



E-cigarettes and Smoking: Correlation, Causation, and Selection Bias

J. E. Prieger¹ · A. Choi²

Received: 29 June 2023 / Accepted: 15 July 2024

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2024

Abstract

Some public health officials discourage smokers from using electronic nicotine delivery systems (ENDS, or “e-cigarettes”) as a cessation aid because ENDS use is positively correlated with smoking. Such correlation does not imply that the causal treatment effect of ENDS use on cessation from smoking is negative, however, due to selection bias. We estimate the treatment effect of ENDS use on cessation. After showing that ENDS use and smoking are positively correlated in data from Korea, we investigate selection bias and show that a tax increase and the government’s negative pronouncements regarding ENDS shifted ENDS use toward those smokers for whom cessation is less likely. After accounting for unobserved confounding characteristics of individuals with regression models for endogenous treatment effects, we find that the evidence suggests that ENDS promote cessation. The average treatment effect on the treated (ATET) is estimated with parametric and moment-based methods and is found to be in the range of 10.1 to 16.4 percentage points from copula models and 17.0 percentage points from a moment-based estimator. The ATET from the results preferred by formal model selection criteria is 16.2 percentage points. The Korean government’s discouragement of ENDS use by smokers may therefore create a massive lost opportunity to reduce smoking and improve public health.

Keywords E-cigarettes · ENDS · Selection bias · Smoking · Copulas

JEL Classification I18 · I12 · I10

Introduction

Many individuals who smoke tobacco want to quit but find it difficult to do so since nicotine is addictive. In the past decade or so, e-cigarettes and heated tobacco products (HTPs) have become popular among some consumers of nicotine, especially those trying to quit

✉ A. Choi
anna.choi@sejong.ac.kr

J. E. Prieger
James.Prieger@pepperdine.edu

¹ School of Public Policy, Pepperdine University, Malibu, CA, USA

² Department of Public Administration, Sejong University, 209 Neungdong-Ro, Gwangjin-Gu, Seoul 05006, South Korea

smoking. Because these electronic nicotine delivery systems (ENDS) mimic the cigarette smoking experience more closely than pharmaceutical cessation aids such as nicotine patches, ENDS may encourage switching from cigarettes and therefore cessation of smoking. Alternatively, ENDS may hinder cessation if they complement the smoking habit. Whether public policy should encourage smokers to use ENDS for cessation, as in the UK (Hampsher 2020), remains an unsettled issue. Many public health officials object to ENDS on the grounds that the use of ENDS and smoking are positively correlated in many surveys, among other reasons¹; some officials incorrectly deduce from such evidence that ENDS discourage cessation from tobacco. However, in these surveys, the use of ENDS is a consumer choice, not a randomly assigned treatment, and that choice is likely to be correlated with unobserved confounding factors related to smoking behaviour. Ignoring this issue of sample selection (or “common liabilities,” as they are known in the public health literature) leads to bias in estimates of the causal effect of ENDS use on smoking. This study investigates this correlation, selection bias, and the causal treatment effect of ENDS use on cessation.

Health ministries in many countries discourage the use of ENDS, even for smokers. South Korea, the country studied here, is among them (Prieger and Choi, 2020). The smoking rate in Korea is higher than in most other countries. The cigarette smoking rate of 20% in 2018 is at the 64th percentile among reporting countries (World Health Organization (WHO) 2019b). Korea has among the highest prevalence of male smoking in the world; the cigarette smoking rate for men of 35% is at the 78th percentile among all countries and in the top decile among high-income countries. Despite large declines in smoking since 1980, when the male smoking rate was almost 80% (Prieger and Choi, 2020), further decline in the smoking rate stalled or reversed between 2005 and 2011. After ENDS began to come into common use in Korea around 2011–2013, the male smoking rate resumed its decline. These trends prompt the question of whether ENDS indeed played a role in cessation from smoking.

Many cohort studies of Korean subjects find that ENDS use and smoking are positively correlated, a finding we replicate below. To show that such findings are due to selection effects from the endogeneity of using ENDS, we show that a tax increase on ENDS and negative pronouncements regarding ENDS by the government in 2015 likely shifted the subpopulation of ENDS users toward those smokers for whom cessation is less likely. Evidence for the shift comes from changes in the composition of ENDS users and increases in the association between ENDS use and current smoking from regression and matching estimators. When the most likely users of ENDS are also those most addicted to smoking, it is unsurprising that smoking and ENDS use are positively correlated even if ENDS help achieve cessation.

We then estimate the treatment effects directly. Our econometric analysis of repeated cross-sectional data suggests that, after accounting for unobserved confounding characteristics of individuals, ENDS promote cessation. The novelty of the study lies in demonstrating that—even without randomized treatment assignment, longitudinal data that would allow individual-level fixed effects to control for selection, or policy and price variation that allows quasi-experimental methods such as difference-in-difference estimation—using appropriate econometric methods for endogenous treatment effects shows that the

¹ The other major objections from the international public health community are that ENDS could renormalize tobacco use (setting back advances in tobacco control), particularly among youth; that ENDS create “another layer of interference by the tobacco industry..., which is ... the most serious barrier to progress” in tobacco control; and that ENDS are not harmless (WHO 2019a).

correlation between ENDS use and smoking is due to selection bias. We use parametric models to estimate the effect of the binary treatment, which is ever use of ENDS, on the binary outcome: recent cessation from smoking. Several versions each of many models for treatment effects with binary dependent variables are estimated: bivariate probit, copula models, and a moment-based control function model. The results from the models controlling for selection effects generally agree that such effects exist, that they are in the direction that explain why ENDS use and smoking are correlated, and that the causal effect of ENDS on cessation is positive. Based on the results preferred by formal selection criteria, we estimate the average treatment effect (ATE) of ever using ENDS on the probability of cessation for ever-smokers in Korea to be between 28.9 and 32.2 percentage points. More importantly, the average treatment effect on the treated (ATET), which represents the expected impact of ENDS use among those choosing to use them (the group of smokers who may find cessation most difficult, as we show below), is estimated from the preferred bivariate copula model to be 16.2 percentage points. The control function estimation yields a similar ATET of 17.0 percentage points. The estimates for the ATET are robust to many different estimators, subsamples of the data, varying definitions of cessation, specification of the time trend, and whether an instrument is used.

Our empirical results finding a positive, significant, and sizeable causal effect of ENDS on cessation link real-world use of ENDS to results from randomized controlled trials (RCTs) on ENDS and cessation. Systematic reviews of RCTs indicate on the whole that ENDS can be a better aid to cessation than nicotine replacement therapy (Chan et al. 2021; Hartmann-Boyce et al. 2022; Hébert-Losier et al. 2020). For ENDS to be beneficial for public health, however, smokers must choose to use them. Adoption of ENDS has been relatively low in Korea, especially in contrast to other Asian countries such as Japan. The reason is largely due to government policy. While officials allowed ENDS into the marketplace for nicotine products in 2007,² the health and drug safety ministries have consistently discouraged the use of ends. Highly publicized pronouncements by the government have convinced many Koreans that ENDS are as harmful to health—or even more harmful—than smoking, that e-cigarettes do not help in quitting smoking, and that they cannot be used as a cessation aid. The pronouncements on harm are contrary to much scientific evidence, as will be discussed in the concluding section. Korea was the first country to tax nicotine-containing ENDS (in late 2010), and taxes on ENDS rose 119% in 2015 (along with increased cigarette taxes) and doubled again in 2021. Governmental agencies have issued several negative statements and recommendations regarding ENDS. A prominent example occurred in January 2015, when the Ministry of Health and Welfare announced that e-cigarettes “contain the same carcinogens as in tobacco cigarettes,” with the tacit implication that vaping is just as harmful as smoking (Ministry of Health & Welfare, 2015). The tax increases and widely publicized health warning from the government led to current ENDS usage to fall from its already low level by almost one-half from 2015 to 2016.

In the next section, we review the literature on ENDS use, smoking, and cessation in Korea and elsewhere. Descriptive empirical results on smoking and ENDS use in Korea are covered next, and then, the positive correlation between smoking and ENDS use is demonstrated with descriptive regressions. The subsequent section discusses the potential for selection bias and presents indirect and direct model-based evidence for it. We conclude with a summary and discussion of the results and their policy implications.

² For a fuller discussion of Korean policy toward ENDS and citations for this paragraph, see Prieger and Choi (2020).

Literature on the Association Between the Use of ENDS and Smoking

Since ENDS start gaining market prominence in the 2000s, a growing body of literature has examined the association between e-cigarette usage, attempts to stop smoking, successful cessation, and other aspects of dual use of cigarettes and ENDS. Although there is relatively limited evidence on the causal impact of e-cigarette usage on smoking, observational studies that investigate e-cigarette use among smokers provide interesting and suggestive results. Several studies find that ENDS use is associated with the intention to quit smoking; the majority of smokers who use ENDS do so to reduce or quit smoking (Adkisson et al., 2013; Maglia et al., 2017). These intentions translate to an increased number of quit attempts among ENDS-using smokers, compared to single-use smokers (Pasquereau et al., 2017; Zhuang et al., 2016). Increasing the number of attempts is important since most attempts to quit smoking fail, regardless of which cessation aids are used.

Are these quit attempts successful? Many studies from around the world find evidence that e-cigarette use is associated with cessation, although the evidence is not unequivocal. Systematic reviews of the literature conclude that most studies find positive relationships between using ENDS and cessation, but also that much of the evidence is correlational (Malas et al. 2016; Villanti et al., 2018). Some meta-analyses of observational studies also find positive and significant effects on cessation from ENDS use, whether in general (Rahman et al., 2015) or only for daily use (Wang et al., 2021). However, some reviews of observational studies find no positive association between vaping and cessation overall (Wang et al., 2021). The mixed findings regarding cessation rates among dual users may be due to differences in the frequency of use; daily e-cigarette users are more likely to quit smoking than non-users, but less-than-daily users are less likely (Wang et al., 2021).

Apart from abstinence, ENDS use can also encourage reducing the intensity of smoking (Pasquereau et al., 2017). On the other hand, other studies find that dual users consume more total nicotine and have higher nicotine dependence than single-use smokers (Kim et al. 2020; Snell et al., 2020). Of course, such findings may say more about the type of smoker who chooses to also use ENDS than about any causal effect of ENDS use on consumption or dependence. This possibility that individuals choosing to be dual users are those less likely to be successful in quitting informs our discussion of selection bias below.

In contrast to the mainly observational cohort studies discussed above, there are far fewer results from research designed to elicit causal effects. Several published studies evaluate the use of ENDS as a cessation aid in randomized controlled trials (RCTs) (Chan et al. 2021; Hébert-Losier et al. 2020). In particular, a recent Cochrane review (a series considered the gold standard of meta-analyses in public health and medicine) covered 22 RCTs and nine uncontrolled intervention studies where every participant received ENDS for cessation. It concluded that ENDS is more effective for smoking cessation than nicotine replacement therapy (with high certainty) and e-cigarettes without nicotine (with moderate certainty) (Hartmann-Boyce et al. 2022).

Other studies use quasi-experimental designs to identify causal effects. Saffer et al. (2020) employ a synthetic control difference-in-difference approach to examine a sharp increase in e-cigarette taxes in Minnesota to find that the adult smoking rate rose and the cessation rate fell in response. Relatedly, Dave et al., (2019a, 2019b) show that television advertising of ENDS has a causal effect on cessation from smoking. These studies are part of a recent set of quasi-experimental studies exploiting variation in taxes and other policies to provide robust evidence that ENDS and cigarettes are economic substitutes instead of complements (e.g., Abouk et al., 2023; Dave et al., 2019a, 2019b; Friedman & Pesko,

2022; Pesko & Warman, 2022), which indirectly supports the cessation hypothesis. Difference-in-difference modeling is unavailable in the present context since taxes and regulation are uniform within Korea.

The studies reviewed above (apart from Kim et al. 2020) were performed outside Korea. Research performed using data from Korean subjects tends to present a negative view of ENDS, mainly because typically no or little attempt is made to address the issue of selection bias. Several studies find that use of e-cigarettes is positively associated with smoking prevalence and intensity, especially among younger people.³ Specific findings include that the dual use of ENDS and cigarettes is more prevalent than single use of e-cigarettes, although some studies do not distinguish concurrent dual use from having ever used ENDS and cigarettes concurrently. Those who smoke more frequently are found to be more likely to be current ENDS users. However, such observational studies have little to say about whether e-cigarette use leads to smoking, vice-versa, or whether both behaviors are caused jointly by other factors. The same critique applies to the association found between e-cigarette use and smoking intensity. Studies on actual or attempted cessation in Korea have only recently begun to examine the role of e-cigarettes. Use of e-cigarettes is associated with more cessation attempts, in contrast to use of HTP (Kang & Cho, 2020).

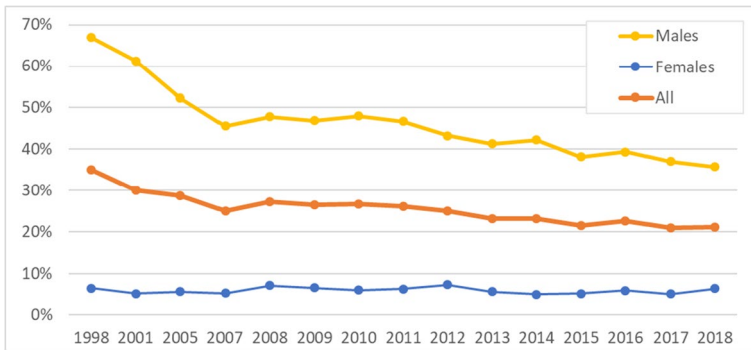
In contrast to the observational studies in Korea, the only RCT in the country involving ENDS came to positive conclusions (Lee et al., 2019). E-cigarettes were found to be about as effective as nicotine gum (a proven cessation aid), and a higher proportion of the e-cigarette users reduced their smoking (i.e., switched for daily to non-daily smoking) at 24 weeks.

Data and Descriptive Results on Smoking and the Use of ENDS in Korea

For our empirical investigation of smoking and vaping, we use data from the 1998 to 2018 Korea National Health and Nutrition Examination Survey (KNHANES), a cross-sectional population-based survey (using multistage stratified cluster sampling). Since questions regarding e-cigarette use are available since 2013, the sample is limited to 2013–2018 when examining ENDS. For 2018, the ENDS usage variable we construct refers to e-cigarette or HTP use; before then, there was no survey question about HTP in particular, although some HTP users may have considered their devices to be e-cigarettes when answering the survey question. The survey was conducted in 1998, 2001, 2005, and annually 2007–2018. The unit of observation is an individual, and all statistics and regressions estimated from the survey data are weighted to reflect the population of non-institutionalized adults of legal smoking age (19 years and older). All inference is based on standard errors computed with the Taylor series linearization method accounting for the complex survey design, including stratification, clustering on the primary sampling units, and the sampling weights.

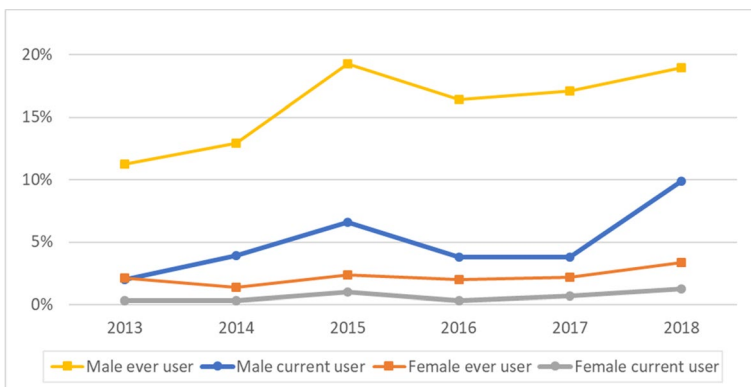
There were large decreases in the smoking rate in Korea since the 1980s until 2007, when it reached 25.0% (95% CI [0.230, 0.273]) (Fig. 1; Prieger and Choi, 2020). After 2007, the prevalence of smoking declined only modestly to its rate in 2018 of 21.1% [0.195, 0.227] overall and 35.8% [0.335, 0.382] for men. All of the recent decline is accounted for by men. The use of ENDS is fairly low in Korea but was increasing through 2015, with a reversal in response to unfavorable policy after 2015 (Fig. 2). Adults in their 20s and 30s are more

³ See Cho et al. (2011), Hwang and Park (2016), Jeon et al. (2016), Lee et al. (2014), Lee and Lee (2019), Lee et al. (2017), and Park et al. (2017) for studies of adolescents; Jeon et al. (2016) for study of college students; and Kim et al. (2018), Lee et al., (2016a, 2016b), and Park and Choi (2019) for studies of adults in general.



Note: data are from the KNHANES (analysis by the authors). Only individuals of age 19+ (the legal smoking age) are included. Note the irregular years of the survey in the early years. Figures are weighted to reflect the adult population.

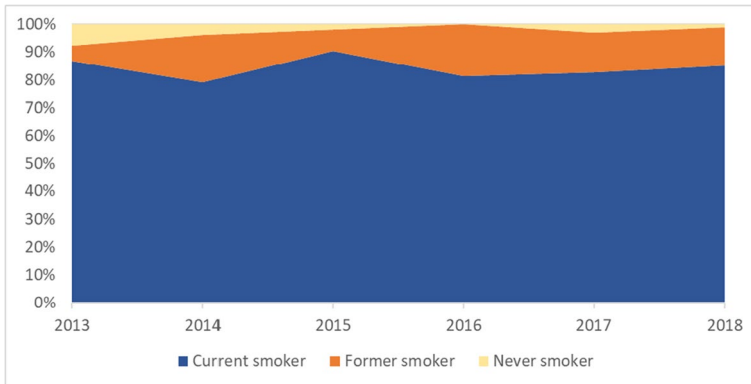
Fig. 1 Smoking prevalence among adults. Note: Data are from the KNHANES (analysis by the authors). Only individuals of age 19+ (the legal smoking age) are included



Note: Data are from the KNHANES (2013 to 2018). Prevalence of ever use is measured from the question “Have you ever used e-cigarettes?” and for 2018 “Have you ever used HTP tobacco?” 2018 data includes both ever e-cigarette use and ever HTP use. Current e-cigarette use is measured from the question “do you use electronic cigarettes now?” (with the inclusion of “do you use HTP tobacco now” in 2018).

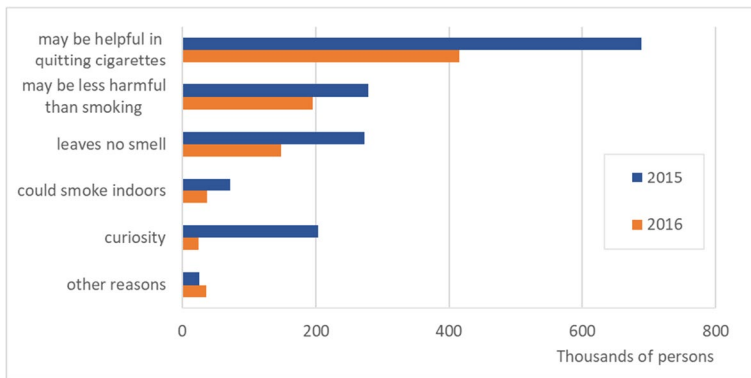
Fig. 2 Percentage of adults who ever used and currently use ENDS

likely than older people to use ENDS, although the share of current ENDS use among adults aged 40–49 years increased threefold in 2018, the first full period that HTPs such as PMI’s IQOS were available (Prieger and Choi, 2020), compared to the year before. Out of all adults, 18.9% of males (95% CI [0.171, 0.209]) and 3.4% [0.027, 0.042] of females report ever using e-cigarettes in 2018. The share of e-cigarette users increased sharply in 2015 (likely because of a large cigarette tax increase and indoor smoking ban that year), but fell the next year, possibly in response to ENDS tax increases and the negative government pronouncements about the health effects of ENDS in 2015. About 9.9% [0.085, 0.115] of males and 1.3% [0.009, 0.017] of females reported that they currently use e-cigarettes in 2018 (Fig. 2). Adult users of ENDS in Korea are nearly exclusively current and former smokers (Fig. 3). Use among never smokers is trivial—less than one-half of 1% of never smokers say that they have ever used ENDS during this period (see Table 2). Little has changed in the composition of ENDS users since 2016, despite the introduction of HTP in 2017.



Note: Data are from the KNHANES (author's calculations).

Fig. 3 Current adult ENDS use by smoking status



Note: Data are from the KNHANES; analysis by the authors. The survey weights are adjusted for nonresponse so that the totals are estimates for the target subpopulations. For comparison, there were about 9 million adult smokers during these years.

Fig. 4 Reasons for using ENDS among adult current users

Why are smokers and others trying e-cigarettes? In 2015, when cigarette taxes rose, there were more adults using ENDS for reasons of curiosity, harm reduction, and as a cessation aid than in 2016 (Fig. 4). In 2015, 44.8% (95% CI [0.357, 0.542]) of current e-cigarette users stated that cessation was their main reason for use, while in 2016, that figure was 48.6% [0.374, 0.599]. More smokers intend to quit than attempt to do so. Table 1 shows how many adult smokers have plans to quit smoking, and when, broken out by usage of ENDS. Over 2016 to 2018, about 61% of smokers stated that they had plans to quit, although more than half of those had no plans within the next six months. The intentions and timing of the plans do not vary greatly depending on whether e-cigarettes have ever been used, except that current and past users are about three percentage points more likely to have cessation plans. Having plans to quit is most likely among former HTP users (perhaps because they tried using HTP to quit in the past) and least likely among current HTP users (perhaps due to the selection effects discussed below). It is interesting to note, however, that more current HTP users have plans to quit more than six months in the future than any other group.

Table 1 ENDS use and plans for cessation among current adult smokers

| ENDS usage | Plans to quit smoking—percentage by category | | | | Total |
|---------------------|--|-----------------|---------------------|----------|-------|
| | Within 1 month | Within 6 months | Not within 6 months | No plans | |
| E-cigarettes | | | | | |
| Never used | 19.9 | 12.2 | 37.3 | 30.6 | 100.0 |
| Formerly used | 15.7 | 20.0 | 37.4 | 27.0 | 100.0 |
| Currently use | 18.4 | 18.3 | 35.7 | 27.7 | 100.0 |
| Total | 18.7 | 14.8 | 37.1 | 29.4 | 100.0 |
| HTP | | | | | |
| Never used | 18.7 | 11.7 | 38.3 | 31.3 | 100.0 |
| Formerly used | 10.2 | 32.5 | 35.5 | 21.8 | 100.0 |
| Currently use | 9.1 | 15.2 | 40.4 | 35.3 | 100.0 |
| Total | 16.6 | 13.4 | 38.5 | 31.5 | 100.0 |

Note: Figures are row percentages. Data are from KNHANES, 2016–2018, for e-cigarettes and 2018 for HTP; analysis by the authors. E-cigarette usage does not include HTP for 2018. Figures are weighted to reflect the subpopulation, which is current adult smokers

Table 2 Adult smoking status and ENDS use

| | Smoking status | | | Total |
|--------------------|----------------|---------------|----------------|-------|
| | Never smoked | Former smoker | Current smoker | |
| | % | % | % | |
| Never used ENDS | 65.1 | 18.6 | 16.4 | 100.0 |
| | 99.6 | 91.1 | 67.3 | 90.9 |
| Used ENDS but quit | 3.2 | 20.3 | 76.5 | 100.0 |
| | 0.3 | 6.9 | 21.7 | 6.3 |
| Current ENDS user | 2.0 | 13.1 | 84.9 | 100.0 |
| | 0.1 | 2.0 | 11.0 | 2.9 |
| Total | 59.4 | 18.5 | 22.1 | 100.0 |
| | 100.0 | 100.0 | 100.0 | 100.0 |

Data are from the 2013–2018 KNHANES. Figures are weighted to reflect the target population. The first row has row percentages and the second row has column percentages. ENDS use includes HTP in 2018

Results: The Positive Association Between ENDS Use and Smoking

Current and past use of ENDS is highly positively correlated with smoking in Korean data, even after controlling for demographic factors. Table 2 shows the cross-tabulation of smoking status and ENDS use in the KNHANES data averaged over 2013–2018. As already seen in Fig. 3, current ENDS users are likely to be current smokers: Out of the 2.9% of adults during the period 2013 to 2018 who were current ENDS users, 84.9% of them reported being current smokers. On the other hand, those who never used ENDS also tended to report that they had never smoked cigarettes. Out of the 90.9% of adults who

reported never using ENDS, 65.1% of them never smoked, 18.6% are former smokers, and only 16.4% of them are current smokers.

The positive association between ENDS use and smoking persists after controlling for observable characteristics of the individuals. Table 3 shows the results from probit regressions of a dependent binary variable for current smoking on current and past use of ENDS using KNHANES data from 2013 to 2018 (summary statistics and complete regression results are in the appendix). The probit model is used to be in accord with the models for endogenous treatment effects explored later. Both marginal effects (more familiar to economists) and adjusted odds ratios (more familiar to health researchers) are shown. The marginal effects and odds ratios are adjusted for the survey year, region of the country, gender, age, urban/suburban/rural location, educational attainment, occupation, household income, home ownership, and insurance status. For each sample, the first regression includes only the fixed effects for year and region, while the second regression includes all the additional control variables. Examples of the unobserved confounders controlled for by the year fixed effects include national tobacco control laws, tobacco and ENDS taxes (which are national in Korea), and the offerings and prices in the national tobacco and ENDS marketplace. The region fixed effects control for systematic differences in products or prices among regional markets, among other unobserved region-specific factors.

As in the Korean cohort studies discussed above, current and ever use of ENDS is highly and statistically significantly associated with current smoking. In the first two estimations in Table 3, the sample includes all adults of legal smoking age. The apparent marginal effect of current use of ENDS on smoking is huge: 60.3 percentage points in the estimation only including fixed effects. After controlling for additional demographic and SES controls, the marginal effect is reduced to a still-large 40.2 percentage points. The adjusted odds ratio for current use of ENDS is 17.1 when the additional controls are excluded in the regression and 10.5 when they are included. The marginal effects and adjusted odds ratios for past use of ENDS are even higher. These very high adjusted marginal effects and odds ratios are due mainly to the fact that very few people use ENDS in Korea unless they were smokers at some point. Thus, ENDS use is a very strong predictor of current smoking.

In the next two estimations in Table 3, the sample is restricted to adults who smoked at some point (in the past or currently). The marginal effects and odds ratios for ENDS use are much smaller than in the all-adult sample, but are still very large and statistically significant. Current ENDS use is associated with a higher probability of current smoking: either a 33 or 24 percentage point increase (and odds ratios of 4.5 or 3.1), depending on whether controls are included. Past ENDS use, as before, is a stronger predictor of current smoking. The final estimation in the table restricts the sample to adult males who smoke or formerly smoked, since most smokers and ENDS users are men and the prevalence of smoking among Korean women may be severely underreported (Kang et al., 2013). The results are similar to the similar regression specification with both sexes.

Results: Selection Bias in the Use of ENDS

As discussed above, the weight of the evidence from RCT studies around the world leans toward a positive treatment effect of ENDS use on cessation. No causal conclusions of any kind can be drawn from regression analyses of data from observational cohort studies that do not correct for bias in the estimates due to the endogeneity of ENDS use. Individuals choose whether to use ENDS, and that choice is highly likely to be correlated with unobserved factors related to smoking behavior.

Table 3 Probit regression results for current smoking and ENDS use

| <i>Y</i> = current smoking | All adults | | Ever-smoking adults | | Ever-smoking males | |
|----------------------------|------------------------------------|------------------------------|------------------------------------|------------------------------|------------------------------------|------------------------------|
| | Year and region fixed effects only | Demographic and SES controls | Year and region fixed effects only | Demographic and SES controls | Year and region fixed effects only | Demographic and SES controls |
| Marginal effects | | | | | | |
| Current ENDS use | 0.603 (0.012) | 0.402 (0.015) | 0.327 (0.013) | 0.244 (0.015) | 0.248 (0.016) | |
| Past ENDS use only | 0.693 (0.015) | 0.499 (0.022) | 0.410 (0.015) | 0.336 (0.018) | 0.326 (0.020) | |
| Adjusted odds ratios | | | | | | |
| Current ENDS use | 17.060 (1.200) | 10.541 (0.785) | 4.461 (0.339) | 3.149 (0.245) | 3.251 (0.277) | |
| Past ENDS use only | 31.111 (3.739) | 17.743 (2.269) | 8.300 (1.074) | 5.696 (0.756) | 5.370 (0.758) | |
| Subpopulation observations | 34206 | 34193 | 12770 | 12763 | 10956 | |
| Total observations | 43431 | 43418 | 44848 | 44841 | 44844 | |

All estimates are significant at the 1% level. Standard errors accounting for survey design effects are in parentheses. Odds ratios are computed at the average values of the regressors. All estimates are adjusted for the survey year and region of residence (17 major regions); the estimations including demographic and SES controls additionally adjust for sex, age, urban/suburban/rural location, educational attainment, occupational group, household income, home ownership, and insurance status. The odds ratios are computed at the average value of the regressors. The full regression results are in the appendix. Data are from the 2013–2018 KNHANES

For purposes of exposition only, assume that there are two types of people. A Type I individual is characterized by unobserved factors that make him more likely to initiate smoking, more likely to become addicted (or more heavily addicted) if he does smoke, and more likely to have weak motivation and low self-efficacy to quit smoking. The same factors may also lead to higher use of ENDS⁴ and thus a positive correlation between use of ENDS and cigarettes. This postulate is in accord with findings that dual users have higher tobacco dependence than single-use smokers (Kim et al. 2020; Snell et al., 2020) and higher urinary cotinine levels (a biomarker for nicotine consumption; Kim et al. 2020). If there is reverse causation between ENDS and cigarettes, because smokers adopt ENDS for the convenience of servicing their nicotine habit when smoking is not possible, then the induced positive correlation between ENDS and smoking would be even stronger. In analysis of cohort data populated with Type I individuals, one may find a large positive coefficient on ENDS use in a regression of smoking status on ENDS use and observable control variables, even if the true causal effect is zero or negative.

Consider now a Type II individual, who is at the other end of the spectrum of unobserved individual factors. He is less likely to initiate smoking, less likely to become addicted to nicotine, and more likely to be able to quit and has little proclivity toward ENDS use. For the same reasons as above but in attenuated form, there will still be positive correlation between observed use of ENDS and cigarettes, but it will be smaller. The same outcome regarding positive bias in the ENDS regression coefficient would be observed as for data populated with Type I individuals, except that (due to the smaller positive correlation between ENDS and cigarette use) the upward bias would be smaller. For both types, ignoring sample selection leads to bias in estimates of the causal effect of ENDS use on smoking.

To assess whether there is selection bias in the regression estimates from the previous section, two approaches are explored. In the first approach, plausible changes in the proportion of Type I and Type II individuals using ENDS are exploited to look for evidence of changes in the amount of bias in the regression coefficient. In the other approach, econometric models designed to account for selection bias directly are employed. Both approaches indicate that there is selection bias in the direction of obscuring the negative effect of ENDS use on smoking (or, equivalent, the positive effect of ENDS on cessation) and the latter approach suggests that the use of ENDS causally increases the probability of cessation.

Indirect Evidence of Selection Bias

Recent work in applied health econometrics notes that marked changes in regression coefficients after officials make a health recommendation regarding an activity can constitute evidence for bias from selection effects (Oster, 2020). For example, if health officials recommend that taking a particular vitamin is conducive to good health, and if generally healthier people are more likely to respond to the recommendation, then cohort studies will reveal that the apparent link between the vitamin and good health becomes spuriously stronger after the announcement.

⁴ That social environmental factors could induce correlation between addiction to nicotine and the use of ENDS seems clear, given the documented role of peer effects in both tobacco initiation and ENDS use (cite some for each). Recent research has also found shared genetic etiology between smoking or smoking intensity and ENDS use, in some cases alongside environmental contributions (Allegrini et al., 2019; Khouja et al., 2021; Prom-Wormley et al., 2020).

The relevant application to the present context is the announcement by the Korean government in 2015 recommending caution in use of e-cigarettes due to their stated harms to health. The difficulty in direct examination of the impact of the announcement is that in 2015, taxes on cigarettes also rose, which spurred much new use of ENDS as an alternative. Here, therefore, data from 2014 will be compared with data from 2016, after much of the new ENDS use had apparently fallen away (as shown in Fig. 2); the number of adult ENDS users is about the same in 2014 as 2016. Due to the health recommendation and the increased use of ENDS by current smokers in 2015 (reverse causality), it is expected that the composition of ever-ENDS users changed between 2014 and 2016. Because of the health recommendations, many Type I individuals may not have adopted ENDS in 2015 and 2016 who otherwise would have. Because of the tax increase, which affected the purchasing decisions of lower income smokers more than others, there also likely was an influx of new lower-income ENDS users who find cessation more difficult. The negative association between income and smoking cessation is well documented in the literature.⁵ For both these reasons, between 2014 and 2016, the mix of ENDS users likely shifted toward Type I individuals and others for whom cessation is less likely.

Before looking for change in the regression coefficients, we show that some key characteristics of e-cigarette users changed over the years 2014 to 2016. It is important to remember that the selection bias postulated to afflict the regression coefficients is caused by *unobserved* confounding factors. However, large changes in observed characteristics may signal that unobserved factors changed as well (Oster, 2020). Table 4 shows how the composition of the subpopulation of current ENDS users in the KNHANES changed with respect to gender, income, age, education, binge drinking, self-reported health, and smoking. The 2015 cigarette tax increase and health warning regarding ENDS changed the profile of e-cigarette users, resulting in users with lower income⁶ and less education, both of which are associated with greater difficulty in cessation.⁷ ENDS use also shifted toward older individuals, who have been found to have higher nicotine dependence, fewer smoking quit attempts, and lower rates of successful abstinence from cigarettes (see Arancini et al. 2021), although the relationship between age and cessation is less clear for Korean smokers (Kim et al., 2021). The incidence of binge drinking rose among ENDS users in 2016, which is highly associated with smoking and greater

⁵ Higher-income smokers are more likely to adopt behavioral and pharmacological treatments for cessation (Shiffman et al., 2008), which can double (for the former type, e.g., with the “rapid smoking” therapy; overall, individual behavioral therapy increases the odds about 60%) or triple (for the latter type of treatment, e.g., with varenicline) the odds of long-term abstinence (Lancaster and Stead, 2017; Tobacco Use and Dependence Guideline Panel, 2008). For Korean smokers, the evidence regarding income and cessation is mixed but also tends in the direction of a negative association. The odds of successful cessation rise with income in the studies of Cho et al., (2017, 2018), but not in the multivariate observational studies of Lee et al., (2016a, 2016b) and Lee and Seo (2020). Lower income has also been found to be associated with a lower probability of cessation in smokers using Korean smoking cessation services (Seo et al., 2019).

⁶ Comparing 2014 to 2016, ENDS users are more likely to be in the lowest and the lowest two income groups in the latter year, but also more likely to be in the highest income group. On net, however, the income distribution in 2014 exhibits second-order stochastic dominance over the income distribution in 2016, and in that sense, income has fallen among ENDS users.

⁷ For evidence regarding income and cessation, see footnote 5. De Walque (2007) estimated the causal effect of education on cessation to be positive. Evidence from observational studies in Korea shows that cessation is most common among college educated adults (Kim, 2014; Kim et al., 2013; Lee, Park et al., 2016a, 2016b) and that educational attainment is positively associated with progress through the stages of cessation (Leem et al., 2017) and maintenance of abstinence among participants in smoking cessation programs (Myung et al., 2011).

Table 4 Characteristics of current ENDS users, 2014 and 2016

| Variable | 2014 % | 2016 % |
|--------------------------------|-----------|-----------|
| Male | 92.1 | 91.9 |
| Household income | | |
| Low income (1st quartile) | 2.7 | 8.5 |
| Low-mid income (2nd quartile) | 18.5 | 22.8 |
| High-mid income (3rd quartile) | 50.4 | 30.0 |
| High income (4th quartile) | 27.4 | 38.6 |
| Age | | |
| 19–25 years | 20.3 | 19.1 |
| 26–35 years | 42.5 | 33.6 |
| 36–45 years | 23.5 | 20.5 |
| 46–55 years | 10.0 | 19.2 |
| Over 55 years | 3.7 | 7.6 |
| Educational attainment | | |
| Less than high school | 6.4 | 9.2 |
| High school | 48.4 | 51.0 |
| College | 42.8 | 33.1 |
| Binge drinking | 75.7 | 80.1 |
| Self-reported health | | |
| Very good | 6.3 | 5.0 |
| Good | 9.7 | 22.7 |
| Normal | 64.8 | 45.9 |
| Bad | 15.8 | 18.9 |
| Very bad | 1.1 | 0.8 |
| Current smoking | 79.2 | 81.4 |

Data are from the KNHANES. Binge drinking = 1 for all individuals not stating that they never binge drink (defined in the KNHANES as consuming ≥ 5 standard drinks (≥ 4 drinks for women) on one occasion. Percentages within a category may not equal 100 due to rounding and survey responses of “don’t know” or nonresponse

difficulty achieving cessation (Kaye et al., 2020). The distribution of self-reported health also changed but in non-monotonic ways; in 2016, there was a lower prevalence of “very good” health and there was more “bad” or “very bad” health combined, but there was also much more “good” health. Although the evidence is not unequivocal in every case, the bulk of these shifts seems to suggest a shift from Type II to Type I ENDS users. Although the timeframe does not exactly match, supplementary evidence for this shift can be found in survey evidence showing that in 2015, 56% of e-cigarette users in Korea did so to stop smoking, while that proportion dropped to 31.7% in 2018 (Mbrain 2018).

Based on the discussion of the selection-into-treatment model above, a shift toward Type I individuals would result in a larger positive coefficient on ENDS use in a regression of current smoking. That is indeed what happens when comparing the results from single-year regressions using data from 2014 to 2016. The first column of results in Table 5 is for probit regressions as in Table 3 performed separately using data from 2014 to 2016. The one change is that instead of including variables for both current and past ENDS use, the single variable for ever use of ENDS is used (to focus attention upon a single measure). The regression

Table 5 Evidence for selection effects from changes in estimated effect of ENDS on smoking, 2014 and 2016

| | Marginal effects | |
|--|-------------------|--------------------------------------|
| | Probit estimation | Propensity score matching estimation |
| 2014 | | |
| Effect of ever use of ENDS on the probability of current smoking | 0.245 | 0.219 |
| Std. error | 0.034 | 0.039 |
| 2016 | | |
| Effect of ever use of ENDS on the probability of current smoking | 0.314 | 0.304 |
| Std. error | 0.027 | 0.032 |
| Difference in the effect of ENDS | 0.070 | 0.129 |
| Std. error | 0.044 | 0.053 |
| 95% confidence interval | (−0.017, 0.155) | (0.025, 0.233) |

Probit results are marginal effects from estimations controlling for sex, age, marital status, metro/small city/rural region, educational attainment, home ownership, occupation, insurance status, and income. Estimations were performed separately by year. Standard errors and confidence intervals account for the design effects in the surveys

results are shown in the form of the predicted effect of ever-ENDS use on the probability of current smoking. The results show that while ever using ENDS is associated with a 24.5 percentage point increase in the probability of current smoking in 2014 (after adjusting for all the control variables), that association rises to 31.4 percentage points in 2016. As expected, the apparent association has risen, as would happen if the selection bias were worse in 2016. The change in the effect between the two years of 7 percentage points is not significant ($p=0.11$), however, despite being an increase in the predicted probability of 28%.

Another way to control for confounding observed factors, instead of regression adjustment, is to match on the observed covariates. A propensity-score matching estimator for the treatment effect of ever-ENDS use on smoking may be better than regression to control for bias from mismatch in covariates between ENDS users and other individuals. (However, a matching estimator does not remove bias from unobserved factors such as those at issue here, any more than regression does.) The results from the matching estimation, in the final column of Table 5, are in accord with those from the logit regressions (see the appendix for details of the estimation procedure). Furthermore, the change in the estimated effect of ENDS, 12.9 percentage points, is statistically significant. These results are consistent with the presence of selection bias that worsened in 2016.

In summary, the evidence in this section shows that the composition of ENDS users changed between 2014 and 2016, suggesting that shifts may also have occurred along unobserved dimensions of confounding factors. The regression results show that, as predicted by the selection-into-treatment theory, the positive association between ever using ENDS and current smoking increased after the tax increase and health recommendation against ENDS. All of this evidence is consistent with there being substantial selection bias in regressions that treat ENDS use as exogenous. Of course, evidence consistent with a hypothesis does not rule out alternative hypotheses, such as that ENDS use actually does have a positive causal effect on smoking. To strengthen the evidentiary case for selection bias, therefore, we turn now to direct econometric evidence.

Direct Evidence of Selection Bias

In the absence of an RCT or longitudinal data on individuals, there remain some weapons in the econometrician's arsenal to detect and correct for bias from endogenous treatment effects on a binary response. There are many approaches to estimating treatment effects on a binary outcome when a binary treatment is endogenously chosen by the individual.⁸ The estimations below include the bivariate probit model, several copula models, and Wooldridge's (2014) control function estimator. Semiparametric models for treatment effects are unavailable for use in the present application.⁹ Models with discrete endogenous regressors generally lack nonparametric point identification (Shaikh & Vytlacil, 2011), and so, estimation typically involves parametric assumptions of some sort, which is the approach taken here. The first is the bivariate probit model for endogenous treatment effects (Heckman, 1978), which has been widely used in health econometrics (e.g., Bhattacharya et al., 2006; Humphreys et al., 2014, among many others). Other parametric approaches to this estimation problem are based on copulas (Han & Vytlacil, 2017). Examples of copula-based approaches to endogenous treatment or sample selection in health economics, albeit not with binary outcomes, include Dancer et al. (2008), Hasebe and Vijverberg (2012), McGovern et al. (2015), Murteira and Lourenço (2011), Prieger (2002), and Winkelmann (2012). Copulas allow separation of the choice for the marginal distributions and the specification of their means from the specification of the dependence of the error structure for the treatment and outcome equations (Nelsen 2006). The Frank and Clayton copulas explored here are Archimedean copulas (Genest & McKay, 1986), which are popular for applied work. Archimedean copulas have closed-form parametric expressions, leading to straightforward maximum likelihood estimation (MLE), and allow radial asymmetry and tail dependence, resulting in a large variety of correlational structures (Nelsen 2006).

While instrumental variable and traditional control function estimators for treatment effects in linear models are unavailable for nonlinear models for binary outcomes (e.g., probit or logit) with binary endogenous variables, Wooldridge (2014) developed a moment-based control function estimator for the problem at hand. His GMM estimator, while resting on the parametric assumption of the normality of the errors, depends only on certain moments implied by that assumption and thus can be viewed as "less parametric" as opposed to fully parametric, MLE-based approaches.

In the endogenous treatment effect models estimated here, the choice of whether to have ever used ENDS is modeled jointly with the smoking cessation decision, with the choices allowed to be correlated due to unobserved confounding factors. If the assumptions of the models are satisfied (to be discussed below), the causal effect of ENDS use on cessation can be recovered even using cross-sectional (or repeated cross-sectional)

⁸ There is a multiplicity of forms for the general estimation problem for endogenous treatment effects. When treatment and outcomes are binary and the latent variables are linear in the regressors, we have the *linear index threshold-crossing model*. When treatment can potentially affect how each exogenous regressor is related to the outcome, the terms *potential outcome specification* and *endogenous switching regression* are used. When the probit model is adapted to the endogenous treatment effect situation, the resulting model is termed the *bivariate probit model with a dummy endogenous regressor* (or with a *binary endogenous explanatory variable*) or the *triangular* (or *recursive*) *bivariate probit model* or variations on these terms. Early expositions used the terms *multivariate probit model with structural shift* (Heckman, 1978) and *Model 6 for recursive dichotomous choice* (Maddala, 1983).

⁹ See section V.C.3 in the Appendix for discussion.

cohort data. The details of the parametric and moment-based control function models are in the appendix, but the idea behind each of the models is as follows.

The dependent variable for individual i for the selection equation, D , which is also the treatment variable in the outcome equation, is ever use of ENDS (including HTP in the 2018 data). Since individuals are observed only once, the time subscript for the KNHANES wave is suppressed in the notation. Ever use instead of current use is chosen to admit the possibility that a smoker switched to ENDS and then also ceased using ENDS. The decision ever to have used ENDS is modeled as in a regression model for binary dependent variables:

$$D_i = \begin{cases} 1 & \text{if } z_i'\gamma + u_i \geq 0 \\ 0 & \text{else} \end{cases} \quad (1)$$

where $D_i = 1$ means that individual i selects into treatment. Vector z_i in the selection equation is a set of variables for observed factors related to the treatment (e.g., demographic and socioeconomic variables of the individual and variables related to the price of ENDS, as those used in the descriptive regressions above), and γ is a vector of coefficients to be estimated. Unobserved factors affecting the individual's choice of treatment are collected into the scalar error term u_i . The outcome, cessation from smoking, is modeled similarly:

$$Y_i = \begin{cases} 1 & \text{if } \alpha_0 + \tau D_i + x_i'\beta + \varepsilon_i \geq 0 \\ 0 & \text{else} \end{cases} \quad (2)$$

where $Y_i = 1$ means subject i quits smoking. The inclusion of D_i in Eq. 2 makes this a triangular (or *recursive*) bivariate system. The linear index in Eq. 2 will be termed the *simple* specification, to distinguish it from the *potential outcome* specification. In the latter, a full set of interactions between D_i and the other regressors is included to allow all β coefficients to take different values depending on treatment status.

The following is known about identification in the triangular bivariate system with binary dependent variables defined by Eqs. 1 and 2 and known joint and marginal distributions of the error terms. Without an excluded instrument (i.e., all elements of z are also in x) but with sufficient variation in the exogenous variables (i.e., full support of at least one regressor in the outcome equation), point identification of all model parameters is obtained (Freedman & Sekhon, 2010). Without enough variation in the exogenous regressors, identification may fail or at best is likely to be weak (Mourifié & Méango, 2014). Thus, contrary to common belief, an excluded instrument is not always necessary, although having one may aid identification and reduce bias (Li et al., 2019). Second, an excluded instrument, even with limited variation (e.g., binary), along with some other conditions, is sufficient for global identification of all parameters (Han & Vytlacil, 2017). Finally, if the exogenous variation is small, the bias in the estimates may be large (as in the usual weak instrument literature for linear models) (Freedman & Sekhon, 2010; Han & Vytlacil, 2017). Nevertheless, the investigation of Denzer (2019) of various IV, parametric, and semiparametric estimation methods for the triangular bivariate probit model in the presence of weak instruments led to a “strong recommendation to primarily rely on the ML estimator.” Furthermore, Li et al. (2019) found that bivariate probit can yield reasonably accurate estimates of treatment effects even when the distributional assumptions are badly violated. In particular, in addition to finding that the bivariate probit model is a “reasonably resilient empirical tool for estimating the effect of an endogenous binary regressor on a binary outcome variable,” Li et al. (2019) also show that when the parametric assumptions are incorrect, the

model yields quasi-ML estimates that constitute a “conservative but rational choice for the practitioner.” For this reason, all except the moment-based estimations performed in the present study are MLE and we use instruments.

For the parametric and control function models, the error terms u_i and ε_i are assumed to be normal. The possibility that the error terms are not independent leads to the endogeneity of the treatment in the outcome equation; in such cases, probit estimation of Eq. 2 alone leads to bias in the estimate of τ and therefore the treatment effect. The various models estimated here differ in the assumed form of the correlation between the error terms and whether estimation is likelihood or moment based (see the appendix for the econometric details). In the parametric models, pseudo-maximum likelihood estimation¹⁰ returns estimates of the population regression coefficients and the joint distribution of the two error terms. Evidence for nonzero correlation between the error terms constitutes evidence of selection effects, implying that simple regression estimation of the outcome equation (e.g., probit) is biased.

The dependent variable for the outcome, Eq. 2, is recent cessation. This binary variable is defined as a formerly smoking male who quit smoking within the past 8 years and maintained cessation to the time of the survey. The outcome variable thus measures past cessation and non-relapse of the most recent cessation attempt. The sample is limited to current and former male smokers who quit within the past 8 years. The time since quitting smoking is by self-report in the KNHANES. Limiting the sample to exclude those who quit smoking long ago was done because the focus of the study is how ENDS may be linked to cessation, and someone who quit smoking in, say, 1990 neither quit with the help of ENDS nor was likely to begin smoking again during the era of ENDS availability. A threshold of 8 years was chosen to roughly match the ENDS period in Korea. The first year of data on ENDS, and thus the first year of data in this estimation, is 2013; cessation for this wave of respondents has to be after 2005, or about 2 years before ENDS became available.¹¹ The final year of data is 2018; cessation for that wave has to be after 2010, the year before ENDS began to be commonly used.

Recent cessation is modeled in the outcome Eq. 2 as a linear function of ENDS use (D_i) and the demographic and socioeconomic variables explored above with the region fixed effects (x_i). In addition to all the elements of x_i , the regressor vector z_i also includes the tax on ENDS products (given the discussion on identification above, this variable may be viewed as an instrument but is not needed as an *identifying* instrument). The tax variable is a composite of e-liquid and HTP heatstick taxes¹²; the two products' taxes are combined as a weighted average based on relative market shares (see Table A-1b in the appendix). The regressors also include the interaction of the tax with Medical Aid status (a public insurance program roughly similar to Medicaid in the U.S.) and categorical household income. The interactions are included as further instruments under the notion that ENDS prices (which are affected by the taxes) may matter most for lower income consumers. In each estimation reported in Table 6 below, the p -value from the Wald test of the coefficients for

¹⁰ The “pseudo” term is applied because of the weights. The weighted likelihood is an estimate of the population likelihood function.

¹¹ Even though cessation by this definition may therefore have occurred for respondents in the 2013 and 2014 survey waves before ENDS were introduced in Korea, someone who quit smoking only one or 2 years prior is still at risk of relapse. Therefore, if ENDS usage either encouraged or discouraged relapse to smoking, then they would be a determinant of (prolonged) cessation even for individuals in these earliest two survey waves.

¹² In the composite, an equivalency of 1 ml of e-liquid to 10 heatsticks is used.

the excluded instruments was below 0.015, and in the copula models, it was below 0.005, indicating the relevance of the instruments. Since the ENDS tax does not vary in the cross-section, it is not possible to use year fixed effects in these regressions, and so instead, a linear time trend is included in x_i and z_i . The analysis sample for the regressions includes male respondents only due to the large differences in ENDS and cigarette use by gender and low self-reported rates of ENDS usage and smoking among females. The econometric results are summarized in Table 6. The average treatment effect on the treated (ATET, defined as $E(Y_1 - Y_0|D = 1)$) is the result of primary interest. The average treatment effect (ATE, or $E(Y_1 - Y_0)$) is also shown. While the ATE yields the expected improvement in cessation from using ENDS across the whole subpopulation of smokers and former smokers, the ATET is more germane because it is the average effect on cessation for those smokers who chose to use ENDS. The latter group is the one which some public health officials and researchers claim have cessation hindered by ENDS use.

To summarize, identification of the treatment effects in these models rests not solely on policy-driven variation in the ENDS taxes but also on variation in the other regressors and the parametric assumptions. Without the additional assumptions, little headway can be made toward estimating or bounding the treatment effect. The Manski nonparametric bounds on the ATE under the minimal assumption of no reporting error are estimated (McCarthy et al., 2015) to be $[-0.47, 0.53]$, so wide as to be virtually uninformative. Under the stronger assumption of monotone negative treatment selection (i.e., that individuals choosing to use ENDS are more likely to not quit smoking, conditional on treatment assignment, which is the direction of the selection bias indicated by the indirect evidence presented above), the nonparametric bounds on the ATE can be narrowed to $[-0.16, 0.53]$, which still do not pin down the sign, hence the parametric assumptions to point-identify the treatment effects.

Table 6 includes results from three families of models: bivariate probit, bivariate copula models for binary dependent variables, and the moment-based control function model for endogenous treatment effects with binary outcomes (details and complete regression results are in the appendix). Two main specifications are explored: a simple triangular specification in which the treatment indicator appears only by itself in the outcome regression equation, and a potential outcome specification in which all the coefficients in the outcome equation are allowed to differ depending on treatment status (estimated by interacting D_i with each element of x_i). The latter is more flexible but introduces a large number of extra parameters to be estimated.

The simple specifications in Table 6 include one from the probit model assuming the exogeneity of ENDS use (for comparison), one from the bivariate probit model, and four from the Frank and Clayton Archimedean family of copula models. The probit model is essentially the same estimation as in the final column of Table 3, except for the change in the ENDS variable to be ever use. The coefficients for the outcome equation, the ATET, and the ATE are the same as would be for those from a bivariate probit model with correlation parameter ρ restricted to zero. Given the assumed exogeneity of ENDS use and correlations between the latter and smoking observed above, it is unsurprising that the ATE and ATET are negative, large, and significant.

However, all the models that correct for endogeneity have positive estimates of the ATET and ATE (and most are statistically significant). For each estimation, evidence for the endogeneity of ENDS use can be gleaned from the significance of θ , the correlation-related parameter in the copula models, the Pearson correlation ρ , and Kendall's rank correlation τ . Kendall's τ is a measure of similarity of the orderings of two random variables that is commonly used with copulas.

Table 6 Estimated treatment effects and evidence for the endogeneity of ENDS usage from the bivariate regression models

| Regression model | ATE† | ATE | Parametric copula parameter θ | Pearson correlation ρ | Kendall's τ | dAIC |
|---|-----------|-----------|--------------------------------------|----------------------------|------------------|----------------------|
| Simple specifications | | | | | | |
| Probit (exogeneity assumed) (S1) | -0.146*** | -0.147*** | NA | 0 (fixed) | 0 (fixed) | 16765.4 |
| Bivariate probit (S1) | 0.053 | 0.077 | NA | -0.411* | -0.269** | 16763.6 |
| Frank copula (S1) | 0.121*** | 0.235** | -4.610*** | -0.581*** | -0.298 | 16760.8 |
| Clayton copula (270° rotation) (S1) | 0.164*** | 0.341*** | 3.406*** | -0.800*** | -0.630*** | 16762.9 |
| Frank copula (S2) | 0.107*** | 0.198** | -3.968*** | -0.527*** | -0.285 | 16707.0 |
| Clayton copula (270° rotation) (S2) | 0.162*** | 0.322*** | 2.933*** | -0.771*** | -0.595*** | 16705.8 [¶] |
| Potential outcome specifications | | | | | | |
| Bivariate probit (S3) | 0.142*** | 0.296** | NA | -0.715** | -0.507*** | 16787.0 |
| Frank copula (S3) | 0.101 | 0.179 | -3.871 | -0.517* | -0.285 | 16792.8 |
| Clayton copula (270° rotation) (S3) | 0.152*** | 0.260** | 2.271 | -0.715*** | -0.532*** | 16789.7 |
| Clayton copula (270° rotation) (S4) | 0.156*** | 0.289*** | 2.523** | -0.739*** | -0.558*** | 16709.4 [‡] |
| Control function model (S3) | 0.170*** | 0.007 | NA | § | NA | NA |

$p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

[†]Best model among the likelihood-based models (which excludes the control function model) indicated by the dAIC, within the same type of regression specification (i.e., simple or potential outcome)

[¶]Best model among the likelihood-based models (excluding the control function models) indicated by the dAIC, across all regression specifications

[§]A hypothesis test for the endogeneity of treatment has p -value 0.180

NA not applicable. The sample is limited to current and past smokers. The data are from the 2013–2018 KNHANES. The parametric copula parameter is θ for the Archimedean copula models (see section 1.D in the Appendix). The dAIC statistic is Lumley and Scott's (2015) design-based, bias-corrected AIC (see section 11.B in the Appendix); lower values indicate less divergence from the true parametric model for the population. Standard errors for ATE† and ATE are unconditional (i.e., they allow for the sampling error of the covariates and thus are for the population). See text for explanation of which variables are included with which specification. See the appendix for the full set of regression results

Given that none of the parametric models may reflect the true data generating process, the estimates can be viewed as yielding quasi-ML estimates that minimize the Kullback–Leibler divergence from the true model (Li et al., 2019). To link this notion to model selection criteria in a manner appropriate for survey data, we compute Lumley and Scott’s (2015) design-based Akaike Information Criterion (dAIC), which is a version of the familiar AIC that is bias-corrected for complex survey samples used to estimate population parameters. Lower values of dAIC, shown in the final column of Table 6, indicate less Kullback–Leibler divergence from the unknown true model.

The first model admitting correlation and endogeneity is the bivariate probit model in the second row of Table 6. The results of the bivariate probit model show insignificant treatment effects and only weak evidence for endogeneity. However, the copula models show strong evidence of both. The Frank copula model with the simple specification estimates the ATET to be 12.1 percentage points. For the Clayton copula, the results shown are for the version of the copula with the highest pseudo-likelihood among the standard and two rotated versions. The standard version admits only positive correlation while the 90° and 270° rotated versions of the Clayton copula assume that correlation is negative.¹³ The 270° rotated Clayton model (hereafter, *Clayton270*) is preferred since it has the highest pseudo-likelihood and the lowest dAIC among the Clayton models. It shows strong negative correlation and a significant ATET of 16.4 percentage points (the other Clayton models are not shown in Table 6 but all results are in the appendix). The tail dependence in the Clayton270 results implies that (above and beyond correlation and observables) a smoker who is highly unlikely to use e-cigs is highly likely to quit smoking. Both copula models are preferred to the bivariate probit models by dAIC.

A more parsimonious version of the simple specification, S2, was also estimated, in which some of the regressors that were generally insignificant in the regressions were dropped. The resulting estimated ATETs were significant: 10.7 from the Frank model and 16.2 percentage points from the Clayton270 model. Per the dAIC, specification S2 outperforms the corresponding Frank and Clayton270 estimations with all the regressors (S1). Of the models for the simple specifications, the Clayton270 model with specification S2 is the preferred estimation per the dAIC. Thus, the best estimate of the ATET is 16.2 percentage points, coupled with a statistically significant estimate of the endogeneity of ENDS use.

To match some of the literature on endogenous treatment effects, the potential outcome specification was also estimated for the various models. The results are in the lower half of Table 6. Specification S3 is the same as S1 but with each coefficient allowed to differ with treatment status. A reduced specification, omitting some insignificant regressors, yields specification S4. The estimated treatment effects are largely similar to those from the simpler specifications and the Clayton270 model is again the best among the fully parametric models. None of these, however, are better (per dAIC) than their counterparts with specifications S1 of S2. Adopting the potential outcome specification allows estimation of Wooldridge’s moment-based control function estimator for endogenous treatment effects with binary treatment and outcomes. This estimation (in the final row of the table), with a specification nearly identical to S3,¹⁴ yields the highest estimate of ATET: 17.0 percentage points. Since this estimator is based on GMM instead of MLE, there is no way to compare it to the other models using the dAIC. It may thus be added to a set of preferred (i.e., non-dominated) estimates.

¹³ Given the asymmetry of the Clayton copula, the 90° and 270° rotations lead to different estimates.

¹⁴ The only difference between specifications S3 and S3’ is that since treating age as categorical led to problems with convergence, the categories were replaced with age and its square.

Note that with each model correcting for sample selection except the control function model, the ATE is much larger than the ATET. This comparison is further evidence in favor of selection bias in the use of ENDS, for it implies that if treatment were randomly assigned across the entire population of smokers, ENDS would have a larger causal effect on cessation than it does for those choosing to use ENDS under the current set of policy, attitudes, and preferences.

To summarize, the reason why the impact of using ENDS switches from being pro-smoking in the simple regressions in Table 3 to being pro-cessation here is due to the strong selection bias in the former. The correlation in the unobserved factors related to cessation and ENDS use is negative, whether measured by the parametric correlation-related coefficient, the Pearson correlation coefficient, Kendall's τ , or the tail dependence in the Clayton copula. That is, there are latent factors of individuals that make them more likely to use ENDS and less likely to quit smoking/maintain cessation. This is exactly as predicted by the conceptual discussion of selection bias from the beginning of this section.

Discussion

The preferred estimates for the treatment effect on cessation of ever use of ENDS and HTP (ATE=0.162 from the Clayton270 model and 17.0 from the control-function model) are sizeable but not unreasonably large. For comparison, one meta-analysis of studies on varenicline (the most effective pharmaceutical for cessation) found a treatment effect of 26 percentage points at 12 weeks after quitting for men. It is to be expected that use of ENDS has a lower estimated treatment effect, since cessation was maintained for a longer period (38 months on average) and in any event, some ENDS users have no intention to quit smoking (refer back to Table 1). The estimated treatment effects are also more modest than the 22.7% cessation rate at 24 weeks found for the ENDS-using group in the Korean RCT of Lee et al. (2019), discussed above, likely for the same two reasons.

There are some limitations to the present study. The outcome studied is cessation maintained to the time of the survey, but some of those individuals may subsequently relapse, which is not captured in the data. On the other hand, relapse to smoking before the survey is counted as “no cessation.” This issue is the same for all studies of cessation, however; see also the robustness test below that varies the length of prolongation of cessation. Perhaps more importantly, the methods employed here depend to an unknown degree on the various modeling assumptions. Due to selection issues, no results from observational, pooled cross-sectional studies can be as convincing as longitudinal studies, RCTs, or quasi-experimental methods using panel data. Given the lack of longitudinal data or sharp subnational policy differentiation and the absence of nonparametric identification in doubly binary treatment effect models, we used parametric assumptions to analyze repeated cross-sectional data to demonstrate the existence of a strong selection effect and a positive causal effect of ENDS use on cessation. To lessen the dependence of the results on any one set of assumptions, a variety of models and regression specifications were presented. Another limitation is the limited variation in the ENDS tax instrument. The relevant identifying variation in the data when the instrument is employed is the change in national ENDS taxes over time (net of the included time trend), “personalized” to each individual only through interactions of this variable with their income. One could certainly wish for a stronger instrument with greater variation in both the cross-section and time dimensions. However, as discussed above, an instrument is not strictly necessary for identification in the parametric models, and as to

be discussed shortly, the results change little if the instrument is dropped altogether and a specification test fails to show evidence that the instrument is invalid.

Given their importance, the parametric assumptions were tested in three ways. Details for all of these are in the appendix. First, the specification test of the bivariate probit model of Acerenza et al. (2023) was adapted to our preferred Clayton model. The null hypothesis of correct model specification for this test can be rejected if any of the identifying assumption discussions above for the parametric models fail (correct choice of marginal distributions for the latent errors, correct choice of copula to bind the two equations into the joint density, and instrument validity [when an instrument is used]). The test rests on the conditional probabilities implied by the model, converting them to moment inequalities that can be tested with Chernozhukov et al.'s (2015) intersection bounds procedure. The test result indicated acceptance of the null hypothesis of correct model specification and instrument validity. Second, we checked whether swapping the probit marginals for logistic errors in the treatment and outcome equations affected the results in the copula model. In all combinations (probit-logit, logit-probit, and logit-logit), the ATET remained about the same as in the main estimation.¹⁵ Third, we explored moving away from fully parametric assumptions with the seminonparametric estimator of Gallant and Nychka (1987) as adapted to the present endogenous treatment effects problem (De Luca, 2008). This estimator uses expansions in Hermite polynomials to flexibly model the joint distribution of the error terms in a manner that yields consistent estimation for a broad class of distributions. The seminonparametric estimator allows for arbitrary skewness and kurtosis to reduce the potentially high sensitivity of the estimates of the treatment effects to the marginal distributional assumptions. The estimate of the ATET produced by this model is 0.103, somewhat lower than our main estimates, yet still significant at the 1% level.¹⁶

Several additional robustness tests were explored; results are in the appendix. First, the linear time trend in the models reported in Table 6 was replaced with a cubic time trend. The results varied little, the same model was preferred by the dAIC, and the estimated ATET was identical to the main result (0.162). Second, given that instruments may not be needed for identification in these models, the preferred model was re-estimated without instruments (and with the cubic time trend). Again, the results were highly similar, the same model was preferred, and the ATET was estimated to be nearly identical at 0.163. Third, we varied the definition of cessation to require prolongation of abstinence for various lengths (three to 24 months). The ATET for the preferred model remained highly statistically significant, although its magnitude naturally fell as the length of abstinence required for cessation increased (from 0.125 at three months to 0.044 at 24 months). Fourth, we varied the window for recent cessation. Whether cessation was restricted to onset in the past three years, the baseline case of eight years, 10 years, or other years in between, the ATET from the preferred model was still significant at the 1% level, with estimates varying between 0.130 and 0.172. Fifth, we limited the sample to various subsets of years due to the fact that the impact of heated tobacco products and e-cigarettes may not be the same on cessation. Whether we excluded 2018 (the first year that questions regarding use of heated

¹⁵ The software used for this test (Hasebe, 2022) did not allow estimating the Clayton model with 270° rotation, and so the second-highest preferred model, the Frank model (specification S2), was used instead.

¹⁶ The estimates from this model proved to be sensitive to initial parameter vectors, and the chosen order of the Hermite polynomials and the pseudo-likelihood function often had little curvature in the dimensions most relevant to the ATET and clearly had many local maxima. Thus, even though most of the results from the attempted seminonparametric estimations bolster the main results, we cannot guarantee that a global optimum was attained.

tobacco were added to the survey) or 2017–2018 (the years in which heated tobacco products were available in Korea), the estimated ATET from the preferred Clayton model and Wooldridge’s model changed little. The same was true when using *only* the data from 2017 to 2018.

Despite the potential limitations of the study, the concordance of the results across models and specifications leads us to draw three general conclusions. First, there is evidence that the same men in Korea who are least likely to quit smoking are those taking up ENDS. This is demonstrated by the consistent finding of negative correlation between the treatment and outcome equations. Second, correcting for the selection bias in using ENDS reveals that they may aid cessation, as shown by the positive and general significant estimates of the ATET and the ATE. Finally, there is therefore potentially severe selection bias in any observational study of the association between ENDS use and smoking.

If ENDS can indeed be an effective aid to cessation, then the Korean government’s discouragement of their use by smokers creates a massive lost opportunity to reduce smoking and improve public health at a time when smoking’s decline has slowed. The implications apply to many other countries as well. Public health officials in most countries take their lead from the World Health Organization (WHO) through the Framework Convention on Tobacco Control (FCTC), which stops short of banning ENDS but encourages parties to “consider prohibiting or otherwise regulating them” (World Health Organization (WHO) 2019b). However, while ENDS are not risk-free, switching completely from cigarettes to e-cigarettes substantially reduces levels of measured carcinogens and toxins in the body (National Academies of Science, Engineering and Medicine 2018; Shahab et al. 2017). Public Health England (2015) encouraged smokers to switch to e-cigarettes with a statement that, based on the evidence, the latter are about 95% safer than cigarettes. Partially due to that reason, there is no excise tax on e-cigarettes in the UK (Hampsher 2020).

Negative governmental attitudes toward ENDS also run the danger of creating a self-fulfilling prophecy. Research shows that the reasons for the use of ENDS greatly affect whether they aid cessation. Smokers who use e-cigarettes for smoking cessation have higher rates of quitting cigarettes in the future, but ENDS users with other reasons can have lower cessation rates than those who do not use ENDS at all (Vickerman et al., 2017). Thus, if public misperception about ENDS causes vapers to use e-cigarettes mainly for reasons other than cessation, then ENDS will in fact be less helpful for cessation. The Korean response to ENDS can be contrasted with Japan’s experience, where ENDS in general (and HTP tobacco in particular) appear to be reducing consumption of cigarettes and perhaps furthering cessation in recent years (Hahn et al. 2020; Stoklosa et al. 2019).

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10603-024-09573-y>.

Funding This report was funded with a grant from the Foundation for a Smoke-Free World, a US nonprofit 501(c)(3) private foundation with a mission to end smoking in this generation. The Foundation accepts charitable gifts from PMI Global Services Inc. (PMI); under the Foundation’s Bylaws and Pledge Agreement with PMI, the Foundation is independent from PMI and the tobacco industry. The contents, selection, and presentation of facts, as well as any opinions expressed herein, are the sole responsibility of the authors and under no circumstances shall be regarded as reflecting the positions of the Foundation for a Smoke-Free World, Inc.

Data Availability Korea National Health and Nutrition Examination Survey (KNHANES) data is available at <https://knhanes.kdca.go.kr/knhanes/eng/index.do>.

Declarations

Competing Interests Prieger received compensation for consulting for a tobacco company on matters of illicit trade, but not for this project. Choi declares no competing interests.

References

- Abouk, R., Courtemanche, C., Dave, D., Feng, B., Friedman, A. S., Maclean, J. C., ... & Safford, S. (2023). Intended and unintended effects of e-cigarette taxes on youth tobacco use. *Journal of Health Economics*, 87, 102720. <https://doi.org/10.1016/j.jhealeco.2022.102720>
- Acerenza, S., Bartalotti, O., & Kédagni, D. (2023). Testing identifying assumptions in bivariate probit models. *Journal of Applied Econometrics*, 38(3), 407–422.
- Adkison, S. E., O'Connor, R. J., Bansal-Travers, M., Hyland, A., Borland, R., Yong, H.-H., ... Fong, G. T. (2013). Electronic nicotine delivery systems. *American Journal of Preventive Medicine*, 44(3), 207–215. <https://doi.org/10.1016/j.amepre.2012.10.018>
- Allegrini, A. G., Verweij, K. J., Abdellaoui, A., Treur, J. L., Hottenga, J. J., Willemsen, G., ... & Vink, J. M. (2019). Genetic vulnerability for smoking and cannabis use: Associations with e-cigarette and water pipe use. *Nicotine & Tobacco Research*, 21(6), 723–730. <https://doi.org/10.1093/ntr/nty150>
- Arancini, L., Borland, R., Le Grande, M., Mohebbi, M., Dodd, S., Dean, O. M., ... & Cummings, K. M. (2021). Age as a predictor of quit attempts and quit success in smoking cessation—Findings from the ITC 4-country survey (2002–2014). *Addiction*, 116(9), 2509–2520. <https://doi.org/10.1111/add.15454>
- Bhattacharya, J., Goldman, D., & McCaffrey, D. (2006). Estimating probit models with self-selected treatments. *Statistics in Medicine*, 25(3), 389–413. <https://doi.org/10.1002/sim.2226>
- Chan, G. C., Stjepanović, D., Lim, C., Sun, T., Anandan, A. S., Connor, J. P., ... & Leung, J. (2021). A systematic review of randomized controlled trials and network meta-analysis of e-cigarettes for smoking cessation. *Addictive Behaviors*, 119, 106912. <https://doi.org/10.1016/j.addbeh.2021.106912>
- Chernozhukov, V., Kim, W., Lee, S., & Rosen, A. M. (2015). Implementing intersection bounds in Stata. *The Stata Journal*, 15(1), 21–44.
- Cho, J. H., Shin, E., & Moon, S. S. (2011). Electronic-cigarette smoking experience among adolescents. *Journal of Adolescent Health*, 4(5), 542–546. <https://doi.org/10.1016/j.jadohealth.2011.08.001>
- Cho, M. H., Kim, S. M., Lee, K., Park, S. M., Chang, J., Choi, S., Kim, K., Koo, H. Y., & Jun, J. H. (2018). Factors associated with continued smoking after the diagnosis of type 2 diabetes: A retrospective study in the Korean cohort. *British Medical Journal Open*, 8(6), e020160. <https://doi.org/10.1136/bmjopen-2017-020160>
- Cho, Y. M., Myong, J. P., Kim, H. R., Lee, H., & Koo, J. W. (2017). Service and sales workers, are they vulnerable to smoking cessation? *Industrial Health*, 55(5), 406–415. <https://doi.org/10.2486/indhealth.2016-0118>
- Dancer, D., Rammohan, A., & Smith, M. D. (2008). Infant mortality and child nutrition in Bangladesh. *Health Economics*, 17(9), 1015–1035. <https://doi.org/10.1002/hec.1379>
- Dave, D., Dench, D., Grossman, M., Kenkel, D. S., & Saffer, H. (2019a). Does e-cigarette advertising encourage adult smokers to quit? *Journal of Health Economics*, 68, 102227. <https://doi.org/10.1016/j.jhealeco.2019.102227>
- Dave, D., Feng, B., & Pesko, M. F. (2019b). The effects of e-cigarette minimum legal sale age laws on youth substance use. *Health Economics*, 28(3), 419–436. <https://doi.org/10.1002/hec.3854>
- De Luca, G. (2008). SNP and SML estimation of univariate and bivariate binary-choice models. *The Stata Journal*, 8(2), 190–220. <https://doi.org/10.1177/1536867X0800800203>
- De Walque, D. (2007). Does education affect smoking behaviors?: Evidence using the Vietnam draft as an instrument for college education. *Journal of Health Economics*, 26(5), 877–895. <https://doi.org/10.1016/j.jhealeco.2006.12.005>
- Denzer, M. (2019). Estimating causal effects in binary response models with binary endogenous explanatory variables: A comparison of possible estimators. Gutenberg School of Management and Economics, Johannes Gutenberg-Universität Mainz Working Papers No. 1916. Retrieved June 1, 2024, from https://download.uni-mainz.de/RePEc/pdf/Discussion_Paper_1916.pdf
- Freedman, D. A., & Sekhon, J. S. (2010). Endogeneity in probit response models. *Political Analysis*, 18(2), 138–150. <https://doi.org/10.1093/pan/mpp037>

- Friedman, A. S., & Pesko, M. F. (2022). Young adult responses to taxes on cigarettes and electronic nicotine delivery systems. *Addiction*, *117*(12), 3121–3128. <https://doi.org/10.1111/add.16002>
- Gallant, A. R., & Nychka, D. W. (1987). Semi-nonparametric maximum likelihood estimation. *Econometrica*, *55*(2), 363–390. <https://doi.org/10.2307/1913241>
- Genest, C., & MacKay, R. J. (1986). Copules archimédiennes et familles de lois bidimensionnelles dont les marges sont données [Archimedean copulas and families of bivariate distributions with given marginals]. *Canadian Journal of Statistics*, *14*(2), 145–159. <https://doi.org/10.2307/3314660>
- Hahn, R., Priege, J. E., & Hunt, E. (2020). A multi-disciplinary study of the drivers of smoking cessation in Japan. BOTE Analysis, LLC, for the Foundation for a Smoke Free World, June 2020. Accessed on SSRN at <https://doi.org/10.2139/ssrn.3773737>
- Hampsher, S. (2020). A multi-disciplinary study into the drivers of smoking cessation in Great Britain. Accessed on SSRN at <https://ssrn.com/abstract=3770858>
- Han, S., & Vytlačil, E. J. (2017). Identification in a generalization of bivariate probit models with dummy endogenous regressors. *Journal of Econometrics*, *199*(1), 63–73. <https://doi.org/10.1016/j.jeconom.2017.04.001>
- Hartmann-Boyce, J., Lindson, N., Butler, A. R., McRobbie, H., Bullen, C., Begh, R., Theodoulou, A., Notley, C., Rigotti, N. A., Turner, T., Fanshawe, T. R., & Hajek, P. (2022). Electronic cigarettes for smoking cessation. *Cochrane Database of Systematic Reviews*, *11*(11):CD010216. <https://doi.org/10.1002/14651858.CD010216.pub7>
- Hasebe, T. (2022). Endogenous models of binary choice outcomes: Copula-based maximum-likelihood estimation and treatment effects. *The Stata Journal*, *22*(4), 734–771. <https://doi.org/10.1177/1536867X221140943>
- Hasebe, T., & Vijverberg, W. P. (2012). A flexible sample selection model: A GTL-copula approach. IZA Discussion Paper No. 7003. Accessed on SSRN at <https://ssrn.com/abstract=2177201>
- Hébert-Losier, A., Fillion, K. B., Windle, S. B., & Eisenberg, M. J. (2020). A randomized controlled trial evaluating the efficacy of e-cigarette use for smoking cessation in the general population: E3 trial design. *CJC Open*, *2*(3), 168–175. <https://doi.org/10.1016/j.cjco.2020.03.006>
- Heckman, J. J. (1978). Dummy endogenous variables in a simultaneous equation system. *Econometrica*, *46*(4), 931–959. <https://doi.org/10.2307/1909757>
- Humphreys, B. R., McLeod, L., & Ruseski, J. E. (2014). Physical activity and health outcomes: Evidence from Canada. *Health Economics*, *23*(1), 33–54. <https://doi.org/10.1002/hec.2900>
- Hwang, J. H., & Park, S.-W. (2016). Association between peer cigarette smoking and electronic cigarette smoking among adolescent nonsmokers: A national representative survey. *PLoS ONE*, *11*(10). <https://doi.org/10.1371/journal.pone.0162557>
- Jeon, C., Jung, K. J., Kimm, H., Lee, S., Barrington-Trimis, J. L., McConnell, R., ... Jee, S. H. (2016). E-cigarettes, conventional cigarettes, and dual use in Korean adolescents and university students: Prevalence and risk factors. *Drug and Alcohol Dependence*, *168*, 99–103. <https://doi.org/10.1016/j.drugalcdep.2016.08.636>
- Kang, H., & Cho, S. I. (2020). Heated tobacco product use among Korean adolescents. *Tobacco Control*, *29*, 466–468. <https://doi.org/10.1136/tobaccocontrol-2019-054949>
- Kang, H. G., Kwon, K. H., Lee, I. W., Jung, B., Park, E. C., & Jang, S. I. (2013). Biochemically-verified smoking rate trends and factors associated with inaccurate self-reporting of smoking habits in Korean women. *Asian Pacific Journal of Cancer Prevention*, *14*(11), 6807–6812. <https://doi.org/10.7314/apjcp.2013.14.11.6807>
- Kaye, J. T., Johnson, A. L., Baker, T. B., Piper, M. E., & Cook, J. W. (2020). Searching for personalized medicine for binge drinking smokers: Smoking cessation using varenicline, nicotine patch, or combination nicotine replacement therapy. *Journal of Studies on Alcohol and Drugs*, *81*, 426–435. <https://doi.org/10.15288/jsad.2020.81.426>
- Khouja, J. N., Wootton, R. E., Taylor, A. E., Davey Smith, G., & Munafò, M. R. (2021). Association of genetic liability to smoking initiation with e-cigarette use in young adults: A cohort study. *PLoS Medicine*, *18*(3), e1003555. <https://doi.org/10.1371/journal.pmed.1003555>
- Kim, C.-Y., Paek, Y.-J., Seo, H. G., Cheong, Y. S., Lee, C. M., Park, S. M., ... Lee, K. (2020). Dual use of electronic and conventional cigarettes is associated with higher cardiovascular risk factors in Korean men. *Scientific Reports*, *10*(5612), 1–10. <https://doi.org/10.1038/s41598-020-62545-3>
- Kim, J., Yu, H., Lee, S., & Paek, Y.-J. (2018). Awareness, experience and prevalence of heated tobacco product, IQOS, among young Korean adults. *Tobacco Control*, *27*(S1), s74–s77. <https://doi.org/10.1136/tobaccocontrol-2018-054390>
- Kim, S. K., Park, J. H., Lee, J. J., Lee, S. B., Kim, T. H., Han, J. W., Youn, J. C., Jhoo, J. H., Lee, D. Y., & Kim, K. W. (2013). Smoking in elderly Koreans: Prevalence and factors associated with smoking

- cessation. *Archives and Gerontology and Geriatrics*, 56(1), 214–219. <https://doi.org/10.1016/j.archger.2012.08.018>
- Kim, Y. J. (2014). Predictors for successful smoking cessation in Korean adults. *Asian Nursing Research*, 8(1), 1–7. <https://doi.org/10.1016/j.anr.2013.09.004>
- Kim, Y., Lee, J. S., & Cho, W. K. (2021). Factors associated with successful smoking cessation according to age group: Findings of an 11-year Korea national survey. *International Journal of Environmental Research and Public Health*, 18(4), 1576. <https://doi.org/10.3390/ijerph18041576>
- Lancaster, T., & Stead, L. F. (2017). Individual behavioural counselling for smoking cessation. *Cochrane Database of Systematic Reviews*, (3). <https://doi.org/10.1002/14651858.CD001292.pub3>
- Lee, B., & Seo, D. C. (2021). Effects of an 80% cigarette price increase on quit attempts, successful quitting and smoking intensity among Korean adult smokers: Results from nationally representative longitudinal panel data. *Tobacco Control*, 30(3), 336–343.
- Lee, J. A., Kim, S. H., & Cho, H.-J. (2016a). Electronic cigarette use among Korean adults. *International Journal of Public Health*, 61(2), 151–157. <https://doi.org/10.1007/s00038-015-0763-y>
- Lee, J., Lee, S., & Cho, H.-J. (2017). The relation between frequency of e-cigarette use and frequency and intensity of cigarette smoking among South Korean adolescents. *International Journal of Environmental Research and Public Health*, 14(3), 305. <https://doi.org/10.3390/ijerph14030305>
- Lee, J. E., Park, E. C., Chun, S. Y., Park, H. K., & Kim, T. H. (2016b). Socio-demographic and clinical factors contributing to smoking cessation among men: A four-year follow up study of the Korean Health Panel Survey. *BMC Public Health*, 16, 1–9. <https://doi.org/10.1186/s12889-016-3583-y>
- Lee, S.-H., Ahn, S.-H., & Cheong, Y.-S. (2019). Effect of electronic cigarettes on smoking reduction and cessation in Korean male smokers: A randomized controlled study. *The Journal of the American Board of Family Medicine*, 32(4), 567–574. <https://doi.org/10.3122/jabfm.2019.04.180384>
- Lee, S., Grana, R. A., & Glantz, S. A. (2014). Electronic cigarette use among Korean adolescents: A cross-sectional study of market penetration, dual use, and relationship to quit attempts and former smoking. *Journal of Adolescent Health*, 54(6), 684–690. <https://doi.org/10.1016/j.jadohealth.2013.11.003>
- Lee, Y., & Lee, K.-S. (2019). Association of depression and suicidality with electronic and conventional cigarette use in South Korean adolescents. *Substance Use & Misuse*, 54(6), 934–943. <https://doi.org/10.1080/10826084.2018.1552301>
- Leem, A. Y., Han, C. H., Ahn, C. M., Lee, S. H., Kim, J. Y., Chun, E. M., Yoo, K. H., & Jung, J. Y., Korean Smoking Cessation Study Group. (2017). Factors associated with stage of change in smoker in relation to smoking cessation based on the Korean National Health and Nutrition Examination Survey II-V. *PLoS ONE*, 12(5), e0176294. <https://doi.org/10.1371/journal.pone.0176294>
- Li, C., Poskitt, D. S., & Zhao, X. (2019). The bivariate probit model, maximum likelihood estimation, pseudo true parameters and partial identification. *Journal of Econometrics*, 209(1), 94–113. <https://doi.org/10.1016/j.jeconom.2018.07.009>
- Lumley, T., & Scott, A. (2015). AIC and BIC for modeling with complex survey data. *Journal of Survey Statistics and Methodology*, 3(1), 1–18. <https://doi.org/10.1093/jssam/smu021>
- Maddala, G. S. (1983). *Limited dependent and qualitative variables in econometrics*. Cambridge University Press.
- Maglia, M., Caponnetto, P., Di Piazza, J., La Torre, D., & Polosa, R. (2017). Dual use of electronic cigarettes and classic cigarettes: A systematic review. *Addiction Research & Theory*, 26(4), 330–338. <https://doi.org/10.1080/16066359.2017.1388372>
- Malas, M., van der Tempel, J., Schwartz, R., Minichiello, A., Lightfoot, C., Noormohamed, A., ... Ferrence, R. (2016). Electronic cigarettes for smoking cessation: A systematic review. *Nicotine & Tobacco Research*, 18(10), 1926–1936. <https://doi.org/10.1093/ntr/ntw119>
- Mbrain. (2018). Geumyeonjeongchaek mit jeonjadambae gwallyeon insik josa [2018 smoking cessation policy and e-cigarette perception survey]. Trendmonitor Report TK_201805_TRY4502. Retrieved June 1, 2024, from trendmonitor.co.kr/tmweb/trend/allTrend/detail.do?bIdx=1673&code=0501&trendType=CKOREA
- McCarthy, I., Millimet, D. L., & Roy, M. (2015). Bounding treatment effects: A command for the partial identification of the average treatment effect with endogenous and misreported treatment assignment. *The Stata Journal*, 15(2), 411–436. <https://doi.org/10.1177/1536867X1501500205>
- McGovern, M. E., Bärnighausen, T., Marra, G., & Radice, R. (2015). On the assumption of bivariate normality in selection models: A copula approach applied to estimating HIV prevalence. *Epidemiology (Cambridge, Mass)*, 26, 229.
- Ministry of Health and Welfare, Republic of Korea. (2015, January 15). E-cigarette is cigarette, not smoking cessation aid! [Press release]. Retrieved May 1, 2020, from mohw.go.kr/eng/nw/nw0101vw.jsp?PAR_MENU_ID=1007&MENU_ID=100701&page=1&CONT_SEQ=316042
- Mourifié, I., & Méango, R. (2014). A note on the identification in two equations probit model with dummy endogenous regressor. *Economics Letters*, 125(3), 360–363. <https://doi.org/10.1016/j.econlet.2014.10.006>

- Murteira, J. M., & Lourenço, Ó. D. (2011). Health care utilization and self-assessed health: Specification of bivariate models using copulas. *Empirical Economics*, *41*, 447–472. <https://doi.org/10.1007/s00181-010-0382-7>
- Myung, S. K., Seo, H. G., Park, E. C., Lim, M. K., & Kim, Y. (2011). An observational study of the Korean proactive quitline service for smoking cessation and relapse prevention. *Public Health Reports*, *126*, 583–590. <https://doi.org/10.1177/003335491112600415>
- National Academies of Science, Engineering, and Medicine. (2018). *Public health consequences of e-cigarettes*. The National Academies Press. <https://doi.org/10.17226/24952>
- Nelsen, R. B. (2006). *An introduction to copulas* (2nd ed.). Springer. <https://doi.org/10.1007/0-387-28678-0>
- Oster, E. (2020). Health recommendations and selection in health behaviors. *AER Insights*, *2*, 143–160. <https://doi.org/10.1257/aeri.20190355>
- Park, M.-B., & Choi, J.-K. (2019). Differences between the effects of conventional cigarettes, e-cigarettes and dual product use on urine cotinine levels. *Tobacco Induced Diseases*, *17*. <https://doi.org/10.18332/tid/10052>
- Park, S., Lee, H., & Min, S. (2017). Factors associated with electronic cigarette use among current cigarette-smoking adolescents in the Republic of Korea. *Addictive Behaviors*, *69*, 22–26. <https://doi.org/10.1016/j.addbeh.2017.01.002>
- Pasquereau, A., Guignard, R., Andler, R., & Nguyen-Thanh, V. (2017). Electronic cigarettes, quit attempts and smoking cessation: A 6-month follow-up. *Addiction*, *112*, 1620–1628. <https://doi.org/10.1111/add.13869>
- Pesko, M. F., & Warman, C. (2022). Re-exploring the early relationship between teenage cigarette and e-cigarette use using price and tax changes. *Health Economics*, *31*(1), 137–153. <https://doi.org/10.1002/hec.4439>
- Prieger, J. E. (2002). A flexible parametric selection model for non-normal data with application to health care usage. *Journal of Applied Econometrics*, *17*, 367–392. <https://doi.org/10.1002/jae.638>
- Prieger, J. E., & Choi, A. (2020). A multi-disciplinary study into the drivers of smoking cessation in South Korea. BOTEK Analysis. Accessed on SSRN at <https://doi.org/10.2139/ssrn.3773245>
- Prom-Wormley, E. C., Clifford, J. S., Cooke, M. E., Cecilione, J., Maes, H. H., Do, E., & Roberson-Nay, R. (2020). The genetic and environmental influences contributing to the association between electronic and conventional cigarette initiation. *Nicotine & Tobacco Research*. <https://doi.org/10.1093/ntr/ntaa201>
- Public Health England (2015, Aug. 19). E-cigarettes around 95% less harmful than tobacco estimates landmark review. [Press release]. Retrieved June 1, 2024, from <https://www.gov.uk/government/news/e-cigarettes-around-95-less-harmful-than-tobacco-estimates-landmark-review>
- Rahman, M. A., Hann, N., Wilson, A., Mnatzaganian, G., & Worrall-Carter, L. (2015). E-cigarettes and smoking cessation: Evidence from a systematic review and meta-analysis. *PLoS ONE*, *10*(3), e0122544. <https://doi.org/10.1371/journal.pone.0122544>
- Saffer, H., Dench, D., Grossman, M., & Dave, D. (2020). E-cigarettes and adult smoking: Evidence from Minnesota. *Journal of Risk and Uncertainty*, *60*, 207–228.
- Seo, Y. G., Paek, Y. J., Jo, M. W., & Choi, J. (2019). Predictors of long-term abstinence rate by income level in the Korean smoking cessation programme. *Addiction*, *114*(11), 2056–2064. <https://doi.org/10.1111/add.14726>
- Shahab, L., Goniewicz, M. L., Blount, B. C., Brown, J., McNeill, A., Alwis, K. U., ... & West, R. (2017). Nicotine, carcinogen, and toxin exposure in long-term e-cigarette and nicotine replacement therapy users: A cross-sectional study. *Annals of Internal Medicine*, *166*(6), 390–400. <https://doi.org/10.7326/M16-1107>
- Shaikh, A. M., & Vytlačil, E. J. (2011). Partial identification in triangular systems of equations with binary dependent variables. *Econometrica*, *79*(3), 949–955. <https://doi.org/10.3982/ECTA9082>
- Shiffman, S., Brockwell, S. E., Pillitteri, J. L., & Gitchell, J. G. (2008). Individual differences in adoption of treatment for smoking cessation: Demographic and smoking history characteristics. *Drug and Alcohol Dependence*, *93*(1–2), 121–131. <https://doi.org/10.1016/j.drugalcdep.2007.09.005>
- Snell, L. M., Barnes, A. J., & Nicksic, N. E. (2020). A longitudinal analysis of nicotine dependence and transitions from dual use of cigarettes and electronic cigarettes: Evidence from waves 1–3 of the PATH study. *Journal of Studies on Alcohol and Drugs*, *81*, 595–603. <https://doi.org/10.15288/jsad.2020.81.595>
- Stoklosa, M., Cahn, Z., Liber, A., Nargis, N., & Drope, J. (2019). Effect of IQOS introduction on cigarette sales: Evidence of decline and replacement. *Tobacco Control*, *29*(4), 381–387. <https://doi.org/10.1136/tobaccocontrol-2019-054998>
- Tobacco Use and Dependence Guideline Panel. (2008). Treating tobacco use and dependence: 2008 update. Rockville (MD): US Department of Health and Human Services. Retrieved June 1, 2024, from <https://www.ncbi.nlm.nih.gov/books/NBK63952/>
- Vickerman, K. A., Schauer, G. L., Malarcher, A. M., Zhang, L., Mowery, P., & Nash, C. M. (2017). Reasons for electronic nicotine delivery system use and smoking abstinence at 6 months: A descriptive study of

- callers to employer and health plan-sponsored quitlines. *Tobacco Control*, 26(2), 126–134. <https://doi.org/10.1136/tobaccocontrol-2015-052734>
- Villanti, A. C., Feirman, S. P., Niaura, R. S., Pearson, J. L., Glasser, A. M., Collins, L. K., & Abrams, D. B. (2018). How do we determine the impact of e-cigarettes on cigarette smoking cessation or reduction? Review and recommendations for answering the research question with scientific rigor. *Addiction*, 113(3), 391–404. <https://doi.org/10.1111/add.14020>
- Wang, R. J., Bhadriraju, S., & Glantz, S. A. (2021). E-cigarette use and adult cigarette smoking cessation: A meta-analysis. *American Journal of Public Health*, 111(2), 230–246. <https://doi.org/10.2105/ajph.2020.305999>
- Winkelmann, R. (2012). Copula bivariate probit models: With an application to medical expenditures. *Health Economics*, 21(12), 1444–1455. <https://doi.org/10.1002/hec.1801>
- Wooldridge, J. M. (2014). Quasi-maximum likelihood estimation and testing for nonlinear models with endogenous explanatory variables. *Journal of Econometrics*, 182(1), 226–234. <https://doi.org/10.1016/j.jeconom.2014.04.020>
- World Health Organization. (2019a). The convention secretariat calls parties to remain vigilant towards novel and emerging nicotine and tobacco products. Retrieved June 1, 2024, from <https://fctc.who.int/newsroom/news/item/12-09-2019-the-convention-secretariat-calls-parties-to-remain-vigilant-towards-novel-and-emerging-nicotine-and-tobacco-products>
- World Health Organization. (2019b). WHO global report on trends in prevalence of tobacco use 2000–2025, third edition. Retrieved June 1, 2024, from <https://apps.who.int/iris/rest/bitstreams/1263754/retrieve>
- Zhuang, Y.-L., Cummins, S. E., Sun, Y., & J., & Zhu, S.-H. (2016). Long-term e-cigarette use and smoking cessation: A longitudinal study with US population. *Tobacco Control*, 25(Suppl 1), i90–i95. <https://doi.org/10.1136/tobaccocontrol-2016-053096>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.