Kim and Mark
Agricultural and Food Economics (2023) 11:5
https://doi.org/10.1186/s40100-023-00245-y

Agricultural and Food Economics

RESEARCH Open Access

What factors make consumers in the USA buy hemp products? Evidence from Nielsen consumer panel data

GwanSeon Kim[1] and Tyler Mark[2*]

*Correspondence: tyler.mark@uky.edu

[1] College of Agriculture, Arkansas State University, P.O. Box 1080, Jonesboro, AR 72404, USA
[2] Department of Agricultural Economics, University of Kentucky, 417, Charles E. Barnhart Bldg., Lexington, KY 40546, USA

Abstract

This study aims to identify critical consumer-demographic characteristics associated with the consumption of products containing hemp and investigate their effect on total expenditure in the USA. To estimate the likelihood of market participation and consumption level, the Heckman selection model is employed utilizing Nielsen consumer panel data from 2008 to 2015. Using a Heckman selection model, this study finds that socio-demographic characteristics, especially income, age, and education, play an important role in explaining purchases of and demand for different categories of hemp products. Noteworthy is the finding that the role of the age of the household head is mixed with respect to consumption decisions and consumption level across the products. In most cases, higher education and income households are more likely to buy hemp products, and those are associated with significantly higher consumption levels. The hemp products used in this study are made of hemp seeds, and our findings show that the hemp seed market could be segmented based on the forms. Results provide a basic understanding of a consumer profile and overall hemp market that has had double-digit growth over the last six years. As the industry continues to move forward, policymakers will need a deeper understanding of the factors driving the industry to create regulations supporting the development of the hemp industry.

Keywords: Heckman selection model, Hemp, Nielsen consumer panel, US

Introduction

Can a market expected to top $6.8 billion in sales by 2025 be based on a feedstock classified as a Schedule 1 narcotic in 2013? Over the last two decades, industrial hemp (also known as hemp) has received a great deal of global interest as an agricultural crop. Industrial hemp is a variety of the *Cannabis sativa* plant species with a delta-9 tetrahydrocannabinol concentration (THC) of no more than 0.3 percent on a dry weight basis.[1] Both industrial hemp and marijuana belong to the plant species Cannabis sativa, even though they are genetically different from a chemical makeup and cultivation practice standpoint (Cherney and Small 2016; Datwyler and Weiblen 2006, Johnson 2017a). The

[1] See more details about the 2014 Farm bill at https://www.gpo.gov/fdsys/pkg/BILLS-113hr2642enr/pdf/BILLS-113hr2642enr.pdf (accessed on September 20, 2019).

Comprehensive Drug Abuse Prevention and Control Act of 1970 classified industrial hemp as a Schedule 1 narcotic. In this regard, growing industrial hemp in the USA for commercial purposes was illegal. Section 7606 of the 2014 US Farm Bill set in motion to protect and generate research into industrial hemp as a potential crop by allowing states to develop research and development in accordance with the Farm Bill. Interest in this crop has continued to gain momentum, with acreage growing to approximately 40,000 acres in 2017. In December 2018, the 2018 Farm Bill was approved by US Congress, legalizing industrial hemp at a federal level by removing it from the Controlled Substances Act.[2]

Industrial hemp has more than fifty thousand uses, ranging from fiber to health products, and more than 30 countries currently grow it (Johnson 2017b). The Kentucky Department of Agriculture reports that approximately 55,700 metric tons of industrial hemp are produced worldwide yearly. Approximately 70 percent of the industrial hemp in the world is produced in China, Russia, and South Korea. According to Fortenbery and Bennett (2004), industrial hemp production has environmental benefits, such as low pesticide and herbicide requirements, a wide range of adaptability for agronomic conditions, increased profit centers for US farmers, and relatively low water needs. In addition, the benefits of industrial hemp on the demand side are increased efficiency compared to other inputs for industrial use, health benefits of both hemp oil and hemp heart consumption, and competitive use in textile manufacture (Fortenbery and Mick 2014). For example, hemp is a substitute for cotton to make textiles, as hemp fiber is 10 times stronger than cotton; in addition, hemp can be used as a building material instead of wood at low costs.[3]

Since there is minimal commercial production in the USA due to production restrictions, most hemp-based products are imported from other countries. The process of commercial production in the USA has started as a pilot program for research purposes since 2014. Growing hemp is only legal if growers receive the requisite permit from the state's department of agriculture. For instance, raw and processed hemp fiber is predominantly imported from China, whereas hemp seed and oilcake are imported from Canada (Johnson 2017b). Figure 1 provides the total value of US hemp imports from 2010 to 2015, showing that the total value of imported hemp is increasing.

Hemp Industries Association estimates annual growth in US hemp retail sales averaged more than 15% from 2010 to 2015. The author also mentions that growth is explained by increased sales of hemp-based body products, supplements, and foods, accounting for more than 60% of the value of US retail sales. Recently, Vote Hemp that is a national, single-issue, nonprofit organization and the nation's leading grassroots hemp advocacy group, estimates the total retail value of hemp products sold in the USA in 2016 at approximately $688 million, including food and body products, clothing, auto parts, building materials, and other products. Vote Hemp is dedicated to accepting the

[2] https://www.fda.gov/news-events/congressional-testimony/hemp-production-and-2018-farm-bill-07252019 (accessed on October 30, 2019).

[3] More detailed environmental and economic benefits of hemp can be found at https://nemeton.com/axis-mutatis/hemp.html (assessed on September 25, 2018).

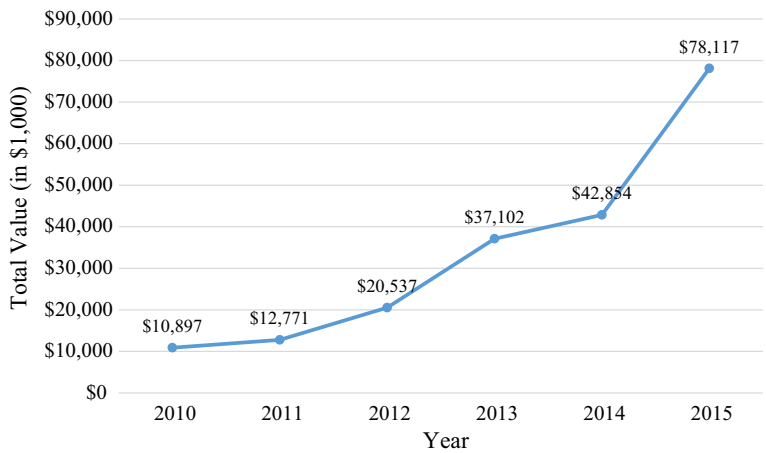

Fig. 1 Total Value of US Hemp Imports, 2010–2015 *Notes*: The main source of the total value for hemp imports is obtained from the US International Trade Commission, and total hemp imports include hemp seed, hemp oil and fractions, hemp seed oilcake and solids, and true hemp. Please see more detailed information on US hemp import at https://fas.org/sgp/crs/misc/RL32725.pdf

free market for industrial hemp and low-THC oilseed and fiber varieties of Cannabis and working to change state and federal laws to allow commercial hemp farming.[4]

Even though there is minimal commercial hemp production in the USA, retail sales for hemp production are increasing over time. To the best of our knowledge, no study has examined the factors that affect the consumption of hemp products. In this study, we investigate the critical economic and demographic characteristics associated with hemp consumption and their effects on US expenditures by utilizing Nielsen's consumer panel data from 2008 to 2015. Due to limited data availability, we only investigated four categories of hemp-containing products: granola, nuts, nutrition, and protein. The term "hemp product" used in this study refers to a product that contains hemp. However, we make no designation regarding the amount of hemp contained in such products: it could be anywhere from 0.0001% to 100%. Later in this paper, we refer to the four different hemp products: hemp granola, hemp nuts, hemp nutrition, and hemp protein, to avoid confusion. This study employs a Heckman selection model in that the model provides different parameters for choice and consumption processes by controlling for non-randomly selected samples. Therefore, we specifically identify the impact of either economics or household characteristics on the probability of purchasing hemp products, and which factors impact total expenditures on these products. Furthermore, we take into account states that have passed regulations allowing the production of industrial hemp. The hypothesis is that the probability of purchasing hemp products is relatively higher in states that have already passed this legislation. Therefore, in this study, hemp legislation refers to any legislation that has been passed or introduced in a state to allow commercial hemp.

[4] More information about their estimates of 2016 annual retail sales for hemp products can be found at https://www.votehemp.com/wp-content/uploads/2018/09/4-14-17-VH-Hemp-Market-Data-2016-FINAL.pdf (accessed on August 18, 2018).

Table 1 The quantity of hemp products sold by region in the USA

Regions	2008	2009	2010	2011	2012	2013	2014	2015
Granola								
Northeast	27,469	17,054	14,508	19,855	23,070	19,023	13,012	14,171
Midwest	36,299	32,427	37,898	42,753	45,648	41,713	37,471	36,763
South	32,387	32,201	34,556	41,673	45,095	42,724	38,065	46,816
West	34,787	37,400	60,834	73,067	75,967	65,929	68,496	77,001
Total	130,942	119,082	147,796	177,348	189,780	169,389	157,044	174,751
Nuts								
Northeast	467	577	970	1873	6,001	18,687	23,220	25,808
Midwest	166	349	751	1189	2902	7607	15,915	19,944
South	222	253	426	613	998	3559	10,271	20,181
West	–	79	214	322	618	1708	26,944	42,806
Total	855	1258	2361	3997	10,519	31,561	76,350	108,739
Nutrition								
Northeast	–	9	26	198	863	10,470	21,753	20,405
Midwest	45	253	94	150	1523	6305	13,531	15,703
South	104	504	32	–	63	553	12,958	14,516
West	4073	3312	2959	6416	9303	12,216	19,194	16,999
Total	4222	4078	3111	6764	11,752	29,544	67,436	67,623
Protein								
Northeast	–	54	47	32	65	185	314	8149
Midwest	253	1540	1018	1367	2230	3256	3403	5027
South	1029	3023	187	153	170	270	276	596
West	660	6106	4806	3235	3813	4845	967	494
Total	1942	10,723	6058	4787	6278	8556	4960	14,266

Background and literature review

Current retail sales of US Hemp products

Nielsen Retail Scanner data provides consumption and accessibility information to allow us to gain a deeper understanding of the US hemp market. The scanner data contains weekly pricing, volume, and store information based on a point-of-sale system with more than 90 participating retail chains in the USA. Table 1 demonstrates the quantity sold for hemp products– granola, nuts, nutrition, and protein–by region in the USA from 2008 to 2015. The regions in Table 1 are based on four statistical regions defined by the US Census Bureau: Northeast, Midwest, South, and West. As shown in Table 1, the total quantity sold in each hemp product category increases over time, regardless of the region. The sales volume of hemp granola is much higher than other hemp products, and about 40% of hemp granola is sold in the West region. This implies that many stores might sell granola hemp, and consumers might have better accessibility in the West region. For the category of hemp nuts, approximately 33% and 31% of hemp nuts are sold in the Northeast and West regions, respectively. Also, about 65% of hemp nutrition is sold mainly in the West and Northeast regions: approximately 38% in the West and 28% in the Northeast. Compared to other hemp products, hemp protein is sold mainly in the West with 43% and Midwest with 31%. For hemp protein, the sales volume in the South region is steadily decreasing since 2008 and rebounding from 2012. To sum, consumption of most hemp products shows

an increasing trend from 2008 to 2015, although there are little variations in four different regions and years. Particularly from 2008 to 2015, the amount of hemp granola sold in the Northeast decreased by approximately 48%, while hemp protein in the Southeast decreased by 42%.

Literature review related to Hemp-based products

The market potential (Cherney and Small 2016) and economic viability (Mark et al. 2020, Ceyhan et al. 2022) for industrial hemp have been actively studied as sales and production of hemp-based products have grown rapidly. In addition, other studies closely related to industrial hemp, such as marijuana (Simkins and Allen 2020; Brown et al. 2020) and cannabis (Jeffers et al. 2021), have been extensively reviewed. However, there is still a lack of literature on consumer demand sectors for hemp products, mainly due to the limited access to market information. Therefore, we review limited recent studies that report and examine the association of consumer demographics with the purchase of hemp-based products.

Cannabidiol (CBD)-based products from industrial hemp, for example, have received great attention and become popular in the USA due to health and wellness benefits. Hemp-derived CBD sales are projected to be increased by 233% in 2022 compared to 2018 (Roach et al. 2019). Especially for CBD-based products, two surveys have been conducted by the Gallup Poll Social Series (GPSS) and Consumer Reports (CR) to find out consumption habits for CBD products. According to Brenan (2019), approximately 14% of Americans use cannabidiol (CBD)-based products. More specifically, younger Americans are more likely to use CBD products and more familiar with the products. In addition, there is regional heterogeneity in the consumption of CBD products: people in the West region tend to use more CBD products compared to people in the South, Midwest, and East regions. Based on the survey by CR, more than 25% of people in the USA tried CBD, and the most popular age group was the 20 s (Gill 2019).

In addition to the literature on CBD products, recent studies by Kolodinsky et al. (2020) and Kolodinsky and Lancasse (2021) investigate consumer response to hemp-based products. Kolodinsky et al. (2020) employ survey data of Vermont residents conducted in February and March of 2019 by the University of Vermont's Center for Rural Studies to examine consumer behavior toward hemp, awareness of hemp products, use of hemp-based products, and support for industrial hemp production in Vermont. By employing bivariate logistic regression, they find that older respondents are less likely to be aware of hemp-based products. However, respondents who are politically liberal are generally more likely to be in the choice set for hemp-based products. Especially they find that younger and more educated people tend to consider the consumption of CBD products more. In contrast, other demographic characteristics such as age, gender, and location are not statistically associated with hemp-based products. Another study by Kolodinsky and Lancasse (2020) examines whether the impact of demographic factors associated with hemp products has changed compared to the study by Kolodinsky et al. (2020) by updating survey data with one additional year. Compared to the previous study, they find that impact of demographic characteristics on support, awareness, and use of hemp products has not significantly changed from 2019 to 2020.

Our study is different from existing studies in that this study is the first paper to use reveled preferences and a national consumer panel to examine the impact of socio-demographic characteristics on the likelihood of buying

Hemp-based products and total expenditure of buying hemp-based products, focusing on food. Therefore, this study contributes to filling the void in the literature on consumer demand for hemp in the USA.

Data description

Nielsen Retail Scanner data provides consumption and accessibility information to allow a deeper understanding of the US hemp market. The scanner data contains weekly pricing, volume, and store information based on a point-of-sale system with more than 90 participating retail chains in the USA. Table 1 demonstrates the quantities sold for hemp products–granola, nuts, nutrition, and protein–by region in the USA from 2008 to 2015. The regions in Table 1 are based on the four statistical regions defined by the US Census Bureau: Northeast, Midwest, South, and West. As shown in Table 1, the total quantity sold in each category of hemp products has been increasing over time, regardless of the region. The sales volume of hemp granola is much higher than that of other hemp products; noticeably, about 40% of hemp granola is sold in the West. This implies that many stores might sell hemp granola and that consumers might have better accessibility in the West. For the hemp nuts category, approximately 33% and 31% of sales are in the Northeast and West, respectively. Also, about 65% of hemp nutrition is sold mainly in the West and Northeast: approximately 38% in the West and 28% in the Northeast. Finally, hemp protein is sold mainly in the West, with 43%, and Midwest, with 31%. Hemp protein sales volume in the South steadily decreased after 2008 but rebounded in 2012. To sum up, consumption of most hemp products shows an increasing trend from 2008 to 2015, although there are variations in the four regions over the years. In particular, from 2008 to 2015, the amount of hemp granola sold in the Northeast decreased by approximately 48%, while sales of hemp protein in the Southeast decreased by 42%.

The consumer panel data started in 2004 and is updated with a 2-year time lag. The database contains information about product purchases made by a representative panel of households, numbering approximately 40,000–60,000, across all retail channels in all US markets, including food, non-food grocery products, health and beauty aids, and general merchandise. The panelist households continuously provide information on what products they purchase and where and when they do so, based on Universal Product Code (UPC) barcodes scanned via in-home scanners. Therefore, the Nielsen Consumer Panel data includes detailed demographic and geographic information about the panelists, products, product characteristics, retail channels, and market locations.

Consumer Panel product data are organized based on the hierarchy, departments, product groups, product modules, and UPC codes. In the first step, we employ a searching index function based on a string of characters that include “hemp” to identify the product hierarchy. Since most hemp products are found in the product groups of cereal, nuts, vitamins, and medications, this study considers only those four product groups. In the second step, we narrow the product groups down to the next hierarchy, product modules, to identify any missing information or irrelevant products associated with the four product groups. In this step, this study excludes product modules if there are few

Table 2 Number of observations for each product with proportion of hemp products

Products	2008	2009	2010	2011	2012	2013	2014	2015	Total
Granola and	727	599	502	559	697	877	1,011	992	5,964
Nature Valley	(25.86)	(32.22)	(38.25)	(37.92)	(25.82)	(17.10)	(14.24)	(17.04)	(23.94)
Nuts (Bags)	1,313	1,304	642	975	1,674	2,859	3,395	3,079	15,241
	(2.74)	(1.99)	(8.10)	(9.74)	(11.95)	(14.20)	(14.02)	(13.51)	(11.20)
Nutritional	1,741	1,744	848	1,211	2,194	3,517	3,966	3,657	18,878
	(0.63)	(1.09)	(1.65)	(2.64)	(1.60)	(2.22)	(3.23)	(3.66)	(2.39)
Protein	105	125	101	132	277	457	569	500	2,266
	(19.05)	(28.00)	(23.76)	(15.15)	(9.75)	(8.97)	(10.19)	(12.60)	(12.71)
Total	3,886	3,772	2,093	2,877	4,842	7,710	8,941	8,228	42,349
	(6.56)	(7.41)	(13.09)	(12.00)	(10.55)	(10.96)	(11.49)	(11.80)	(10.66)

Observations in Table 2 represent the total number of observations for each product type, regardless of whether it contains a hemp product. Parentheses represent the proportion of these observations for which a hemp product is included in the ingredients

or no observations to represent product groups identified in the first step. In the third step, we collect all households from the Nielsen Consumer Panel data and limit the panelists to four main product categories: granola, nut, nutrition, and protein. In the final step, we exclude households based on each household's uniquely assigned store code. This is because some households may not have access to hemp products, if stores do not sell them. In this case, we cannot identify and differentiate factors that make consumers more likely to buy products containing hemp than conventional products. We explicitly classify hemp consumers through these steps, then estimate their probability of purchasing hemp products and the impact of household characteristics on total hemp expenditure. Table 2 shows the number of observations for each product, with the proportion of hemp products. For nuts, for example, 15,241 households consumed nuts from 2008 to 2015, and 11.20 percent of them consumed hemp nuts.

The demographic and socioeconomic characteristics in Nielsen's consumer data, especially education level, age, race, and ethnicity, contain both male and female heads of households. According to Dettmann (2008) and Alviola and Capps (2010), they used either the female or the male head of household for the demographic information by assuming that females carry out the majority of grocery shopping. Therefore, information on the male head of the household was used if the female head of the household was not present in the household. This study follows the previous studies to use male household information if no female head exists in the household.[5] Table 3 shows the summary statistics of the variables used in the analysis.

Many of the demographic and socioeconomic variables in Nielsen's consumer data can be classified into different categories. This study reclassifies some of them for use as explanatory variables. The reclassification of these explanatory variables is as follows. Income in Nielsen is initially classified into 16 different categories, ranging from less than $5,000 to above $200,000. We reclassify these 16 income categories into three

[5] In the estimations for each hemp product, more than 90% of the households are female head households. Specifically, the proportions of female households for granola, nuts, nutrition, and proteins are 90.60%, 92.56%, 91.49%, and 90.01%, respectively.

Table 3 Summary statistics of variables used in the analysis

Variable	Description	Granola		Nuts		Nutrition		Protein	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD
Hemp Exp	Total monthly expenditure for Hemp product in log	1.40	0.48	2.45	0.54	2.44	0.47	2.69	0.39
Hemp	=1 if HH consume Hemp by product	0.24	0.43	0.11	0.32	0.02	0.15	0.13	0.33
Low Income*	=1 if HH income is less than $30,000	0.10	0.30	0.14	0.35	0.14	0.35	0.11	0.31
Median Income	=1 if HH income is between $30,000 and $70,000	0.37	0.48	0.39	0.49	0.39	0.49	0.33	0.47
High Income	=1 if HH income is above $70,000	0.53	0.50	0.47	0.50	0.47	0.50	0.56	0.50
Age1*	=1 if HH age is less than 40	0.14	0.34	0.09	0.28	0.07	0.26	0.13	0.34
Age2	=1 if HH age is between 40 and 64	0.65	0.48	0.65	0.48	0.63	0.48	0.67	0.47
Age3	=1 if HH age is above 64	0.21	0.41	0.26	0.44	0.29	0.46	0.20	0.40
HH Size	Size of Households	2.52	1.26	2.39	1.19	2.28	1.16	2.45	1.29
Married	=1 if HH married	0.73	0.44	0.71	0.46	0.67	0.47	0.68	0.47
Edu1*	=1 if HH education is High School or less	0.15	0.36	0.20	0.40	0.18	0.39	0.13	0.34
Edu2	=1 if HH education is Some College	0.27	0.44	0.30	0.46	0.31	0.46	0.32	0.47
Edu3	=1 if HH education is College Graduate	0.36	0.48	0.34	0.47	0.35	0.48	0.36	0.48
Edu4	=1 if HH education is Post-Collegiate	0.22	0.41	0.16	0.36	0.16	0.36	0.18	0.39
White	=1 if HH is White	0.84	0.37	0.81	0.40	0.81	0.40	0.78	0.42
Black	=1 if HH is African American (Black)	0.06	0.23	0.09	0.29	0.09	0.28	0.11	0.31
Asian	=1 if HH is Asian	0.05	0.21	0.05	0.22	0.05	0.23	0.05	0.21
Other Race*	=1 if HH is other races	0.06	0.23	0.06	0.23	0.05	0.23	0.07	0.25
Hispanic	=1 if HH is Hispanic	0.08	0.27	0.08	0.27	0.07	0.26	0.08	0.27
Employ	=1 if HH is employed	0.64	0.48	0.57	0.50	0.56	0.50	0.64	0.48
Hemp State	=1 if HH is living in State with Hemp Legislation	0.37	0.48	0.40	0.49	0.40	0.49	0.48	0.50
Midwest	=1 if HH is living Midwest region	0.20	0.44	0.21	0.41	0.15	0.36	0.17	0.38
South	=1 if HH is living South region	0.26	0.44	0.31	0.46	0.31	0.46	0.29	0.45
West	=1 if HH is living West region	0.35	0.48	0.32	0.47	0.41	0.49	0.41	0.49
East*	=1 if HH is living East region	0.19	0.39	0.16	0.37	0.13	0.33	0.13	0.33
Observations		5959		15,233		18,871		2263	

S.D represents the standard deviation. HH represents the head of household. A variable with an asterisk symbol represents a reference (base) category

categories: low, if household income is less than $30,000; middle, if household income is between $30,000 and $70,000; and high, if household income is above $70,000. Next, the age of the household head is reclassified from nine categories into three: less than 40 years, between 40 and 64 years, and over 64 years. Finally, the education of the household head is reclassified from six categories into four: high school or less, some college, college graduate, and post-collegiate.

In addition to demographic and socioeconomic characteristics, this study incorporates a new variable called hemp legislation if a state has enacted any legislation allowing for hemp production in that state. This study hypothesizes that households in states where hemp bills and resolutions are introduced are more likely to be exposed to hemp products than households in other states.

The variable *Hemp* is a dependent variable for the probit model and is defined as 1 to represent the purchase of hemp products and 0 otherwise. The percentages of

households purchasing hemp products in granola, nuts, nutrition, and protein categories are approximately 24%, 11%, 2%, and 13%, respectively, from 2008 to 2015. The percentages of low-, medium-, and high-income households across all products are approximately 10%, 37%, and 50%, respectively. On average, households have approximately 2.4 members, 70% include a married couple, and more than 50% are between 40 and 64 years old, across all products. As to other demographic characteristics, more than 50% of household heads are employed, and over 80% have completed at least some college, on average across the products. This study also includes race, and the sample is classified into White, Black, Asian, and other races, with percentages of 84%, 6%, 5%, and 5%, respectively. Additionally, about 8% of the sample are classified as Hispanic. Finally, this study includes four regional dummies, Midwest, South, West, and East. The majority of households, on average across the products, are in the West (about 37%), followed by the South, Midwest, and East. The Nielsen Consumer Panel originally classified nine different regions. However, we reclassified these into four major regional divisions: East includes New England and Middle Atlantic; Midwest includes East North Central and West North Central; South includes South Atlantic, East South Central, and West South Central; and finally, West includes Mountain and Pacific. Even though this study includes year dummies to avoid and control for potential heterogeneity across years, we do not report them in Table 3.

Empirical methodology

This paper employs the Heckman sample selection approach (also called a two-step model) developed by Heckman (1979) to correct for sample selection bias from non-randomly selected samples. Therefore, this study estimates the likelihood of market participation and consumption level. The Heckman selection model is different from other approaches such as Tobit model and Cragg's model (also known as the hurdle model) for the censored data (i.e., truncated sample) in that the Heckman model is based on incidental truncation rather than truncation. The Heckman approach takes place in two stages as follows.

First stage of the Heckman model

The first stage is estimated by the probit model (i.e., selection model), assuming that error terms are normally distributed. The probit model is defined as follows:

$$Pr(z_i = 1) = \Phi(W_i\gamma) \tag{1}$$

where z_i is an indicator that takes on the value of 1 if the household i buys hemp product and 0 otherwise, Φ is the standard normal cumulative distribution function, and W_i is the vector of explanatory variables for the decision to buy hemp products. In the first stage, we obtain estimates of γ by Maximum Likelihood Estimation (MLE), and the inverse Mills ratio (IMR) for each household in the selected sample can be estimated as follows:

$$\text{IMR} = \widehat{\lambda}_i\left(W_i\widehat{\gamma}\right) = \frac{\phi(W_i\widehat{\gamma})}{\Phi(W_i\widehat{\gamma})} \tag{2}$$

where $\phi(W_i\widehat{\gamma})$ is the estimated probability density function (pdf), and $\Phi(W_i\widehat{\gamma})$ is the cumulative density function (cdf). The calculated IMR indicates the probability that the

household i decided to buy hemp products over the cumulative probability of the household's decision. In addition, the IMR captures all the effects of the omitted variables (Alviola and Capps 2010).

Second stage of the Heckman model

In the second stage of the Heckman model, we include estimated IMR as an additional explanatory variable to control the endogeneity since the part of the error term for which the decision to buy hemp products influences the total expenditure. Therefore, the regression model for the selected sample in the second stage is mathematically formed as

$$E(Y_i|z_i = 1) = X_i\beta + \alpha\widehat{\lambda}_i\left(W_i\widehat{\gamma}\right) \tag{3}$$

where Y_i represents the total expenditure of hemp products by the ith household, W is the vector of variables that explain the decision to purchase hemp products, X is the vector of explanatory variables associated with the total expenditure of the hemp products, and α is the parameter related to the IMR.

Marginal effects of the Heckman model

The following discussion about the marginal effects of the Heckman model is based on Saha et al. (1997) and Alviola and Capps (2010). Let X_{ij} denote the jth regression, and it is common for both W_i and X_i. Then, estimated marginal effect (ME) of a change in the regressor is defined as

$$\widehat{\mathrm{ME}}_{ij} = \frac{\partial E(Y_i|z_i = 1)}{\partial X_i} = \beta_j + \alpha\frac{\partial \mathrm{IMR}_i}{\partial X_{ij}} \tag{4}$$

Therefore, the marginal effect of the independent variables on Y_i in the observed sample is composed of two parts. First, there is a direct effect of the expected expenditure on hemp products captured by β_j. Second, the indirect effect is captured by a change in the IMR with respect to a unit change in X_{ij}. The equation above can be simplified and rewritten as

$$\widehat{ME}_{ij} = \widehat{\beta}_j - \widehat{\alpha}\widehat{\gamma}\left(W_i\widehat{\gamma}\widehat{\lambda}_i + (\widehat{\lambda}_i)^2\right) \tag{5}$$

where $\widehat{ME}_{ij}$ represents the marginal effect of the jth explanatory variable for the ith household, $\widehat{\beta}_j$ is a parameter estimates for the jth explanatory variable in the second stage of the Heckman model, $\widehat{\alpha}$ is an estimated parameter for the IMR variable, $\widehat{\gamma}$ is an estimated parameter of the jth explanatory variable in the first stage of the Heckman model, $W_i\widehat{\gamma}$ is the prediction from the probit model for the ith household, and $\widehat{\lambda}_i$ is an estimated IMR for the ith household who purchase hemp products. Saha et al. (1997) and Alviola and Capps (2010) argue $\widehat{ME}_{ij} \neq \widehat{\beta}_j$ in general, but $\widehat{ME}_{ij} = \widehat{\beta}_j$ if and only if $\widehat{\alpha} = 0$, implying covariance of two error terms between first- and second-stage equations are equal to

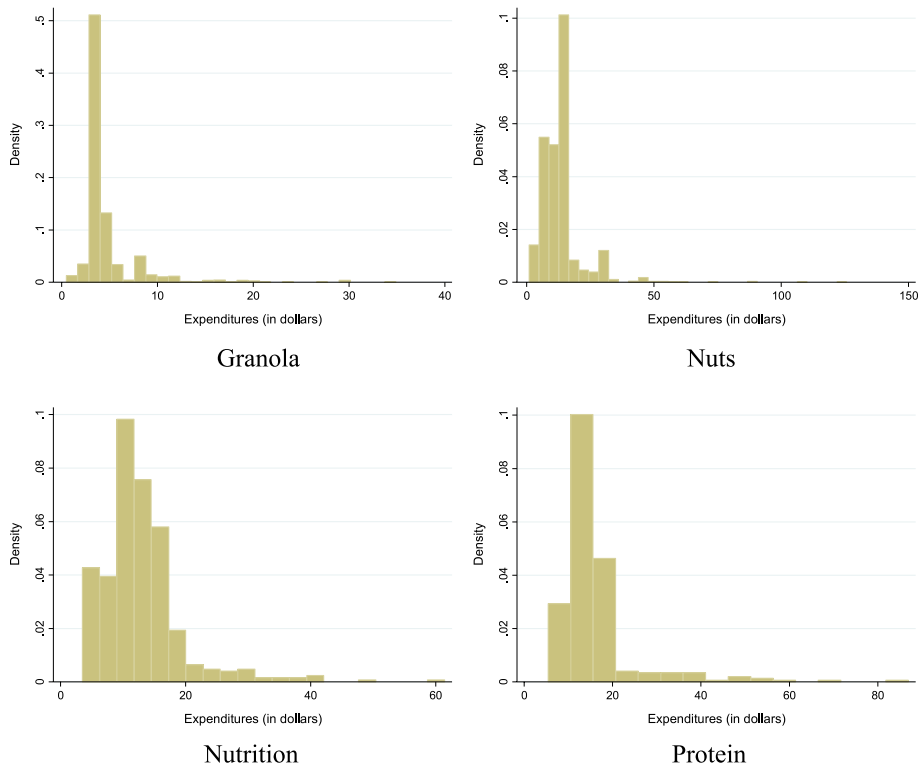

Fig. 2 Distributions of hemp products expenditures in original scale

zero. Since this case is not common, and the ME is different across the observation (i.e., observation dependent), this paper evaluates the ME at the sample mean as follows:

$$\widehat{ME}_{ij}|_{\mathbf{samplemean}} = \widehat{\beta}_j - \widehat{\lambda}\widehat{\gamma}_j\left(\left(\overline{W}\widehat{\gamma}\right)\overline{\widehat{\lambda}} + \overline{\widehat{\lambda}}^2\right) \tag{6}$$

where $\overline{W}$ denote the vector of regressor sample mean and $\overline{\widehat{\lambda}} = \frac{\phi(\overline{W}\widehat{\gamma})}{\Phi(\overline{W}\widehat{\gamma})}$ is the IMR evaluated at the means.

Empirical specification

For the model specification, the first-stage Heckman model, probit model, is hypothesized as a function of the socioeconomic and demographic characteristics including household income, household size, marital status, age, education, race and ethnicity of the household head, employment, and hemp state.[6] The mathematical expression of the probit model for the decision to purchase hemp products is written as follows:

[6] We test multicollinearity between the variables based on variance inflation factor (VIF) across all different categories of hemp products, and we find there is no strong evidence of multicollinearity.

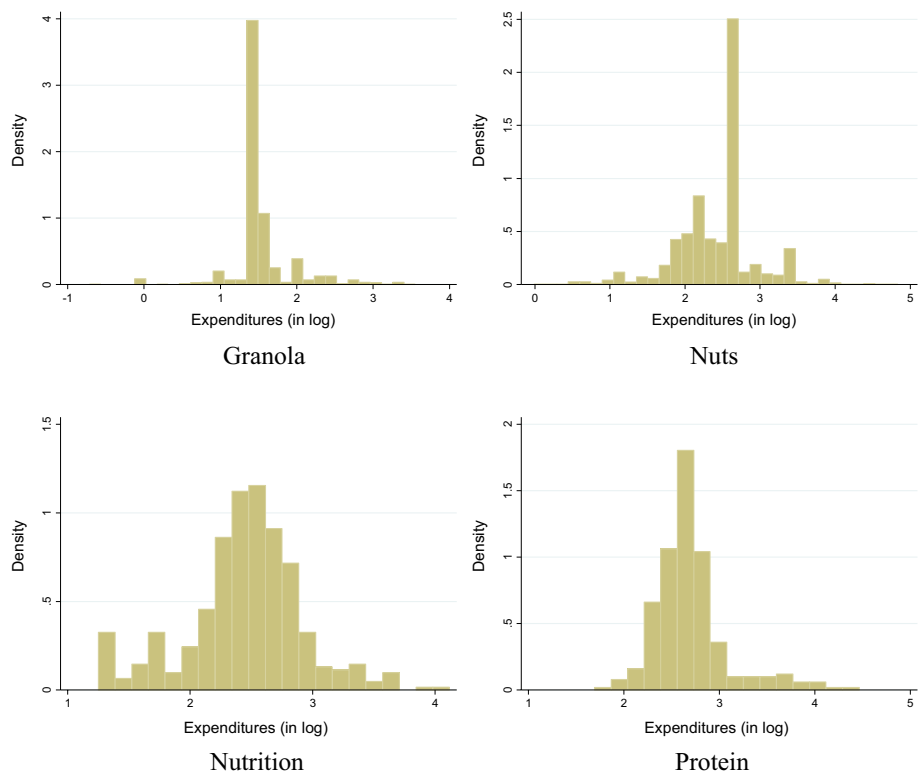

Fig. 3 Distributions of hemp products expenditures in natural logarithm scale

$$
\begin{aligned}
Pr(z_i = 1) = \gamma_0 &+ \gamma_1 M_{\text{Income}} + \gamma_2 H_{\text{Income}} + \gamma_3 \text{Age2} \\
&+ \gamma_4 \text{Age3} + \gamma_5 \text{HHSize} + \gamma_6 \text{Married} + \gamma_7 \text{Edu2} \\
&+ \gamma_8 \text{Edu3} + \gamma_9 \text{Edu4} + \gamma_{10} \text{White} + \gamma_{11} \text{Black} \\
&+ \gamma_{12} \text{Asian} + \gamma_{13} \text{Hispanic} + \gamma_{14} \text{Employ} \\
&+ \gamma_{15} \text{Hemp\_State} + \epsilon_i
\end{aligned} \tag{7}
$$

A description of the variable names in the equation above is based on Table 3 with associated descriptive statistics. The reference category for each dummy variable scheme is excluded and reported with an asterisk symbol in Table 3. Regional and year dummies are also included in the estimation but are not reported in the equation above.[7]

$$
\begin{aligned}
\ln(\text{Expenditure}) = \beta_0 &+ \beta_1 M_{\text{Income}} + \beta_2 H_{\text{Income}} + \beta_3 \text{Age2} \\
&+ \beta_4 \text{Age3} + \beta_5 \text{HHSize} + \beta_6 \text{Married} \\
&+ \beta_7 \text{Edu2} + \beta_8 \text{Edu3} + \beta_9 \text{Edu4} + \beta_{10} \text{White} \\
&+ \beta_{11} \text{Black} + +\beta_{12} \text{Asian} + \beta_{13} \text{Hispanic} \\
&+ \beta_{14} \text{Employ} + \lambda_1 \text{IMR} + u_i
\end{aligned} \tag{8}
$$

This study uses aggregated monthly expenditure for the dependent variable, and we transform the dependent variable into the logarithm form. Econometrically, more outliers in the linear dependent variable reflect high variance and increase the risk of heteroskedasticity. Also, the distribution of consumption data is commonly highly skewed

[7] For the time dummies, this paper employs year fixed effects instead of monthly fixed effects due to the fact we find there is no significant variation among months.

Table 4 First-stage probit estimation results

Variable	Granola		Nuts		Nutrition		Protein	
	Coef	M.E	Coef	M.E	Coef	M.E	Coef	M.E
M_Income	0.154**	0.047	0.151***	0.026	0.041	0.002	0.035	0.007
	(0.069)	(0.021)	(0.047)	(0.008)	(0.065)	(0.003)	(0.128)	(0.025)
H_Income	0.129*	0.038	0.167***	0.029	−0.107	−0.005	−0.077	−0.015
	(0.072)	(0.021)	(0.050)	(0.009)	(0.070)	(0.003)	(0.126)	(0.025)
Age2	−0.199***	−0.061	−0.083	−0.014	−0.298***	−0.016	0.084	0.016
	(0.054)	(0.017)	(0.051)	(0.009)	(0.066)	(0.004)	(0.107)	(0.020)
Age3	−0.294***	−0.082	−0.141**	−0.023	−0.403***	−0.016	−0.113	−0.021
	(0.069)	(0.018)	(0.059)	(0.009)	(0.078)	(0.003)	(0.148)	(0.026)
HH Size	−0.084***	−0.025	−0.021	−0.004	0.005	0.0002	−0.030	−0.006
	(0.018)	(0.005)	(0.015)	(0.002)	(0.021)	(0.001)	(0.034)	(0.007)
Married	−0.066	−0.020	−0.157***	−0.028	0.098*	0.004	0.063	0.012
	(0.050)	(0.015)	(0.037)	(0.007)	(0.054)	(0.002)	(0.096)	(0.018)
Edu2	−0.053	−0.016	0.121***	0.021	0.163***	0.008	0.241**	0.049
	(0.060)	(0.018)	(0.042)	(0.008)	(0.063)	(0.003)	(0.117)	(0.025)
Edu3	−0.039	−0.012	0.100**	0.017	0.178***	0.009	0.237**	0.048
	(0.059)	(0.017)	(0.042)	(0.007)	(0.064)	(0.003)	(0.119)	(0.025)
Edu4	0.029	0.009	0.184***	0.034	0.006	0.0003	−0.035	−0.007
	(0.065)	(0.020)	(0.049)	(0.010)	(0.081)	(0.004)	(0.138)	(0.026)
Employed	0.134***	0.040	0.015	0.003	−0.006	−0.0003	0.159*	0.030
	(0.043)	(0.013)	(0.031)	(0.005)	(0.045)	(0.002)	(0.086)	(0.016)
White	−0.259	−0.082	−0.206***	−0.038	−0.024	−0.001	0.095	0.018
	(0.084)	(0.028)	(0.061)	(0.012)	(0.098)	(0.005)	(0.158)	(0.029)
Black	−0.013	−0.004	−0.281***	−0.041	0.150	0.008	0.019	0.004
	(0.114)	(0.034)	(0.076)	(0.009)	(0.114)	(0.007)	(0.195)	(0.038)
Asian	−0.504***	−0.123	−0.321***	−0.045	−0.283	−0.010	0.167	0.035
	(0.123)	(0.024)	(0.088)	(0.010)	(0.149)	(0.004)	(0.223)	(0.051)
Hispanic	−0.011	−0.003	0.147***	0.027	−0.098	−0.004	0.031	0.006
	(0.073)	(0.022)	(0.053)	(0.010)	(0.089)	(0.004)	(0.148)	(0.030)
Midwest	0.159***	0.049	−0.097*	−0.016	−0.192***	−0.008	0.147	0.030
	(0.059)	(0.019)	(0.051)	(0.008)	(0.067)	(0.002)	(0.124)	(0.027)
South	−0.202***	−0.058	0.260***	0.047	−0.321***	−0.014	0.092	0.018
	(0.059)	(0.016)	(0.046)	(0.009)	(0.060)	(0.002)	(0.115)	(0.023)
West	0.268***	0.082	0.269***	0.049	−0.307***	−0.014	−0.279**	−0.052
	(0.053)	(0.017)	(0.044)	(0.008)	(0.057)	(0.002)	(0.118)	(0.021)
Hemp State	0.086*	0.026	−0.055*	−0.009	0.001	0.0001	−0.027	−0.005
	(0.044)	(0.013)	(0.032)	(0.005)	(0.057)	(0.003)	(0.082)	(0.016)
Constant	−0.302**		−1.940***		−2.090***		−1.068***	
	(0.145)		(0.125)		(0.189)	(0.002)	(0.297)	
Log Likelihood	−3054.221		−5035.573		−2027.504		−812.334	
McFadden R^2	0.069		0.058		0.046		0.058	
Observations	5,959		15,233		18,871		2,263	

Significance levels are indicated by ***, **, * for 1, 5, and 10 percent significance levels, respectively. Robust standard errors are reported in parentheses. This study also includes time dummy variables (year dummies in the model) to control for time-specific fixed effects, but we do not report the results in Table 5

Table 5 The goodness of fit measures from the Probit model

Categories	Sensitivity (%)	Specificity (%)	Cutoff value	% of Correct Predictions
Granola	63.77	64.41	0.239	61.97%
Nuts	72.47	51.28	0.112	53.65%
Nutrition	67.11	59.6	0.024	59.77%
Protein	63.19	61.62	0.127	61.82%

Sensitivity represents the percentage of correctly predicting the choice of hemp products, whereas specificity represents the percentage of correctly predicting the choice of choosing non-hemp products

(Zhang et al. 2008), and estimators might be inconsistent with a dependent variable without transformation (Newman et al. 2003). According to Newhouse (1987), Wagner and Hanna (1983), and Zhang et al. (2008), the transformation of the natural logarithm will control for positively skewed expenditures. Figure 2 shows the histograms for hemp products. The distributions of the values are positively skewed across all categories, implying that most consumers spend small amounts of money to purchase hemp products. On the other hand, the distributions of the expenditures in the natural logarithm of hemp products are normally distributed, as shown in Fig. 3. For both Figs. 2and 3, we only use the positive value of the expenditures that head of household who purchase hemp products. Therefore, we use the logarithm dependent variable to reduce problems resulting from a non-normal distribution. In the second-stage estimation, the variable of IMR calculated from the probit model is included to test the selection bias.

Results and discussion

First-Stage estimation

The results of the first-stage probit model for the four different categories are reported in Table 4, including the maximum log-likelihood estimates and McFadden R^2. The marginal effects associated with the estimates of the parameters are also reported in Table 4, since the magnitude of the coefficients does not provide direct interpretation. Looking at the marginal effects in Table 4, households are more likely to consume hemp granola and nuts than low-income households. Interestingly, we find that number of households is negatively associated with the likelihood of buying hemp granola. In general, the likelihood of buying production increases as the number of household members increases. However, the negative relationship might be explained due the lack of consumer awareness and preference toward the new emerging products in hemp market. For marital status, the likelihood of buying hemp products is positively related to hemp nutrition but negatively related to hemp nuts. The mixed results might be due to no clear logic or theory to explain and support the relationships (Kim et al. 2018). Older households are less likely to consume all hemp products except hemp protein than younger households below 40 years old. For education level, we find that most categories of hemp products except hemp granola are more likely to be consumed as education increases, indicating the lack of knowledge or awareness toward hemp granola. Employed households are more likely to buy hemp granola and protein than unemployed households. Also, we find significant regional effects on the probability of buying hemp products, but these regional effects vary across the categories of hemp products. This suggests that

Table 6 Second stage estimation results

Variable	Granola		Nuts		Nutrition		Protein	
	Coef	M.E	Coef	M.E	Coef	M.E	Coef	M.E
M_Income	0.080*	0.085	−0.039	0.058	0.102	0.115	0.089	0.073
	(0.044)	(0.044)	(0.053)	(0.044)	(0.070)	(0.066)	(0.095)	(0.058)
H_Income	0.094**	0.098	0.004	0.106	0.119	0.087	0.056	0.083
	(0.045)	(0.046)	(0.055)	(0.046)	(0.078)	(0.074)	(0.097)	(0.061)
Age2	−0.110***	−0.117	0.091*	0.042	0.226***	0.136	−0.027	−0.073
	(0.040)	(0.040)	(0.053)	(0.042)	(0.071)	(0.060)	(0.079)	(0.048)
Age3	−0.010	−0.021	0.113*	0.025	0.206**	0.083	−0.247**	−0.192
	(0.049)	(0.048)	(0.066)	(0.054)	(0.091)	(0.075)	(0.106)	(0.063)
HH Size	−0.064***	−0.067	0.028*	0.014	−0.035	−0.034	−0.013	0.001
	(0.014)	(0.014)	(0.015)	(0.012)	(0.023)	(0.022)	(0.022)	(0.015)
Married	0.096***	0.094	0.049	−0.055	0.077	0.107	−0.036	−0.063
	(0.033)	(0.033)	(0.043)	(0.036)	(0.060)	(0.058)	(0.071)	(0.045)
Edu2	0.063	0.061	−0.072	0.008	−0.072	−0.023	0.077	−0.023
	(0.042)	(0.042)	(0.048)	(0.038)	(0.073)	(0.066)	(0.086)	(0.055)
Edu3	0.003	0.002	−0.070	−0.007	−0.031	0.023	0.062	−0.041
	(0.040)	(0.040)	(0.046)	(0.037)	(0.071)	(0.065)	(0.089)	(0.057)
Edu4	−0.029	−0.028	−0.117**	0.006	0.002	0.004	−0.114	−0.064
	(0.044)	(0.043)	(0.056)	(0.044)	(0.085)	(0.081)	(0.102)	(0.065)
Employed	0.072***	0.077	−0.061	−0.052	−0.042	−0.044	0.093	0.002
	(0.027)	(0.027)	(0.037)	(0.030)	(0.050)	(0.048)	(0.064)	(0.039)
White	−0.071	−0.081	0.158**	0.039	0.060	0.053	0.038	−0.034
	(0.055)	(0.055)	(0.064)	(0.052)	(0.090)	(0.088)	(0.113)	(0.056)
Black	−0.119*	−0.119	0.164**	0.000	0.122	0.167	0.029	−0.010
	(0.065)	(0.065)	(0.083)	(0.067)	(0.112)	(0.111)	(0.144)	(0.085)
Asian	−0.135	−0.154	0.186**	−0.014	0.181	0.096	−0.091	−0.179
	(0.084)	(0.083)	(0.092)	(0.073)	(0.175)	(0.166)	(0.161)	(0.098)
Hispanic	−0.144***	−0.145	−0.040	0.056	0.183*	0.153	−0.014	−0.016
	(0.049)	(0.049)	(0.060)	(0.049)	(0.099)	(0.095)	(0.107)	(0.056)
Midwest	0.029	0.035	0.064	−0.001	0.254***	0.195	0.237***	0.148
	(0.045)	(0.044)	(0.060)	(0.049)	(0.073)	(0.074)	(0.088)	(0.052)
South	−0.105**	−0.113	−0.083	0.073	0.328***	0.229	0.247***	0.173
	(0.041)	(0.041)	(0.053)	(0.042)	(0.073)	(0.071)	(0.081)	(0.049)
West	−0.076*	−0.066	−0.054	0.115	0.391***	0.297	0.067	0.184
	(0.039)	(0.038)	(0.052)	(0.042)	(0.074)	(0.071)	(0.092)	(0.059)
Lambda	−0.046*	–	−0.765***	–	−0.340*	–	0.657***	–
	(0.026)	–	(0.047)	–	(0.194)	–	(0.059)	–
Constant	1.493***	–	3.698***	–	3.101***	–	1.811	–
	(0.097)	–	(0.195)	–	(0.530)	–	(0.209)	–
Log Likelihood	−3,976.466		−6,340.948		−2,278.849		−894.196	
Censored	4,532		13,526		18,421		1,975	
Uncensored	1,427		1,707		450		288	
Observations	5,959		15,233		18,871		2,263	

Significance levels are indicated by ***, **, * for 1, 5, and 10 percent significance levels, respectively. Robust standard errors are reported in parentheses. This study also includes time dummy variables (year dummies in the model) to control for time-specific fixed effects, but we do not report the results in Table 6

consumers may have different preferences for hemp products across the regions, regardless of the quantity sold. For example, the likelihood of buying hemp granola is lower in the South but higher in the Midwest and West than in the Northeast, even though the quantity of hemp granola sold in the South is about 13% higher than in the Northeast, based on Table 1. States that have enacted industrial hemp legislation are more likely to consume hemp granola but less likely to consume hemp nuts, relative to states with no hemp legislation. This could be due to a lack of hemp processing and its minimal advertising in states without hemp legislation.

After the estimation of the probit model, prediction success is evaluated to assess the usefulness of the probit model to test whether any households in the model that purchased hemp products are correctly classified, as suggested by other studies such as Alviola and Capps (2010), Capps et al. (1999), and Park and Davis (2001). Table 5 shows the goodness of fit measures from the probit model for all four categories. To generate the classification statistics, especially the percentage of correct predictions, we employ different cutoff values for each category rather than the default value of 0.5. This is because the classification of households who purchased hemp products is incorrect if the default value is used instead of market penetration (Alviola and Capps 2010). Therefore, the cutoff value represents market penetration or the proportion of households who purchase hemp products. As shown in Table 5, the percentage of correct predictions of hemp and non-hemp granola, nut, nutrition, and protein products are 61.97%, 53.65%, 59.77%, and 61.82%, respectively. Based on the sensitivity in Table 5, our models correctly predict the decision to buy granola, nut, nutrition, and protein hemp products approximately 63%, 72%, 67%, and 63%, respectively. Furthermore, the decision to purchase non-hemp granola, nut, nutrition, and protein products is correctly predicted approximately 64%, 51%, 59%, and 61% of the time, respectively.

Second-stage estimation

The results of the second-stage estimation are reported in Table 6. Within the second stage of results, we estimate the lambda (i.e., Inverse Mills Ratio) to test sample selection bias. This is statistically significant for the granola, nut, nutrition, and protein categories at the 0.10, 0.01, 0.10, and 0.01 levels, respectively. This indicates evidence of sample selection bias, and justifies the use of the Heckman selection model. In Table 6, we also reported the marginal effects that are evaluated at the mean due to the observation dependence. The marginal effect used in this study is the partial effect on the truncated mean. It is calculated based on consumers with an observed value by excluding consumers who do not purchase hemp products. For the second-stage estimation, once households have decided to buy hemp products (i.e., hemp buyers), higher-income households are positively associated with total expenditure, but only hemp granola. If the household income is above $70,000, the total expenditure on hemp granola increases by 9.8% compared to households whose income is less than $30,000, ceteris paribus. Total expenditure for hemp nuts and nutrition are positively associated with households in higher age groups. Specifically, on average, the total expenditure on hemp nuts increases by 2.5%. In comparison, the total expenditure on hemp protein decreases by 19.2% if the age of the household is above 74, as compared to the household whose age is less than 40. This finding suggests that the younger age group may be looking for healthier protein sources

than the older age group. Noteworthy is the finding that married and employed households positively influence the total expenditure for hemp granola. Additionally, the total expenditure on hemp granola decreases by 6.7%, whereas the total expenditure on hemp nuts increases by 2.8% for one additional household. Across all hemp product categories, we find higher education level is not statistically related to the total expenditure across most categories. For the different regions, households in the South and West consume less hemp granola than households in the East. In contrast, households in the Midwest and South consume more hemp nutrition and protein than households in the East. Two potential factors can explain these findings. First, households in the South and West have less access to stores to purchase hemp granola and fewer available products than households in the East. Likewise, more stores might carry hemp nutrition and protein in the Midwest and South than in the East. Second, households in the Midwest and South might have a stronger perception of hemp products as nutritious and protein-rich than households in the East.

Concluding remarks

Industrial hemp, as a variety of the Cannabis sativa plant species, has received great interest in the last two decades since it has many benefits for the environment, production, and health. The passage of the 2014 Farm Bill only accelerated interest in this crop and its potential. In global markets, industrial hemp is an agricultural crop used in textiles, automotive paneling, furniture, food, personal care, construction, paper. In December 2018, the 2018 Farm Bill was passed and approved by the US Congress, legalizing industrial hemp production. Ratification of this legislation would open up the opportunity for commercial hemp production and increase the supply of hemp available in the US market.

In the USA, retail sales for hemp production are increasing over time even though there is no commercial hemp production due to government restrictions. This study investigates the critical socio-demographic factors associated with increasing hemp consumption and measures their effects on total expenditure in the USA by utilizing Nielsen consumer panel data from 2008 to 2015. The study identifies a more objective consumer profile for this developing industry by analyzing the retail data. Knowing this consumer profile can contribute to the viability of the hemp product market in the USA.

By employing the Heckman selection model, this study finds that socio-demographic characteristics, especially income, age, and education, play an important role in explaining purchases of and demand for different categories of hemp products. Specifically, a higher income is positively associated with the probability of consuming hemp products and with the level at which they are consumed. The role of the age of the household head is mixed with respect to consumption decisions and consumption level across the products: it has a negative and significant effect on the probability of buying hemp products, except for hemp protein, but a positive and significant effect on total expenditure on hemp nuts and nutrition, once households decide to buy hemp products. In most cases, better educated households are more likely to buy hemp products, and those households are associated with significantly higher consumption levels, except in the case of hemp nuts. These findings will provide insights into a more targeted marketing strategy for hemp industries seeking to attract new consumers and increase sales from current

consumers. Many different markets, such as hemp seed, hemp fiber, and hemp CBD, can be derived from the hemp industry, since hemp has more than fifty thousand uses. The hemp products used in this study are made of hemp seeds; however, our findings show that the hemp seed market could be segmented based on the forms: hemp cereal, hemp nuts, hemp nutrition, and hemp protein.

Industrial hemp was recently removed from the Schedule 1 narcotic list and is now legal to produce in the USA, according to the 2018 Farm Bill. As of 2021, over 50 states already have hemp legislation in place to allow for industrial hemp production and processing. To the best of our knowledge, there is no empirical study of hemp in the USA. Thus, the findings of this study will begin to fill the knowledge gap about a crop that is increasing in consumption and production in the USA. As the industry continues to move forward, the findings of this study may also open the door to create business and marketing plans that allow the adoption of goals and strategies by marketers, retailers, and other stakeholders. Findings from this study contribute to existing but limited industrial hemp literature, but it also has the potential to generate significant discussion. Little is known about modern industrial hemp, and many unknowns exist, from its production to its marketing channels. A basic understanding of consumer profiles is a starting point for these discussions.

Acknowledgements
The authors would like to thank the Marketing Data Center at the University of Chicago Booth School of Business. Information on accessing this data can be found at http://research.chicagobooth.edu/nielsen/. Researcher(s) own analyses calculated (or derived) based in part on data from Nielsen Consumer LLC and marketing databases provided through the NielsenIQ Datasets at the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business. The conclusions drawn from the NielsenIQ data are those of the researcher(s) and do not reflect the views of NielsenIQ. NielsenIQ is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein.

Author contributions
GK acquired the data and performed statistical analysis with guidance from TM. GK drafted the manuscript, and TM critically reviewed it. All authors read and approved the final manuscript.

Funding
This research was supported by National Science Foundation (Award # 1355438).

Availability of data and materials
The data that support the findings of this study are available from the Kilts Center for Marketing Data Center at the University of Chicago Booth School of Business but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are, however, available from the authors upon reasonable request and with permission of the Kilts Center for Marketing Data Center at the University of Chicago Booth School of Business.

Declarations

Competing interests
The authors declare that they have no competing interests.

Received: 22 February 2022 Revised: 28 October 2022 Accepted: 1 February 2023
Published online: 24 February 2023

References

Alviola PA, Capps O (2010) Household demand analysis of organic and conventional fluid milk in the United States based on the 2004 Nielsen Homescan panel. Agribusiness 26:369–388
Brenan M (2019) 14% of Americans Say They Use CBD Products. Gallup. Accessed October 05, 2022, at https://news.gallup.com/poll/263147/americans-say-cbd-products.aspx
Brown JD et al (2020) Characteristics of older adults who were early adopters of medical cannabis in the Florida medical marijuana use registry. J Clin Med 9(4):1166
Capps O et al (1999) Examining packer choice of slaughter cattle procurement and pricing methods. Agric Resour Econ Rev 28:11–25

Ceyhan V et al (2022) Economic viability of industrial hemp production in Turkey. Ind Crops Prod 176:114354
Cherney JH, Small E (2016) Industrial hemp in North America: production, politics and potential. Agronomy 6:58
Datwyler SL, Weiblen GD (2006) Genetic variation in hemp and marijuana (*Cannabis sativa* L.) according to amplified fragment length polymorphisms. J Forensic Sci 51:371–375
Dettmann RL (2008) Organic produce: Who's eating it? A demographic profile of organic produce consumers. In: American Agricultural Economics Association Annual Meeting, Orlando. pp 27–29
Johnson R (2017a) Defining "industrial hemp: a fact sheet. In: Library of congress Washington DC congressional research service
Johnson R (2017b) Hemp as an agricultural commodity. In: Library of congress Washington DC congressional research service
Fortenbery TR, Mick BT (2014) Industrial Hemp: opportunities and challenges for Washington
Fortenbery TR, Bennett M (2004) Opportunities for commercial hemp production. Rev Agric Econ 26:97–117
Gill LL (2019) CBD goes mainstream. Consumer reports. Available at https://www.consumerreports.org/cbd/cbd-goes-mainstream/ (Accessed October 05, 2022)
Heckman J (1979) Sample specification bias as a selection error. Econometrica 47:153–162
Jeffers AM et al (2021) Socio-demographic characteristics associated with and prevalence and frequency of cannabis use among adults in the US. JAMA Netw Open 4(11):e2136571–e2136571
Kim G, Seok JH, Mark TB (2018) New market opportunities and consumer heterogeneity in the US organic food market. Sustainability 10(9):3166
Kolodinsky J, Lacasse H, Gallagher K (2020) Making hemp choices: evidence from Vermont. Sustainability 12(15):6287
Kolodinsky J, Lacasse H (2021) Consumer response to hemp: a case study of Vermont residents from 2019 to 2020. GCB Bioenergy 13(4):537–545
Mark T et al. (2020) Economic viability of industrial hemp in the United States: a review of state pilot programs, EIB-217, U.S. Department of Agriculture, Economic Research Service
Newhouse JP (1987) Health economics and econometrics. Am Econ Rev 77:269–274
Newman C, Henchion M, Matthews A (2003) A double-hurdle model of Irish household expenditure on prepared meals. Appl Econ 35:1053–1061
Park J, Davis GC (2001) The theory and econometrics of health information in cross-sectional nutrient demand analysis. Am J Agr Econ 83:840–851
Roach A, Milhollin R, Horner J (2019) Market opportunities for industrial hemp: guide to understanding markets and demand for various industrial hemp plant products. Available at https://extension.missouri.edu/publications/mx72 (Accessed October 05, 2022)
Saha A, Capps O, Byrne PJ (1997) Calculating marginal effects in dichotomous-continuous models. Appl Econ Lett 4:181–185
Simkins TJ, Allen BJ (2020) Illicit to legal: marijuana as a de-stigmatising product and the role of social acceptability in new product adoption. Innovation 22(4):447–468
Wagner J, Hanna S (1983) The effectiveness of family life cycle variables in consumer expenditure research. J Consumer Res 10:281–291
Zhang F et al (2008) Modeling fresh organic produce consumption with scanner data: a generalized double hurdle model approach. Agribusiness 24:510–522

Publisher's Note