فصلنامهٔ اقتصاد و الگوسازی
دانشگاه شهید بهشتی، تابستان و پاییز ۱۳۹۲

Quarterly Journal of Economics and Modelling
Shahid Beheshti University

استفاده از الگوهای متغیر وابسته گسسته در تحلیل رفتار مصرفی خانوار (مورد مصرف دخانیات)

دکتر محمد امین کوهبر*

تاریخ دریافت ۹۳/۸/۱۵ تاریخ پذیرش ۹۳/۱۲/۱۲

چکیده

هدف از نگارش این مقاله، ارائه سیر تکاملی برآوردکننده‌های مناسب متغیر وابسته گسسته، ارائه نکاتی در مورد تحقیقات صورت گرفته مشابه در داخل کشور، خطاهای متداول در برآورد توابع تقاضا و رویکردهای اقتصاد سنجی نوین در استفاده بهینه از اطلاعات داده‌هایی با وجود صفرهای مکرر است. بدین منظور، از داده‌های مصرف دخانیات در سال ۱۳۹۰ برگرفته از مرکز آمار ایران برای ۳۱۱۲۰ خانوار شهری و روستایی استفاده شده است. ممکن است در مورد مصرف سیگار، حتی بیش از نیمی از مشاهدات صفر و مابقی آنها عددی مثبت باشند. در واقع، بسیاری از افراد اصلاً مصرف‌کننده بالقوه نبوده و مشارکتی در مصرف دخانیات ندارند. در این مطالعه با استفاده از داده‌های هزینه درآمد خانوارهای شهری روستایی ایران در سال ۱۳۹۰ نشان داده شده که برآوردهای حداقل مربعات، توبیت و حتی هکمن از قدرت توضیح دهندگی پایینی نسبت به مدل دابل هاردل (دو مانعی) برخوردار بوده و از این حیث موافق نتایج مطالعات معتبر اخیر صورت گرفته در دنیا بوده است.

کلید واژه‌ها: *تقاضای دخانیات، متغیر وابسته گسسته، دو مانعی، توبیت، هکیت، اقتصاد سنجی داده‌های خرد*

طبقه‌بندی JEL: *C24, D12*

* استادیار گروه اقتصاد دانشگاه علوم و فنون دریایی خرمشهر aminkuhbor@yahoo.com

۱- مقدمه:

امروزه تحقیقات فراوانی در موضوع رفتار مصرف‌کننده و تخمین تقاضای گروه‌های مصرفی انجام شده و برپایه نتایج بدست آمده، پیشنهاداتی صورت گرفته و در مواردی هم به اجرا رسیده است. بنابراین خواص مطلوب برآوردکننده‌ها، از جمله سازگاری و دقت آنها حائز اهمیت بوده و می‌بایست به آن توجه شود. پیشرفتهای اخیر تکنیکهای اقتصاد سنجی به بهبود دقت برآوردها کمک نموده و استفاده از الگوهای قدیمی تر به مرور زمان شایستگی خود را از دست داده است. در این مقاله، سیر تکاملی مدلهای اقتصاد سنجی استفاده از داده‌های خرد و متغیر وابسته گسسته را مرور نموده و آخرین دستاوردهای استفاده از داده‌هایی با مشاهدات صفر فراوان معرفی شده است. رفتار مصرفی دخانیات نیز بعنوان شاهد آماری مورد بررسی قرار گرفته است. قبل از پرداختن به بدنه اصلی بحث، لازم است تعاریف مختصری از واژگان تخصصی مورد استفاده در مقاله ارائه شود.

به طور عمومی، توزیع داده‌هایی از خانواده توزیع‌های سانسور شده[1] و یا بریده شده[2] در صفر به شمار می‌روند که مشاهدات متغیر وابسته در جامعه دارای صفرهای فراوان بوده[3] یا در نمونه‌گیری هم از این مشاهدات استفاده کرده، و با آنها برخورد مناسب صورت گیرد و یا اینکه از کل جامعه صفرها را حذف کرده و از زیر بخش مثبت نمونه‌گیری می‌شود[4]. (گرین صص۷۵۷-۷۶۶). توزیع بریده یا شاخ و برگ زده، عبارتست از بخش خاصی از توزیع جامعه که دارای ویژگی مورد نظر است. مثلاً اگر در ایران بخواهیم جامعه خانوارهایی با درآمد بیشتر از ۶ میلیون تومان در سال را در نظر بگیریم، این زیر بخش از جامعه، دارای توزیع بریده است. پس توزیع بریده شده، بخش خاصی از توزیع اصلی است که مقادیر آن از یک و یا دو طرف محدوده شده باشند[5]. طبعاً اگر از این زیر بخش

1. Censored
2. Truncated
۳. در حالت کلی، میتوان بجای صفرها بازه خاصی از داده را حذف کرد. اما در مقاله حاضر بیشتر تمرکز روی صفرها است.
4. Greene, Econometric analysis, 5th ed, pp 757-766.
5. Truncation

نمونه‌ای تصادفی گرفته شود، همه اعضای نمونه دارای ویژگی مورد نظر بوده و مشاهده خارج از محدوده در آنها وجود نخواهد داشت. (همان منبع)

در مقابل، در توزیع سانسور شده، از جامعه اصلی[1]، نمونه‌گیری شده ولی تنها از اطلاعات بخش خاصی از ان نمونه استفاده می‌شود. در این مواقع داده‌های خارج از محدوده مورد نظر را با استفاده از کددهی، به عددی معین (مثلاً صفر) تبدیل می‌کنیم[2]. باید باید توجه شود که هدف محقق، همواره استنتاج آماری در مورد کل جامعه است و نه محدوده مورد نظر. در حالیکه، معمولا توزیع سانسور شده تنها دارای مشاهدات مثبت بوده و صفرهای موجود در نمونه حذف شده اند. در حالت توزیع بریده، توزیع اولیه به هم نمی‌خورد و ما تنها از بخش خاصی از آن توزیع استفاده می‌کنیم، درحالی‌که در حالت سانسور شده شکل توزیع اولیه حفظ شده، اما توزیع نمونه ای، دیگر ترکیبی از دو ناحیه گسسته و پیوسته می‌شود. برای داده‌های خارج از محدوده مورد نظر، توزیع احتمال صفر شده و در مقابل، چگالی داده‌های غیر صفر افزایش میابد.[3] نمونه‌های بارز این توزیع، خرید مواد غذایی و دخانیات، درخواست بیمه، بازپرداخت وام و غیره می‌باشد(همان منبع). در ادامه، خلاصه‌ای از تحقیقات انجام شده در زمینه مورد نظر مرور شده، سپس مبانی اقتصاد سنجی و تکامل تدریجی این الگوها ارائه شده است. نهایتاً داده‌های مصرف دخانیات را در کشور با تمامی الگوهای رقیب برآورد کرده و الگوی بهینه را انتخاب خواهیم نمود.

۲- پیشینه تحقیق

هوانگ و لین درمقاله‌ای[4] با عنوان تخمین تابع تقاضای مواد غذایی، با استفاده از اطلاعات مقطعی ۴۵۰۰ خانوار ایالات متحده، ضمن به کارگیری سیستم تقاضای تقریباً ایده آل، دست به تخمین تقاضای ۱۳ گروه غذایی زده و کشش‌های قیمتی و در آمدی ۲۰ مادهٔ غذایی را استخراج کرده اند. مدل مورد استفاده، حداقل مربعات تعمیم یافته بوده و علیرغم

1. Untrancated

۲. به معادله شماره ۴ همین مقاله بعنوان نمونه رجوع گردد.

۳. تا همچنان سطح زیر توزیع چگالی ۱ باقی بماند.

4. Huang & Lin

وجود مصرف صفر در بسیاری از گروه‌های غذایی، اریب ناشی از عدم توجه به این داده‌ها لحاظ نشده است. اغلب پارامترها معنی دار شدند و نتایج بدست آمده، به لحاظ علامت و مقدار مطلق موافق انتظار به دست آمد. ون چرن به همراه سه محقق دیگر[1] در مقاله ای، الگوی مصرف مواد غذایی خانوارهای ژاپنی را تحلیل نموده اند. هدف اصلی آن‌ها در این مطالعه پاسخ به این سئوال بود که اولاً آیا برنج در ژاپن یک کالای پست محسوب می‌شود یا خیر؟ ثانیاً آیا الگوی مصرف مواد غذایی ژاپنی‌ها غرب گرایی داشته است؟ آن‌ها با استفاده از داده‌های سالانه هزینه درآمد بیش از ۹۵ هزار خانوار ژاپنی از مرکز آمار FIES[2]، در سال ۱۹۹۷ میلادی به تخمین تقاضای ۱۱ گروه عمدهٔ غذایی پرداختند. محققان محققان مسئله داده‌های صفر را لحاظ کرده و از الگوهای توبیت و هکمن بعنوان مدلهای رقیب حداقل مربعات استفاده کرده اند. نتایج نشان می‌دهد که الگوی توبیت بهترین تصریح را در میان سایر مدلها داشته، بسیاری از کششهای مخارج غذایی، بیش از حد انتظار به دست آمده اند. کوهبر و کیانی(۱۳۸۷) در مقاله‌ای تحت عنوان بررسی کشش‌های مواد مغذی، تقاضای مواد غذایی را در کشور ایران با استفاده از داده‌های هزینه درآمد خانوارهای شهری- روستایی در سال ۱۳۸۳ برآورد و برای مقابله با مشکل داده‌های صفر از الگوی توبیت استفاده کرده اند. اغلب ضرایب به‌دست آمده معنی دار بوده و به لحاظ علامت و مطلق[3] موافق انتظار بدست آمدند.

۳- مدلهای مختلف برخورد با داده‌های صفر:

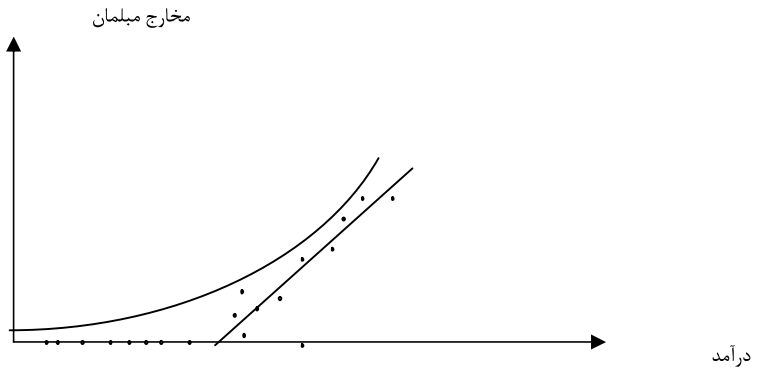
در ابتدا نشان داده می‌شود که استفاده از برآوردکننده‌های حداقل مربعات چه مشکلاتی را به دنبال خواهد داشت. به همین منظور کار را با مطالعه توبین شروع می کنیم.

۳-۱- الگوهای لاجیت، پروبیت و توبیت: توبین (۱۹۵۸) با الهام از الگوی پروبیت، مشکل استفاده از برآوردکننده‌های خطی را در هنگامی که صفرهای مکرر (به هر

1. Chern
2. Family Income and Expenditure Survey.(FIES)

۳. به لحاظ مطلق، یعنی ضرایب، لوکس نرمال یا ضروری بودن کالاها و نیز کششهای قیمتی متقاطع را طبق انتظار نشان دادند.

دلیلی) وجود دارند نشان داد.



منبع: توبین؛ ص ۲۵

نمودار۱- پراکنش نقاط در حالت وجود صفرهای مکرر

وی نشان داد اگر صرفاً از داده‌های غیر صفر استفاده شود، $E(u)$ الزاماً مساوی صفر نشده و حتی تضمینی بر سازگاری ضرایب حاصل از روش حداقل مربعات نیست. به‌عبارتی برآورد فقط برای یک دسته از مشاهدات انجام شده، درحالی‌که بدنبال استنباط آماری برای کل جامعه هستیم. اگر هم برای کل داده‌ها استفاده کنیم، بوضوح نیازمند یک برآوردکننده غیر خطی هستیم. در واقع در این مسئله، داده‌ها دارای توزیع سانسور شده هستند، (در برخی داده‌ها همه متغیرها قابل مشاهده اند و در برخی دیگر، فقط متغیرهای توضیحی). درحالی‌که استفاده از الگوی بریده شده، بدلیل اینکه نمونه برداری فقط از یک دسته از مشاهدات صورت می‌گیرد، در برخی سطرها هر دو متغیر وابسته و مستقل قابل مشاهده اند و در مورد برخی دیگر هیچکدام. قطعاً دور ریختن این اطلاعات موجود نیز دارای کارایی نخواهد بود[۱].

فرض می‌شود که رابطه $y = X\beta + u$ مورد نظر محقق باشد و مشاهدات از فرم سانسور شده باشند. همچنین فرض میشود که : $u \sim IN(0, \sigma^2)$ ، با فرض اینکه برای

۱. بخاطر عدم کفایت

y^*های کوچکتر از ۰ این متغیر سانسور شده و عدد ۰ را اختیار کند، تابع درستنمایی این مشاهدات به صورت زیر درخواهد آمد:

$$L = \prod_{y_i>0} \frac{1}{\sigma} \varphi\left(\frac{y-X\beta}{\sigma}\right) . \prod_{y_i=0} \Phi\left(\frac{-X\beta}{\sigma}\right) \quad (۱)$$

با برازش مدل رگرسیونی بالا برای تنها مشاهدات مثبت، داریم:

$$E(u|y>0) = \sigma\varphi\left(-{}^{X\beta}/_{\sigma}\right)/(1-\varphi(-{}^{X\beta}/_{\sigma})) \quad (۲)$$

که الزامی به صفر بودن آن نیست. بنابراین علاوه بر عدم کفایت، برآوردکننده‌های حداقل مربعات اریب هم هستند. که هکمن از اریب نشان داده شده در بالا بعنوان یک متغیر توضیحی استفاده کرده و بنام نسبت معکوس میل در مدل بکار می‌برد[۱].

آمیمیا (۱۹۸۴)[۲] بدنبال برآورد تقاضای کالاهای بادوام، حالات متعددی از فرمهای مناسب متغیرهای وابسته گسسته را مطرح کرده و به عقیده بسیاری از محققین، اغلب الگوهای نوینی که در ادامه مطرح شده اند، بنوعی از طبقه‌بندی صورت گرفته در مقاله آممیا منشعب شده اند. از نظر آمیمیا(۱۹۸۴)، زمانی از توزیع بریده شده صحبت می‌کنیم که کل داده‌های متغیر وابسته و توضیحی برای مقادیر خارج از محدوده مفقود باشد، در حالی‌که در داده‌های سانسور شده، تنها داده‌های متغیر وابسته در این محدوده از بین می‌رود و برای این مشاهدات مقادیر متغیر توضیحی مانند قبل است. با این فرض که y^* بیانگر میزان مصرف کالای بادوام و حاصل از فرایند حداکثر مطلوبیت باشد، و نیز این فرض که قیمت‌ها دارای اطلاعاتی نظیر کیفیت نیز باشند، برای بررسی رفتار مصرف‌کننده می‌توان عوامل موثر بر مصرف این کالا را استخراج کرد:

$$y^* = \beta_1 + \beta_2 X + u \quad (۳)$$

---

۱. برگرفته از هکمن (۱۹۷۹) و بخش ۲-۲ از لینک زیر

http://www.wanchuanlin.org/Policy%20Evaluation%20of%20Social%20Policies/Lecture3typed.pdf

2. Amemiya(1984), Tobit models: A survey, Journal of Econometrics, 24(1-2):3-61.

بردار x شامل متغیرهای توضیحی موثر بر مصرف کالای مورد نظر و u عمدتاً بیانگر سایر متغیرهای موثر بر مطلوبیت اند که خارج از مدل نگه داشته شده اند. اما در واقعیت y^* قابل مشاهده نبوده و مصرف‌کننده y را از خود به نمایش می‌گذارد که :

$$y = y^* \quad if \quad y^* > 0$$
$$y = 0 \quad if \quad y^* \le 0 \qquad (۴)$$

با این نماد گذاری، تابع درستنمایی مشاهدات میزان مصرف به‌صورت زیر بیان می‌شود:

$$L = \prod_0 F_i(y_{oi}).\prod_1 f_i(y_{1i}) \qquad (۵)$$

که درآن، F و f به ترتیب توابع توزیع و چگالی احتمال متغیر پنهان y^* هستند. البته در این مطالعه فرض شده که مشاهدات دارای توزیع نرمال مستقلند: $u_i \sim N(0,\sigma^2)$

به این ترتیب معادله بالا به صورت زیر مشخص می‌شود :

$$L = \prod_0 \left[1 - \Phi\left(\frac{x\beta}{\sigma}\right)\right] \prod_1 \frac{1}{\sigma}\left[1 - \phi\left(\frac{y_i - x\beta}{\sigma}\right)\right] \qquad (۶)$$

$\Phi(.)$ و $\phi(.)$ به ترتیب توابع توزیع و چگالی احتمال متغیر نرمال استاندارد اند. باید توجه داشت که این تابع، معادله اصلی الگوی سانسور شده توبیت است چراکه برای مشاهدات صفر هم متغیر توضیحی قابل مشاهده است. در مقابل معادله توبیت با تابع درستنمایی احتمال بریده شده[1] عبارتست از:

$$L = \prod_1 \frac{\frac{1}{\sigma}\phi(\frac{y_i - x\beta}{\sigma})}{\Phi(x\beta)} \qquad (۷)$$

که درآن بخشی از توزیع نرمال که داده‌های درآن مثبت اند، (۱ها) فقط مد نظر بوده، و برای اینکه بعد از حذف بخشی از داده‌ها، سطح زیر چگالی ۱ باقی بماند، تقسیم بر یک منهای بخشی از داده‌ها شده که حذف شده اند؛ $\Phi(x\beta)$. به هر حال در داده‌هایی با چنین توزیعی می‌توان از برآوردهای زیر استفاده نمود:

1. Truncated tobit

۱- حداقل مربعات و عدم توجه به مشاهدات صفر، که در این صورت توبین نشان داد برآوردها ناسازگار خواهند بود.

۲- ساده‌ترین راه برآورد سازگار، استفاده از مدلهای دو حالتی درمورد متغیر وابسته است.

$$E\left[y\middle|x\right] = p\left(y_i = 1\middle|x_i\right) = F(x_i) \qquad (۸)$$

ماهیت $F(x_i)$ بسته به توزیع جملات اخلال متغیر بوده که به تناسب هر فرض مدل ویژه‌ای بوجود آمده است. الگوهای احتمال خطی، لاجیت، پروبیت همگی شاخه‌هایی از مدلهای دوحالتی متغیر وابسته به شمار می‌روند. در این مقاله قصد ارائه چنین مدلهایی نیست و بنظر می‌رسد که چنین مدلهایی در منابع اقتصاد سنجی به وفور یافت می‌شود و تنها به ارائه تابع درستنمایی اکتفا شده است(وودریج؛ ص ۵۳۴):

$$p(y_i = 1\middle|x_i) = p(y_i^* \succ 0\middle|x_i) = p(e_i \succ -x_i\beta) = F(x_i\beta)$$

$$\log L = \sum_{i=1}^{n}\left[\left(1-y_i\right)\log\left(1-F(x_i\beta)\right) + y_i\log\left(F(x_i\beta)\right)\right] \qquad (۹)$$

استفاده از تابع درستنمایی در مواقعی که صفرها را از جامعه حذف و سپس نمونه‌گیری کنیم، به مدل توبیت با تابع احتمال بریده شده مشهور شده است. از این رو در مسائل پژوهشی استفاده نامناسب از توبیت با تابع بریده شده و سانسور شده توبیت، علیرغم حداکثر کردن تابع راستنمایی، ما را به حداکثر راستنمایی نمی‌رساند و تخمینهای ناسازگار و غیر کارا بدست می‌دهد. دلیل این امر تصریح نامناسب الگوی درستنمایی است.

۳-۲- الگوهای انتخاب نمونه(هکیت)

هکمن(۱۹۷۹) برای اولین بار این الگو را مطرح کرده، به همین دلیل گاهاً از آن بعنوان الگوی هکیت هم یاد می‌شود. مناسبت چنین الگویی برای مواردی است که وجود صفر در متغیر وابسته به دلیل پدیده عدم مشاهده باشد. با این ترتیب در این دسته از داده‌ها، صفرها در برآورد الگوی اصلی y^* نقشی ندارند. تابع الگوی انتخاب نمونه به صورت زیر است(گرین؛ ۷۸۲):

$$y_i = \begin{cases} y_i^* & if\ y_i^* \succ 0 \\ Unobserved & O.W. \end{cases} \qquad (۱۰)$$

روش مقتضی برای برآورد این الگوها، به صورت زیر است:

الف- برآورد الگوی پروبیت در مرحله اول با استفاده از کل نمونه

ب- محاسبه نسبت معکوس میل از معادله توبیت بدست آمده در مرحله قبل

پ- قرار دادن نسبت معکوس میل بعنوان یک متغیر توضیحی در معادله عوامل موثر بر مصرف و برآورد آن به روش حداقل مربعات معمولی (گرین؛ ص۷۸۴)

به عقیده جونز(۲۰۰۰)، مشاهده داده صفر برای متغیر وابسته در داده‌های تقاضا در سطح اقتصاد خرد، می‌تواند ناشی از چند عامل باشد:

- عدم انتخاب کالای مورد نظر توسط فرد، راه حل گوشه‌ای(مصرف‌کننده بالقوه نبودن)
- مصرف‌کننده بالقوه است، ولی در این دوره میزان مصرف اش برابر صفر است.
- مصرف‌کننده بالقوه است، در این دوره مصرف مثبت هم داشته، اما به دلیل فاصله زیاد بین دو دفعه خرید، در این دوره هزینه خریدش صفر گزارش شده است. (به نقل از آریستی و ال پیرونی). جونز متغیر پنهان دیگری مانند d تعریف می‌کند که درصورت جواب گوشه‌ای و مصرف‌کننده بالقوه نبودن، صفر و در غیر این‌صورت ارزش یک را اختیار می‌کند. ضمناً فرض بر این است که دو دسته عوامل مجزا بر احتمال مشارکت و نیز بر میزان یا هزینه مصرف صورت گرفته روی کالا و خدمات تاثیر گذاشته و هریک از این دو متغیر از فرایند خاص خود تبعیت می‌کنند. $d_i = \mathrm{X}_1\beta_1 \quad \text{and} \quad y_i^* = \mathrm{X}_2\beta_2$

با در نظر گرفتن $y_i = d_i.y_i^*$، تنها در یک صورت مشاهده مثبت حاصل می‌شود و آنهم این است که اولاً، فرد یا خانوار مورد نظر، مصرف‌کننده بالقوه این کالای مورد مطالعه باشد. ثانیاً، در انتخاب خود، (بهینه یابی مطلوبیت)میزان مثبتی از مصرف را انتخاب کند.

جدا در نظر گرفتن دو مرحله و تخصیص مجزای دو متغیر مخفی، وجه تمایز مدلهای انتخاب نمونه(هکمن) و دو مانعی[1] است.

۳-۳- مدل دو بخشی

مدلهایی که صفرهای متغیر وابسته آنها کاملاً هوشمندانه و انتخاب شده فرض شده است.

1. Double Hurdle

فرض می‌شود که ابتدا مصرف‌کننده تصمیم مشارکت در مصرف را گرفته و سپس میزان مصرف را تعیین می‌کند. بنابراین روش مورد استفاده، به این صورت است که ابتدا مرحله اول با استفاده از تمامی متغیرهای موجود در مدل توسط برآورد گرهای لاجیت یا پروبیت برآورد شده، سپس با استفاده از حداقل مربعات، مرحله دوم فقط برای مقادیر مثبت مصرف انجام میشود. نقطه شباهت این الگو با هکیت این است که برای هر مرحله از مصرف از یک معادله جداگانه استفاده می‌کنند. اما در مدل دو مرحله ای، دیگر از نسبت معکوس میل بعنوان تصحیح‌کننده خطای تورش انتخاب نمونه استفاده نمی‌شود. یعنی بجای توزیع بریده شده در صفر، از توزیع اصلی داده‌ها استفاده میشود. در تفسیر ضرایب باید دقت نمود که ضرایب مدل هکیت برای کل نمونه تفسیر می‌شود، اما در مدلهای سانسور شده دو بخشی و مدلهای ذکر شده در ادامه این مطالعه، ضرایب مرحله اول برای کل نمونه و ضرایب معادله مصرف تنها برای نمونه سانسور شده قابل تفسیر اند.

$$p(y_i \succ 0) = F(x_{1i}\alpha)$$
$$E\left(y_i \middle| y_i, x_{i2}\right) = x_{2i}\beta \qquad (۱۱)$$

تمایز این مدل با مدلهای دیگر این است که در اینجا هیچگونه متغیر مخفی تعریف نمیشود. در واقع، این الگوی ساده مقدمه‌ای به سوی الگوهای بعدی مانند دو مانعی[1] است.

۳-۴- مدلهای هاردل:

مانند الگوی دومرحله ای، فرض این مدل بر این است که صفرها هوشمندانه انتخاب شده و مصرف‌کننده به دلخواه خود مصرفی از کالای مورد مطالعه نداشته است. مطالعات نشان می‌دهند که فرض مدل دو مرحله‌ای مبنی بر پیاپی و در امتداد هم بودن تصمیم دو مرحله (طولی بودن دو تصمیم) محدودکننده و غیر واقعی بوده و ممکن است که مصرف‌کننده دو تصمیم را به طور موازی و همزمان اتخاذ کند. مثلاً زمانی که فرد در حال انتخاب بین مصرف ساندویچ و کباب است، به میزان مصرف آنها هم فکر می‌کند. پس وابستگی جملات اخلال دو مرحله، اطلاعاتی است که نادیده گرفتن آن باعث عدم کفایت مدلهایی

1. Double Hurdle

مانند دومرحله میشود. در مدلهای دو مانعی، صفرهای مشاهده شده متغیر وابسته، به سهم خود، در توضیح معادله مصرف نقش دارند. گارسیا و لابیگا[1] نشان داده اند که در مورد یکسری کالاها (به خصوص کالاهای عادت آور مانند سیگار، مشروبات الکلی و یا حتی غذاهای خارج از خانه)، چنین وابستگی ای وجود نداشته و بعد از تصمیم به مصرف، مدت زمانی طول می‌کشد تا مصرف‌کننده تصمیم به میزان مصرف بگیرد[2](عرضی بودن دو تصمیم). در چنین مدلهایی شاید مدلهای دو مرحله به نتایجی مشابه با مدلهای دابل هاردل برسند. (آریستی و ال پیرونی؛ ۲۰۱۳) الگوی دابل هاردل به‌عنوان آخرین دستاورد اقتصاد سنجی مدلهای متغیر وابسته گسسته و اقتصاد سنجی داده‌های خرد رفتار مصرف کننده، با الهام از دلایل واقعی بروز داده‌های صفر، با ظرافت خاصی از تلفیق الگوهای توبیت، پروبیت و دو مرحله ای، یک مدل آشیانه‌ای برای این مدلها طراحی کرده است. طبق این الگو، مشاهده مصرف صفر می‌تواند معلول دلایل متفاوتی باشد: جواب گوشه‌ای و عدم مشارکت در مصرف، مصرف غیر مکرر و فاصله زیاد میان دو خرید و داده‌های مشاهده نشده و یا صفرهای اتفاقی و مواردی که مصرف‌کننده به طور بالقوه مصرف‌کننده است، اما در بهینه یابی مطلوبیت ممکن است مقدار مصرف کالا اتفاقاً صفر یا منفی شده باشد. مثلاً فردی که به غذای خارج از خانه علاقه دارد، اما به دلایل اقتصادی، در ماه معینی رستوران نمی‌رود. بنابراین برای مشاهده مصرف مثبت باید دو مانع متفاوت پشت سر گذاشته شود. اولاً مشارکت در مصرف وجود داشته باشد، ثانیاً متغیر مخفی(مصرف تعادلی حاصل از بهینه یابی مطلوبیت) عدد مثبتی را اختیار کند. با این استدلال، تابع رگرسیون، به صورت زیر قابل تعریف است:

$$\begin{aligned} &Participation\ Decision: \quad y_1^* = x_1\alpha + u_i \\ &Consumption\ Decision: \quad y_2^* = x_2\beta + v_i \end{aligned} \qquad (۱۲)$$

با توجه به گفته‌های بالا، شرط مشاهده مقدار مثبت برای متغیر وابسته عبارتست از:

1. Garcia, J and Labeaga,J, M(1996)
2. Ttime Cconsuming Process of Consumption

$$y_i = \begin{cases} y_{2i}^* & if: y_{2i}^* \succ 0 \ \ \& \ \ y_{1i}^* \succ 0 \\ 0 & Otherwise \end{cases} \qquad (۱۳)$$

همچنین وابستگی جملات اخلال ایجاب می‌کند که از توزیع مشترک دو متغیره پیروی کنند که جونز باز هم در فرمول نویسی خود فرض نرمالیتی را قائل شده است:

$$\left(\varepsilon_1, \varepsilon_2\right) \square N(0, \Omega) \quad ,$$
$$\Omega = \begin{bmatrix} 1 & \rho \\ \rho & \sigma^2 \end{bmatrix} \qquad (۱۴)$$

به هر حال، با فرض توزیع گفته شده برای جملات اخلال، تابع درستنمایی مدلهای دابل هاردل به صورت زیر به دست می‌آیند:

$$L = \prod_0 \left[1 - \Phi\left(X_1\beta_1, \frac{X_2\beta_2}{\sigma}, \rho\right)\right] * \prod_+ \left[\Phi\left(\frac{X_1\beta_1 + \frac{\rho}{\sigma}(y - X - 2\beta_2)}{\sqrt{1-\rho^2}}\right)\frac{1}{\sigma}\phi\left(\frac{y - X_2\beta_2}{\sigma}\right)\right] \qquad (۱۵)$$

بنابراین، مدل دابل هاردل الگوی آشیانه‌ای برای مدلهای قبل(بجز هکمن) بوده و درستنمایی الگوهای دیگر در بهترین حالت[1]، برابر با درستنمایی الگوی دو مانعی است.

۴- آزمون‌های تصریح و انتخاب الگو:

در انتخاب تخمین‌زن مناسب از میان برآوردکننده‌های گفته شده، معمول این است که حالتهای ممکن را دو بدو مقایسه کنیم و یک مدل را از میان آنها انتخاب کنیم. هنگامی که یکی از الگوها برای دیگری آشیانه‌ای و به لحاظ آماری تعمیم یافته باشد[2]، آزمون متداول نسبت راستنمایی بکار میرود. (مثلاً توبیت که همان دابل هاردل است که درآن پارامترهای دومرحله با هم یکسان اند.)

$$LR = -2(LL_{DH}^* - LL_T^*) \qquad (۱۶)$$

۱. یعنی با فرض صحت محدودیتهای اعمال شده در هر الگو

2. Nested Model

جدول۱- خلاصه‌ای از الگوهای متفاوت برخورد با متغیر وابسته با صفرهای فراوان

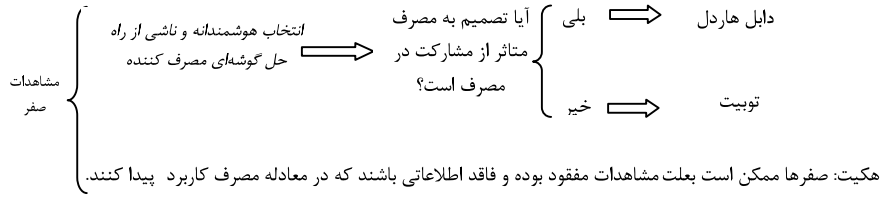
الگو	متغیر وابسته	وجه افتراق
کلاسیک	بدون محدودیت	عدم توجه به مشاهدات صفر
لاجیت و پروبیت	۰ یا یک	صرفاً احتمال مشارکت در مصرف بررسی می‌شود.
توبیت	۰ یا عدد مثبت	برآوردها ترکیبی از دو مرحله مشارکت و هزینه مصرف است و عوامل موثر بر دو تصمیم و ضرایب مربوطه با هم تلفیق می‌شوند.
دو بخشی	۰ یا عدد	دو مرحله جداگانه بررسی شده و امکان تفکیک عوامل وجود دارد. اما در تابع درستنمایی هزینه، از توزیع اولیه استفاده میشود. صفرها هوشمندانه است.
هکیت	۰ یا عدد	دو مرحله جدا بررسی میشوند. ضریب معکوس میل برای رفع اریب استفاده از توزیع اولیه در تصمیم هزینه بکار می‌رود. صفرهای هوشمندانه در مدل تعبیه نشده است.
دو مانعی	۰ یا عدد	در همه موارد مانند هکیت است. صفر هوشمندانه هم تعبیه شده است. دو مانع می‌توانند به هم وابسته هم باشند.(عرضی بودن تصمیم)

آماره آزمون کای دو بوده و درجه آزادی آن مساوی با تعداد محدودیتها است. در مسئله حاضر، تعداد محدودیتها، تعداد پارامترهایی است که با هم مساوی قرار داده ایم و همانا تعداد متغیرهای توضیحی (k-1) است. در حالی‌که در غیر اینصورت، مثلا در انتخاب بین هکمن و دابل هاردل، این آزمون مناسب نبوده و می‌بایست تعدیل شود. در این حالت، می‌بایست از یک تبدیل مشخصی از نسبت درستنمایی استفاده کرد:(وونگ؛ ۱۹۸۹)

$$\omega_n = \left(\frac{1}{n}\right)[LR_1]^2 - \left[\left(\frac{1}{n}\right)LR_1\right]^2 \qquad (۱۷)$$

که $LR_1 = (LL^*_{DH} - LL^*_T)$ اختلاف لگاریتم درستنمایی دو الگو است. وونگ نشان داد که $\sqrt{n}\frac{LR_1}{\omega_n}$ دارای توزیع نرمال استاندارد بوده و مقدار محاسباتی، باید با مقادیر

بحرانی این جدول مقایسه شود. زمانی مدل دو مانعی تصریح مقتضی است که آماره وونگ، بالاتر از مقدار بحرانی یک دامنه باشد، وگرنه دلیلی برای ترجیح دو مانعی نسبت به هکمن وجود ندارد.



۵- آزمونهای تشخیص و عیب یابی:

همان‌طور که در بخش‌های گذشته اشاره شد، تمامی برآوردکننده‌های گفته شده، با روش حداکثر درستنمایی محاسبه شده و از همین رو صحت برآورد و خواص مطلوب آنها شدیداً در گرو نرمالیتی جزء اخلال است. آزمونهای مقتضی را می‌توان از مطالعات جاراک و لی(۱۹۸۴) و پاگان و ولا(۱۹۸۹)[1] استخراج کرد. از سوی دیگر در صورت بروز ناهمسانی واریانس، می‌بایست جزء اخلال موزون را در تابع درستنمایی وارد کرد. که به پیروی از ین (۱۹۹۳)، این تعدیل در معادله ۱۵ لحاظ شده و اجازه داده شده تا واریانسها در میان مشاهدات مختلف، به صورت تابعی از متغیرهای توضیحی پیوسته مدل تغییر نمایند.

۶- متغیرها و داده‌های بکار رفته در مدل

مطالب این تحقیق، با استفاده از داده‌های آماری مقطع زمانی برگرفته از مرکز آمار ایران مشتمل بر ۲۹۳۹۰ خانوار شهری و روستایی در سال ۱۳۹۰ انجام شده است. متغیر وابسته، مصرف دخانیات و دو دسته از متغیرها، تحت عنوان اقتصادی و دموگرافیکی به عنوان عوامل توضیح دهنده مشارکت و میزان مصرف دخانیات مورد استفاده قرار گرفته اند. این متغیرها شامل درآمد خانوار، بعد خانوار، تعداد افراد باسواد موجود در هر خانوار، تعداد

1. Pagan, V and Vella, p(1989)

مردان موجود در خانواده، متوسط سن خانوار، شهری بودن یا روستایی بودن، و جنوبی بودن، شمالی بودن و یا زندگی در مرکز کشور است. طبعاً هزینه دخانیات بسیاری از خانوارها در طول دوره آمارگیری صفر گزارش شده که می‌تواند به دلایل متعددی مانند عدم اعتیاد به دخانیات، عدم خرید سیگار در دوره مورد نظر و استفاده از خریدهای گذشته، راه حل گوشه‌ای و ... باشد. جدول ۲ آماره‌های توصیفی را نشان می‌دهد:

جدول۲- شاخصهای آماری متغیرهای بکار رفته در الگو

نام متغیر	تعریف	کل نمونه انتخابی		نمونه سانسور شده	
		میانگین	انحراف معیار	میانگین	انحراف معیار
Tobacco	هزینه دخانیات	۴۸۲۲۱	۱۵۱۳۲۷	۱۹۱۲۸۴	۲۵۱۹۴۰
Lnincm	لگاریتم درآمد خانوار	۱۷/۹۳	۱۹/۳۳	۱۷/۸۷	۱۹/۰۱
Size	بعد خانوار	۴/۲۹	۱/۶۷	۴/۴۸	۱۹/۰۱
Age	متوسط سن خانوار	۳۰/۰۳	۱۲/۱۲	۳۰/۰۵	۱۱/۱۱
Educatednum	تعداد افراد تحصیلکرده در خانوار	۳/۱۵	۱/۶۱	۳/۲۷	۱/۶۲
Diplom	تعداد افراد دیپلم موجود در خانوار	۰/۱۱	۰/۳۶	۰/۰۷	۰/۲۹
Ba	تعداد افرادی که دارای تحصیلات لیسانس یا بالاتر اند.	۰/۶۹	۰/۰۷	۰/۵۸	۰/۰۵
South	متغیر مجازی: برای خانوارهای ساکن جنوب ۱ است.	۰/۱۸	۰/۳۹	۰/۱۷	۰/۳۸
North	متغیر مجازی: برای خانوارهای ساکن شمال ۱ است	۰/۵۳	۰/۵	۰/۵۷	۰/۴۹
west	متغیر مجازی: برای خانوارهای ساکن غرب ۱ است	۰/۲۲	۰/۴۱	۰/۱۴	۰/۳۵
East	متغیر مجازی: برای خانوارهای ساکن شرق کشور ۱ است	۰/۴۱	۰/۴۹	۰/۵	۰/۴۹
Urban	متغیر مجازی: برای خانوارهای شهری ۱ است	۰/۵۲	۰/۴۹	۰/۵۵	۰/۴۹

بر طبق اطلاعات خام هزینه درآمد خانوارهای شهری و روستایی، حجم کل نمونه مشتمل بر ۳۱۱۲۰ خانوار بوده و برخی از آنها آمار غیر قابل قبول ارائه کرده اند. از این میان، ۲۹۳۹۰ خانوار شهری و روستایی سطر مربوط به دخانیات را تکمیل نموده که در مجموع، برای ۷۴۱۰ خانوار شهری و روستایی مصرف مثبت حداقل یکی از دخانیات گزارش شده است. طبق جدول بالا می‌توان متوجه شد که ۵۵ درصد این خانوارها یعنی کمتر از ۴۱۰۰ خانوار، شهری و مابقی روستایی بوده اند. ضمناً اطلاعات خام بیان میکند که متوسط درآمد ماهانه خانوارهای نمونه کل ۶۱۱۶۳۹۰۰ ریال برای هر ماه و ۵۷۸۶۶۹۵۹ ریال در نمونه‌ای بود که مصرف‌کننده دخانیات بوده اند. هرچند که در این مشاهدات درآمد خانوارهای مصرف‌کننده دخانیات اندکی کمتر از متوسط خانوارهای ایرانی است، اما با توجه به انحراف معیار بالای دو نمونه، آزمون t مقایسه میانگین معادل ۰/۰۵۵ بدست آمده و با استناد به آن نمی‌توان گفت تفاوت معنی داری میان درآمد خانوارهای مصرف‌کننده دخانیات با سایر خانوارها وجود دارد. در تمام برآورد کننده‌ها، متغیر وابسته، در مرحله اول، مشارکت یا عدم مشارکت در مصرف دخانیات بوده و در مرحله دوم، میزان مصرف این گروه مصرفی.

همانطور که در بالا هم بدان اشاره شد، متغیرهای توضیحی در سه دسته اقتصادی (لگاریتم درآمد خانوار)، جمعیت شناختی (مشخصات سن و تحصیلات خانوار و بویژه سرپرست آن) و ویژگی‌های جغرافیایی محل سکونت مانند جنوبی یا شمالی بودن، شرقی یا غربی بودن و نیز شهری و یا روستایی بودن است. عمده متغیرهای دموگرافیکی و جغرافیایی به صورت مجازی بوده و در بسیاری از مطالعات، تاثیر آنها فقط بر احتمال مشارکت در مصرف سنجیده شده است؛ که البتّه در این مطالعه در هر دومرحله مشارکت و میزان مصرف به عنوان متغیر توضیحی مورد استفاده قرار گرفته اند. آمارهای جدول بالا نشان می‌دهند که ۱۴ درصد خانوارهای مصرف‌کننده دخانیات ساکن غرب بوده، درحالی‌که ۵۰ درصد آنها در شرق کشور زندگی می‌کنند. البته ضرایب تاثیر آنها در احتمال مشارکت در مصرف دخانیات و میزان مصرف دخانیات نشان خواهد داد که این متغیر مجازی چه تاثیری بر مصرف دخانیات خواهد داشت. به همین شکل، ۱۷ درصد مصرف

کنندگان دخانیات ساکن جنوب کشور و ۵۷ درصد آنها در شمال کشور زندگی می‌کرده اند. متوسط سن اعضاء خانوارهای کل کشور ۳۰/۰۵ سال بوده و در مورد خانوارهای مصرف‌کننده دخانیات ۳۰/۰۳ سال می‌باشد. همچنین، با توجه به اینکه متوسط بعد خانوارها در نمونه کل و نیز نمونه خانوارهای مصرف‌کننده دخانیات به ترتیب ۴/۲۹ و ۴/۴۸ نفر است، ظاهراً مصرف سیگار و سایر مواد دخانی بیشتر در خانوارهای پرجمعیت تر صورت می‌گیرد. نهایتاً آمارها نشان می‌دهد که در مورد هر خانوار، به طور متوسط، خانوارهای مصرف‌کننده دخانیات ماهانه بیش از ۱۹ هزار تومان صرف خرید دخانیات کرده اند.

۷- برآورد پارامترهای رفتار مصرف دخانیات

همانگونه که در مبحث تشخیص و عیب‌یابی بحث شد، مرحله قبل از پذیرش برآوردهای سه الگوی بالا، بررسی نرمال بودن و همسانی واریانسها است. همانگونه که مادالا و نلسون(۱۹۷۵) و عرب مازار و اسمیت(۱۹۸۲) نیز اشاره کردند، حتی خواص نمونه‌های بزرگ برآوردکننده‌ها مانند سازگاری، به شدت تحت تاثیر تامین نرمالیتی و واریانس همسانی اجزای اخلال است. به همین دلیل با استفاده از آزمون پاگان و ولا[1]، این آزمونها انجام شده و نتایج حاصله در جدول ۳ ارائه شده است.

جدول۳- آزمونهای نرمالیتی و همسانی در الگوهای مختلف

الگو	نرمال بودن	Prob	واریانس همسانی	Prob
توبیت	(۲)۲۳۵/۴۲	0.00	(۳)۱۲۰/۲۹	0.00
هکیت (مرحله اول)	(۲)۱۰۱/۱۱	0.00	(۱)۱۴۳/۰۱	0.00
هکیت (مرحله دوم)	(۲)۹۱/۷۸	0.00	(۴)۱۰۳/۹۱	0.00
دابل هاردل(مرحله اول)	(۲)۶۸/۰۷	0.00	(۱)۹۰/۷۱	0.00
دابل هاردل (مرحله دوم)	(۲)۲۱۵/۵۰	0.00	(۴)۷۴/۳۴	0.00

نتایج بدست آمده نشان می‌دهد که هر سه الگو، دارای مشکل ناهمسانی و عدم نرمال

1. Pagan and Vella's LM

بودن جملات اخلال هستند. به همین دلیل به پیروی از ین (۱۹۹۳ و ۲۰۰۵)، آریستی و ال پیرونی (۲۰۱۳) و نیز ین و جونز (۲۰۰۰) در همه الگوها از تبدیل باکس کاکس استفاده شده و جملات اخلال را با وزن ناهمسانی تصحیح کرده ایم. با این تعدیلات، برآوردکنندهای توبیت، دابل هاردل و هکیت به همراه لگاریتم درستنمایی با استفاده از نسخه ۱۲ نرم افزار استاتا محاسبه و در جدول ۴ ارائه شده است. دستور برآوردکنندههای توبیت و هکمن در بسته نرم افزاری استاتا از قبل فراهم شده و محاسبه ضرایب دابل هاردل کامل با فروض ناهمسانی و عدم توزیع نرمال هم با استفاده از برنامه نویسی در این محیط میسر شد.[۱]

جدول۴- نتایج حاصل از برآورد تقاضای دخانیات با استفاده از الگوهای رقیب

نام متغیر	برآوردکننده ها									
	توبیت		هکیت				برآوردهای دو مانعی			
	ضریب	P	هزینه	P	مشارکت	P	هزینه	P	مشارکت	P
بعد خانوار	-۲/۲۹	۰	-/۱۲	/۰۰۱	-۰/۱۷	۰	-/۱۲	/۰۰۱	-/۱۷	۰
زیر دیپلم	۲/۸۹	۰	-/۶۰	/۰۹	۰/۲۳	۰	-/۰۶	۰/۸۹	۰/۲۳	۰
کارشناسی	-/۲۶	۰/۵	-/۰۹	/۰۲۸	-/۰۲	/۵۹	-/۰۹	۰/۰۳	-/۰۱۷	۰/۵۹
تعداد مردان	۰/۵۵	۰	/۰۵	/۰۰۱	۰/۰۴	۰	/۰۴۷	/۰۰۱	۰/۰۴	۰
شهری بودن	-/۱۴	۰/۵۱	-/۰۰۳	/۸۷۵	-/۰۱	/۵۱	-/۰۰۵	۰/۸۸	-/۰۱	۰/۵۱
متوسط سن	۰/۰۴	۰	/۰۰۶	۰	/۰۰۲	۰	۰/۰۰۶	۰	/۰۰۲	۰
تحصیلات	۰.۰	۰/۶۹	-/۰۰	/۱۹	-/۰	/۶۷	۰	۰/۱۹	۰	۰/۶۷۱
شمالی بودن	۱/۸۸	۰	/۰۹۸	۰	۰/۱۴	۰	۰/۰۹۸	۰	۰/۱۵	۰
جنوبی بودن	۰/۰۱	۰/۹۷	-/۱۷	۰	/۰۰۷	/۷۹۸	-/۱۷	۰	/۰۰۷	/۷۹
شرقی بودن	-۴/۳۱	۰	-/۱۹	۰	-/۳۳	۰	-/۱۸	۰	-/۳۳	۰
غربی بودن	۲/۱۷	۰	-/۱۴	۰	/۱۷۱	۰	/۱۳۸	۰	/۱۷۱	۰
Lincm	-۰/۰۹	۰/۱۵	/۰۳	۰	-/۰۱	/۰۶	۰/۰۳	۰	-/۰۰۹	۰/۰۶
_cons	-۳۹۴۳۱/۷۸		-۲۶۲۵۵				-۲۶۲۵۳/۰۳			
Log pseudo likelihood										

۱. حجم بالای نمونه باعث شده تا توزیع نمونهای پارامترهای مدل به توزیع نرمال استاندارد گرایش پیدا نماید.

پس از برآورد ضرایب، آماره والد توسط نرم افزار برابر با ۴/۷۲ محاسبه شده که در مقایسه با مقدار جدول کای دو با درجه آزادی (۱)، نشان دهنده رد شدن استقلال دو مرحله تصمیم از یکدیگر می‌باشد. سطح احتمال بدست آمده برابر با ۰/۰۲۹۸ به این معنا است که در ناحیه بحرانی قرار داشته و دو تصمیم از یکدیگر مستقل نیستند. این مسئله در بسیاری از مطالعات به ماهیت زمانبر[1] بودن تصمیم به مصرف معروف بوده و به این حقیقت اشاره دارد که خانوار در همان لحظه‌ای که تصمیم به مصرف سیگار (یا انتخاب میان خرید سیگار و آبمیوه) می‌گیرد، به میزان مصرف سیگار هم فکر می‌کند.

Wald test of indep. eqns.	(rho=0):	chi2(1)	4.72	Prob>chi2	0.0298

پس از تخمین ضرایب با هر سه الگوی فوق، نوبت به انتخاب بهترین مدل می‌رسد. به عبارتی بهترین تصریح می‌بایست انتخاب و بر آن پایه، کششها محاسبه و تفسیر شوند. مطابق مطالب آمده در بخش مربوط به آزمونهای تصریح، برای این منظور باید رویه‌ای را دنبال نماییم که بهترین مدل با بهترین برازش و توضیح دهندگی رفتار مصرف‌کننده انتخاب شود. بنابراین برای انتخاب باید میان توبیت و دابل هاردل یکی را انتخاب نمود. با توجه به اینکه الگوی دابل هاردل، شکل آشیانه‌ای و عمومی مدل توبیت است، آزمون LR معمول قابل استفاده است. نتایج جدول۴ نشان می‌دهد که مقدار محاسباتی نسبت راستنمایی 12986.75 بوده و برآوردهای دابل هاردل، قدرت توضیح دهندگی بسیار بیشتری نسبت به برآوردکننده‌های توبیت دارند. البته این مسئله با یک نگاه کلی به لگاریتم راستنمایی دو برآوردکننده در جدول هم قابل انتظار است. همچنین آماره وونگ براساس معادله ۱۷ محاسبه شد که در مقایسه با مقدار بحرانی توزیع استاندارد نرمال، در ناحیه بحرانی قرار گرفته و نشان می‌دهد که فرض صفر، یعنی برتری الگوی هکیت در مقابل دابل هاردل قابل پذیرش نمی‌باشد.

1. Time Consuming

جدول۴- آزمون تصریح و انتخاب مدل

الگو	نوع آزمون	Statistics	Prob
الگوی دابل هاردل در مقابل توبیت	LR	۱۲۹۸۶/۷۵	۰/۰۰۰
الگوی دابل هاردل در مقابل هکیت	Voung	۵۸۱۵/۵۱	۰/۰۰۰

بنابراین علیرغم تشابه بالای نتایج حاصل از هر سه برآوردکننده، در کل می‌توان نتیجه گرفت که تصریح دابل هاردل بر سایر الگوها مرجح بوده و قدرت توضیح دهندگی بیشتر در رفتار مصرفی مصرف کنندگان دخانیات خواهد داشت؛ هرچند که تئوریهای اقتصاد خرد و اقتصاد سنجی هم برتری این الگو را قبلاً تایید کرده بودند. به همین دلیل، در ادامه، نتایج حاصل از الگوی دابل هاردل تفسیر شده و کشش‌های مختلف، از روی همین الگو استخراج شده اند.

با استفاده از اثرات حاشیه‌ای و ضریب حاشیه‌ای هر یک از متغیرها، کششهای مصرف دخانیات نسبت به تمام عوامل تعیین‌کننده محاسبه شده است. ابتدا کششهای مشارکت در مصرف یا به نوعی کششهای اعتیاد به مصرف سیگار، محاسبه شده است. سپس، برای مصرف کنندگان دخانیات و سیگار، کشش میزان هزینه این کالای مصرفی محاسبه و تحت عنوان کشش شرطی هزینه دخانیات ارائه شده است. در نهایت با استفاده از قاعده بیز، کششهای غیرشرطی هزینه دخانیات با استفاده از حاصلضرب دو کشش بالا محاسبه شده و در جدول شماره ۵ قابل مشاهده‌اند.

در الگوی دابل هاردل و هکمن، ممکن است برخی از عوامل، تاثیر مثبت بر احتمال مشارکت بر مصرف داشته، در حالیکه بر میزان هزینه مصرف کنندگان بالقوه، تاثیر منفی داشته باشد. در الگوی توبیت، چنین تجزیه‌ای برای دو اثر ممکن نبوده و ترکیب دو اثر را می‌توان مشاهده نمود. به‌عبارتی انتظار می‌رود که نتایج توبیت تا حدود بسیاری مشابه کششهای غیر شرطی الگوی دابل هاردل باشند.

جدول۵- کششهای بدست آمده بر اساس برآوردهای دابل هاردل

نام متغیر	کشش مشارکت	کشش شرطی مصرف	کشش غیر شرطی مصرف
بعد خانوار	-۰/۷۴۰۸	-۰/۵۲۲۲	-۱/۲۶۳۱
تعداد افراد زیر دیپلم	۰/۸۸۳۶	-۰/۲۳۴۲	۰/۶۴۹۵
تعداد افراد دیپلم خانوار	-۰/۰۰۴۸	-/۰۲۷۲	-۰/۰۳۲۰
تعداد افراد مذکر خانوار	۰/۰۹۲۴	۰/۱۰۳۷	۰/۱۹۶۱
سکونت در شهر	-/۰۰۵۲	-۰/۰۰۲۲	-۰/۰۰۷۴
متوسط سن اعضای خانوار	۰/۰۸۷۹	۰/۱۶۷۵	۰/۲۵۵۵
تعداد افراد با سواد خانوار	۰/۰۰۰۱	۰/۰۰۰۴	۰/۰۰۰۵
سکونت در شمال کشور	۰/۰۷۸۰	۰/۰۵۲۲	۰/۱۳۰۲
سکونت در جنوب کشور	۰/۰۰۰۱۲	-/۰۳۰۷	-۰/۰۲۹۴
سکونت در شرق کشور	۰/۰۷۲۶	-۰/۰۴۰۹	-۰/۱۱۳۵
سکونت در غرب کشور	۰/۰۷۰۴	۰/۰۵۶۸	۰/۱۲۷۱
درآمد خانوار	-۰/۱۴۳۳	۰/۵۳۲۳	۰/۳۸۸۹۰

در مورد دخانیات، یک درصد افزایش در درآمد خانوار، بیش از ۰/۱۴ درصد احتمال اعتیاد به دخانیات را کاهش داده، اما برای مصرف کنندگان این مواد، بیش از ۰/۵۳ درصد هزینه مصرف دخانیات را افزایش می‌دهد. بر اساس کشش غیر شرطی، ترکیب این دو اثر، موجب می‌شود که در کل نمونه، اعم از مصرف کنندگان بالقوه دخانیات و خانوارهایی که هرگز مصرف‌کننده این مواد نیستند، افزایش یک درصد درآمد خانوار به حدود ۰/۳۸۸ درصد افزایش هزینه دخانیات ختم شود. اثرگذاری دوگانه‌ای را می‌توان در تعداد افراد دیپلمه موجود در هر خانوار دید. افزایش تعداد افراد دیپلمه موجب افزایش احتمال مشارکت در مصرف دخانیات شده، درحالی‌که هزینه مصرف دخانیات را برای هر خانوار کاهش می‌دهد. به همین شکل، افراد جنوبی با احتمال ۰/۰۰۱ درصد بیشتر از افراد ساکن در مرکز کشور مصرف‌کننده دخانیات بوده، حال آنکه مصرف کنندگان دخانیات در این منطقه از کشور، ۰/۰۳ درصد هزینه کمتری صرف خرید این مواد می‌کنند.

سایر متغیرها، اثر همسویی بر مشارکت در مصرف و هزینه مصرف بر جا گذاشته اند. یکی از متغیرهای مهمی که طبق نتایج الگوی دابل هاردل، تاثیر شدیدی بر احتمال مصرف سیگار در خانوار داشته و هزینه آن را هم افزایش می‌دهد، تعداد مردان موجود در هر خانواده است. طبعاً با توجه به اینکه مصرف دخانیات و بویژه سیگار در مردان رایج تر از زنان می‌باشد، چنین پدیده‌ای طبیعی و قابل انتظار می‌باشد. تعداد افراد باسواد هرچند به لحاظ مطلق اثر کوچکی دارد، اما به لحاظ علامت تاثیر مشابه با تعداد مردان داشته و هم احتمال مصرف دخانیات و هم هزینه آنرا افزایش می‌دهد.

در مقایسه با مرکز کشور، خانوارهای ساکن مناطق شمالی، ۰/۰۷ درصد احتمال مشارکتشان در مصرف سیگار و دخانیات بالاتر بوده و بیش از ۰/۰۵ درصد نیز هزینه بیشتری صرف خرید این مواد می‌کنند. بنابراین در مجموع حدود ۱۳ درصد هزینه بیشتری نسبت به خانوارهای مرکزی کشور بابت خرید دخانیات پول پرداخت می‌کنند. هر درصد افزایش در متوسط سنی خانوار، ۰/۸۷ درصد بر احتمال مشارکت در مصرف اثرگذار بوده و حدوداً ۰/۱۷ درصد هم هزینه افراد سیگاری را افزایش می‌دهد. بنابراین در مجموع، بیش از ۰/۲۵ درصد هزینه کل دخانیات را افزایش می‌دهد.

۸- نتیجه‌گیری و پیشنهادات:

رفتار مصرفی دخانیات توسط مدلهای مختلف متغیر وابسته گسسته بررسی شد. الگوی توبیت، توانایی تفکیک احتمال مشارکت در مصرف و میزان مصرف را نداشته، ضمن آنکه به نحو محسوسی دارای لگاریتم درستنمایی پایینتری نسبت به سایر الگوها بود. این واقعیت توسط آزمون LR هم تایید شد. الگوی هکیت نیز علیرغم اینکه پارامترهای دو مرحله را از هم جدا می‌کند، در مورد دلیل وقوع صفرها فرض درستی نداشته و راه حل گوشه‌ای را در مدلسازی تعبیه نکرده است. (لوکا و پیرونی ۲۰۱۲)به همین دلیل دارای لگاریتم درستنمایی پایینتری بود. لذا براساس آماره وونگ، مدل دابل هاردل بر هکمن ترجیح داده شده و به عنوان الگوی مقتضی و مناسب در رفتار مصرفی دخانیات بکار گرفته شد. این در حالیست که نه تنها در تمامی مطالعات صورت گرفته در کشور ایران، بلکه در

اغلب تحقیقات انجام شده در دنیا، به دلایلی مانند سهولت و محدودیتهای نرم افزاری یا جدید و ناشناخته بودن، از مدلهای توبیت و یا نهایتا هکمن استفاده شده و ضرایب تورش داری براورد می‌شود. چنین خطاهایی حتی در سایر تحقیقات اقتصادی و حتی علوم مشابه تکرار شده است.

منابع:

۱. هژبر کیانی، کامبیز و کوه بر، محمد امین؛(۱۳۸۷) « برآورد کششهای مواد مغذی: کاربردی از الگوی توبیت»، فصلنامه بررسیهای اقتصادی، دوره ۵، شماره ۱، صص ۸۹-۱۱۰.

2. Amemiya, T. (1984) Tobit models: asurvey, Journal of Econometrics, 84,3-61.
3. Arabmazar, A. and Schmidt, P.(1982) an investigation of the robustness of the tobit estimator to non-normality, Econometrica,50 , 1055-1063.
4. Aristei, D and Pieroni, L.(2008) A double hurdle approach to modelling tobacco consumption in Italy, Applied economics, 40:19, 2463:2476.
5. Bascketael, N; Strand,J; McConnell, K, E and Arsanjani, F, (1990) Sample Selection bias in the estimation of recreation demand function: An application of Sport fishing, Land Economics, 66(1), 40-49.
6. Bleylock, J and Blizard, W, N(1992)Cigarette consumption: the case of low income women, American Journal of Agricultiral Economics, 74, 698-705.
7. Blundell, R. and Meghir, C. (1987) Bivariate Alternatives to the UnivariateTobit Model,Journal of Econometrics, 34, 179-200.
8. Cragg, J.(1971), Some Statistical models for limited dependent variables with applivation to the demand for the duarable goods. , Econometrica, 23,59-80.
9. Dettman, R, L and Dimitri, C(2007), Organic consumers: A demographic portrayal of organic vegetables consumption within the United States, , the 105th EAAE Seminar of international marketing and international trade of quality food products, Bologna, Italy.
10. Garcia, J and Labeaga,J, M(1996) Alternative approaches to modelling zero expenditures: an application to Spanish demand for tobacco, Oxford bulletin of econometrics and statistics, 58, 589-506.

11. Greene, W, H, (2000), Econometrics analysis(Forth Ed.). N, J, Prentice Hall, Englewood cliffs.
12. Heckman, J. (1979) Sample selection bias as a specification error, Econometrica, 47, 153-61.
13. Hovhannisyan, V and Gold, W (2011), Quantifying the structure of food demand in china: An econometric approach, Agricultural economics, 42, 1-17.
14. Huang,K,S.(1996).Nutrient Elasticities in A Complete Food Demand System. American Jouurnal of Agricultural Economics. 72:145-168
15. Jabarin, A, S and Al-Karablieh,E, K, (2011), Estimating the fresh vegetables demand system in Jordan: A linear approximate Almost Ideal Demand System, Journal of Agricultural science and technology,5(3), 322-331.
16. Jones, A. M. (1989a), A Double-Hurdle Model of Cigarette Consumption, Journal of Applied Econometrics, 4, 23-39.
17. Jones, A. M. (1989b), The UK Demand for Cigarettes 1954-86, A Double-Hurdle Approach, Journal of Health Economics, 8, 133-141.
18. Jones, A, M. (1992), A note on computation of the double hurdle model with dependence with an application of Tobacco expenditures, bulletin of economic research, 44, 67-74.
19. Jones, A, M and Yen, S, T.(1996), Individual cigarette consumption and addiction: a flexible limited dependent variable approach, Health economic,5, 105-117.
20. Jones, A, M and Yen, S, T.(2000), A box-cox double hurdle model, Manchester school, 68, 203-221.
21. Labeaga, J, M. (1999) A double hurdle rational addiction model with heterogeneity: estimating the demand for tobacco, Journal of Econometrics, 93, 49-72.
22. Kiani, K,H and Kouhbor,M,A(2008), Estimating nutrients elasticities: An application of the Tobit model, *Journal of quantitative Economics*, 2, 42-60.
23. Maddala, G, S (1983) Limited Dependent and qualitative variables in econometrics, Cambridge university press, Cambridge.
24. Newman, C., Henchion, M and Mathews, A,(2003), A double hurdle model of Irish household expenditure on prepared meals, Applied Economics, Taylor and Francis Journals, 35(9),1053-1061.
25. Pagan, V and Vella, p(1989)Diagnostic tests for models based on

individual data: a survey, Journal of Applied Econometrics.4, 529-559.
26. Pudney, S(1989), Modelling individual choice: the econometrics of corner Kinks and Holes. Basil Blackwell, New York.
27. Reynolds, A. and Shonkwiler, J. (1991) Testing and Correcting for DistributionalMisspecifications in the Tobit Model: An Application of the Information Matrix Test,Empirical Economics, 16, 313-323
28. Smith, M, D (2003), On dependency in double hurdle models, statistical papers, 44, 581-595.
29. Sue, S, J and Yen, S, T(1996), Microeconometric models of infrequently purchased goods: an application to household pork consumption, Empirical economics, 21, 513-533.
30. Sue, S, J and Yen, S, T (2000), A censored system of cigarette and alcohol consumption, Applied Economics, 32, 729-737.
31. Tiffin, R and Arnoult, M,(2008), Baysian Estimation of the infrequency of purchase model with an application to food demand in the UK.
32. Tobin, J, (1958), Estimation of relationship for limited dependent variables, Econometrica, 26, 24-36.
33. Voung, Q, H, (1989), Likelihood ratio test for model selection and non-nested hypothesis, Econometrica, 57,303-333.
34. Yen, S, T, (1995), Alternative transformation in a class of limited dependent variable models: alcohol consumption by US women. Applied Economics Letters, 2, 258-262.
35. Yen, S, T and Jensen, H, H,(1996) Determinants of household expenditures on alcohol, The Journal of Consumers affairs, 30, 48,67.
36. Yen, S, T and Huang, C, L, (1996), Household demand for finfish: A generalized double hurdle model, Journal of Agricultural and Resource Economics, 21(2), 220-234.