WARSAW SCHOOL OF ECONOMICS
COLLEGIUM OF ECONOMIC ANALYSIS

Disaggregated analysis of demand for cigarettes in Poland

Bartosz Olesiński

Summary of the doctoral dissertation
written under the supervision of
Professor Emilia Tomczyk (supervisor) and
Professor Andrzej Torój (supporting supervisor)

WARSAW, 2020

1 Introduction

Effective government policy towards tobacco should be based on an economic impact assessment of the considered regulations, taking into account different goals such as raising government revenue, reducing tobacco consumption or minimising the illicit trade. The critical role in such analyses is played by understanding determinants of demand, including price, non-price factors and various regulations affecting tobacco consumption.

Some regulations affect selected product categories more strongly than others. The quantitative assessment of the effects of those targeted policy measures requires developing such models that account for the product-specific demand effects. The models of tobacco demand that account for product diversity have already attracted much attention in the literature. An important strand of this research is related to discrete choice models, with the most notable example of the multinomial logit model, popularized by McFadden (1973). Discrete choice models treat each product as a bundle of its characteristics, with the utilities drawn by consumers from particular characteristics explained with a well-designed regression. In discrete choice models, the option with the highest utility is selected by the consumer out of the product space. Such models can explain the related probabilities of choice (or market shares, if the analysed dataset includes the market-level information), depending on specification of the consumer utility and consumer heterogeneity.

The available literature (for example, Ciliberto and Kuminoff, 2010, Park, 2010, Min, 2011, Pham and Prentice, 2013, Liu et al., 2015, Tuchman, 2019) provides hints that discrete choice models constitute a versatile framework that allows for diverse economic impact assessments in the case of the tobacco market. However, one of the fundamental issues that can render some discrete choice models inadequate to the analysis of the substitution patterns is the independence of irrelevant alternatives (*IIA*) property. In a model bearing that property, all the cross-price elasticities of demand with respect to price of good 1 are equal. This is unrealistic because in reaction to an increase in price of good 1, demand for *similar* products should increase by a larger percentage than the demand for products that differ from good 1 to a larger extent.

While the analyses that account for diversity of the tobacco products can cover many topics, there is a gap in the literature as regards the *ex ante* impact of bans of particular products on the tobacco market. An interesting example of such a targeted policy is included in the Tobacco Products Directive (2014/40/EU), mandating a withdrawal of menthol cigarettes (also those with capsules that change the taste) and menthol fine-cut tobacco from the EU market in 2020.

This dissertation applies a well-known random coefficient logit model proposed by Berry, Levinsohn and Pakes (1995) and developed, among others, by Nevo (2000, 2001) to the Polish tobacco market data. The model is referred to as "BLP" thereafter. It is

a widely-used, market-level tool that can solve the *IIA* problem. The marginal utilities are allowed to differ across consumers and are calculated via simulation, basing on two groups of essential parameters, estimated using aggregate information about consumers:

- mean (over all consumers) utility parameters,
- consumer-specific parameters.

2 Research objectives and hypotheses

In this study, the impact of two regulations on the Polish tobacco market are considered: the withdrawal of menthol products from the Polish tobacco market and the excise tax hike. Additionally, this work discusses many aspect of estimation, including the income-related parameters, strength-related parameters, the economic interpretation of price effects and feasibility of estimation.

The estimation in this work is based on Nielsen data (Nielsen Retail Index for Cigarettes and Tobacco categories, including sales value, sales volume, average price, covering Total Poland monitored channels for 2004-2017), representative for the Polish retail market (purchases done by the final customers). The Polish data provides an interesting basis for policy impact analyses, because the tobacco market in Poland has been strongly influenced by the regulations initiated at the level of the European Union. For instance, a 2011 Directive (2011/64/EU) increased the minimum level of excise per 1000 cigarettes, regulated the structure of cigarette excise tax and obliged Poland to implement the new standards by the end of 2017. These requirements were met by Poland three years ahead of schedule, following a series of excise hikes that put the local tobacco prices more in line with the prices elsewhere in the EU. The excise hikes and changes in the taxation structure have produced unique patterns of retail price variation. In this study, it is discussed whether this variance makes it possible to identify the product-specific effects for a very rich portfolio of 865 individual products, observed over the 2004-2017 period. All the variables used throughout the estimations are summarized in Table 1.

Table 1: Summary of variables used in the estimations in this study.

Variable name	Description
	Product-level dataset
menthol	a dummy variable equal to 1 if a given product has menthol flavour
caps	a dummy variable equal to 1 if a given product contains a capsule that changes the taste
slim	a dummy variable equal to 1 for slim cigarettes
light	a dummy variable equal to 1 for light cigarettes
super_light	a dummy variable equal to 1 for super light cigarettes
soft	a dummy variable equal to 1 for cigarettes in soft packs
low_price	a dummy variable equal to 1 for products belonging to the low-price segment
mid_price	a dummy variable equal to 1 for products belonging to the mid-price segment
high_price	a dummy variable equal to 1 for products belonging to the high-price segment
sticks_per_pack	a numerical variable indicating the number of cigarettes in a single pack of a given product
temp	a numerical variable indicating the average temperature in Poland observed in a given month
disp_income	a numerical variable describing the disposable income of households in Poland
adults	the number of adult (15+) residents in Poland
volume	the volume of retail sales of factory-made cigarettes
price	retail price per 20 factory-made cigarettes
	Time-series dataset
fine_cut_share	the share of fine cut tobacco retail sales volume in the sum of the factory-made cigarettes and fine-cut tobacco retail sales volumes
cpi	CPI inflation index (average for 2010=100)
income_mean	aggregate level mean of the disposable income, expressed per adult
income_sd	aggregate level standard deviation of the disposable income, expressed per adult

Note: The segments are defined by author based on the information obtained from British American Tobacco about price positioning of particular brands on the Polish market. The terms 'light' and 'super light' are purely descriptive and based on the product characteristics. Essentially, following the respective regulations, such terms related to strenght are not put on packs of cigarettes in Poland.
Source: the variable names were chosen based on the own analysis of the available data.

This dissertation is based on the following Thesis Statement: **the BLP model is a tool that allows one to understand the mechanisms of demand for cigarettes in Poland (including both the level and structure of demand) and draw practical conclusions for different types of economic policy**. The thesis focuses on the practical aptitude of that model to actually explain brand-switching behaviour of consumers under different policies with a dataset that covers a large portfolio of 865 products. Such number of products makes the BLP estimation difficult and the possible issues are related to numerical problems and economic interpretation of some of the results. The usefulness of the BLP model in the context of the dataset used in this dissertation is assessed using five hypotheses.

Hypothesis 1: BLP model can be estimated for cigarettes market in Poland in such a way that the parameter estimates imply lack of the independence

of irrelevant alternatives (*IIA*) property of that model.
One of important features of the BLP model, as suggested by its proponents, is the fact that it does not have the *IIA* property. However, in practice, the consumer-specific parameters that are responsible for relaxing the *IIA* property are not always easy to estimate. This hypothesis is verified both by interpreting the consumer-specific parameters of the BLP model and by analysing the cross-price elasticities of demand that are implied by that model. In addition, the practical issues related to the estimation of those crucial parameters are considered.

Hypothesis 2: Consumer preferences towards particular cigarette characteristics depend on the income level.
In the BLP models estimated here, the income is included in two forms – as the denominator of the real price variable that explains utility (which is, however, not interpretable in the context of this hypothesis) and as an interaction variable that affects the marginal utilities of individual consumers with respect to product characteristics. Is is an important feature of the BLP model that it allows to include the information about income in a macro model, even without income information for individual consumers. The sings and significance of the consumer-specific BLP parameters related to income are used to verify this hypothesis.

Hypothesis 3: Consumers in reaction to cigarette price increases switch towards stronger tobacco products.
The strength of cigarettes is an important element in the BLP models considered here because it is one of the product characteristics from which consumers draw utility. This hypothesis is verified by interpreting the consumer-specific parameters of the BLP model that are related to the strength of particular tobacco products. Hypotheses 1-3 jointly play an important role in the economic interpretation of the BLP results.

Hypothesis 4: The policy of withdrawal of menthol products from the Polish tobacco market leads to stronger reduction of cigarettes retail sales than a 5% increase in excise tax rate per 1000 cigarettes (the specific rate).
This hypothesis is important because different kinds of policies can have an impact on the same outcome variable, such as government revenues or consumption level. In particular, the policymaker can consider tax hikes and product bans as alternative solutions to achieve the same goal. When distinct models are used to analyse those policies, they cannot be compared with each other and therefore, the choice between those kinds of policy cannot be based on quantitative results. Verification of this hypothesis is based on a simulation analysis using the estimated BLP parameters in the case of two instances of

policy: (i) the ban of menthol products from the Polish market and (ii) a 5% increase in excise tax rate per 1000 cigarettes (the specific rate).

Hypothesis 5: BLP model better explains the actual developments in the cigarettes market level and structure than simpler, conditional logit models.

This is verified using a cross-validation procedure, focusing on the predictive ability of particular models. Such an exercise can be helpful in studying the consequences of analysing a big product space, which is related to many challenges in the estimation. Many of those issues are not present in the case of simple conditional logit models because they are linear in parameters. Therefore, if BLP fares better in terms of predictive ability, this would show that even big products spaces can be considered in that framework. From that perspective, verifying this hypothesis has two benefits for other researchers analysing disaggregated demand. Firstly, it provides guidance in selecting the most suitable model to analyse the cigarettes market in Poland (and possibly, in other countries). Secondly, it demonstrates whether big product space has indeed severe consequences for the BLP estimation.

3 The empirical results

The estimation results for BLP and a corresponding conditional logit model are demonstrated in Tables 2 and 3. The considered BLP specifications are rather parsimonious as they include consumer-specific parameters only for the *real price*, *menthol*, *slim*, *light*, *super light*, *mid-price segment* and *high-price segment* variables (with a maximum of 4 consumer-specific coefficients in a single specification). The mean (over all consumers) utility components in the BLP model can be compared with a simpler conditional logit model coefficients estimated using 2SLS (Table 2). Those parameters are interpreted as the marginal utilities. In the case of coefficients that differ statistically from zero, positive (negative) sign means that a given product characteristic (dummy variable equal to 1 or a higher value in the case of non-binary variables) increases (reduces) utility of the consumer, *ceteris paribus*, which translates into higher (lower) market share. The signs of the estimated marginal utilities are largely consistent among the specifications with most notable exceptions for the BLP Model 4 (see the *light*, *mid-price segment* and *high-price segment* variables in particular). Apart from that, the magnitude of coefficients differ, for example for the *real price* variable. The decisions with respect to which coefficients are allowed to depend on consumer characteristics are thus clearly influential for the average marginal utilities.

Important results of the BLP modeling are related to the consumer-specific components of the utility coefficients (Table 3). According to the Wald test, the consumer-specific components are jointly significant (under all standard significance levels) in all

the BLP models with the exception of the BLP Model 1, which means that most models should include additional consumer-specific components of utility parameters to better describe the data. However, when particular parameters are considered, the estimation results are not straightforward to interpret. The standard error components (σ parameters) of the coefficients that depend on consumer characteristics do not differ statistically from zero in any of the cases which suggests that unobservable consumer characteristics might be hard to capture with normal distribution with constant mean and standard deviation parameters that are used in the estimation. The parameters turn out to be statistically different from zero in the case of the interactions of the random coefficients with *log income* variable (governed by the π parameters). The distribution of the *log income* variable is normal but with the mean parameter that increases over time (due to an increase in disposable income of households in Poland over the 2004-2017 period). The results for the interactions with *log income* are quite different across specifications demonstrating the difficulties with estimation of the BLP model.

Table 2: Results of the BLP estimation – mean (over consumers) utility parameters as compared to the conditional logit (2SLS) results.

	Conditional Logit (2SLS)	BLP Model 1	BLP Model 2	BLP Model 3	BLP Model 4	BLP Model 5
			Mean utility parameters			
Real price	-18.558***	-19.110***	-12.098***	-26.280***	-33.199***	-27.768***
	(0.000)	(0.000)	(0.002)	(0.000)	(0.002)	(0.000)
Menthol	-0.646***	-0.647***	-0.769***	-0.859***	-6.185**	-0.969**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.012)	(0.021)
Slim	0.310***	0.312***	0.430***	0.576***	1.452***	0.553***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)
The number of sticks in pack	-0.036***	-0.036***	-0.035***	-0.063***	-0.046***	-0.070***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Flavour capsules	0.956***	0.961***	1.463***	1.521***	1.095**	1.509***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.048)	(0.000)
Soft pack	-0.739***	-0.738***	-0.719***	-0.884***	-0.229*	-0.905***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.096)	(0.000)
Light	-0.014	-0.014	0.038	1.968***	-1.261***	1.980***
	(0.617)	(0.611)	(0.228)	(0.000)	(0.000)	(0.000)
Super light	-0.742***	-0.743***	-0.637***	-0.904***	-25.928***	-1.158***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.002)
Mid-price segment	-1.760***	-1.761***	-11.412***	-1.830***	0.747	-1.690***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.107)	(0.000)
High-price segment	-0.593***	-0.588***	-18.959***	-1.175***	1.825***	-1.118***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)
Temperature	0.023***	0.023***	0.034***	0.032***	0.036***	0.033***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Trend	-0.034***	-0.034***	-0.040***	-0.029***	-0.039***	-0.029***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Intercept	-9.284***	-9.281***	-9.738***	-9.661***	-8.719***	-9.439***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Consumer-specific coefficients for:	(-)	*real price*	*real price, mid-price segment, high-price segment*	*real price, menthol, lights*	*real price, menthol, lights, super lights*	*real price, menthol, slims, lights*

Note: p-values in parentheses (* p<0.10, ** p<0.05, ***p<0.01). The number of observations: 57622 (unbalanced panel of 865 products over 168 months). The number of simulations of consumer characteristics equals 5000. 210 brand dummies are omitted for clarity. Source: own calculations.

Interaction between *log income* and *real price* variables is statistically different from zero only in the BLP Model 3 and 5, with a negative sign of the respective estimate which suggests that consumers with higher income are more price sensitive. One explanation might be fact that more affluent consumers consider smoking a less attractive habit and are more inclined to quit smoking in the analysed period, as compared to less affluent consumers. This inclination could have increased over time, making it correlated with prices (which were in an upward trend), so more frequent quitting among more affluent consumers could have been captured by the *real price* consumer-specific coefficient. However, this result does not hold (i) in the case BLP Model 1 and 2 wherein the parameter estimate is negative, but the parameter itself is statistically not different from zero and (ii) in BLP Model 4, in which this parameter estimate has the opposite sign, but the parameter itself is, again, statistically not different from zero.

Table 3: Results of the BLP estimation – the consumer-specific components of the utility coefficients and the additional results.

	BLP Model 1	BLP Model 2	BLP Model 3	BLP Model 4	BLP Model 5
Unobserved component parameters for the consumer-specific coefficients					
σ for *real price*	0.005 (1.000)	-0.024 (1.000)	-0.056 (0.999)	-0.003 (1.000)	0.129 (0.998)
σ for *menthol*			0.001 (1.000)	-0.723 (0.940)	-0.017 (1.000)
σ for *slim*					-0.006 (1.000)
σ for *light*			0.001 (1.000)	0.022 (0.999)	-0.017 (0.999)
σ for *super light*				4.146 (0.758)	
σ for *mid-price segment*		-1.467 (0.914)			
σ for *high-price segment*		1.505 (0.892)			
Interaction with log income for the consumer-specific coefficients					
π for *real price*	-2.335 (0.811)	-2.262 (0.754)	-95.666*** (0.000)	5.307 (0.714)	-96.647*** (0.000)
π for *menthol*			0.805 (0.446)	-14.205*** (0.000)	1.192 (0.414)
π for *slim*					-1.371 (0.407)
π for *light*			11.746*** (0.000)	-0.762 (0.638)	12.155*** (0.000)
π for *super light*				33.522*** (0.000)	
π for *mid-price segment*		18.861*** (0.000)			
π for *high-price segment*		-23.326*** (0.000)			
Additional results					
Wald p for joint significance of the consumer-specific components	0.972	0.000	0.000	0.000	0.000
Total elasticity in December 2017	-1.20	-0.59	-3.11	-1.65	-3.21
Positives among own-price elasticities	0.0%	0.0%	43.3%	0.0%	43.3%
Negatives among cross-price elasticities	0.0%	0.0%	33.8%	0.0%	30.3%
The value of GMM objective in the BLP estimation	7 051.5	4 980.2	6 478.0	5 167.3	6 444.7

Note: p-values in parentheses (* p<0.10, ** p<0.05, ***p<0.01). The number of observations: 57622 (unbalanced panel of 865 products over 168 months). The number of simulations of consumer characteristics equals 5000. Source: own calculations.

Another set of important results are the substitution matrices. The formulas presented in the dissertation allows one to calculate the aggregate elasticity for each period – December 2017 is selected throughout this work. The BLP models imply quite diverse total price elasticities of demand, with the results ranging from -0.59 (BLP Model 2) to -3.21 (BLP Model 5).

The preferred BLP model should produce substitution matrix that have (i) negative diagonal elements, (ii) positive non-diagonal elements and (iii) demonstrates that the underlying model does not have the *IIA* property. The requirements (i) and (ii) are satisfied only in the case of BLP Model 1, 2 and 4, while in the case of BLP Model 3 and 5, there is a considerable share of positive own-price elasticities and negative cross-price elasticities (see the Additional results section of Table 3). When it comes to the requirement (iii), BLP Model 1 is the only specification in which the consumer-specific components, responsible for relaxing the *IIA* property, do not differ statistically from zero even jointly and thus should be rejected.

Both BLP Model 2 and 4 satisfy the requirements discussed above so to make a final choice of the baseline specification, an additional cross-validation exercise focusing on the predictive ability of those models was conducted. In that procedure, both BLP models were compared with the conditional logit models estimated using OLS and 2SLS. According to the cross-validation results, the BLP Model 2 fares substantially better the BLP Model 4. In addition, it produces better forecasting results (on average) for the market shares than the simpler conditional logit models. However, when the the log of relative market share is predicted, the baseline BLP model is worse than the conditional logit models. In such case, the best results were obtained for the conditional logit model estimated with OLS. It is worth noting that log of relative market share is not the most important form of the explanatory variable in this analysis of demand for tobacco products – the results for the simple market shares should be treated as more relevant because they are more closely related to the market volumes. Still, the results suggest that if prediction is the sole purpose of the discrete choice modeling exercise and if the substitution patterns are not as important, conditional logit model can be treated as an easily-applicable rule-of-thumb model, because its estimation is much less involved than in the case of the BLP model. In addition, the BLP Model 2 is subject to a sensitivity analysis with respect to some crucial parameters, values of which need to be assumed for the purposes of the estimation. This analysis has confirmed that BLP Model 2 with the baseline values of the key parameters is the most appropriate for the purposes of a more detailed scrutiny.

The level of detail included in the BLP model and its focus on the characteristics space rather than product space makes it possible to carry out various post-estimation simulation analyses that can greatly support policymakers in making decisions. To illustrate the versatility of the BLP model, the baseline results (BLP Model 2) are used

to analyse two quite different instances of policies towards tobacco: the ban of menthol products from the Polish market and a cigarette excise hike.

The first post-estimation simulation considers a counterfactual withdrawal of the menthol cigarettes from the duty-paid market in December 2017 in order to demonstrate the impact of a ban on menthol cigarettes from the EU market from 2020 on under the Tobacco Products Directive. This policy was implemented in May 2020, before this research project was finished. Therefore, the results of that simulation exercise should be interpreted as an *ex ante* impact of withdrawal of menthol products on the Polish tobacco market. According to the simulation analysis, the ban on menthol cigarettes would lead to a reduction of the duty-paid market volume by 10.02%, yet consumers to a large extent would switch to other brands of cigarettes in the duty-paid market.

The BLP model also allows for tax policy simulations based on the estimated substitution matrix. The cigarette excise tax in Poland includes three rates: (i) the specific rate, which is calculated per cigarette, (ii) the *ad valorem* rate that is calculated as a % of gross retail price and (iii) the minimum rate which is applied per cigarette as long as the taxation implied by (i) the specific and (ii) *ad valorem* rates is too low. The simulation scenario of tax policy considers a counterfactual 5% increase in specific excise rate in December 2017. As a result of the tax hike, the average retail price grows by 2.56%, which is related to a drop in retail sales volume by 1.50%. The joint, excise- and VAT-related government revenues are higher by 1.82% as compared to the *status quo* scenario. The magnitude of impact of the considered policies on the cigarettes market volume show that the menthol ban is a relatively strong measure, as compared to the excise hike, despite considerable switching effects towards non-menthol cigarettes, as demonstrated by the BLP model.

Taking into account the results provided above, hypothesis 1 has been verified positively. In this study, a reasonable baseline BLP model has been estimated (referred to as BLP Model 2) in which the parameters responsible form relaxing the *IIA* property are statistically different from zero and produce results with useful economic interpretation. In addition, this model fares relatively well in the cross-validation procedure and produces reasonable results in the sensitivity analysis, that is described in the dissertation. This confirms that the BLP model can indeed be a useful tool to overcome the *IIA* property in the discrete choice models for the tobacco market.

Hypothesis 2 has not been confirmed. While the interaction terms of some explanatory variables with the *log income* were statistically different from zero, mixed results across different BLP specifications suggest that BLP models might not be a very useful device to verify hypotheses about individual parameters driving the consumer behaviour in the Polish tobacco market – other methods are likely to be more informative. Perhaps more robust results could be obtained by using additional micro-level data about consumer behaviour – those were, however, not available for the purposes of this study.

Hypothesis 3 verification was based on interpretation of those consumer-specific parameters of the BLP model that are related to the strength of particular tobacco products. As indicated in the comment to the Hypothesis 2, BLP results with respect to individual parameters should be interpreted with caution. In addition, the BLP Model 4 that included consumer-specific parameters for the *light* and *super light* variables and had acceptable economic interpretation in terms of price elasticities of demand, fared substantially worse in terms of its predictive ability. This was demonstrated by the cross-validation exercise and better results were produced not only by the BLP model that was chosen as baseline (BLP Model 2), but also by the simple conditional logit models. In sum, this hypothesis has not been confirmed. The results for the hypotheses 1-3 suggest that the economic interpretation of the baseline BLP model is reasonable, but some results should be interpreted with caution.

Hypothesis 4 has been confirmed in the dissertation – the simulations results suggest that the menthol ban, despite considerable switching of tobacco consumers towards non-menthol cigarettes, has a considerable impact on the reduction of the cigarettes duty-paid retail sales volume, when compared to the results of a 5% increase in cigarettes specific rate. Of course, the cigarette excise hike can be made big enough so the impact on the total duty-paid market is as strong as the impact of the menthol ban – however, some tax hikes might not be feasible from the political point of view while the menthol ban is a policy measure that has already been imposed across the EU by the EU-level regulations. As a device to verify this hypothesis, an approach to analyse large-scale products bans before any data is available (*ex ante*) has been proposed, which can be used for future analyses of consequences of such kind of policy.

Hypothesis 5 has been verified positively using a cross-validation procedure. Such an exercise is helpful in studying the consequences of a big product space used here, which is related to many challenges in the estimation. It turns out that the baseline BLP model estimated as a part of this study produces better forecasting results, at least for the market shares, than the simpler conditional logit models. One caveat is that when an alternative form of the explained variable is used, namely log of the relative market share, the baseline BLP model is worse than the conditional logit models (with the best result for the conditional logit model estimated with OLS). However, this is not the most important form of the explanatory variable in this analysis of demand for tobacco products – the results for the simple market shares should be treated as more relevant. Good predictive ability of the baseline BLP model suggests that the consequences of applying the BLP method to a big product space (not present in the case of the simple conditional logit model, which is linear in parameters) are not detrimental to the estimation results.

Apart from this quantitative criterion, many economic arguments have been discussed throughout the dissertation that a model that is free from the the *IIA* property should be used whenever possible. Following the results of verification of Hypothesis 1, the BLP

model is more reasonable in economic terms than conditional logit model, despite the BLP estimation issues. Those issues were not avoided in this study, yet some solutions to address those problems were applied, such as high frequency of data. On the other hand, the results also suggest that if prediction is the sole purpose of the discrete choice modeling exercise and if the economic interpretation of the results is of lesser importance, conditional logit model can be treated as an easily-applicable rule-of-thumb model, because its estimation is much less involved (especially in the case of OLS) than in the BLP model. In business applications, a model that is fast and easy to estimate might be more appealing unless producing absurd predictions.

4 Conclusions

This dissertation uses a well-known BLP model to verify various hypotheses regarding the capacity of this econometric tool to analyse a large portfolio of 865 cigarette products in the Polish tobacco market over the 2004-2017 period and to carry out *ex ante* measurements of the impact of different policies. The simulations suggest that the menthol ban, despite switching of tobacco consumers towards non-menthol cigarettes, has a considerable impact on the reduction of the cigarettes duty-paid retail sales volume, as compared to the cigarette excise hike. Also, the simulation exercises demonstrate that BLP is a versatile framework that allows for diverse policy analyses that can help inform future policy decisions. In particular, the BLP model includes different market segments, making it a natural framework to analyse changes in the level and structure of excise taxation, possibly in combination with other kinds of policies.

The estimation in this study took advantage of complex developments in cigarettes retail prices in Poland over the 2004-2017 period. The related price variability that was shaped by both EU-level and local-level tobacco regulations made it possible to estimate product-specific demand effects for a very large product portfolio. The results presented here are based on the Polish data, but the approach can essentially be applied in any other country in which product-level data for the tobacco market is available and the government considers diverse policies towards tobacco.

Some of estimation issues are not avoided in this study. In particular, inclusion of all the important consumer-specific parameters is not feasible in a single BLP specification – a large number of such parameters aggravates the estimation problems. Such problems suggest that BLP models might not be a very useful device to verify hypotheses about individual parameters driving the consumer behaviour in the Polish tobacco market. Particular hypotheses covered (i) the impact of income on consumer behaviour and (ii) the switching behaviour of consumers to products of different strength in reaction to changes in policy.

This study uses a cross-validation procedure to compare the performance of different BLP specifications and to verify if they perform better than the simpler conditional logit models. The baseline BLP model indeed produced more precise out-of-sample forecasts for the market shares as compared to the simpler approach. The cross-validation procedure put forward here can be used by other researchers either to find the best BLP specification, or as an argument to use a simpler approach – the results are likely to be highly specific to particular research context and the quality of data.

It must be noted that the actual impact of the considered policies on the tobacco market and the government revenues could be different from the results presented above because the dataset used in this study ends in December 2017. A more up-to-date policy analysis should use a dataset covering periods closer to the actual policy change. In addition, taking into account that this study has been finished after the actual withdrawal of the menthol products in May 2020, particular simulations should not be treated as equivalent to an *ex post* assessment of the impact of menthol ban on the Polish tobacco market. Such an *ex post* assessment should use the actual market data, covering many periods after the withdrawal of the menthol products to account for any delays in the market adjustment. However, the insights provided in this study can be used for future *ex ante* analyses if regulators consider withdrawal of other large groups of tobacco products from the market.

The contribution of this study is the following. Firstly, it provides product-level evidence for Poland which has not been available before (especially with such level of representativeness that is assured by the Nielsen data), adding to the generally limited tobacco discrete choice literature published globally. While the *ex ante* impact of large-scale product bans in an analysis with such level of representativeness is a novelty, this work adds to the literature on the impact of increasing the tobacco taxation on the market by using a very detailed, product-level demand model, which is generally rarely done. Secondly, this dissertation uses a cross-validation procedure as a solution to BLP estimation issues. Thirdly, this dissertation provides the results that fill the gap in relatively limited literature when it comes to up-to-date price elasticities of demand for tobacco products in Poland (the published examples include Florkowski and McNamara, 1992, Ross et al., 2014 and Olesiński et al., 2020).

Further research in this area could benefit from including information about electronic cigarettes and heated tobacco products that could be potential substitutes for the banned menthol cigarettes. In addition, the precision of the BLP estimates could be increased by including the regional split of the product-level data and by including additional variables describing the availability of the analysed products across the retail points (if such data is available to the researchers). Finally, other specifications of the BLP model and even other classes of models could be considered in future research of policies towards tobacco market that have different impact on particular product groups. Inconclusive results

with respect to some of the hypotheses considered in this dissertation suggest that an important role in those future analyses of the tobacco market should be played by data on individual behaviour of consumers.

5 The structure of dissertation and the related publications

The dissertation is structured as follows. Section 1 contains the Introduction in which the areas of contribution of this work to the available literature are explained, along with the exposition of the Thesis Statement and the related Hypotheses. Section 2 provides more detailed overview of the literature relevant to the hypotheses verification, starting with a general discussion of the economic literature on tobacco demand and regulations, which is then narrowed towards the empirical analyses that account for tobacco products diversity. Further, this section outlines the fundamentals of BLP and other discrete choice models. Section 3, in turn, covers the empirical strategy taken in this dissertation, including a description of the dataset, an overview of the regulations of the Polish tobacco market that affected the critical variables and a discussion of the issues related to econometric specification and selection of the discrete choice models in this work. Section 4 describes the empirical results and the final section provides conclusions with respect to the hypotheses put forward in the Introduction.

As a part of this research project, two papers have been published. In Olesiński (2020), the impact of the menthol tobacco product ban on the Polish tobacco market is assessed using a BLP model. In the dissertation, most of the contents of that paper are reiterated, but more detail and additional results obtained at later stages of that research are provided. In Olesiński et al. (2020), the issue of the optimum tax policy in Poland, taking into account diverse cigarettes categories (but without a split into particular products) and the aggregate fine-cut tobacco market is discussed. Most of the contribution of that second paper lies beyond the scope of the dissertation, but some parts are also recapitulated in the main text as they play an important role in the economic interpretation of the BLP results.

Acknowledgements

The author is grateful to Nielsen and British-American Tobacco for providing access to monthly data on tobacco market in Poland. The author is also grateful to participants of 18. Doctoral Workshop in Statistics and Econometrics in Uniejów in May and June 2017 for fruitful discussions.

References

Ciliberto F., Kuminoff N.V. (2010): *Public Policy and Market Competition: How the Master Settlement Agreement Changed the Cigarette Industry*, The B.E. Journal of Economic Analysis & Policy (Frontiers), 10(1), Art. 63.

Florkowski W.J., McNamara K.T. (1992): *Policy implications of alcohol and tobacco demand in Poland*, Journal of Policy Modeling, 14(1), 93–98.

Liu H., Rizzo J.A., Wu Q.S.F. (2015): *How do smokers respond to cigarette taxes? Evidence from China's cigarette industry*, Health Economics, 24(10), 1314–1330.

McFadden D. (1973): *Conditional Logit Analysis of Qualitative Choice Behavior*, in: **P. Zarembka** (ed.), *Frontiers of Econometrics*, 105–142, Academic Press, New York.

Min H. (2011): *Reform in a Differentiated-Product Industry: The Case of the Korean Cigarette Manufacturing Industry*, The Korean Economic Review, 27(1), 57–74.

Nevo A. (2000): *Mergers with differentiated products: the case of the ready-to-eat cereal industry*, RAND Journal of Economics, 31(3), 395–421.

Nevo A. (2001): *Measuring Market Power in the Ready-to-Eat Cereal Industry*, Econometrica, 69(2), 307–342.

Olesiński B. (2020): *The Analysis of the Tobacco Product Bans Using a Random Coefficients Logit Model*, Central European Journal of Economic Modelling and Econometrics, 12(2), 113–144.

Olesiński B., Rozkrut M., Torój A. (2020): *How Time-Varying Elasticities of Demand Translate into the Excise Related Laffer Surface*, Argumenta Oeconomica, 1(44), 259–299.

Park S. (2010): *Essays on Time-varying Consumer Preferences*, chapter Preference Evolution in the South Korean Cigarette Market, 40–100, A Dissertation Presented to the Faculty of the Graduate School of Cornell University In Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy.

Pham V., Prentice D. (2013): *A Random Coefficients Logit Analysis of the Counterfactual: A Merger and Divestiture in the Australian Cigarette Industry*, mimeo.

Ross H., Kostova D., Stoklosa M., Leon M. (2014): *The Impact of Cigarette Excise Taxes on Smoking Cessation Rates From 1994 to 2010 in Poland, Russia, and Ukraine*, Nicotine & Tobacco Research, 16(Supplement 1), S37–S43.

Tuchman A.E. (2019): *Advertising and Demand for Addictive Goods: The Effects of E-Cigarette Advertising*, Marketing Science, 38(6), 994–1022.

Appendix. Research and academic activity

Education

1. Master Studies, SGH Warsaw School of Economics, 2012-2014
   Programme: Quantitative Methods in Economics and Information Systems
   Title of the master thesis: *US GDP forecasting using Bayesian Structural Vector Autoregression models*
   Supervisor: dr hab. Michał Rubaszek, SGH Warsaw School of Economics professor
   Additional achievement: finalist of Econometric Game 2014 as a member of SGH Warsaw School of Economics Representation

2. Bachelor Studies, SGH Warsaw School of Economics, 2009-2012
   Programme: Economics
   Title of the bachelor thesis: *The dynamics of Bertrand model with multi-agent demand and bounded-rational firms*
   Supervisor: dr hab. Michał Ramsza, SGH Warsaw School of Economics professor

Publications

1. **Olesiński B. (2020)**: *The Analysis of the Tobacco Product Bans Using a Random Coefficients Logit Model*, Central European Journal of Economic Modelling and Econometrics, 12(2), 113-144

2. **Olesiński B., Rozkrut M., Torój A. (2020)**: *How Time-Varying Elasticities of Demand Translate into the Excise Related Laffer Surface*, Argumenta Oeconomica, 1(44), 259-299

3. **Dybka P., Kowalczuk M., Olesiński B., Rozkrut M., Torój A., (2019)**: *Currency demand and MIMIC models: towards a structured hybrid method of measuring the shadow economy*, International Tax and Public Finance, 26 (1), 4-40

4. **Dybka P., Olesiński B., Pękała P., Torój A. (2017)**: *To SVAR or to SVEC? On the transmission of capital buffer shocks to the real economy*, Bank & Credit, 48(2), 119-148

5. **Olesiński, B. (2017)**: *When data is not enough. Bayesian approach to linear regression (Gdy dane to za mało. Regresja liniowa w ujęciu bayesowskim)*, in Torój A. (ed.), *Applications of econometrics. 10 harmless examples (Zastosowania ekonometrii. 10 niegroźnych przykładów)*, Oficyna Wydawnicza SGH (Warsaw School of Economics Press), Warsaw

6. **Olesiński B., Rozkrut M., Torój A. (2016)**: *Measuring the consequences of short-termism in business – the econometric evidence for a sample of European*

companies, Annals of the Collegium of Economic Analyses, Warsaw School of Economics, 41, 63–78

7. **Olesiński B. (2012)**: *Proposition of a simple leading industrial confidence indicator (Popozycja prostego, wyprzedzającego wskaźnika koniunktury w przemyśle)*, in Walczyk K. (ed.) Business survey - mirror of an economy. Part I. Works and Materials of Institute of Economic Development at Warsaw School of Economics (Badania koniuktury - zwierciadło gospodarki. Część I. Prace i Materiały Instytutu Rozwoju Gospodarczego SGH), 90, 215-227

Working paper

1. **Dybka P., Olesiński B, Rozkrut M., Torój A. (2020)**: *Measuring the uncertainty of shadow economy estimates using Bayesian and frequentist model averaging*, SGH KAE Working Papers Series 2020/046

Conferences, workshops and seminars

1. 6th Shadow Economy Conference – Tax Evasion and Economic Inequality, University of Trento, Italy, July 2019, *Measuring the uncertainty of model-based shadow economy estimates using Bayesian and frequentist model averaging* (material co-authored with Piotr Dybka, Marek Rozkrut, Andrzej Torój, presented by Bartosz Olesiński)

2. 18. Doctoral Workshop in Statistics and Econometrics in Uniejów, Poland, May/June 2017, *Disaggregation of demand for cigarettes using random coefficients logit model (Dezagregacja popytu na papierosy z wykorzystaniem modelu logitowego z losowymi współczynnikami)*

3. Econometric modelling seminar SENAMEK, SGH Warsaw School of Economics, Poland, February 2017, *Disaggregation of demand for cigarettes using random coefficients logit model – continued (Dezagregacja popytu na papierosy z wykorzystaniem modelu logitowego z losowymi współczynnikami - kontynuacja)*

4. Econometric modelling seminar SENAMEK, SGH Warsaw School of Economics, Poland, January 2017, *Disaggregation of demand for cigarettes using random coefficients logit model (Dezagregacja popytu na papierosy z wykorzystaniem modelu logitowego z losowymi współczynnikami)*

5. Panel data modelling: theory and practice (Modelowanie danych panelowych: teoria i praktyka) conference, SGH Warsaw School of Economics, Poland, May 2016, *The measurement of consequences of short-termism in business. The results of an econometric study for a sample of European enterprises (Pomiar skutków short-termismu dla biznesu. Wyniki badania ekonometrycznego na próbie europejskich*

przedsiębiorstw, material co-authored with Marek Rozkrut and Andrzej Torój, presented by Bartosz Olesiński)

6. Business survey – mirror of an economy (Badania koniunktury – zwierciadło gospodarki) conference, organized by Institute of Economic Development at SGH Warsaw School of Economics, December 2012, *Proposition of a simple leading industrial confidence indicator (Popozycja prostego, wyprzedzającego wskaźnika koniunktury w przemyśle)*

Bartosz Olesiński