19)
- Jason LA, Ji PY, Anes MD, Brown J, Birkhead SH. Active enforcement of cigarette control laws in the prevention of cigarette sales to minors. *The Journal of the American Medical Association*. 1991; 36:3159–3161.
- Jason LA, Pokorny SB, Schoeny ME. Evaluating the effects of enforcements and fines on youth smoking. *Critical Public Health*. 2003; 13:33–35.
- Jones SE, Sharp DJ, Husten CG, Crossett L. Cigarette acquisition and proof of age among US high school students who smoke. *Tobacco Control*. 2002; 11:20–25. [PubMed: 11891364]
- Krevor BS, Lieberman A, Gerlach K. The application of consumer protection authority in preventing tobacco sales to minors. *Tobacco Control*. 2002; 11:109–111. [PubMed: 12035002]
- Ling PM, Landman A, Glantz SA. It is time to abandon youth access programmes. *Tobacco Control*. 2002; 11:3–6. [PubMed: 11891349]
- McGuire, WJ. Theoretical foundations of campaigns. In: Rice, RE.; Atkin, CK., editors. *Public communication campaigns*. 2. Newbury Park, CA: Sage; 1989. p. 43-65.
- Public Health Service Act. (1992). Section 1926(b)(2) [42 USC 300x-26(b)(2)].
- Richardson L, Hemsing N, Greaves L, Assanand S, Allen P, McCullough L, et al. Preventing smoking in young people: A systematic review of the impact of access interventions. *International Journal of Environmental Research and Public Health*. 2009; 6(4):1485–1514. [PubMed: 19440530]
- Rigotti NA, DiFranza JR, Chang YC, Tisdale T, Kemp B, Singer D. The effect of enforcing tobacco sales laws on youth's access to tobacco and smoking behavior: A controlled trial. *New England Journal of Medicine*. 1997; 337:1044–1051. [PubMed: 9321533]
- US Department of Health and Human Services. *Best Practices for Comprehensive Tobacco Control*. Atlanta, GA: Centers for Disease Control and Prevention, Office on Smoking and Health; 1999.
- US Department of Health and Human Services. *Preventing tobacco use: A report of the Surgeon General*. Atlanta, GA: Centers for Disease Control and Prevention, Office on Smoking and Health; 2000.

US Department of Health and Human Services. FFY 2007 Annual Synar Reports: Youth Tobacco Sales. Rockville, MD: Substance Abuse and Mental Health Services Administration; 2008. URL <http://prevention.samhsa.gov/tobacco/synarreportfy2007.pdf>

Wilson, AM. The use of mystery shopping in the measurement of service delivery. In: Hogg, G.; Gabbott, M., editors. Service industries marketing: New approaches. Portland, OR: Frank Cass Publishers; 1998. p. 148-163.

Figure 1.
Average age-verification rates by experimental condition and survey quarter

Table 1

Descriptive Statistics

Variable	Mean	Std Dev	Minimum	Maximum
Verification of Shopper's Age	0.759	0.428	0	1
Age of Mystery Shopper (years)	20.944	0.680	20	22
Male Mystery Shopper	0.650	0.477	0	1
Located in Houston	0.329	0.470	0	1
Located in St. Louis	0.332	0.471	0	1

Note: Outcome measure is coded 1 if the clerk refused to sell tobacco without presentation of valid ID, 0 otherwise. N = 859 store visits.

Table 2

Logit Results Explaining Age Verification

Variable	No Store Fixed Effects		With Store Fixed Effects	
	Model 1	Model 2	Model 3	Model 4
1 or 2 Reports/Month Treatment	0.689 *** (3.57)		0.514 (1.73)	
2 Reports/Month Treatment		0.607 ** (2.60)		0.097 (0.26)
1 Report/Month Treatment		0.776 ** (3.21)		0.981 * (2.48)
Age of Mystery Shopper	-0.462 *** (-3.54)	-0.464 *** (-3.55)	-0.485 *** (-3.55)	-0.491 *** (-3.58)
Male Mystery Shopper	-0.160 (-0.83)	-0.162 (-0.84)	-0.161 (-0.80)	-0.170 (-0.84)
Linear Trend	0.006 *** (5.60)	0.006 *** (5.61)	0.007 *** (5.62)	0.007 *** (5.62)
Located in Houston	-0.439 * (-1.99)	-0.438 * (-1.99)		
Located in St. Louis	-0.390 (-1.83)	-0.390 (-1.83)		
Intercept	10.048 *** (3.70)	10.083 *** (3.71)	(45 Store-Specific Intercepts Not Shown)	

Notes: Numbers in parentheses are t-statistics associated with each regression coefficient. Outcome measure coded 1 if purchaser's age was verified by store personnel, 0 otherwise. City effects in Model 1 are relative to Milwaukee, which is the excluded category.

N = 859 store visits (up to 20 semi-monthly visits to each of 45 stores).

p ≤ .001;

**
p ≤ .01;

*
p ≤ .05 (two-sided)