Marketing Science

Assessing the Sales Impact of Plain Packaging Regulation for Cigarettes: Evidence from Australia

André Bonfrer, Pradeep K. Chintagunta, John H. Roberts, David Corkindale

Please scroll down for article—it is on subsequent pages

Assessing the Sales Impact of Plain Packaging Regulation for Cigarettes: Evidence from Australia

André Bonfrer,[a] Pradeep K. Chintagunta,[b] John H. Roberts,[c] David Corkindale[d]

[a] Department of Marketing, Deakin Business School, Deakin University, Burwood, Victoria 3125, Australia; [b] Booth School of Business, University of Chicago, Chicago, Illinois 60637; [c] School of Marketing, University of New South Wales, Randwick, New South Wales 2052, Australia; [d] Ehrenberg-Bass Institute for Marketing Science, University of South Australia, Adelaide, South Australia 5000, Australia
Contact: andre.bonfrer@deakin.edu.au, https://orcid.org/0000-0002-3061-7103 (AB); pradeep.chintagunta@chicagobooth.edu,
https://orcid.org/0000-0003-2854-5216 (PKC); johnr@agsm.edu.au, https://orcid.org/0000-0002-8662-584X (JHR);
david.corkindale@unisa.edu.au, https://orcid.org/0000-0002-0106-2954 (DC)

Abstract. We assess the impact of legislation mandating the plain packaging of cigarettes in 2012 in Australia on both primary and secondary demand. We first examine the causal impact of the legislation at the cigarette category level by comparing the changes in sales before and after legislation with the corresponding changes in sales in a comparable market, New Zealand, where the plain packaging mandate (PPM) was not imposed. Our results suggest a decline in sales due to the PPM of around 67 million units (sticks) per month, representing around 7.5% of the market. Our results on the mechanism using brand-level sales data from Australia suggest reduced differentiation after the PPM, with higher price sensitivity. Premium and mainstream brands' price sensitivities are most affected after the PPM, but we also find channel-specific differences, with grocery (convenience) channels showing an increase (a decline) in post-PPM short-term price sensitivity. Because the government has some control over price through excise taxes, understanding changes in price sensitivities provides guidance to health authorities on the relative impacts of price- and non-price-related policy on cigarettes sales. We also explore other public policy implications of our results, such as the expected reduction in sales per month we might see in New Zealand due to their instituting a PPM.

Keywords: cigarette smoking • plain packaging • difference in differences • error correction models

1. Introduction

The tobacco industry has seen a progressive tightening of regulatory interventions in various countries, with the primary aim to reduce and/or control tobacco consumption. Interventions include over-the-counter display and sales controls, advertising bans of tobacco products, introduction of health-warning labels on tobacco packaging, smoking bans in public places, and sales tax increases. A number of researchers suggest that, historically, many of these initiatives have been successful in reducing tobacco consumption (Levy et al. 2004, Capella et al. 2011). However, the degree of success achieved by these measures in terms

of reducing smoking rates has also been questioned by others (Capella et al. 2011). Overall, regulatory bodies have tended to accept the effectiveness of advertising controls and price increases as mechanisms by which to reduce cigarette consumption (e.g., U.S. Department of Health and Human Services 2010, p. 652), leading to what is described as a "dark market" (Dewhirst 2012, p. 516), that is, one largely devoid of manufacturer communication stimuli (Burton et al. 2015).

Despite the regulatory interventions that have greatly limited marketing activities, tobacco marketers still have some control over one effective marketing communication tool: packaging. Given the dark market,

the branding elements on packaging are said to be key marketing communication tools for tobacco products (Ford et al. 2012), and packaging is considered a source of added value to the consumers (Wakefield et al. 2002). As part of this, trademarks on tobacco packaging have played a major market role as the legal properties that protect distinctiveness and thus help identify and differentiate competing products. A leading example is the Marlboro brand, regularly identified as the most highly valued tobacco brand in the world (Millward Brown 2018) and of significant visibility and recognition for both the smoking and nonsmoking public. Unsurprisingly, therefore, recent policy propositions for further restricting available marketing communication tools have focused on regulating trademarks and the packaging of tobacco products (Ford et al. 2012).

The primary objective of this paper is to examine empirical evidence on the impact of the plain packaging mandate (PPM) on the demand for tobacco products in Australia. From December 1, 2012, Australia became the first country to introduce plain packaging for all tobacco brands sold at retail. The PPM requires that marketers in this category adopt package designs that make use of a uniform background color ("drab matte brown"), use a standard font for the brand and variant names, and to contain graphic health warnings that cover 75% of the pack (enlarged from 30%). This is in the context of an industry that has already been restricted from using other branding elements (including point-of-sale displays).[1]

We study the effects of the PPM on cigarette sales using a quasi experiment and based on a difference-in-differences (DiD) analysis (see Angrist and Pischke 2008) with the neighboring country of New Zealand (NZ) as a control market over the time period from January 2011 to December 2013. Whilst Australia instituted the PPM, NZ maintained the status quo and continued to allow marketers to leverage the benefits of packaging as a branding element. As we shall show later, NZ shares many of the features of Australia. Thus, NZ serves as a plausible control group for Australia. Our key identifying assumption is that sales of cigarettes in NZ, not being subject to the PPM, are a valid counterfactual for the sales that would have been obtained in Australia in the absence of the PPM (conditional on the variety of controls that we include in the analysis).

As our central result, an estimated decline in monthly baseline sales of around 67 million sticks is attributable to the PPM. In the Australian market at the time of the PPM implementation, this represents a 7.5% decline in monthly baseline sales. This result is robust to the following: (1) alternative functional form assumption for our analysis, (2) a shorter time period, (3) a placebo treatment, and (4) potential endogeneity of one of our control variables, category price. We estimate that the decline in smoking of 67 million sticks per month is the rough equivalence of the amount smoked by around 135,000 average usage smokers (in a population of approximately 23 million people in 2012).[2] With an estimated two-thirds of smokers dying prematurely of smoking-related diseases (Banks et al. 2015) and the world population of smokers being approximately 1 billion,[3] a global rollout of plain packaging policies with success similar to that which we find for Australia could indicatively deliver a reduction in premature human deaths due to smoking of well over 50 million.

Going beyond our product category analyses, we aim to understand possible mechanisms underlying our above results. Our hypothesis is that plain packaging reduces differentiation between the various brands, thereby raising price sensitivity (and affecting sensitivity to other marketing activities, such as variants being offered by the brands). In testing this hypothesis, we run into challenges associated with our data. Micro-level (consumer-level) data are not available for our two countries, so heterogeneous effects across consumers or segments cannot be explored. Furthermore, brand- and channel-level aggregate data are not available from NZ. Instead, we turn to aggregate brand-level data from Australia to examine whether sensitivity to prices and to the shares of variants accounted for by the various brands changed after the PPM. We recognize that any findings from these data are likely to be suggestive rather than causal because of the absence of an explicit control group.

The results also indicate that after the PPM, particularly for premium and mainstream products, both short- and long-term price sensitivity increase. In terms of brand variant shares, our evidence does not suggest that brands change in their responsiveness to the variant share after the PPM. Furthermore, our evidence suggests differences across *channels* in how much price sensitivity increases, with price sensitivity for convenience channels actually declining in the short term. Given that the grocery channel accounts for most (over 60%) of the category share, average price sensitivity increases across these channels. The higher price sensitivity we observe after the PPM carries important implications for public policy. It is well known that price levels are heavily influenced by excise tax policy (Wang et al. 2015), meaning that after the PPM, the effect of the public policy sword of excise taxes on sales may be considerably sharper. Although our findings do not permit us to attribute this observation to the PPM, this effect is potentially of considerable importance, because public policy typically relies on multiple instruments. This paper, we believe, is the first to document a possible synergistic impact of such policies with those associated with packaging. In 2018, NZ also introduced a PPM. We use our results to provide a benchmark for policy makers in that

country on the reduction in cigarette sales due to such a mandate.

Our above findings that pertain to measuring the impact of the removal of packaging-based differentiation in this dark market allows us to make a number of contributions:

1. Because many other countries since have either adopted or are considering the adoption of plain packaging for tobacco products, providing evidence on its likely impact is of high value to policy makers and marketers. The debate about whether to mandate plain packaging or not will ultimately be driven by the accumulating evidence on its impacts, assessed by challenges in judicial institutions and trade organizations (e.g., the World Trade Organization). The body of knowledge informing these debates is complemented by our findings.

2. In terms of the data and methodology used, those in our study differ from those used in previous research in important ways. Our study is the first to assess the impact on sales (rather than, for example, self-reports). The closest published evidence is an event study by Diethelm and Farley (2015) that examines the impact of plain packaging on smoking prevalence measured with survey data.[4] Although that study is an important step in contributing to our overall understanding of the PPM, the use of survey data could lead to biased estimates of the effects of PPM because respondents' stated behavior could differ from their actual actions. As part of our contribution, and in contrast with Diethelm and Farley (2015), we examine actual sales data to compare the difference in sales before and after the PPM to that in an appropriate "control" market, that is, NZ, that did not have the PPM at that time. In doing so, we also control for the effects of retail prices. With survey data, one might also be concerned about heterogeneity in scale usage (see Gilula et al. 2006). Furthermore, it is useful to study actual sales behavior, rather than self-reported prevalence data, to ascertain the effect on actual demand for cigarettes and the corresponding price elasticities.

3. Our findings provide insights into the role packaging plays in differentiated consumer goods. Packaging (and package communication) has long been understood to play a significant role in the consumer evaluation and decision process (Meyers and Lubliner 1998), particularly in the absence of other manufacturer communication vehicles. While packaging has been argued to be a critical driver of choice, empirical studies evaluating the role that packaging plays in sales response are limited (for an exception, see Vanclay et al. 2011). This study thereby augments the limited empirical evidence available on sales response to packaging changes. In addition, that plain packaging has been adopted in an environment where there is no confound from other branding elements presents a rare opportunity to isolate the role of packaging from that of other branding elements.

This paper is structured as follows. We first discuss the evidence available to regulatory authorities prior to the PPM as to whether plain packaging would be likely to affect brand and category sales and price sensitivity. From there, we move to the data we have for the window of time during which plain packaging was introduced. Our empirical analysis involves a difference-in-differences analysis using data from Australia and having NZ as a control group. In the second half of our empirical analysis, we focus on the Australian market and delve deeper into quality segment sales response models before and after the PPM. Our discussion and conclusion sections reflect upon our findings and draw some possible implications for public policy and for marketers more generally.

2. Literature Review
2.1. Evidence of the Likely Efficacy of Plain Packaging Prior to Its Introduction

We briefly review the weight of evidence available prior to the introduction of plain packaging, which, in the absence of data, suggests the types of impact that the regulator and the regulated might expect. This evidence is both general (what we know about packaged goods as a whole) and specific (what we know about the cigarette category specifically).

Packaging may have a number of features including slogans (Elder and Krishna 2009); colors (Morrot et al. 2001); physical attributes such as size, images, or logos (e.g., Bruce et al. 2013); fonts (e.g., Henderson et al. 2004); and brand names, including descriptors (e.g., adjectives like smooth, low tar, etc.; see Gilpin et al. 2002). A consumer may value these features directly (i.e., find the package attractive) or may infer other attributes from them. Valued attributes influence the likelihood that those contemplating starting smoking will start, a smoker who wishes to quit will quit, and a quitter tempted to relapse will relapse. As discussed, there is little empirical evidence in the academic literature of sales response to packaging changes, but industry sources suggest that it may be influential (e.g., Nielsen Corporation 2015).

Past work in this area has largely focused on two information theory–based themes regarding how packaging can influence category demand and choice among brands: the role in communication and the role in social influence.

2.1.1. The Role of Packaging in Communication. Packaging can be a powerful form of communication of a brand's value proposition, product personality, or function (e.g., see Klimchuk and Krasovec 2006, p. 33) Keller (1993) describes "Customer-based brand equity" as the value added to the product by knowledge

of the brand. Much of this knowledge may be inferred, and one major form of this inference occurs between a product's features, packaging and other marketing communication messages, and the attributes and benefits that the consumer sees it providing (e.g., Huber and McCann 1982). For example, Garber et al. (2000) demonstrate the associations evoked by different colored packaging in four packaged good categories.

2.1.2. Social Influence of Packaging.
In addition to direct effects on the consumer, packaging can also play a role in consumer learning via social influence. Packaging has been shown to play a role in adding value in the user's possession, and also as a mechanism for social influence (Argo and White 2012). The social information role is likely to affect competing brands in the market differentially. With market share driving the probability that a person observes a specific brand, the likelihood of choosing the market leader may be partially determined by inferred popularity signals, driving contagion not linked to any objective quality of the brand (e.g., Tucker and Zhang 2011). With plain packaging, the number of touch points with the company's branding elements reduces dramatically. This shuts down an important mode of communication among consumers: observable quality signals (e.g., Zhang 2010), as it relies on consumers discussing product attributes without tangible reinforcements. Thus, information transfer through social channels may be reduced, further eroding the signal value available through brand equity.

2.1.3. Response of Cigarette Category Specifically to Package Redesign.
There is some direct econometric evidence of the effect of changes in pack design in the cigarette category specifically with respect to the introduction of graphic health warnings (GHWs). For example, Huang et al. (2014), using the United States as a control for Canada, estimated a 2.9 to 4.7 percentage point decline in smoking prevalence as a result of the introduction of GHWs. However, although the majority of studies tend to support the conclusion that they have an effect, others suggest that those effects may be minor (e.g., Bardsley and Olekalns 1999). There are also some laboratory experiments that test the effect of plain packaging on perceptions and appeal. These studies tend to conclude that plain packaging is associated with negative associations (e.g., less attractive taste) and lower appeal (e.g., Wakefield et al. 2008).

2.1.4. Response of Cigarette Brand Shares and Elasticities to Package Redesign.
Perhaps because most of the literature on the PPM intervention is focused on the public health perspective, and thus category demand, there is relatively little literature or empirical research on how brand shares, market structure, and brand prices might be impacted by it. The general consensus seems to be that plain packaging is likely to lower perceived differentiation between brands, and thus lead to lower prices and higher price elasticities (e.g., Freeman et al. 2008).

2.2. Past Work Examining Plain Packaging
In attempting to assess the effect of the Australian government's plain packaging measure on cigarette usage, we first examine prior research on this topic. We review three sources: (1) a monthly survey commencing eight months prior to the measure's introduction and continuing for two years after (the Australian National Tobacco Plain Packaging Tracking Survey [NTPPTS]), (2) a monthly omnibus tracking survey conducted by the market research company Roy Morgan, and (3) sales (scanner) data provided by Nielsen Research. The NTPPTS was designed to measure the effectiveness of the measure in achieving intermediate goals such as decreasing the appeal of cigarettes, increasing the visibility of graphic health warnings, etc.), rather than its effect on consumption. Although Wakefield et al. (2002) show that PPM will be likely to be largely effective in achieving those goals, they are silent as to whether it was associated with any change in tobacco consumption.

The Australian market research firm Roy Morgan ran a syndicated survey of cigarette smoking prevalence rates on a sample of approximately 4,500 respondents aged 14 and older prior and subsequent to the plain packaging measure (as well as other category usage). These data were first analyzed by Kaul and Wolf (2014) under contract for Philip Morris, finding no evidence of a decrease in smoking prevalence attributable to plain packaging. This working paper has attracted considerable controversy partly because of process (it did not acknowledge that Philip Morris had the right to vet its contents or that the terms of the contract with Philip Morris should be kept secret) and partly because of methodological issues (see, e.g., Doward 2015). Although a review of the paper commissioned by the University of Zurich suggested that the working paper not be withdrawn (Jann 2015), the review's author did add "Although I am not happy with all aspects of the papers (see, e.g., Section 2), I do not think that the papers are fundamentally flawed from a methodological point of view. I do not suggest their retraction. There is some space for improvement and some of the interpretations by Kaul and Wolf might be challenged" (Jann 2015, p. 45).

In a peer-reviewed paper reanalyzing Kaul and Wolf's (2014) data, Diethelm and Farley (2015) modified some of Kaul and Wolf's (2014) assumptions (e.g., that the trend of smoking prevalence in

Australia was occurring independent of previous policy changes) and reached a different conclusion. Using what they considered to be more realistic assumptions, they identified a statistically significant decrease in smoking prevalence of 3.7% coincident with the introduction of plain packaging. In a report for the Australian government, an independent econometric consultant (Chipty 2016) found similar results using these data, a decrease in prevalence of 0.55 percentage points (which, given smoking prevalence at the time of 17.77% amounts to a 3.1% decrease in prevalence rates coincident with the change (0.0055/0.1777 = 0.031). This report has since been challenged by Davidson and Silva (2018). Rather than looking for changes in prevalence at the time of the measure, Davidson and Silva (2018) looked for evidence of a break in the series and failed to find one in December 2012. Perhaps, given the gradual effect of any intervention on the addictive behavior of smokers, these two studies may not be as at odds as they appear.

The controversy about the effectiveness of the plain packaging measure using Roy Morgan data gives us the opportunity to consider the issue using a further set of data, point-of-sale retail scanner data from Nielsen Research. This also allows us to move beyond the limitations of self-reports to objectively measure sales and to look at usage volumes rather than just number of users. For reasons we will discuss, our data window is limited to the dates surrounding the PPM date of implementation, from January of 2011 up to December of 2013, with the PPM being implemented in December of 2012. Notwithstanding this limited window of time, the Australian government (and other governments) needs to make policy decisions with respect to plain packaging and its interaction with other policy instruments. In keeping with the objectives of this special issue, our aim is to give the best possible analysis on which to base that policy, within the limitations of the data available.

3. Data

3.1. Australian Data

We have scanner data available from AC Nielsen representing retail transaction sales in Australia. Our retail sales data represent the total amount that smokers buy in any given period for the two largest distribution channels, grocery stores (supermarkets) and convenience stores/forecourt retailers combined. In 2012, these channels collectively accounted for 72.3% of total sales volume per year (Euromonitor Passport 2018a). The remaining 28% is accounted for primarily by specialist tobacco retailers and "independent small grocers," for which detailed scanner data are not available.

Australia comprises six states, and the data are aggregated across these states. There are 42 brands in our database, and for each brand we have unit sales

volume (measured as the number of cigarette sticks), revenues, number of variants per brand, and average transacted retail prices per brand for each four-week period spanning from January 2011 to the end of December 2013 (a total of 39 four-week periods). Henceforth, we will refer to the four-week periods as "months," as distinct from the use of the term "calendar months." The PPM required all tobacco products to be in the new packaging format after December 1, 2012. All monetary amounts are measured in Australian currency (AUD),[5] unless stated otherwise, and are deflated by the Australian Consumer Price Index (CPI, all products), with a base year set to 2009, at 100.

3.2. New Zealand Data

We collected a comparable time series for category-level demand of cigarettes and prices for NZ, which, as we will discuss, is a country similar to Australia on a number of key dimensions but wherein no plain packaging policy existed at the time. This second data set, collected from the Tobacco Control Data Repository[6] in NZ, is also Nielsen data and represents cigarette sales in units from channels comparable to those contained in the Australian data set (supermarkets, smaller convenience outlets). The NZ data include sales units and retail prices on a four-week basis from January 2011 to the end of December 2013. For NZ data, currency values (including excise taxes) are deflated by the CPI and converted to Australian dollars to allow for comparison. In July of 2012, NZ implemented a display ban—the absence of special displays of tobacco products in stores. Such a ban was already in effect in Australia over the entire duration of our data. In the analysis, we explicitly control for the effects of this ban and also assess the robustness of our results to the inclusion of only postban data.

4. Plain Packaging as a Quasi Experiment

Our setting involves a quasi experiment, starting in December of 2012, when Australia implemented a ban on all branding elements on the packaging of cigarettes. This is an exogenous shock to demand, as it was set by the government, rather than, for example, in response to consumer demand for an absence of packaging. It also affected all brands simultaneously, across all channels of distribution and product types. This was against the backdrop of long-standing policy against any form of marketing communication other than packaging, including advertising, displays at point of sale, and digital communications (e.g., websites).

4.1. New Zealand as a Control for Australia

We assess how New Zealand is a suitable control for Australia in several ways. First is a set of qualitative assessments on similarities between Australia and NZ. The use of NZ as a control for Australia is well

established (e.g., see Murphy 1999). Hofstede (2001, exhibit 2.9) shows that Australia and NZ are very similar with respect to a number of economic conditions, including per-capita gross domestic product (GDP), economic growth, and population. Hofstede's (2001) clustering of nations based on cultural similarities uncovers 12 clusters, with Australia and NZ falling into the same cluster (along with the United Kingdom, United States, Canada, and Ireland). An examination of cultural dimensions (power distance, uncertainty avoidance, individualism, masculinity/femininity, and long/short-term orientation; see Hofstede 2001, exhibit A5.1) reveals a similar closeness between Australia and the United States, United

Kingdom, Canada, NZ, and (to a lesser extent) Ireland relative to the other 44 countries calibrated. We also compared Australia and NZ on Schwartz's dimensions of conservatism, hierarchy, mastery, affective autonomy, intellectual autonomy, egalitarian commitments, and harmony (included in Hofstede 2001, exhibit 5.17). Of the 26 countries reported, NZ (along with the United States, Netherlands, Brazil, Mexico, and Poland) is considerably more culturally similar to Australia than are the other 20 countries.

Second, Table 1 more formally compares a variety of metrics to assess the similarity. In particular, the table shows remarkable similarities in prevalence of smoking as well as the nature of regulation (other

Table 1. NZ is a Suitable Control Country for a Difference-in-Differences Model

Descriptor	Australia	New Zealand
Smoking prevalence (Euromonitor Passport 2018c)		
Males (%)	19.3	19.4
Females (%)	16.7	18.3
Tobacco sale regulation	No advertising, promotion, sponsorship; no smoking in public places; health warning on packs; no display of packs at retail; heavy use of excise taxes	
Annual rate of population change (%)	1.1	0.86
Life expectancy, male (at birth)	79.5	79.3
Life expectancy, female	84.5	83
% population +65 years	14.7	14.2
% population <15 years	18.1	19.9
CPI annual change	2.4	0.7
Unemployment (% of labor force)	5.2	5.4
Employment (% of those aged 15–64)	72.3	72.1
Average annual hours worked	1,762	1,683
Shopping access, supermarkets	6 days a week (less on Sunday)	6 days a week (less on Sunday)
Convenience	7 days a week	7 days a week
Fraser's World Rankings[a]		
% home ownership	65	66
Overall economic freedom ranking	8.2	8.48
Size of government	6.57	6.48
Legal system and property rights	8.02	8.72
Sound money	9.26	9.46
Freedom to trade internationally	7.71	8.64
Regulation	8.52	9.16
Hofstede (2001) dimensions[b]		
Power distance	36	22
Individualism	90	79
Masculinity	61	58
Uncertainty avoidance	51	49
Long-term orientation	21	33
Indulgence	71	75
Schwartz scores[b]		
Conservatism	4.06	3.73
Hierarchy	2.36	2.38
Mastery	4.09	4.23
Affective autonomy	3.50	3.98
Intellectual autonomy	4.12	4.36
Egalitarian commitment	4.98	5.15
Harmony	4.05	3.99

  [a]Data are from the Fraser Institute's World Rankings of countries' economic freedom and contributory factors (Fraser Institute 2018).
 [b]Hofstede (2001) dimensions and Schwartz scores are based on product-moment correlations of Schwartz's culture-level scores on seven value dimensions of teachers (Hofstede 2001).

than PPM) across the two countries, and along a series of demographic and economic conditions. The one key distinction appears to be in the CPI metric across countries. To the extent that we control for prices in the analysis, we are able to accommodate this particular deviation. Taken together, the qualitative and the more quantitative similarities provide us with some reassurance that our choice of NZ as a control group is reasonable.

Third, using available data, we assess pre-PPM demand patterns to establish evidence of the common trend assumption. First, we note that from Figure 1, visual inspection suggests that the time series for the two countries before PPM exhibit very similar patterns, including trend and seasonality. A possible exception is where the lines diverge prior to the PPM date, and prior to the 12.5% excise tax late in 2013. We speculate that this may be due to other factors. We did look for other possible influential factors in these periods, but were unable to identify any specific omitted variables.

To assess the strength of linear correlation, we estimate the Pearson correlation (ρ) between the two countries' sales series in the pre-PPM period, that is, from January 2011 up to December 2012. The estimated correlation is $r = 0.68$, with a 95% confidence interval (CI) of [0.39, 0.85]. The estimated correlations using shorter data series are higher. Using data prior to the NZ display ban it is 0.76 with 95% CI = [0.47, 0.90],

Figure 1. (Color online) NZ and Australian Sales Volume (in Sticks) Time Series from 2011 to 2014, with Data Prior to the PPM Showing Similar Patterns Across the Two Countries
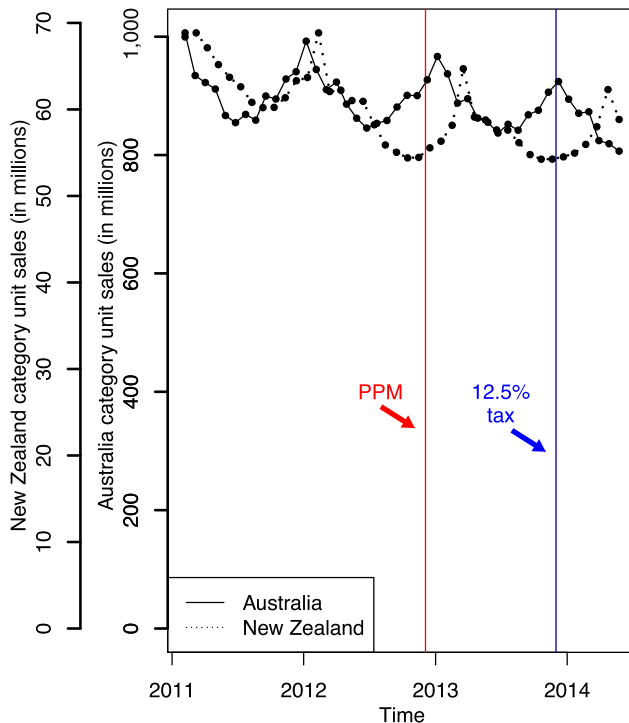
and using data from the display ban until PPM it is 0.86 with 95% CI = [0.15, 0.98]. Although these numbers suggest strong correlation, unsurprisingly they reject the null hypothesis that $\rho = 1$. Therefore, the raw data per se do not exhibit common linear trends. To further investigate this issue, we test the association after controlling for other concomitant factors in the data. This more focused hypothesis test of the presence of a common *linear* trend in the data prior to the PPM is given in Table 2, columns (1)–(5). We return to a discussion of column (6) in the robustness checks section. In particular, we regress the stacked Australia and NZ sales series in the pretreatment period on a number of controls, a time trend, a dummy variable for Australia (*OZd*), and the interaction between the time trend and the Australia dummy. NZ invoked a display ban in July of 2012. Because this represents a significant shift in policy relevant to the PPM it must be accommodated in the analysis as one of the controls. If the common linear trends assumption is satisfied, then the coefficient on the interaction term *Time trend* × *OZd* would not be statistically different from zero. In the absence of any controls, there seems to be some (weak) evidence of an interaction (although not significant at $p < 0.05$). The inclusion of controls makes the effect decline and weakens the statistical significance. With a full set of controls (model (5) in Table 2), the estimate of the interaction *Time trend* × *OZd* also declines in magnitude (and is statistically insignificant), which assures us that conditional on the various controls, the common linear trends assumption is satisfied in the data (Angrist and Pischke 2008). Note the small negative (but not statistically significant) effect of the October–December DiD term, suggesting that the manufacturing ban implemented in that period had a negligible impact on sales prior to the PPM date of December 2012.

4.2. Difference-in-Differences Analysis

We need to measure the sales difference (if any) due to the Australian PPM policy implemented in December 2012 compared with the sales difference (before versus after) of the control market (NZ). Our final data series therefore spans from January 2011 to the end of December 2013. To formally estimate the DiD effect, we can use linear regression (see, e.g., Angrist and Pischke 2008, Goldfarb and Tucker 2014) allowing us to control for several effects. The full specification that provides us with a DiD estimator is

$$Sales_{it} = \lambda_0 + \lambda_1 OZd_i + T(i, t, \Gamma)$$
$$+ \lambda_2 PP_t + \lambda_3 DB_t + \lambda_4 (OZd_i \times DB_t)$$
$$+ \lambda_5 (OZd_i \times PP_t)$$
$$+ \lambda_6 Price_{it} + \lambda_7 Sales_{it-1} + \epsilon_{it}, \quad (1)$$

Table 2. Common Trend Tests for Australia vs. NZ Using Only Data Prior to the Australian PPM

	(1)	(2)	(3)	(4)	(5)	(6)
	Dependent variable: *Unit sales* (millions)					
Time trend	−1.56***	−1.03	2.64*	−0.83	−0.33	−0.39
	(0.44)	(0.83)	(1.14)	(0.74)	(0.44)	(0.40)
OZd	−176.47*	−106.77	−138.50	−96.82	24.38	9.36
	(93.58)	(126.31)	(231.07)	(131.46)	(62.77)	(81.35)
Placebo						12.33
						(14.18)
October–December 2012					29.65*	30.96*
					(15.11)	(15.75)
DB				−43.05	−32.39**	−27.46***
				(27.27)	(11.54)	(9.15)
Category price		−488.87	−377.97	−567.17**	−1,001.38**	−1,039.71**
		(347.87)	(578.73)	(267.02)	(383.80)	(397.26)
Time trend × OZd	1.09*	0.75	1.07	0.71	0.19	0.30
	(0.62)	(0.76)	(1.16)	(0.82)	(0.35)	(0.49)
OZd × Placebo						−15.14
						(34.60)
OZd × October–December 2012					−32.42*	−33.82*
					(17.54)	(18.61)
OZd × DB				35.87	45.51***	38.96
				(33.53)	(13.75)	(22.99)
Constant	1,151.92***	1,323.69***	518.77**	1,335.36***	1,537.92***	1,567.05***
	(65.21)	(68.94)	(217.54)	(42.76)	(135.90)	(147.74)
Monthly fixed effects	No	No	No	No	Yes	Yes
R^2	0.32	0.34	0.96	0.53	0.97	0.97
Adjusted R^2	0.26	0.26	0.93	0.47	0.91	0.91

Notes. Common trend tests are reported based on data only prior to the PPM. Column (1) is a model with only a time trend interacted with the *OZd* dummy. Column (2) adds category price. Column (3) uses data only since the display ban and prior to the PPM. Column (4) uses the full pre-PPM period but adds the display ban difference-in-differences effect. Column (5) adds monthly fixed effects and a difference-in-differences effect to account for the October–December restriction on manufacturing of branded packaging. Column (6) tests the effect of a placebo difference-in-differences effect in the pre-PPM data. Robust autocorrelation and heteroscedastic adjusted standard errors are in parentheses.

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

with $Sales_{it}$ in month t being category unit sales volume for country $i \in \{NZ, OZ\}$, with NZ being NZ and OZ being Australia. The term OZd_i is an indicator variable for Australia (equal to 1 if $i = OZ$), PP_t an indicator for the PPM (equal to 1 after PPM), and DB_t an indicator for the NZ display ban (equal to 1 after the display ban). The parameter measuring the main effect of the OZd_i variable is λ_1 and is to be measured relative to the intercept λ_0 (representing expected baseline sales for NZ). The parameter λ_2 is the main effect of the time dummy representing the common effect on sales coinciding with the period after December 2012. Similarly, the common effect of the post–display ban period is captured by the parameter λ_3 on the DB_t dummy variable. We adopt a set of controls, $T(i, t, \Gamma)$, with Γ a corresponding set of parameters. The term $Price_{it}$ is the category price per stick of cigarettes in country i in period t. We will refer to the variable $Sales_{it-1}$ henceforth as *Lag sales*, with parameter λ_7.

The interaction of OZd_i and PP_t (the parameter λ_5) represents our empirical test of the plain packaging effect on baseline sales quantity in Australia relative to NZ. Similarly, the interaction of OZd_i and DB_t (the parameter λ_4) is the DiD effect of the display ban. We ran a sequence of DiD analyses, the results of which are reported in Table 3. The difference-in-differences estimate for the plain packaging event (λ_5) is given by the row $OZd \times PP$. We also include the DiD value for the display ban enforced in NZ in July of 2012. We begin with an ordinary least squares (OLS) estimation of the basic DiD model with no controls included [model (1)]. Although negative, the DiD effect is not statistically significant even at $p < 0.10$. Adding prices to this specification, however, we find that the DiD effect magnitude is −94.99 and statistically significant at $p < 0.05$. In the remaining columns of Table 3, we successively add more controls. In model (2), we added category price; in model (3), we added monthly fixed effects; and in models (4)

Table 3. Difference-in-Differences Baseline Model, Estimated Using OLS, with Robust (Heteroscedastic and Autocorrelation Corrected) Standard Errors in Parentheses

	Dependent variable: *Unit sales* (millions)					
	(1)	(2)	(3)	(4)	(5)	(6)
OZd	−20.52	15.67	33.58	49.87*	49.86*	33.62
	(17.31)	(11.12)	(23.67)	(28.66)	(29.07)	(24.15)
PP	25.07*	109.14***	80.20***	66.59***	66.60**	80.86***
	(14.10)	(17.60)	(18.38)	(22.40)	(28.42)	(24.17)
DB	−106.61***	−74.55***	−63.65***	−43.13***	−43.14**	−64.18***
	(15.09)	(9.27)	(14.26)	(14.29)	(16.28)	(8.91)
Category price		−831.98***	−707.92***	−574.88**	−574.98**	−708.72***
		(183.93)	(196.50)	(234.95)	(247.26)	(207.74)
Lag sales				0.29	0.29	
				(0.24)	(0.24)	
October–December 2012					−0.001	1.11
					(15.82)	(17.28)
OZd × PP	−25.95	−94.99***	−81.74***	−66.32***	−67.26**	−83.62***
	(18.86)	(17.86)	(16.29)	(21.66)	(27.86)	(21.81)
OZd × DB	85.49***	64.56***	57.48***	39.43***	40.37**	59.26***
	(17.46)	(10.13)	(13.76)	(13.65)	(16.01)	(9.60)
OZd × October–December 2012					−1.87	−3.58
					(17.79)	(19.38)
Constant	928.56***	1,350.37***	1,330.12***	963.99***	964.16**	1,330.52***
	(16.56)	(92.60)	(104.21)	(352.21)	(364.46)	(110.05)
Effect size (%)	−2.9	−10.7	−9.2	−7.4	−7.5	−9.4
Monthly fixed effects	No	No	Yes	Yes	Yes	Yes
Lagged sales	No	No	No	Yes	Yes	No
October–December 2012	No	No	No	No	Yes	Yes
R^2	0.39	0.52	0.89	0.9	0.9	0.89
Adjusted R^2	0.35	0.48	0.83	0.85	0.84	0.82

Notes. All models report OLS results for the DiD model in Equation (1). Column (1) reports the OLS regression results with only DiD effects. Column (2) adds category prices. Column (3) includes monthly fixed effects. Column (4) adds lag sales. Column (5) is the full specification. Column (6) removes lag sales from the model in column (5). Robust autocorrelation and heteroscedastic adjusted standard errors are in parentheses.

$^*p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

and (5) we add lagged sales. In model (5), we also control for a difference in timing between the manufacturing ban and sales ban of highly branded packs. Column (6) is the same as column (5) but without the lagged sales variable.

Our findings hold across these various specifications: The plain packaging effect lowers the sales of cigarettes in Australia. Furthermore, relative to the constant term in the regressions, the magnitude is quite stable and changes (although not statistically significantly) only when we include lagged sales in the analysis. Computing the effect size, we find that the absolute drop in cigarette sticks for the full control specification including lag sales is −67.3 million (or −7.5% of average monthly sales observed in 2012), and without lagged sales included it is estimated at around −83.6 million (or −9.4%). The difference between the estimates without lagged sales versus with can be reconciled on the basis of the longer-term effect on baseline sales carried through by the lagged sales component on future baseline sales.

4.3. Robustness Tests

Several robustness tests were carried out. As we next report, we find that the size of the base estimate [we used model (5) from Table 3] is stable across these different robustness tests.

4.3.1. Accounting for Potential Price Endogeneity. We tested robustness by instrumenting for price using an instrumental variables (two-stage least squares [2SLS]) approach. Although price is not the focus of interest in this analysis, a concern about bias in this parameter may spill over to the parameter of interest. For instruments valid for prices, we use excise taxes on cigarettes (typically measured in dollars per cigarette stick) and the U.S. retail price index. Data for excise taxes from NZ are available from the Treasury of NZ,[7] and those for Australia they are available from Scollo and Bayly (2018). Excise taxes are valid instruments (the "exclusion" restriction) due to their being set by the government for the purpose of revenue generation, assuring their exogeneity. Excise taxes

also represent a large proportion of retail price. There is also considerable precedent for using excise taxes for prices of cigarettes (e.g., Gruber et al. 2003). Furthermore, Stock and Watson 2012, chapter 12) use excise taxes for endogeneity of prices in cigarette demand as a textbook illustration for instrumental variable selection, arguing that "[t]hose choices about public finance are driven by political considerations, not by factors related to the demand for cigarettes" (Stock and Watson 2012, p. 480). Although it cannot be denied that excise taxes are designed to raise prices and increase the cost (and therefore reduce demand) of smoking, we found evidence to suggest that excise tax increases are implemented to raise funds for other purposes as well. For example, as noted in a press release by Australian member of parliament and then treasurer Chris Bowen (Bowen and Plibersek 2013), such purposes are particularly to cover health costs, but also to help federal budgets increase their surpluses (or reduce deficits).

The other instrument used is a seasonally adjusted U.S. retail price index for cigarettes, city average for all urban consumers, available from the U.S. Bureau of Labor Statistics. The validity of the U.S. retail price index as an instrument is based on this being a Hausman-style instrument (see, e.g., Nevo 2001) that captures common cost shocks experienced globally by cigarette brands, typically arising from common global energy, labor and capital cost, transportation, and commodity price shocks. There is some possibility they may have a more direct effect on demand because of correlated global demand, so we tested that and found a weak correlation of 0.22 in annual GDP growth (data from the World Bank) between Australia and the United States, with that from 1961 to 2017 being not statistically significantly different from zero ($p < 0.05$). Further testing of just the period from 2001 to 2012 (before the PPM) also does not reject the test of zero correlation among the countries.

In the three columns of Table 4, we report the first-stage regressions (i.e., category price regressed onto instruments and exogenous variables) to assess the strength and face validity of these instruments. In terms of strength, the F-statistics of these instruments in the 2SLS regression are $F_{(1,46)} = 159.47$ for the excise tax variable alone, $F_{(1,46)} = 22.59$ for the U.S. cigarette price index alone, and $F_{(2,45)} = 88.54$ for both instruments combined. Each of these F-statistics is well above the rule of thumb of 10 suggested by Stock et al. (2002), as a test to assure that minimal bias effect remains. Exogeneity of these instruments can be assessed on the basis of the Sargan overidentification test [only for both instruments combined and under 2SLS; see model (3) in Table 5], which is $\chi^2 = 0.717$ not rejected for these instruments ($p > 0.1$). In terms of

Table 4. First-Stage Regressions for 2SLS Demonstrating Strength for Each of the Instruments in Isolation and Combined

	(1)	(2)	(3)
	Excise tax	*U.S. price*	Both
OZd	0.0265	0.0339	0.0292
	(0.0195)	(0.0291)	(0.0196)
PP	0.0138*	0.0699***	0.0148*
	(0.00795)	(0.0122)	(0.00847)
OZd × PP	0.00121	−0.101***	−0.0109
	(0.0100)	(0.0108)	(0.0111)
October–December 2012	0.00808***	0.0287**	0.0115***
	(0.00260)	(0.0126)	(0.00392)
OZd × October–December 2012	0.00295	−0.0219	0.000125
	(0.00964)	(0.0152)	(0.00903)
DB	−0.000867	−0.0116	−0.00252
	(0.00917)	(0.0166)	(0.00977)
OZd × DB	0.0144*	−0.00919	0.00782
	(0.00850)	(0.0178)	(0.00923)
Lag sales	0.000141	−0.0000371	0.000170
	(0.000124)	(0.000190)	(0.000129)
Excise tax	1.597***		1.401***
	(0.126)		(0.142)
U.S. price		1.784***	0.459***
		(0.375)	(0.111)
Constant	−0.127	−1.364**	−0.587***
	(0.167)	(0.533)	(0.204)
Monthly fixed effects	Yes	Yes	Yes
N	78	78	78
F-statistic	159.47	22.59	88.54

Notes. The first two columns report first-stage results for each instrument in isolation. The third column reports the results for both instruments combined. Robust autocorrelation and heteroscedastic adjusted standard errors are in parentheses. The dependent variable is category price.
  *$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.

face validity of the first stage, both excise tax and U.S. price have the expected positive impact on retail prices, with excise tax appearing to have a larger effect size than U.S. price changes. U.S. price has a smaller size impact than excise taxes. Excise taxes for Australia (the parameter for *Excise tax × OZd*) have a smaller absolute impact on prices than those of NZ.

Using these instruments, in Table 5 we find that the size of the plain packaging effect changes by a small amount when accounting for potential price endogeneity. The Wu–Hausman χ^2 test statistics are 14.90 ($p < 0.01$) for both of the instruments (excise taxes and U.S. prices), 7.526 ($p < 0.01$) for excise taxes alone, and 12.41 ($p < 0.01$) for the U.S. cigarette price index alone. These tests do indicate the presence of endogeneity remaining in the price variable after we control for many observables. We do see a substantial decline (i.e., attenuated toward zero) in price sensitivity when using instruments. However, none of these tests resulted in a result for the DiD effect qualitatively different

Table 5. Results for the Baseline Model Estimated with 2SLS and LIML, Instrumenting for Category Price, Suggest that Some Endogeneity Bias Exists but Affects Mainly the Price Sensitivity Parameter

	(1)	(2)	(3)	(4)	(5)
	2SLS tax	2SLS U.S. price	2SLS both	LIML	Reduced form
Category price	−493.1***	−582.3***	−505.8***	−505.1*	−1123.3***
	(174.2)	(151.5)	(168.7)	(100.3)	(651.8)
OZd	48.23	50.01*	48.48	48.47***	65.05
	(30.61)	(28.25)	(30.26)	(15.08)	(26.64)
PP	57.28***	67.43***	58.72***	58.64***	66.00***
	(19.29)	(17.63)	(18.76)	(16.73)	(26.64)
OZd × PP	−58.78***	−68.02***	−60.09***	−60.02***	−58.60*
	(18.82)	(17.19)	(18.34)	(20.17)	(29.54)
October–December 2012	−1.793	0.159	−1.516	−1.532	3.47
	(11.84)	(12.14)	(11.87)	(16.62)	(17.99)
October–December 2012 × OZd	0.0698	−2.040	−0.230	−0.212	1.79
	(14.45)	(14.63)	(14.46)	(23.58)	(17.48)
DB	−42.39***	−43.20***	−42.50***	−42.50***	−43.01**
	(14.37)	(13.49)	(14.24)	(13.99)	(14.22)
DB × OZd	38.89***	40.50***	39.12***	39.10**	47.64**
	(14.04)	(13.15)	(13.90)	(18.25)	(16.01)
Lag sales	0.324*	0.287*	0.319*	0.319***	0.41*
	(0.178)	(0.161)	(0.175)	(0.0903)	(0.179)
Excise tax					997.00
					(889.1)
U.S. price					21.9
					(438.44)
Constant	888.5***	970.9***	900.2***	899.5***	786.3
	(262.1)	(233.2)	(256.8)	(124.6)	(474.9)
Monthly fixed effects	Yes	Yes	Yes	Yes	Yes
N	78	78	78	78	78
R^2	0.903	0.904	0.904	0.904	0.910
Adjusted R^2	0.838	0.840	0.839	0.839	0.840
Sargan χ^2			0.717		
Sargan *p*-value			0.397		
Wu–Hausman χ^2	7.526***	12.41***	14.90***		
Wu–Hausman *p*-value	0.00608	0.000427	0.000582		

Notes. The first two columns report regression results correcting for price endogeneity using each instrument in isolation. The third column reports the results for both instruments combined. The fourth column reports results for the limited information maximum likelihood model. The last column reports reduced form OLS estimates, adding excise tax and U.S. price as covariates to the baseline model. Robust autocorrelation and heteroscedastic adjusted standard errors are in parentheses.

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

from that of the baseline model. The limited information maximum likelihood (LIML) estimate [model (4) in Table 5] indicates that the effect of plain packaging is very similar to that in the baseline model, and price estimates are similar to those using 2SLS estimation, which does not suggest any weak instrument issue. Finally, the last column runs the DiD regression with covariates added for excise tax and U.S. prices. Although the size of the effect drops a little, and its statistical significance is reduced (but still significant at $p < 0.10$), we get some assurance of the expected sign for the U.S. price and excise tax instruments.

We report the results of the next few robustness tests in Table 6:

1. *Functional form assumptions.* Included in these robustness tests, we redid our baseline analysis using a log-linear model, where the PPM DiD estimate is interpreted as a percentage value. We also tried adding a polynomial (cubic) to allow for time trends. As Table 6 indicates, our results are robust to these changes.

2. *Using only post–NZ display ban data.* Next, we shortened the time series to include only data from after the display ban in NZ. Testing a DiD specification using only data since the NZ display ban [see results for models (4) and (5) in Table 6] increases the size of the effect (and naturally also the standard error). However, the effect size remains statistically significant.

Table 6. Results from Robustness Checks for the Difference-in-Differences Model

	(1)	(2)	(3)	(4)	(5)
	Time trend	Cubic	Logged cubic	DB = 1	DB = 1, logged
Category price	−324.4	−646.9***	−0.390***	−657.7***	−0.481***
	(209.9)	(217.0)	(0.131)	(158.5)	(0.103)
OZd	34.07	33.11	16.73***	71.95***	16.51***
	(33.72)	(25.49)	(2.914)	(11.06)	(2.497)
PP	55.83***	54.42***	0.0623***	159.1***	0.178***
	(19.10)	(15.99)	(0.0170)	(26.42)	(0.0282)
OZd × PP	−43.01**	−78.06***	−0.0817***	−74.35***	−0.0918***
	(21.91)	(22.64)	(0.0232)	(19.08)	(0.0208)
October–December 2012	−5.946	−6.021	−0.00167	27.36	0.0328*
	(12.14)	(12.96)	(0.0139)	(17.15)	(0.0185)
October–December 2012 × OZd	4.212	−1.622	−0.00433	13.24	0.00862
	(14.89)	(15.89)	(0.0172)	(18.06)	(0.0197)
DB	−39.10***	−57.86***	−0.0617***		
	(13.83)	(19.10)	(0.0203)		
DB × OZd	50.25***	60.63***	0.0674***		
	(16.01)	(13.56)	(0.0146)		
Time trend	−0.427**	12.62	0.00515	235.3**	0.269**
	(0.214)	(22.28)	(0.0234)	(104.6)	(0.117)
Time trend²		15.75***	0.0144***	−689.7***	−0.758***
		(5.525)	(0.00533)	(173.9)	(0.197)
Time trend³		−7.518	−0.00500	376.7***	0.413***
		(9.911)	(0.0104)	(84.72)	(0.0955)
Lag sales	0.193	−0.0268	−0.0108	−0.0439	0.00554
	(0.197)	(0.184)	(0.177)	(0.155)	(0.151)
Constant	997.0***	1317.3***	4.195***	1189.9***	3.969***
	(265.3)	(277.7)	(0.753)	(94.20)	(0.603)
Monthly fixed effects	Yes	Yes	Yes	No	No
N	78	78	78	40	40
R2	0.90	0.92	0.99	0.88	0.99
Adjusted R2	0.84	0.86	0.99	0.84	0.99

Notes. Including a polynomial time trend component for both linear and log models yields similar findings, and shortening the data series to only after the NZ display ban increases the size of the plain packaging effect. In each case, the dependent variable is unit sales. Column (1) includes a linear time trend, and column (2) has a cubic time trend. Column (3) uses a log-linear model where both unit sales and prices are logged, and also includes the cubic time trend. Columns (4) and (5) both use data only since the NZ display ban was introduced. Column (4) is the linear model (unlogged unit sales) with the cubic trend, and column (5) is the log-linear model. Robust autocorrelation and heteroscedastic adjusted standard errors are in parentheses.
   $*p < 0.1$; $**p < 0.05$; $***p < 0.01$.

3. *Placebo intervention.* A potential issue associated with using time series variation to establish a causal effect is the presence of other temporal factors that could be masquerading as the effect of the intervention. In our case, this is mitigated by the inclusion of flexible time effects (see above robustness results). Furthermore, the common trend analysis reported previously is also reassuring in this regard. Nevertheless, we further investigate this issue. Specifically, using data only on the preintervention period, we introduced a placebo intervention for the months April to June of 2012. The results are reported in column (6) of Table 2. Once again, this analysis did not reveal the presence of non-intervention-related temporal affects; that is, we find the placebo intervention to be statistically insignificant and also not large in magnitude.

4.4. Potential Mechanisms for Changes in Behavior

As noted in the literature review, product packaging serves as a means of differentiating products in the market. The PPM then lowers differentiation, which can make consumers more price elastic. Another dimension along which brands are differentiated is the number of variants they carry; PPM might have also affected consumer sensitivity to these variants. Such responses to marketing mix variables can be short term or long term, or a combination of the two. The short-term effects reflect the impact on the immediate purchase, suggesting an effect at the point of sale, whereas the long-term effects are indicative of a more persistent effect on sales. In this section, we turn to look at data at the brand (not the category) level to shed some light on the potential mechanisms driving the results identified in the previous section.

Before proceeding with such an analysis, however, several issues are worth highlighting. First, if price sensitivity changed with the PPM, then, in principle, we can include a three-way interaction effect in our previous DiD analysis between price, the Australia dummy, and the PPM variable to reflect this possibility. Although we did include such an interaction, and found that it directionally supported our hypothesis, the short time series and aggregate country-level variation did not afford us the statistical power to find a statistically significant effect across various model specifications. With brand level data, we can leverage the additional cross-sectional price variation to better assess our hypothesis of increased price sensitivity. The second issue pertains to brand variants. If these influence brand shares, the issue arises as to why we did not test their effect at the category level. One reason is the absence of appropriate brand-specific data for the NZ market. Instead, we assessed whether the total number of variants (i.e., the sum of variants across each of the brands in the cigarette category) was a significant driver of category sales using the above category data only for Australia. We found a very small and statistically insignificant effect, so our category analyses excluded this variable. A related issue we face when looking at the number of variants at the brand level is that some of the brands show no changes over time in the number of branded variants they have, so they cannot be directly included in the analysis. Therefore, we operationalize this variable as the proportion (or share) of variants for each brand in each time period. This generates some variation, although, as we show below, the number of variants continues to have a very small effect even at the brand level.

To understand how sensitivity to price and variant share may have changed after the PPM, we build a brand-level unit sales response model, calibrated on data for Australia. Because data are available at the channel level (grocery and convenience), we study brand sales in each channel over time. In the brand response model, we study the interaction of the plain packaging dummy variable with each of the price and variant share sensitivity parameters. We considered two model specifications: the first is an error-correction model (ECM; for details, see Fok et al. 2006, van Heerde et al. 2013). This model allows us to measure each of the long- and short-term effects of the variables. The second model is a more conventional linear sales response model along the lines that we used in the difference-in-differences analysis above. Because the ECM generalizes the linear model (in log terms), we focus on that specification here and report the results from the other functional form in the online appendix.

For each channel, we specify a set of B equations ($b = 1, \dots, B$), with sales for brand b at time t being represented by (omitting channel subscripts)

$$
\begin{aligned}
\Delta \ln Sales_{bt} \\
= \beta_{0b} &+ T(b, t, \Lambda) + \beta_{1b} \Delta \ln Price_{bt} \\
&+ \beta_{2b} \Delta \ln Comp\, price_{bt} + \beta_{3b} \Delta \ln Variants_{bt} + \varphi_b \\
&\times \begin{pmatrix} \ln Sales_{b,t-1} - \beta_{4b} \ln Price_{b,t-1} \\ \beta_{5b} \ln Variants_{b,t-1} - \beta_{6b} \ln Comp\, price_{b,t-1} \end{pmatrix} + \epsilon_{bt},
\end{aligned}
$$
(2)

where Δ is a first difference operator (e.g., $\Delta \ln Price_t = \ln Price_t - \ln Price_{t-1}$), *Price* is the deflated price, and *Variants* is the number of variants expressed as a share of the total variants available at that point in time. The term "comp" is used to denote average competing brands' values for the same marketing mix variable. For example, *Comp price* is the average value of the competing brands' prices. Because we use brand variant share, we do not require any variable for competing brands' variants. The error term is jointly normally distributed ($\epsilon_{bt} \sim N(0, \Sigma)$), capturing cross-brand covariation. (For more complete details, see Fok et al. 2006 or van Heerde et al. 2013.) The "error correction" component is in the parentheses and is premultiplied by φ. The φ parameter represents a "speed of adjustment" toward a long-term equilibrium in any change of the variables included in the parentheses. This parameter value is identified from the lagged sales covariate $\ln Sales_{bt-1}$. While we initially obtain values for the product of φ and various β parameters, we obtain the latter by dividing through by φ and applying the delta method to calculate standard errors of this ratio. For the time component, $T(b, t, \Lambda)$, we include a trend component and the first of the harmonic basis functions (a more parsimonious representation than fixed effects for individual months). This first harmonic function will account for intrayear cyclical effects (e.g., seasonality). The parameters β_{1b} and β_{3b} account for short-term price and variant share, and β_{4b} and β_{5b}, respectively, represent the corresponding long-term effects. Short- and long-term sales responses to competitive prices are captured by the parameters β_{2b} and β_{6b}, respectively.

We selected the top brands (a total of 13), and we added to this the quality-tier-specific composites of all other brands. (Cigarettes are considered by manufacturers to be classifiable into three quality tiers: value, mainstream, and premium.) This means we have a total of 13 + 3 = 16 different brands inclusive of the composites. The brand-specific data represent around 99.2% of the total sales of cigarettes across the two channels. The top three brands (Winfield, Longbeach, and Peter Jackson) represent collectively around

50% of the volume share of the category, with leading share brand (Winfield) having an average of 21% share. The 13 noncomposite brands collectively represent 93.7% of the market share (the composites with 5.5%).

An assumption justifying the use of error correction methods is stationarity of the time series. We therefore conducted several tests for stationarity. Consistent with evidence available in past literature on packaged goods, and particularly for error correction models (e.g., van Heerde et al. 2013), we find evidence for stationarity for each of the sales, variant share, and price time series variables. The tests we conducted were as follows. Unit root tests for individual time series [the test by Phillips and Perron (1988) and that by Kwiatkowski et al. (1992) with intercept, lags, and trend included] suggest stationarity for all brands in the convenience channel, and for 15 out of the 16 sales series in the grocery channel. Panel-based unit root tests have more power, so we apply these tests to the system of equations in (3) set up as a panel. These include the Levin et al. (2002) test with brand-specific intercepts and trends, as suggested by van Heerde et al. (2013), rejected for grocery ($z = -3.73, p = 0.00019$) and convenience ($z = -5.19, p = 2.1e-07$) channels. A panel-based unit root test that takes into account that errors may be correlated between brands is the Im et al. (2003) test. For this test, we find that, with brand-specific intercepts and trends, for the grocery channel, $z = -3.44$ and $p = 0.00058$, and for the convenience channel, $z = -4.07$ and $p = 4.7e-05$. Both are rejected, finding evidence for stationarity. For additional robustness, we also ran the model in (3) using a simple log-linear but non-error-correcting specification. The only "dynamics" included in that model were control variables for time trend and lagged unit sales.

The main results for the brand response models are presented in Table 7.[8] All estimates have considerable face validity, with positive (but low) response to variant share, low negative own price elasticities, and positive cross price elasticities. Prices are inelastic (i.e., below one in absolute value), consistent with past work on tobacco and other vice goods (e.g., Baltagi and Griffin 1995, Gallet and List 2003, Chen et al. 2009). Although low, these estimated price elasticities are consistent with past meta-analytic work (e.g., Bijmolt et al. 2005). The low price elasticities before the PPM indicate that brands were relatively impervious to price changes, indicative of a likely combination of high levels of perceived differentiation and low primary demand sensitivity to price. The latter we established earlier in our DiD results.

The bottom four rows of Table 7 present results for how much each of the short- and long-term pricing and variant share effectiveness changed after the PPM. We discuss only the last two columns and focus attention more on the grocery channel as indicative of any major changes in the market, owing to the larger share of sales volume this channel accounts for. From these results we see that price sensitivity, both in the long term and short term, increased significantly ($p < 0.05$). We do not see any change in short- or long-term sensitivity to variant share. Indeed, there appears to be little effect of variants on share except perhaps in the convenience channel. Although the convenience channel accounts for a smaller fraction of total sales, there is a similar increase in price sensitivity for the long run. However, the short-term price effectiveness declines (toward zero) in that channel.

To understand where the main changes may be coming from, in Table 8 we present results split by type of brand based on the quality tiers. We focus here on baseline sales effects and sales responses to prices and variant share. We first note (see row *PP*) that baseline sales declined in premium and mainstream brands (particularly in the grocery channel), but *increased* toward value brands. This suggests a shift in sales toward such value brands. This is consistent with the removal of the ability for consumers to distinguish between such classifications other than by price. We further observe that both long- and short-term price sensitivity for premium and mainstream brands increased substantially.

Overall, the brand results suggest that consumers became more price sensitive after the PPM, consistent with a lower perceived differentiation among brands. The results averaged by product type suggest that consumers became more price sensitive, especially in the premium and mainstream segments of the market (Table 8, last two rows). Premium products have higher perceived differentiation before the PPM, so such products are likely to see a greater reduction in perceived differentiation, and therefore a larger increase in price sensitivity. Another explanation for the differential impact of premium and mainstream brands lies in attracting new smokers to the category. New customers tend to be drawn to this category because of the glamor and visibility of consumption of products in social settings (e.g., Machado and Sinha 2007).[9] In such settings, we suspect that premium and mainstream brands are more likely to be influential than value brands. However, after the PPM, the ability for brands to transfer this brand equity by social contagion is likely to be significantly thwarted, simply because brands are not as distinctive and memorable as they were before the PPM.

Two possible bright spots for tobacco marketers appear to be the value brands and the convenience channel, where price sensitivities did not go up. For value brands, however, with low prices and preexisting higher price elasticities (particularly long term), it may be more difficult to bring about profitable price reductions because of lower margins (e.g., see

Table 7. Price Sensitivity Increased After the PPM Both in the Short and Long Term, but Is Not Sensitive to Variant Share

	Dependent variable: *Brand unit sales*					
	No *PP* interaction		*PP* for price		Full *PP* interaction	
	Grocery	Convenience	Grocery	Convenience	Grocery	Convenience
Intercept	−0.03***	0.02*	−0.04***	0.02	−0.05***	0.01
	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)
PP	−0.01**	−0.00	−0.04***	0.02**	−0.02***	0.03***
	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
October–December 2012	−0.01	−0.00	−0.02**	0.01	−0.02**	0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Time trend	−0.01	−0.01	−0.04***	−0.03***	−0.05***	−0.03**
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Lag sales	−0.35***	−0.16***	−0.51***	−0.25***	−0.57***	−0.28***
	(0.02)	(0.02)	(0.03)	(0.03)	(0.04)	(0.03)
Short-term price	−1.52***	−1.52***	−0.75***	−0.15	−0.76***	−0.22
	(0.17)	(0.21)	(0.10)	(0.15)	(0.13)	(0.19)
Short-term variants	−0.01	0.06***	0.01	0.03	0.02	0.04
	(0.03)	(0.02)	(0.03)	(0.02)	(0.04)	(0.03)
Short-term competitive price	0.81***	0.78***	0.54***	0.87***	0.64***	1.06***
	(0.17)	(0.22)	(0.17)	(0.25)	(0.21)	(0.33)
Long-term price	−3.16***	−4.61***	−1.76***	−2.93***	−1.88***	−3.45***
	(0.42)	(0.96)	(0.29)	(0.69)	(0.31)	(0.80)
Long-term variants	−0.05	0.45***	0.00	0.15*	−0.01	0.11
	(0.06)	(0.14)	(0.04)	(0.09)	(0.05)	(0.13)
Long-term competitive price	2.42***	3.52***	1.48***	2.26***	1.55***	2.78***
	(0.40)	(0.92)	(0.28)	(0.64)	(0.30)	(0.77)
PP × Short-term price			−0.51***	1.02***	−0.47***	0.97***
			(0.13)	(0.19)	(0.15)	(0.22)
PP × Short-term variants					0.07	0.07
					(0.08)	(0.06)
PP × Long-term price			−0.89***	−0.64**	−0.73***	−0.56*
			(0.15)	(0.28)	(0.16)	(0.31)
PP × Long-term variants					0.07	0.23
					(0.13)	(0.22)
Seasonal harmonic effects	Yes	Yes	Yes	Yes	Yes	Yes
McElroy's R^2	0.68	0.68	0.75	0.70	0.72	0.69
Likelihood ratio	1,078.3***	939.6***	1,088.5***	862.7***	1,047.4***	854.6***
J-statistic	4.81	4.95	4.51	4.50	3.87	3.72
χ2	0.96	0.96	0.95	0.95	0.95	0.96
Number of observations	928	928	928	928	928	928

Notes. Estimation results for the error correction model in Equation (3) are shown. Reported estimates are averages across brands within channels. The first two columns are models estimated without interactions with the plain packaging dummy variable. The middle two columns include the PPM interaction only for prices, and the last two columns include full interaction for both long- and short-term variants and prices. Robust autocorrelation and heteroscedastic adjusted standard errors are in parentheses.

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

Dreze et al. 1994), and these segment effects do not outweigh the negative effects at the higher end of the market and in the larger grocery channel. One interesting potential explanation for the decline in short-run price sensitivity for the convenience channel is as follows. The convenience channel is where new smokers incubate—they tend to be "social" smokers at first, and because they do not know the category, they use higher prices as a signal more when the brand elements are taken away. This explanation is consistent with the low price sensitivity in this channel prior to the PPM, particularly

for premium brands (see the columns for convenience in Table 7).

4.5. Implications for Public Policy

Our analysis suggests that the plain packaging measure was an effective policy intervention in terms of decreasing cigarette sales, and our results also suggest a post-PPM increase in price sensitivity. Furthermore, any policy that results in increases in prices (e.g., excise tax) may have increased potency in its ability to reduce the quantity demanded after the PPM,

Table 8. Sales Shift from Premium to Value Products, Particularly in the Grocery Channel

| | Dependent variable: *Brand unit sales* | | | | | |
| | Grocery | | | Convenience | | |
	Value	Mainstream	Premium	Value	Mainstream	Premium
Intercept	−0.00	0.39***	−0.28***	−0.02	0.35***	−0.00
	(0.01)	(0.11)	(0.04)	(0.02)	(0.09)	(0.03)
PP	0.02*	−0.08***	−0.06***	0.08***	−0.01	0.01
	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)
October–December 2012	0.01	−0.04***	−0.02*	0.01	−0.00	0.01
	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)
Time trend	0.06***	−0.09***	−0.09***	−0.05*	−0.03*	−0.02
	(0.02)	(0.02)	(0.02)	(0.03)	(0.02)	(0.02)
Lag sales	−0.25***	−1.01***	−0.59***	−0.15***	−0.45***	−0.29***
	(0.04)	(0.08)	(0.05)	(0.03)	(0.06)	(0.05)
Short-term price	−0.75***	−0.62***	−0.82***	−0.53*	−0.13	0.20
	(0.17)	(0.22)	(0.16)	(0.27)	(0.25)	(0.27)
Short-term variants	−0.03	0.10	0.00	0.08	−0.01	0.03
	(0.05)	(0.07)	(0.04)	(0.05)	(0.05)	(0.02)
Short-term competitive price	1.32***	0.06	0.10	1.86***	0.40	0.09
	(0.28)	(0.33)	(0.28)	(0.41)	(0.40)	(0.52)
Long-term price	−6.44***	−0.35	−0.67	−7.07***	−1.69**	−1.03
	(1.25)	(0.28)	(0.40)	(2.15)	(0.68)	(1.07)
Long-term variants	0.00	0.09*	−0.09	0.70**	−0.04	0.06
	(0.12)	(0.05)	(0.06)	(0.34)	(0.13)	(0.09)
Long-term competitive price	6.23***	0.13	0.45	6.82***	1.15*	0.09
	(1.24)	(0.26)	(0.39)	(2.07)	(0.64)	(1.00)
PP × Short-term price	−0.02	−0.68**	−0.86***	1.35***	0.77**	0.94***
	(0.21)	(0.28)	(0.19)	(0.37)	(0.31)	(0.29)
PP × Long-term price	−0.09	−0.66***	−1.11***	0.54	−0.61**	−0.72*
	(0.51)	(0.13)	(0.22)	(1.17)	(0.24)	(0.38)

Notes. Only for grocery did price sensitivity increase more for premium and mainstream brands. For convenience, short-term price sensitivity decreased for all brand types, and long-term price sensitivity increased for only value and premium products. Estimation results are based on Table 7. The dependent variable here is brand-level unit sales within each channel. The averages for these results are taken across brands within product segments. Robust autocorrelation and heteroscedastic adjusted standard errors are in parentheses.
  *$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.

certainly at the top end of the market and through the grocery channel. We speculate that this is a plausible outcome of the PPM.

To put the main effect of the PPM into perspective, we consider the retail price change it would take to decrease sales by the expected amount of 67 million (7.5%). With price elasticity defined as $\eta_p = \%\Delta Sales/\%\Delta Price$, and given the pre-PPM price elasticity, a target price change $\%\Delta Price^*$, required to bring about a target percentage change in sales of $\%\Delta Sales^*$, is given by $\%\Delta Price^* = \%\Delta Sales^*/\eta_p$. Given the estimated category sales price elasticity of −0.39 [see the results for the log-linear model in column (5) of Table 6], the target price change to bring about a 7.5% decline in demand is equal to $0.075/0.39 = 19.2\%$. This is a substantial increase in price, likely requiring major increases in excise taxes to it bring about. However, there may also be other qualitative reasons that plain packaging brings about a more favorable outcome than that of price increases. For example, consumer

welfare arguments could be made that the higher prices tend to hit consumers with lower socioeconomic status more severely, consumers considered to be more vulnerable, with higher prevalence rates and lower capability to quit smoking. Another reason why plain packaging may be a more appealing mechanism in the public policy arsenal is that it is less likely than excise taxes to affect votes by smokers at election time.

Starting on March 14, 2018, NZ implemented a similar plain packaging policy, meaning that tobacco companies in NZ were subsequently not able to manufacture and sell cigarettes in branded packaging. Can we use the results from our Australia analysis to predict the potential implications of PPM for NZ? From Euromonitor's Passport database (Euromonitor Passport 2018b), sales of cigarettes in NZ in 2017 were 1.8 billion sticks. From our results, and under conditions comparable with Australia at the time they introduced the PPM, if NZ were to implement plain

packaging, we expect a decline of 7.5% in category sales volume. In particular, the prevalence rate in Australia was 0.194 at the time the Australian PPM was introduced in December of 2012. Given that prevalence in NZ was 0.166 (Euromonitor Passport 2018c) at the time of the NZ PPM implementation, we should adjust for the lower prevalence rate. We accomplish this by multiplying the anticipated decline by the lower fraction of smokers (a factor of 0.166/0.194). Thus, the anticipated decline in NZ sales volume is $0.166/0.194 \times 0.075 \times 1.8$ billion, which equates to 115 million annually, or around 9.6 million sticks per month. In 2017 excise tax rates of NZD 0.74 per stick, the annual excise tax revenue from cigarettes would be expected to decline by around NZD 85.4 million. More importantly, assuming a similar proportionate decline in incidence rates, from a base of 603,000 smokers in NZ in 2017 (Euromonitor International 2017), we would expect a decline in the smoking population of around $0.166/0.194 \times 0.075 = 0.064$ (6.4%). This further translates into 38,600 fewer smokers ($0.064 \times 603,000$). The annual decline in excise taxes must be offset against the lower societal cost imposed by 38,600 fewer smokers such as lower health costs. Although we recognize that our results are not intended as precise estimates, they nevertheless give an indicative metric of what the impact might be in that country.

4.5.1. Implications for Manufacturers. Although we cannot necessarily attribute causality to changes in manufacturer actions after the introduction of the PPM, it is interesting to see how they changed the use of the major marketing tools still at their disposal. These data may provide interesting insights to regulators. For example, it may be used to forecast whether changes in excise are fully passed through to consumers, partially absorbed, or amplified. Overall, weighted (by volume) prices in the three months after the introduction of the PPM were 1.9% higher than those in the three months before. This varies by market segment with, grocery and convenience stores being 2.1% and 0.4% higher, and the premium, mainstream, and value quality segments being 1.2%, 2.2%, and 1.2% higher, respectively. The other major marketing tool available to manufacturers, the number of brand variants, *decreased* after the PPM, with a decrease of 8% in the three months after, relative to that before. Corresponding changes in the number of variants by segment are 7.3% and 12% lower for grocery and convenience stores, and 3.8%, 4.8%, and 14.0% lower for premium, mainstream, and value, respectively.

Manufacturers of cigarettes have repeatedly stated that their objective is not to grow category sales, but to maximize profits by trying to grow/maintain price and margins (Chaloupka et al. 2002). Our results

indicate that the bright spots for manufacturers are the value segment in the market and the convenience channel. Going forward, we would expect companies to focus their efforts on these aspects of the market. By the same token, policy makers should also be vigilant regarding potential moves that firms may make in the value segment and in the convenience store channel, perhaps explicitly aimed at gaining share that may end up having category demand effects.

5. Conclusions

Policy makers and regulators have to make and evaluate decisions based on imperfect information, both before policies are introduced and often soon after they have been implemented. The implications of these evaluations may cross jurisdictions, forming the basis of other countries' actions. For example, based on early analysis of PPM effects, the United Kingdom, France, New Zealand, Hungary, Ireland, and Slovenia have followed Australia's lead. This paper aims to show how, using the best data available in a given policy setting, we can ascertain the most accurate view of the effects of a regulatory intervention. The mandated plain packaging of tobacco products by Australia gives us a rare opportunity to learn more about what role packaging can play in differentiated product markets. Based on data we obtained on sales, we find a significant reduction of around 60 to 70 million sticks monthly in quantity sold. We also report several important effects of the plain packaging regulation on sales response to marketing mix elements remaining under the control of marketers of tobacco products: pricing, variants, and distribution channels. Our main insight is that of an increase in price sensitivities, borne largely by the grocery channel and in the longer-term effectiveness of pricing. We find some evidence that price sensitivities across different segments of products converge after the PPM, with differences among brand types typically seen in differentiated products less prevalent. Collectively, our findings are consistent with the valuable role that packaging plays in a differentiated goods market, and that removal of the ability for firms to differentiate via distinctive branding elements via distinctive branding elements on packaging has a deleterious effect on brands' ability to garner higher sales overall. We find this mechanism to be at play and likely to underlie the decline in post-PPM baseline sales that we observe across our analyses.

We acknowledge that the richness of our conclusions and insights are limited by data constraints, but that is the problem facing regulators who do not control the products in the market. Nevertheless, we believe valuable insights can already be gleaned from the Australian experience. One advantage of our study is the use of sales data, both before and after

the introduction of a PPM and for a similar market. A corresponding disadvantage is the limited length of time window in which to examine changes in cigarette sales. Given the addictive nature of cigarettes, once consumers enter the market, it is difficult for them to leave. This suggests that long-term effects may be larger than short-term effects. Sales volume will vary somewhat from actual consumption because of possible changes in inventory levels, but we do not expect this to be a major factor. Finally, although using NZ as a control for Australia has a long pedigree, and we were careful to control for changes in the NZ market, it may not represent a perfect control for omitted factors in the Australian market. We expect interesting future research to come out of other data sets, including those derived from future countries adopting plain packaging. To the extent that our study can be replicated in these contexts, interesting empirical generalizations about the effectiveness of public policy can be drawn, in particular, on how different instruments can work together to achieve socially beneficial outcomes.

Acknowledgments

Endnotes

[1] For details, see http://www.health.gov.au/internet/main/publishing .nsf/content/tobacco-plain (accessed August 20, 2018).

[2] We assume a consumption average of 500 sticks per month (Euromonitor International 2017), and that the decline in smoking after the PPM is mainly due to lower incidence rates.

[3] See http://www.who.int/en/news-room/fact-sheets/detail/tobacco (accessed December 1, 2018).

[4] Prevalence is the percentage of the population that are self-reported regular smokers.

[5] As of December 1, 2012, AUD 1 = USD 1.04.

[6] See http://www.tcdata.org.NZ (accessed August 16, 2016).

[7] For example, see https://treasury.govt.NZ/publications/tax-outturn -data/tax-outturn-data-september-2018 (accessed November 30, 2018).

[8] As with prior work, we aggregate and calculate weighted means of the coefficients, and standard errors are calculated with Rosenthal's (1991) method of adding Z values. The delta method to obtain the standard error (s.e.) for the ratio of two estimates (θ_X and θ_Y, with corresponding s.e. σ_X and σ_Y) is as follows:

$$s.e.\left(\frac{\theta_X}{\theta_Y}\right) = \frac{\theta_X}{\theta_Y}\sqrt{\frac{\sigma_X^2}{\theta_X^2} + \frac{\sigma_Y^2}{\theta_Y^2} - 2\frac{\sigma_{X,Y}}{\theta_X \theta_Y}}.$$

[9] Having not fully developed into a habit comprising of considerably higher levels of consumption and its commensurate financial budget,

such new customers are more able to purchase or afford the premium products.

References

Angrist JD, Pischke JS (2008) *Mostly Harmless Econometrics: An Empiricist's Companion* (Princeton University Press, Princeton, NJ).

Argo JJ, White K (2012) When do consumers eat more? The role of appearance self-esteem and food packaging cues. *J. Marketing* 76(2):67–80.

Baltagi BH, Griffin JM (1995) A dynamic demand model for liquor: The case for pooling. *Rev. Econom. Statist.* 77(3):545–554.

Banks E, Joshy G, Weber MF, Bette Liu B, Grenfell R, Egger S, Paige E, Lopez AD, Sitas F, Beral V (2015) Tobacco smoking and all-cause mortality in a large Australian cohort study: Findings from a mature epidemic with current low smoking prevalence. *BMC Medicine* 13(1):38.

Bardsley P, Olekalns N (1999) Cigarette and tobacco consumption: Have anti-smoking policies made a difference? *Econom. Record* 75(3):225–240.

Bijmolt THA, van Heerde HJ, Pieters RGM (2005) New empirical generalizations on the determinants of price elasticity. *J. Marketing Res.* 42(2):141–156.

Bowen C, Plibersek T (2013) Government to increase tobacco excise. Press Release No. 015, Department of the Treasury, Canberra, ACT, Australia.

Bruce AS, Lepping RJ, Bruce JM, Cherry JBC, Martin LE, Davis AM, Brooks WM, Savage CR (2013) Brain responses to food logos in obese and healthy weight children. *J. Pediatrics* 162(4):759–764.

Burton S, Hoek J, Nesbit P, Khan A (2015) "Smoking is bad, it's not cool...Yet I'm still doing it": Cues for tobacco consumption in a "dark" market. *J. Bus. Res.* 68(10):2067–2074.

Capella ML, Webster C, Kinard BR (2011) A review of the effect of cigarette advertising. *Internat. J. Res. Marketing* 28(3):269–279.

Chaloupka FJ, Cummings KM, Morley CP, Horan JK (2002) Tax, price and cigarette smoking: Evidence from the tobacco documents and implications for tobacco company marketing strategies. *Tobacco Control* 11(Suppl 1):i62–i72.

Chen T, Sun B, Singh V (2009) An empirical investigation of the dynamic effect of Marlboro's permanent pricing shift. *Marketing Sci.* 28(4):740–758.

Chipty T (2016) Study of the impact of the tobacco plain packaging measure on smoking prevalence in Australia. Accessed March 9, 2018, http://www.health.gov.au/internet/main/publishing.nsf/ Content/study-pp-measure.

Davidson S, Silva A (2018) Did recent tobacco reforms change the cigarette market? *Econom. Papers* 37(1):55–74.

Dewhirst T (2012) Price and tobacco marketing strategy: Lessons from "dark" markets and implications for WHO framework convention on tobacco control. *Tobacco Control* 21(6):519–523.

Diethelm PA, Farley T (2015) Refuting tobacco-industry funded research: Empirical data shows a decline in smoking prevalence following the introduction of plain packaging in Australia. *Tobacco Prevention Cessation* 1(6):1–7.

Doward J (2015) Row over Marlboro-funded research that undermined plain cigarette packs. *The Guardian* (February 14), https:// www.theguardian.com/business/2015/feb/14/ow-over-arlboro -funded-research-that-undermined-plain-cigarette-packs.

Dreze, X, Hoch SJ, Purk ME (1994) Shelf management and space elasticity. *J. Retailing* 70(4):301–326.

Elder RS, Krishna A (2009) The effects of advertising copy on sensory thoughts and perceived taste. *J. Consumer Res.* 36(5):748–756.

Euromonitor International (2017) Cigarettes in New Zealand. Report, Euromonitor International, London.

Euromonitor Passport (2018a) Distribution. Euromonitor International. Accessed May 12, 2018, http://www.portal.euromonitor.com .ezproxy-b.deakin.edu.au/portal/statisticsevolution/index.

Euromonitor Passport (2018b) Market sizes. Euromonitor International. Accessed May 12, 2018, http://www.portal.euromonitor.com.ezproxy-b.deakin.edu.au/portal/statisticsevolution/index.

Euromonitor Passport (2018c) Smoking population. Euromonitor International. Accessed November 15, 2018, http://www.portal.euromonitor.com.ezproxy-b.deakin.edu.au/portal/statisticsevolution/index.

Fok D, Horváth C, Paap R, Franses PH (2006) A hierarchical Bayes error correction model to explain dynamic effects of price changes. *J. Marketing Res.* 43(3):443–461.

Ford A, Moodie C, Hastings G (2012) The role of packaging for consumer products: Understanding the move toward "plain" tobacco packaging. *Addiction Res. Theory* 20(4):339–347.

Fraser Institute (2018) Economic freedom of the world: 2018 annual report. Accessed November 16, 2018, https://www.fraserinstitute.org/economic-freedom/map?geozone=world&year=2016&countries=AUS,NZL&page=map.

Freeman B, Chapman S, Rimmer M (2008) The case for the plain packaging of tobacco products. *Addiction* 103(4):580–590.

Gallet, Craig A, List JA (2003) Cigarette demand: A meta-analysis of elasticities. *Health Econom.* 12(10):821–835.

Garber LL Jr, Burke RR, Jones JM (2000) The role of package color in consumer purchase consideration and choice. Working paper, Marketing Science Institute, Cambridge, MA.

Gilpin EA, Emery S, White MW, Pierce JP (2002) Does tobacco industry marketing of "light" cigarettes give smokers a rationale for postponing quitting? *Nicotine Tobacco Res.* 4(Suppl 2):S147–S155.

Gilula Z, McCulloch RE, Rossi PE (2006) A direct approach to data fusion. *J. Marketing Res.* 43(1):73–83.

Goldfarb A, Tucker CE (2014) Conducting research with quasi-experiments: A guide for marketers. Rotman School of Management Working Paper No. 2420920, Rotman School of Management, University of Toronto, Toronto.

Gruber J, Sen A, Stabile M (2003) Estimating price elasticities when there is smuggling: The sensitivity of smoking to price in Canada. *J. Health Econom.* 22(5):821–842.

Henderson PW, Giese JL, Cote JA (2004) Impression management using typeface design. *J. Marketing* 68(4):60–72.

Hofstede G (2001) *Culture's Consequences—Comparing Values, Behaviors, Institutions, and Organizations Across Nations*, 2nd ed. (SAGE Publications, Newbury Park, CA)

Huang J, Chaloupka FJ, Fong GT (2014) Cigarette graphic warning labels and smoking prevalence in Canada: A critical examination and reformulation of the FDA regulatory impact analysis. *Tobacco Control* 23(Suppl 1):i7–i12.

Huber J, McCann J (1982) The impact of inferential beliefs on product evaluations. *J. Marketing Res.* 19(3):324–333.

Im KS, Pesaran MH, Shin Y (2003) Testing for unit roots in heterogeneous panels. *J. Econometrics* 115(1):53–74.

Jann B (2015) Methodological report on Kaul and Wolf's working papers on the effect of plain packaging on smoking prevalence in Australia and the criticism raised by OxyRomandie. Technical report, Institute of Sociology, University of Bern, Bern, Switzerland.

Kaul A, Wolf M (2014) The (possible) effect of plain packaging on smoking prevalence in Australia: A trend analysis. Working Paper No. 165, University of Zurich, Zurich.

Keller KL (1993) Conceptualizing, measuring, and managing customer-based brand equity. *J. Marketing* 57(1):1–22.

Klimchuk MR, Krasovec SA (2006) *Packaging Design: Successful Product Branding from Concept to Shelf* (John Wiley & Sons, Hoboken, NJ).

Kwiatkowski D, Phillips PCB, Schmidt P, Shin Y (1992) Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? *J. Econometrics* 54(1–3):159–178.

Levin A, Lin CF, Chu CSJ (2002) Unit root tests in panel data: Asymptotic and finite-sample properties. *J. Econometrics* 108(1):1–24.

Levy DT, Chaloupka F, Gitchell J (2004) The effects of tobacco control policies on smoking rates: A tobacco control scorecard. *J. Public Health Management Practice* 10(4):338–353.

Machado FS, Sinha RV (2007) Smoking cessation: A model of planned vs. actual behavior for time-inconsistent consumers. *Marketing Sci.* 26(6):834–850.

Meyers H, Lubliner MJ (1998) *The Marketer's Guide to Successful Package Design* (NTC, Lincolnwood, IL).

Millward Brown (2018) BrandZ top 100 global brands. Accessed April 29, 2019, http://www.millwardbrown.com/brandz/rankings-and-reports/top-global-brands/2018.

Morrot G, Brochet F, Dubourdieu D (2001) The color of odors. *Brain Language* 79(2):309–320.

Murphy WH (1999) Hofstede's national culture as a guide for sales practices across countries: The case of a MNC's sales practices in Australia and New Zealand. *Australian J. Management* 24(1):37–58.

Nevo A (2001) Measuring market power in the ready-to-eat cereal industry. *Econometrica* 69(2):307–342.

Nielsen Corporation (2015) Packaging sells. Accessed July 14, 2017, http://innovation.nielsen.com/packaging-sells-report.

Phillips PCB, Perron P (1988) Testing for a unit root in time series regression. *Biometrika* 75(2):335–346.

Rosenthal R (1991) *Meta-Analytic Procedures for Social Research*, vol. 6 (SAGE Publications, Inc., Newbury Park, CA).

Scollo M, Bayly M (2018) Tobacco in Australia: Facts and issues. Technical report, Cancer Council Victoria, Melbourne, Australia.

Stock JH, Watson MW (2012) *Introduction to Econometrics*, updated 3rd ed., global ed. (Pearson Boston).

Stock JH, Wright JH, Yogo M (2002) A survey of weak instruments and weak identification in generalized method of moments. *J. Bus. Econom. Statist.* 20(4):518–529.

Tucker C, Zhang J (2011) How does popularity information affect choices? a field experiment. *Management Sci.* 57(5):828–842.

U.S. Department of Health and Human Services (2010) Preventing tobacco use among youth and young adults: A report of the surgeon general. Technical report, Office on Smoking and Health, National Center for Chronic Disease Prevention and Health Promotion, Centers for Disease Control and Prevention, U.S. Department of Health and Human Services, Atlanta.

Vanclay JK, Shortiss J, Aulsebrook S, Gillespie AM, Howell BC, Johanni R, Maher MJ, Mitchell KM, Stewart MD, Yates J (2011) Customer response to carbon labelling of groceries. *J. Consumer Policy* 34(1):153–160.

van Heerde H, Gijsenberg MJ, Dekimpe MG, Steenkamp JB (2013) Price and advertising effectiveness over the business cycle. *J. Marketing Res.* 50(2):177–193.

Wakefield MA, Germain D, Durkin SJ (2008) How does increasingly plainer cigarette packaging influence adult smokers' perceptions about brand image? An experimental study. *Tobacco Control* 17(6):416–421.

Wakefield MA, Morley C, Horan JK, Cummings KM (2002) The cigarette pack as image: New evidence from tobacco industry documents. *Tobacco Control* 11(Suppl 1):i73–i80.

Wang Y, Lewis M, Singh V (2015) The unintended consequences of countermarketing strategies: How particular antismoking measures may shift consumers to more dangerous cigarettes. *Marketing Sci.* 35(1):55–72.

Zhang J (2010) The sound of silence: Observational learning in the U.S. kidney market. *Marketing Sci.* 29(2):315–335.