راهنمای اقتصاددانان تجربه‌گرا،

اقتصادسنجیِ تقریباً بی‌ضرر

این متن، ترجمه کتاب: «اقتصادسنجی تقریباً بی‌ضرر؛ کتابچه راهنمای اقتصاددانان تجربه‌گرا» (Mostly Harmless Econometrics: An Empiricist's Companion) نوشته جوشوآ. دی. آنگریست (Joshua D. Angrist) و جورن استفن پیشکه (Jörn-Steffen Pischke)، منتشر شده توسط Princeton University Press است.

کد موضوعی: ۳۴۰

شماره مسلسل: ۱۶۴۵۷

خردادماه ۱۳۹۸

معاونت پژوهش‌های اقتصادی

به نام خدا

فهرست مطالب

راهنمای اقتصاددانان تجربه‌گرا، اقتصادسنجیِ تقریباً بی‌ضرر

مقدمه مرکز پژوهش‌های مجلس شورای اسلامی

تقریباً مهم‌ترین ابزار سنجش و برآورد در اختیار محققان اقتصادی، مجموعه‌ای از مفاهیم و روش‌های آماری است که از آن به اقتصادسنجی یاد می‌کنیم. استفاده از روش‌های اقتصادسنجی برای سنجش و پیش‌بینی متغیرها تا آنجا در حوزه‌های مختلف علوم اجتماعی، و نه تنها اقتصاد، عمومیت پیدا کرده است که به قولی، تقریباً هر محقق کاربردی علوم اجتماعی در عمل یک اقتصادسنجی‌دان به شمار می‌رود.

از سوی دیگر، با توسعه اقتصادسنجی در شاخه‌های مختلف آن اعم از اقتصادسنجی خُرد، کلان‌سنجی و... تسلط بر تمام حوزه‌های سنجش متغیرهای اقتصادی برای تمام محققان اقتصادی عملاً ناممکن و یا حتی غیرضروری شده است و محققان حوزه‌های مختلف، بسته به نیاز و حوزه کاری خود، جعبه ابزار تحلیلی سفارشی شده خود را تهیه و استفاده می‌کنند. با این حال، برخی روش‌ها و مفاهیم، به خصوص در حوزه سنجش علیت و میزان اثر، همانند آزمایش‌های تصادفیده،[1] الگوهای تفاضلِ تفاضل[2] و الگوهای دودویی[3] و چندکی،[4] تقریباً مورد نیاز قریب به اتفاق محققان کاربردی و تجربی حوزه‌های مختلف اقتصادی، اعم از اقتصاد کار، اقتصاد فقر، اقتصاد درمان، خط‌مشی‌گذاری عمومی و... است.

اثر حاضر با عنوان «اقتصادسنجیِ تقریباً بی‌ضرر»[5] تلاش می‌کند، از بین ابزارهای متعدد توسعه داده شده در اقتصادسنجی، پُرکاربردترین و چندکاره‌ترین ابزارهای ضروری مشترک را انتخاب و آنها را توضیح دهد. به دلیل وسواس مؤلفان در انتخاب ابزارهای کاربردی اقتصادسنجی در تدوین اثر، از زمان انتشار، این کتاب با استقبال قابل توجهی به ویژه در سطح اندیشکده‌های کاربردی سیاستگذاری عمومی مواجه شده و به راهنمای مروری این روش‌ها تبدیل شده است.

نکته قابل توجه در مورد آثار ترجمه‌ای، به طور عام، این مسئله است که اصولاً ترجمه، ولو با بهترین کیفیت، متن را از روح اصلی آن جدا می‌کند و در واقع گونه‌ای از تألیف است و در وضعیتی که امکان مطالعه اصل اثر به راحتی وجود دارد، مطالعه ترجمه توصیه نمی‌شود. این مسئله در مورد متون فنی و زیاده از حد فنی، حساسیت بیشتری پیدا می‌کند، به گونه‌ای که به دلیل انبوه ترکیب‌های جدید ترجمه شده که برای خواننده متن، ناآشنا جلوه می‌کند، متن ترجمه شده، ولو در قالب کلمات ظاهراً آشنا، غریب

1. Randomized Trials
2. Difference in Difference Models
3. Binary Regression
4. Quantile Regression
5. Mostly Harmless Econometrics: An Empiricist's Companion

می‌نماید و این چنین است که معمولاً، ترجمه متون فنی، از بابت اینکه خروجی کار تقریباً به اندازه متن اصلی، ثقیل است، کاری **عبث** یا صرفاً کمکی است. به این ترتیب، با توصیفات ارائه شده، به نظر می‌رسد، شاید اگر در استفاده از این کتاب نیز، متن اصلی مورد استفاده قرار گیرد و از این متن صرفاً در صورت نیاز به عنوان **تسهیل‌گر** استفاده شود، عایدی بیشتری برای مخاطب خواهد داشت.

در ترجمه این راهنما، از واژه‌نامه و اصطلاحات آماری پژوهشکده آمار برای ارائه استانداردی از ترکیب‌های ترجمه‌ای آماری استفاده شده است. با وجود این، در برخی موارد که ترکیب‌های جدیدی استفاده شده است، ترکیب اصلی در پاورقی هر صفحه ذکر شده است. این اثر نیز مانند هر اثر دیگری بی‌نقص نیست و اتفاقاً از بابت ترکیب‌های متعدد استفاده شده در آن و نیز انقطاع مکرر پیش آمده در فرایند ترجمه در طول زمان و مشورت و کمک افراد مختلف، ویرایش اول به شمار می‌رود و نیازمند ویرایش‌های آتی خواهد بود.[1]

پیش‌درآمد

دنیای اقتصادسنجی دائماً در حال گسترش است. در نتیجه، روش‌های اقتصادسنجی و کاربرد آن پیشرفت زیادی کرده است. منوی مدرن روش‌های اقتصادسنجی ممکن است، حتی برای یک کارشناس مجرب کار با داده‌ها، گیج‌کننده به نظر برسد. خوشبختانه، تمام اقلام موجود در این منو، به یک اندازه ارزشمند یا مهم نیستند. برخی از موارد عجیب موجود در این منو، به طور غیرضروری پیچیده شده و حتی ممکن است مضر باشند. از سوی دیگر، در حالی که تفسیر ابزارهای اساسی اقتصادسنجی نزدیک به هم و پیچیده‌تر شده‌اند، روش‌های اصلی اقتصادسنجی کاربردی، تا حدود زیادی بدون تغییر مانده‌اند. کتاب ما، راهنمای روش‌های اساسی اقتصادسنجی برای تجربه‌گراهاست... اقتصادسنجی تقریباً بی‌ضرر.

مهم‌ترین اقلام موجود در جعبه ابزار یک اقتصادسنجی‌دان عبارتند از:

۱. الگوهای رگرسیونی برای کنترل متغیرهایی که ممکن است اثر علّی مورد نظر را پنهان کنند،

۲. روش‌های متغیرهای ابزاری برای تحلیل آزمایش‌های واقعی و طبیعی،

۳. راهبردهایی از نوع تفاضلِ تفاضل[2] که از مشاهده‌های تکراری برای کنترل عوامل حذف شده مشاهده نشده استفاده می‌کنند.

کاربرد مفید این فنون نیازمند مبنای مفهومی محکم و درک خوب ساختارهای استنباط آماری است. در اینجا، هر دو جنبه اقتصادسنجی کاربردی پوشش داده می‌شوند.

---

۱. مطالعه‌کنندگان کتاب یا علاقمندان یا گروه‌های دانشجویی می‌توانند پیشنهادهای خود برای رفع این نقایص یا ارائه ویرایش دوم از این ترجمه را به این ایمیل ارسال فرمایند.

2. Differences-in-difference

دیدگاه ما در خصوص آنچه مهم است، توسط تجربه ما به عنوان محققان تجربی و به ویژه توسط کار ما در زمینه آموزش و ارائه مشاوره به دانشجویان دکتری اقتصاد شکل گرفته است. این کتاب با در نظر داشتن این دانشجویان نوشته شد. در عین حال، امیدواریم که این کتاب در بین گروه‌های دیگری از محققان که نیاز مبرمی به پاسخ‌های کاربردی راجع به انتخاب فن و تفسیر یافته‌های تحقیقات دارند مخاطب پیدا کند. دغدغه‌های اقتصادسنجی کاربردی اساساً با دغدغه‌های علوم اجتماعی دیگر یا همه‌گیرشناسی (اپیدمیولوژی) متفاوت نیستند. هر شخصی که علاقه‌مند به استفاده از داده‌ها برای شکل دادن به سیاستگذاری عمومی یا ارتقای سلامت عمومی باشد، باید نتایج آماری را خلاصه و از آنها استفاده کند. می‌توان گفت که هر شخصی که علاقه‌مند به استخراج استنباط‌های مفید از داده‌ها در مورد مردم باشد، یک اقتصادسنج کاربردی است.

بسیاری از کتاب‌های درسی راهنمایی را برای روش‌های تحقیق ارائه می‌دهند و بین این کتاب و کتاب‌های دیگری که به طور گسترده استفاده می‌شوند تا حدی هم‌پوشانی وجود دارد، اما کتاب راهنمای ما از چند باب مهم با متون اقتصادسنجی متفاوت است. نخست اینکه، ما معتقدیم که تحقیقات تجربی، زمانی بسیار ارزشمند هستند که از داده‌ها برای پاسخ به سؤالات علّیتی خاص استفاده کنند، همانند وضعیتی که در آزمایه بالینی تصادفی[1] برقرار است. این دیدگاه، رویکرد ما را در خصوص تمام سؤالات تحقیق شکل می‌دهد. در نبود آزمایش واقعی، ما به دنبال مقایسه‌های به خوبی کنترل شده[2] و/یا شبه ـ آزمایش‌های[3] طبیعی می‌رویم. البته برخی از طرح‌های تحقیق شبه ـ آزمایشی نسبت به طرح‌های دیگر قانع‌کننده‌تر هستند، اما روش‌های اقتصادسنجی مورد استفاده در این مطالعات تقریباً همواره نسبتاً ساده هستند. در نتیجه، کتاب ما کوتاه‌تر و متمرکزتر از نحوه پرداختن کتاب‌های درسی به روش‌های اقتصادسنجی است. ما بر موضوعات مفهومی و تکنیک‌های آماری ساده‌ای تأکید می‌کنیم که در تحقیقات کاربردی که می‌خوانیم و انجام می‌دهیم ظاهر می‌شوند، و این ایده‌ها و تکنیک‌ها را با تعداد زیادی نمونه تجربی شرح می‌دهیم. اگرچه دیدگاه‌های ما در خصوص آنچه مهم است به صورت کلی در میان اقتصاددانان کاربردی، مشترک نیست، اما هیچ بحثی بر سر این واقعیت که طرح‌های تحقیقاتی آزمایشی و شبه ـ آزمایشی به طور فزاینده‌ای در قلب مؤثرترین مطالعات تجربی در اقتصاد کاربردی قرار دارند، وجود ندارد.

دومین تمایزی که مدعی‌اش هستیم فقدان مشخص جدیت است. اکثر متون اقتصادسنجی ظاهراً مدل‌های اقتصادسنجی را بسیار جدی می‌گیرند. به طور معمول، این کتاب‌ها توجه زیادی به نقص‌های فرضی مفروضات مدل‌سازی کلاسیک از قبیل خطی بودن و همپراشی دارند. گاهی اوقات هشدارهایی صادر می‌شوند.

---

1. Randomized Clinical Trial
2. Well-controlled Comparisons
3. Quasi-experiments

ما رویکرد متساهل‌تر و مبنایی‌تری را اتخاذ می‌کنیم. اصلی که بحث ما را هدایت می‌کند این است که برآوردکننده‌های رایج تقریباً همیشه تفسیر ساده‌ای دارند که خیلی وابسته به الگو نیست.

اگر برآوردهایی که به دست می‌آورید برآوردهایی نیستند که می‌خواهید، خطا در اقتصادسنج است نه در اقتصادسنجی! یک مثال برجسته رگرسیون خطی است که اطلاعات مهمی را در مورد تابع میانگین شرطی، صرف‌نظر از شکل این تابع، ارائه می‌کند. به همین ترتیب، روش‌های متغیرهای ابزاری یک اثر علّی میانگین را برای یک جامعه آماری به خوبی تعریف شده برآورد می‌کنند، حتی اگر این ابزار بر همه تأثیر نگذارد. استواری مفهومی ابزارهای اقتصادسنجی اساسی به طور شهودی توسط بسیاری از محققان کاربردی درک می‌شود، اما نظریه‌ای که پشت این استواری قرار دارد در اکثر متون نمایان نمی‌شود. کتاب راهنمای ما نیز با اکثر کتاب‌های اقتصادسنجی متفاوت است، از این باب که از نظر استنباط، ما خیلی با اثربخشی مجانبی سروکار نداریم.

در عوض، بحث ما در مورد استنباط بیشتر مربوط به دشواری‌های نمونه ـ محدود است که متخصصان را آزار می‌دهند.

پیش‌نیازهای اصلی برای موضوع مطرح شده در اینجا، آموزش پایه‌ای احتمال و آمار است. ما به طور خاص امیدواریم که خوانندگان با ابزارهای ابتدایی استنباط آماری، مانند آماره t و خطاهای معیار، آشنا باشند. آشنایی با مفاهیم بنیادی احتمال مانند امید ریاضی نیز مفید است، اما به تخصص خیلی زیاد در مورد ریاضی نیاز نیست. اگرچه شواهد مهمی ارائه می‌شوند، اما بحث‌های فنی خیلی طولانی یا پیچیده نیستند. برخلاف بسیاری از کتاب‌های اقتصادسنجی سطح بالا، ما به سادگی به جبر خطی می‌پردازیم. به همین دلیل و به دلایل دیگر، خواندن کتاب راهنمای ما راحت‌تر از کتاب‌های رقیب است. در نهایت، همانند سری کتاب مفرح داگلاس آدامز (راهنمای مسافران پیاده ـ بدون سفینه ـ در کهکشان،[1] تقریباً بی‌ضرر،[2] و غیره) که ما مدام از آنها الهام می‌گیریم، کتاب راهنمای ما ممکن است گاهاً بی‌دقتی‌هایی داشته باشد، اما بسیار ارزان‌تر از بسیاری از نسخه‌های E-cyclopedia Galactica Econometrica است که بازار را در چنگ خود دارند. از انتشارات دانشگاه بسیار سپاسگزارم که با توزیع کتاب ما با این شرایط موافقت کرد.

1. The Hitchhiker's Guide to the Galaxy
2. Mostly Harmless

قدردانی

به موازات پیشرفت پروژه نوشتن این کتاب، از نظرات بسیاری از دوستان و همکاران بهره بردیم. از آلبرتو آبادی، دیوید اتور، آمیتاب چاندرا، مونیکا چن، جان دی ناردو، جو دیل، جری هاسمن، آندریا ایچینو، گیدو ایمنز، رافائل لالیو، آلن منینگ، کارن نوربرگ، باربارا پتررونولو، جیمز رابینسون، تاونیت سوری و جف وودلدریج، که در مراحل مختلف نظرات خودشان در مورد پیش‌نویس متن کتاب را با ما در میان گذاشتند، تشکر ویژه داریم. آنها مسئول اشتباهات و نواقص کار نیستند. همچنین از دانشجویانمان در LSE و MIT تشکر می‌کنیم. آنها نسخه‌های اولیه کتاب را در ابتدا مشاهده کرده و به ما کمک کردند که بتوانیم تصمیم بگیریم چه چیزی مهم است. از دستیاران تحقیق، بریگهام فرنسن، سینتیا کینان و کریس اسمیت که ماهرانه و خستگی‌ناپذیر به ما کمک کردند، تشکر ویژه داریم. ما همچنین از حمایت و هدایت تیم سولیوان و ست دچیک، ویراستاران ما در انتشارات دانشگاه پرینستون سپاسگزاریم و در انتها ما از همسرانمان برای عشق و حمایتشان سپاسگزاریم؛ آنها بهتر از هر کسی می‌دانند که معنای همدم تجربه‌گرا بودن یعنی چه.

ساختار کتاب

ما با دو فصل مقدماتی شروع می‌کنیم. اولین فصل، برنامه پژوهشی‌ای که مطالب فصل‌های بعدی برای استفاده آنها مفید است را توصیف می‌کند و فصل دوم نکاتی در مورد این موضوع که آزمایش‌های آماری مورد نظر ما، یعنی آزمایشات تصادفیده[1] تحقیقات درمانی، معیارهایی ایدئال برای سؤالاتی که از نظر ما جالب هستند ارائه می‌کنند. پس از این مقدمه، سه فصل بخش دوم، مواد اصلی رگرسیون، متغیرهای ابزاری و تفاضلِ تفاضل را ارائه می‌دهند. این فصول، بر ویژگی‌های عمومی برآوردگرهای مربوطه (به عنوان مثال، رگرسیون همیشه تابع میانگین شرطی را تقریب می‌زند) و مفروضات لازم برای تفسیر علّی نتایج (فرض استقلال شرطی، ابزارهایی که به خوبی ابزارهای تصادفاً تخصیص داده شده هستند و دنیاهای موازی) تأکید دارند. سپس به قسمت‌های مهم بخش سوم می‌رویم. فصل ۶، طرح‌های گسست رگرسیونی را بررسی می‌کند که می‌توان به عنوان یک مشتقه از استراتژی‌های کنترل رگرسیونی و یا یک نوع استراتژی متغیرهای ابزاری به آن نگریست. در فصل هفتم، در مورد استفاده از رگرسیون چندکی[2] برای تخمین تأثیرات توزیع بحث خواهیم کرد. در فصل آخر، مشکلات مهم استنباط آماری که از طریق رویکرد مجانبی کتب درسی دست‌نیافتنی است بررسی می‌شود. برخی فصل‌ها شامل بخش‌های تخصصی فنی یا تخصصی می‌شوند که عدم مطالعه آنها به ایده‌های اصلی مورد توجه کتاب آسیبی

---

1. Randomized Trials
2. Quantile Regression

نمی‌زنند و می‌توانید از آنها رد شوید. این فصل‌ها با ستاره (*) مشخص می‌شوند. نشانه‌شناسی، واژه‌نامه اختصارات و فهرست نمونه‌های تجربی در انتها آمده است.

الف) مقدمه

۱. سؤال‌هایی در مورد سؤال‌ها

رایانه گفت «آن را کاملاً بررسی کردم و پاسخ قطعاً همان است. فکر می‌کنم اگر با شما صادق باشم، مشکل در این است که شما هرگز نمی‌دانستید که سؤال چیست؟»

داگلاس آدامز، راهنمای کهکشان‌گردی برای مسافران پیاده[1] (۱۹۷۹)

بسیاری از دوره‌های آموزشی اقتصادسنجی، با انتخاب موضوعات به همان صورتی که در ادبیات اقتصادسنجی ارائه می‌شود، به جزئیات تحقیقات تجربی توجه دارند، اما آنچه مبنای یک تجزیه و تحلیل آماری مفید قرار می‌گیرد، یک برنامه تحقیقاتی منسجم، جالب و قابل اجراست. اقتصادسنجی خوب نمی‌تواند یک برنامه تحقیقاتی ضعیف را نجات دهد، اما استفاده بی‌قاعده از تکنیک‌های اقتصادسنجی تجملی (تفننی)،[2] گاهی اوقات یک برنامه خوب را تضعیف می‌کند. در این فصل به طور خلاصه مبنای یک پروژه تحقیقاتی موفق مورد بحث قرار می‌گیرد. همانند داستان مهاجرت قوم یهود در انجیل، یک برنامه تحقیقاتی را می‌توان حول چهار سؤال تنظیم کرد. ما این سؤالات را سؤالات پرسیده شده به صورت مکرر (FAQ)[3] می‌نامیم، زیرا باید به صورت مکرر پرسیده شوند. سؤالات مکرر، در مورد رابطه مورد نظر یا علاقه،[4] آزمایش ایدئال،[5] استراتژی شناسایی[6] و شیوه استنباط[7] سؤال می‌کنند.

در ابتدا باید بپرسیم: رابطه علیت مورد علاقه چیست؟ اگرچه تحقیق صرفاً توصیفی، نقش مهمی در تحلیل ایفا می‌کند، اما ما معتقدیم که جالب‌ترین تحقیق در علوم اجتماعی، مانند تأثیر اندازه کلاس بر نمرات آزمون کودکان که در فصل ۲ و ۶ مورد بحث قرار گرفته است، در مورد علت و اثر است. رابطه علیت برای پیش‌بینی‌هایی درباره عواقب تغییر شرایط یا سیاست‌ها مفید است و به ما می‌گوید که در دنیاهای جایگزین (یا غیر از حالت کنونی) چه تصادفی می‌افتد. به عنوان مثال، به عنوان بخشی از برنامه تحقیقاتی که ظرفیت تولید انسانی را بررسی می‌کند ـ که اقتصاددانان کار آن را سرمایه انسانی می‌نامند ـ ما اثر علّی تحصیل بر دستمزد را مورد بررسی قرار دادیم (کارد، ۱۹۹۹، تحقیقات در این زمینه را

1. Douglas Adams, The Hitchhiker's Guide to the Galaxy (1979)
2. Fancy
3. Frequently Asked Questions
4. Relationship of Interest
5. Ideal Experiment
6. Identification Strategy
7. Mode of Inference

بررسی کرده است). تأثیر علّی آموزش بر دستمزد به صورت افزایش دستمزد دریافتی فرد متناسب با تحصیلات بیشتر او بروز می‌کند. طیفی از مطالعات نشان می‌دهد که مدرک دانشگاهی حدود ۴۰ درصد دستمزد را نسبت به میانگین بالاتر می‌برد که نتیجه قابل توجهی است. رابطه علّی تأثیر آموزش بر دستمزد برای پیش‌بینی پیامدهای آن بر درآمد، مثلاً در نتیجه تغییر هزینه‌های تحصیل در دانشگاه یا تقویت قوانین حضور اجباری مفید است. این رابطه از نظر تئوری نیز جالب است، زیرا می‌تواند از یک الگوی اقتصادی نیز حاصل شود.

ما به عنوان اقتصاددانان حوزه کار به احتمال زیاد آثار علّی را در نمونه‌های کارگری بررسی می‌کنیم، اما واحد مشاهده در این مطالعات نباید فرد باشد. سؤالات علیت می‌تواند در مورد شرکت‌ها و یا در مواردی در مورد کشورها پرسیده شود. نمونه‌ای از مورد اخیر تحقیق عاصم اغلو، جانسون و رابینسون (۲۰۰۱) در مورد تأثیر نهادهای استعماری بر رشد اقتصادی است. این مطالعه مربوط به این مسئله است که آیا کشورهایی که نهادهای دموکراتیک‌تری را از حاکمان استعماری خود به ارث برده‌اند، بعدها از رشد اقتصادی بالاتری برخوردار شده‌اند یا نه. پاسخ به این سؤال حاوی دلالت‌هایی بر درک ما از تاریخ و پیامدهای سیاست توسعه‌ای معاصر است. برای مثال، امروزه ممکن است بپرسیم که آیا نهادهای نوظهور دموکراتیک برای توسعه اقتصادی در عراق و افغانستان اهمیت دارند یا نه. پاسخ به این سؤال در مورد دموکراسی فاصله زیادی با جواب روشن دارد؛ چراکه در حال حاضر چین بدون بهره‌مندی از آزادی سیاسی کامل از رشد قدرتمندتری برخوردار است و این در حالی است که بسیاری از کشورهای آمریکای لاتین اگرچه دموکراتیک شده‌اند، اما رشد اقتصادی بزرگی به عنوان نتیجه آن تجربه نکرده‌اند.

دومین سؤال مکرر تحقیقاتی، مربوط به آزمایشی است که می‌تواند به صورت ایدئال برای ثبت اثر علّی مورد نظر استفاده شود. برای مثال، در مورد تحصیل و دستمزد می‌توانیم تصور کنیم که به ترک تحصیل‌کنندگان بالقوه جایزه‌ای برای فارغ‌التحصیلی پیشنهاد شود و سپس پیامدهای آن را مطالعه کنیم. در واقع، انگریست و لوی (۲۰۰۷) دقیقاً چنین آزمایشی را انجام دادند. اگرچه این مطالعه به آثار کوتاه‌مدتی مانند ثبت نام در کالج می‌پردازد و پیامد طولانی‌مدت‌تر آن می‌تواند نگاهی هم به دستمزدها بیاندازد. در مورد نهادهای سیاسی، ممکن است بخواهیم که به گذشته برگردیم و به طور تصادفی ساختارهای دولتی مختلفی را برای مستعمره‌های سابق در روزهای استقلالشان اعمال کنیم (آزمایشی که به احتمال زیاد به جای اینکه توسط بنیاد ملی علوم تأمین بودجه شود تبدیل به یک فیلم می‌شود).

آزمایش‌های ایدئال اغلب فرضی هستند. با این حال، آزمایشات فرضی ارزش در نظر گرفتن دارند، زیرا به ما کمک می‌کنند تا موضوعات پژوهشی مفیدی را انتخاب کنیم. ما این طور از این ادعا دفاع می‌کنیم؛ از شما می‌خواهیم که خود را به عنوان محققی فرض کنید که با محدودیت بودجه‌ای مواجه نیستید و هیچ کمیته موضوعات انسانی[1] برای صحیح بودن کار شما از لحاظ اجتماعی نظارت ندارد.

1. Human Subjects Committee

چیزی شبیه به استنلی میلگرم،[1] که به خوبی تأمین مالی شده است. میلگرم روانشناسی بود که در دهه ۱۹۶۰ کارهای تحولی‌ای در مورد واکنش به قدرت با استفاده از طرح‌های تجربی بسیار بحث‌برانگیز انجام داد که به احتمال زیاد آن کارها امروزه منجر به از دست دادن شغلش می‌شد.

میلگرم (۱۹۶۳) به دنبال درک چگونگی واکنش به قدرت، نشان داد که می‌تواند افراد مورد آزمایش را متقاعد کند تا شوک‌های الکتریکی دردناکی را به قربانیانی که به صورت ترحم‌برانگیزی اعتراض می‌کردند اعمال کنند (شوک‌ها موهومی و قربانیان بازیگر بودند). معلوم شد که این آزمایش بحث‌برانگیز و همچنین هوشمندانه است ـ برخی از روانشناسان ادعا کردند که افرادی که شوک‌ها را اعمال کردند در این آزمایش از نظر روانشناختی صدمه دیدند. با این حال، مطالعه میلگرم این نکته را نشان می‌دهد که آزمایشات بسیاری وجود دارد که می‌توانیم در مورد آن فکر کنیم حتی اگر بهتر باشد که برخی از آنها در حالت طراحی باقی بمانند.[2] اگر نمی‌توانید آزمایشی طراحی کنید که به سؤال شما در جهانی که در آن هر چیزی اتفاق می‌افتد پاسخ دهد، پس احتمال تولید نتایج مفید با بودجه متوسط و داده‌های آمارگیری غیرتجربی بسیار کم به نظر می‌رسد. همچنین توصیف یک آزمایش ایدئال به شما کمک می‌کند که سؤالات مرتبط با علیت را به طور دقیق تنظیم کنید.

سازوکار یک آزمایش ایدئال، نیروهایی را که می‌خواهید دستکاری کنید و عواملی که می‌خواهید ثابت نگه دارید را برجسته می‌کند.

سؤالاتی از تحقیق که با هر آزمایشی قابل پاسخ دادن نیستند، سؤالات مشخص نشده به طور اساسی (FUQ'd) هستند.[3] یک سؤال «مشخص نشده به طور اساسی» دقیقاً به چه شکلی است؟ در نگاه اول، سؤالاتی در مورد اثر علیت نژاد یا جنسیت گزینه‌های خوبی به نظر می‌رسد، زیرا دستکاری این موارد به طور ایزوله شده دشوار است («تصور کنید کروموزوم‌های شما در هنگام تولد تعویض شده است»). از سوی دیگر، موضوعی که بیشتر مورد توجه اقتصاددانان در قلمرو نژاد و جنسیت است تبعیض در بازار کار است که مشخص می‌کند که آیا فردی به این دلیل که *معتقد* است شما سیاه‌پوست یا سفیدپوست و یا مرد یا زن هستید با شما رفتار متفاوتی دارد. مفهوم دنیای جایگزین که در آن مردان به عنوان زنان شناخته می‌شوند یا برعکس، دارای سابقه‌ای طولانی است و نیاز به سبک عجیب داگلاس آدامز برای معرفی ندارد (روزالیند در نمایشنامه «*هر طور شما دوست دارید*» اثر شکسپیر خودش را به جای گینمد جا زده است). ایده تغییر نژاد نیز به همین مسئله نزدیک است: در «*لکه انسانی*»، فیلیپ روث، جهان متعلق به کلدمن سیلک ـ یک استاد ادبیات سیاه‌پوست که در زندگی حرفه‌ای خود یک سفیدپوست در

---

1. Stanley Milgram

۲. نقش میلگرم بعدتر توسط ویلیام شانتر در یک فیلم تلویزیونی بازی شد، افتخاری که هیچ اقتصاددانی تاکنون دریافت نکرده است، هرچند انگریست (Angirst) هنوز امیدوار است.

3. Fundamentally Unidentied Questions

نظر گرفته می‌شود ـ را تصور می‌کند. اقتصاددانان کار، همیشه چنین چیزهایی را تصور می‌کنند. بعضی اوقات حتی چنین سناریوهایی را برای پیشرفت علم ایجاد می‌کنیم، همان طور که در برخی مطالعات ارزیابی، از متقاضیان شغل ساختگی و رزومه کاری موهومی استفاده می‌کنیم.[۱]

کمترین تصورات هم در مورد طراحی تحقیقات تأثیر زیادی دارد، اما تصورات نمی‌تواند هر مشکلی را حل کند. فرض کنید که علاقمندیم بدانیم که اگر کودکان کمی بزرگ‌تر از سن کنونی شروع به تحصیل کنند عملکرد بهتری خواهند داشت یا نه؟ شاید مغز ۷ ساله نسبت به مغز ۶ ساله آمادگی بیشتری برای یادگیری داشته باشد. این سؤال جنبه خط‌مشی‌ای دارد که از این واقعیت به دست می‌آید که برخی از مدارس در حال حاضر برای تلاش برای افزایش نمرات آزمون در حال بررسی بالا بردن سن شروع تحصیل هستند (که موجب ناراحتی بسیاری از مادران شاغل است). برای ارزیابی تأثیر تأخیر در شروع تحصیل بر یادگیری ممکن است تعدادی کودک را به طور تصادفی انتخاب کنیم که مهدکودک را در سن ۶ سالگی شروع کنند، در حالی که دیگران در ۵ سالگی شروع می‌کنند، همان طور که اکنون معمول است. ما علاقمندیم بدانیم که آیا کسانی که عقب می‌مانند همان گونه که در نمرات آزمون مدرسه ابتدایی آنها نشان داده شده است یادگیری بیشتری در مدرسه دارند یا نه؟ برای اطمینان به نمرات آزمون در کلاس اول نگاه می‌کنیم.

مشکل این سؤال ـ اثرات سن شروع بر نمرات آزمون کلاس اول ـ این است که گروهی که مدرسه را در ۷ سالگی شروع کرده‌اند، بزرگ‌تر هستند و بچه‌های بزرگ‌تر معمولاً نمرات بهتری در آزمون کسب می‌کنند و این اثر خالص بزرگ شدن است. اکنون ممکن است به نظر برسد که می‌توانیم این مشکل را با ثابت نگه داشتن سن به جای کلاس حل کنیم. فرض کنید ما کسانی که در سن ۶ سالگی تحصیل را شروع کردند و کلاس دوم هستند را با کسانی که در سن ۷ سالگی تحصیل را شروع کردند و کلاس اول هستند مقایسه کنیم، در این حالت همه در ۷ سالگی آزمایش می‌شوند، اما گروه اول زمان بیشتری را در مدرسه گذرانده‌اند؛ واقعیتی که اگر مدرسه چیز ارزشمندی باشد پیشرفت را افزایش می‌دهد. تا زمانی که بچه‌ها هنوز در مدرسه هستند هیچ راهی برای جدا کردن اثر سن شروع، از اثرات بلوغ و زمان بودن در مدرسه وجود ندارد. مشکل اینجاست که سن شروع برابر با سن فعلی منهای زمان بودن در مدرسه است. این ارتباط قطعی در نمونه‌های متعلق به بزرگسالان از بین می‌رود، بنابراین ممکن است به بررسی این نکته امید داشته باشیم که آیا تغییرات در سیاست‌های سن ورود به تحصیل، بر نتایج متعلق به بزرگسالان مانند درآمد و یا بالاترین مدرک دریافت شده تأثیر دارد یا نه، اما تأثیر سن شروع بر نمرات آزمون ابتدایی به احتمال زیاد FUQ'd است.

۱. مثال اخیر مطالعه Bertrand و Mullainathan (۲۰۰۴) است که واکنش کارفرمایان را به رزومه‌هایی با نام‌هایی که به نظر سیاه‌تر و سفیدتر می‌رسید مانند Lakisha و Emily مقایسه کرد (اگرچه Fryer و Levitt، ۲۰۰۴ خاطرنشان کردند که نام، ممکن است علاوه بر نژاد اطلاعاتی در مورد وضعیت اجتماعی و اقتصادی نیز داشته باشد).

سؤالات مکرر پرسیده شده سوم و چهارم تحقیق مربوط به عناصر ریز و مهم کاربردی است که یک مطالعه خاص را تولید می‌کند. سؤال شماره ۳ می‌پرسد: *استراتژی شناسایی شما چیست؟* کروگر و انگریست (۱۹۹۹) از اصطلاح *استراتژی شناسایی*[1] برای توصیف شیوه‌ای استفاده کردند که محقق از داده‌های مشاهداتی (یعنی داده‌هایی که توسط یک آزمایش تصادفی تولید نشده است) برای تقریب یک آزمایش واقعی استفاده می‌کند. باز به مثال تحصیل بازمی‌گردیم، انگریست و کروگر (۱۹۹۱) از تعامل بین قوانین اجباری حضور در مدارس آمریکایی و فصل تولد دانش‌آموزان به عنوان یک آزمایش طبیعی برای برآورد آثار فارغ‌التحصیلی از دبیرستان بر دستمزد استفاده کردند (فصل تولد بر میزانی که دانش‌آموزان دبیرستانی مجبور به حضور در مدرسه، در روز تولد خود ترک تحصیل می‌کنند مؤثر است). فصل‌های ۳ تا ۶ عمدتاً در مورد چارچوب مفهومی استراتژی‌های شناسایی است.

اگرچه تمرکز بر استراتژی‌های شناسایی معتبر نشانه‌ای از کار تجربی مدرن است، همبستگی آزمایش‌های ایدئال و طبیعی دارای تاریخچه طولانی در اقتصادسنجی است. در اینجا پیشگام اقتصادسنجی‌مان، تریگوه هاولمو[2] (۱۹۴۴، ص ۱۴) را می‌بینیم که درخواست بحث صریح‌تری در مورد هر دو نوع طرح‌های آزمایشی را دارد:

طراحی آزمایش (نسخه‌ای از آنچه فیزیکدانان «آزمایش حیاتی» می‌نامند) ضمیمه‌ای ضروری برای هر تئوری مقداری است و ما معمولاً وقتی چنین تئوری‌هایی را می‌سازیم چند مورد از چنین آزمایش‌هایی را در ذهن داریم، اگرچه ـ متأسفانه ـ اکثر اقتصاددانان طراحی‌شان از آزمایش‌ها را با صراحت توصیف نمی‌کنند. اگر این کار را انجام دهند می‌بینند که آزمایش‌هایی که در ذهن دارند ممکن است به دو دسته متفاوت تقسیم شود: (۱) آزمایش‌هایی که *بهتر است بخواهیم انجام دهیم* تا ببینیم آیا پدیده اقتصادی واقعی مشخصی ـ زمانی که به طور مصنوعی از «تأثیرات دیگر» جدا می‌شود ـ فرضیات مشخصی را تأیید می‌کند و (۲) جریان آزمایش‌هایی که طبیعت به طور مداوم در آزمایشگاه عظیم خود انجام می‌دهد و ما فقط به عنوان ناظران غیرفعال مشاهده می‌کنیم. در هر دو مورد، هدف نظریه یک چیز است، تبدیل شدن به استاد وقایع زندگی واقعی.

سؤال مکرر چهارم تحقیق، ادبیات خود را از روبین (۱۹۹۱) اقتباس می‌کند: *شیوه استنباط آماری شما چیست؟* پاسخ این سؤال، جامعه مورد مطالعه، نمونه مورد استفاده و فرضیه‌های ایجاد شده در زمان ساخت خطاهای استاندارد را توصیف می‌کند. گاهی اوقات استدلال ساده است، مانند زمانی که از نمونه‌های داده‌های خُرد سرشماری برای مطالعه جامعه آمریکا استفاده می‌کنید. در اغلب موارد استنباط آماری پیچیده‌تر است، اگرچه به طور خاص در مورد داده‌هایی که خوشه‌بندی یا گروه‌بندی شده‌اند این

1. Identification Strategy
2. Trygve Haavelmo

وضعیت پیش می‌آید. فصل آخر مسائل عملی را مطرح می‌کند که بعد از پاسخ دادن به سؤال مکرر شماره ۴ پیش می‌آید. اگرچه مسائل مرتبط با استنباط به ندرت خیلی هیجان‌انگیز و اغلب کاملاً فنی هستند، اما موفقیت نهایی حتی یک پروژه به خوبی تهیه شده و با مفهومی هیجان‌انگیز، وابسته به جزئیات استنباط آماری است. این واقعیت که گاهی هم گمراه‌کننده است، الهام‌بخش این هایکوی اقتصادسنجی است که کیسوکه هیرانو[1] که در آن زمان دانشجوی دکتری اقتصادسنجی بود به مناسبت تکمیل پایان‌نامه‌اش نوشته شده است:

آماره **t** به نظر خیلی خوب می‌رسد.
از خطاهای استاندار استوار استفاده کنید ـ
دیگر معنادار نیست

همان طور که از بحث فوق باید مشخص شده باشد، چهار سؤال مکرر تحقیق که ذکر شد، بخشی از فرایند توسعه پروژه است. فصل‌های بعد بیشتر به سؤالاتی از اقتصادسنجی مربوط می‌شود که پس از پاسخ دادن به سؤالات مکرر تحقیق پیش می‌آید. به عبارت دیگر، مربوط به مسائلی است که پس از تنظیم برنامه تحقیق ایجاد می‌شود. با این حال، قبل از وارد شدن به اجزای کاربردی کار تجربی، بحث را با توضیح دقیق‌تری از اینکه چرا معیار ما با استفاده از آزمایشات تصادفی به دست می‌آید آغاز می‌کنیم.

۲. ایدئال تجربی

همه چیز آن طور که به نظر می‌رسد، نیست و این حقیقتی مهم و معروف است. برای مثال، روی سیاره زمین، انسان همواره تصور کرده بود که به دلیل دستاوردهای زیادی که کسب کرده است، از دلفین باهوش‌تر است ـ اختراع چرخ، ایجاد شهری مانند نیویورک، جنگ‌ها و غیره ـ در حالی که تمام چیزی که دلفین‌ها تا به حال انجام داده‌اند، انجام کارهای سرگرم‌کننده در آب و خوشگذرانی بوده است، اما برعکس، دلفین‌ها هم همیشه باور داشتند که از انسان‌ها باهوش‌ترند ـ آن هم دقیقاً به همان دلیل. در حقیقت، فقط یک گونه روی زمین بود که از دلفین‌ها باهوش‌تر بود، و آنها (این گونه) زمان زیادی را در آزمایشگاه‌های پژوهشی رفتارشناسی می‌گذراندند؛ در حالی که آزمایش‌هایی به طور دلهره‌آور پیچیده و ظریف روی انسان انجام می‌دادند. این حقیقت که انسان این ارتباط را به اشتباه تفسیر کرد، کاملاً طبق این برنامه‌های این موجودات بود.

داگلاس آدامز، راهنمای کهکشان‌گردی برای مسافران بدون سفینه (۱۹۷۹)

معتبرترین و مؤثرترین طراحی‌های پژوهشی از تخصیص تصادفی[2] استفاده می‌کنند. پروژه پیش‌دبستان پِری[3] مثالی خوب در این مورد است. این پروژه آزمایشی ـ تصادفی در سال ۱۹۶۲ بود که

1. Keisuke Hirano
2. Random Assignment
3. Perry

برای ارزیابی اثر برنامه آموزش پیش‌دبستانی طراحی شده بود و شامل ۱۲۳ دانش‌آموز سیاه‌پوست در ایپسیلانتی میشیگان بود. گروه مورد آزمایش تحت دوره فشرده آموزش از جمله آموزش پیش از دبستان و ملاقات در خانه قرار گرفتند. به ندرت می‌شود درباره اثر آزمایش کوچک، اما به خوبی طراحی شده پری مبالغه کرد که داده‌های تکمیلی را روی همین شرکت‌کنندگان در آزمایش در سن ۲۷ سالگی در سال ۱۹۹۳ به دست آورد. تعداد زیادی از پژوهش‌های آکادمیک از پژوهش پری استفاده کردند یا از آن نام بردند (برای مثال بارنت، ۱۹۹۲ را ببینید). مهم‌تر اینکه، پروژه پری بنیادی عقلانی برای برنامه آموزشی بزرگ پیش‌بینی هد استارت[1] که در سال ۱۹۶۴ شروع شد، فراهم کرد که در نهایت به میلیون‌ها کودک آمریکایی خدمت کرد (و هنوز هم خدمت می‌کند).[2]

۱-۲. مشکل گزینش

پیش از پرداختن به بحث رسمی درباره نقشی که آزمایش‌ها در آشکار کردن اثر عِلّی دارند، زمانی کوتاه توقف می‌کنیم. تصور کنید که شما به سؤالات معمول «اگر ـ سپس» علاقه‌مندید. برای اطمینان، مثالی ساده را در نظر بگیرید. آیا بیمارستان همه را سالم‌تر می‌کند؟ برای هدف ما، این سؤال نمادین است، اما به شکلی شگفت‌انگیز از آن دسته سؤالاتی است که اقتصاددانان سلامت به آن توجه زیادی دارند. برای واقع‌بین‌تر کردن این سؤال، تصور کنید که جامعه سالمند فقیری را مطالعه می‌کنیم که از اورژانس بیمارستان برای تیمار بهداشتی ضروری استفاده می‌کنند. تعدادی از این بیماران در بیمارستان بستری می‌شوند. این دسته از بیماران در بیمارستان ازدحام ایجاد می‌کنند و احتمالاً زیاد ثمربخش نیست (گرومبچ، کین، و بایندمن،[3] ۱۹۹۳ را ببینید). در حقیقت، افراد آسیب‌پذیر در کنار افراد بیمار قرار می‌گیرند و اثر منفی بر سلامت آنها می‌گذارد. از آنجا که افرادی که در بیمارستان بستری می‌شوند، خدمات زیاد باارزشی دریافت می‌کنند، پاسخ به سؤال اثربخشی بیمارستان، به احتمال زیاد مثبت است، اما آیا داده‌ها از این موضوع پشتیبانی می‌کند؟ رویکرد طبیعی برای فردی با ذهنیت تجربی این است که وضعیت سلامت آنهایی که در بیمارستان بستری شدند و آنها که بستری نشدند را مقایسه کند. سازمان بهداشت و سلامت ملی[4] اطلاعاتی دارد که برای این مقایسه مورد نیاز است. به ویژه، این مقایسه این سؤال را دربر می‌گیرد که «آیا در طول ۱۲ ماه گذشته، پاسخ‌دهنده بیماری بوده است که فقط یک شب در بیمارستان اقامت داشته است؟». ما می‌توانیم از این سؤال برای شناسایی مراجعان به بیمارستان استفاده کنیم. همچنین پرسشنامه سازمان مذکور می‌پرسد «آیا اعتقاد دارید

---

1. Head Start Pre-school Program

۲. داده‌هاي پري همچنان توجه‌ها را به خود جلب مي‌كند، به ويژه آنكه علاقه به سياستگذاري در مورد تحصيلات اوليه بار ديگر جلب شده است. يك تحقيق اخير مجدد توسط مايكل اندرسون (۲۰۰۶) بيشتر يافته‌هاي تحقيق پري را تأييد مي‌كند، اگرچه اندرسون نشان مي‌دهد كه عمده تأثيرات مثبت نشان داده شده در تحقيق پره، تماماً ناشي از تأثير در مورد دختران بوده است و به نظر مي‌رسد، مداخله پري، تأثير خاصي روي پسرها نداشته است.

3. Grumbach, Keane, and Bindman

4. National Health Interview Survey (NHIS)

که وضعیت سلامت شما به طور کل عالی، خیلی خوب، خوب، نسبتاً خوب، ضعیف است؟» جدول زیر میانگین وضعیت سلامت را (۱ وضعیت عالی و ۵ وضعیت ضعیف) در میان آنهایی که در بیمارستان بستری شدند و آنهایی که بستری نشدند (گرفته شده از داده‌های **NHIS** از سال ۲۰۰۵) نشان می‌دهد:

گروه	اندازه ی نمونه	میانگین وضعیت سلامت	خطای استاندارد
بیمارستان	7774	2.79	0.014
غیر بیمارستان	90049	2.07	0.003

تفاوت میانگین‌ها ۰/۷۱ است، اختلافی بزرگ و قابل ملاحظه به نفع بستری نشده‌ها با آماره t ۵۸٫۹. در ظاهر امر، نتایج نشان می‌دهد که رفتن به بیمارستان مردم را بیمارتر می‌کند. ممکن است که جواب صحیح باشد: بیمارستان پُر از بیمارانی است که ممکن است بیماری آنها به ما سرایت کند و دستگاه‌ها و مواد شیمیایی خطرناکی که ممکن است به ما آسیب بزند، اما دلیل اینکه چرا این مقایسه نباید به راحتی پذیرفته شود، به راحتی قابل تشخیص است: آنهایی که بیمارستان می‌روند احتمالاً از سلامتی کمی برخوردارند. به علاوه، حتی بعد از بستری شدن بیمارانی که جویای درمان بودند، به طور میانگین نسبت به آنهایی که هیچگاه بستری نشدند، از سلامت کمتری برخوردارند، گرچه ممکن است وضعیت سلامتی‌شان در صورت بستری شدن نسبت به بستری نشدن بهتر شود. برای شرح چنین مشکلی به طور دقیق به درمان بیمارستان به عنوان متغیر تصادفی باینری (دو حالتی) $D_i=\{0,1\}$ در نظر بگیریم. خروجی مورد نظر، یعنی میزان وضعیت سلامت، با Y_i مشخص می‌شود. سؤال این است که آیا Y_i تحت تأثیر درمان است یا خیر؟ برای پاسخ به این سؤال ما فرض می‌کنیم که می‌توانیم تصور کنیم اگر شخصی به بیمارستان نرفته بود، ممکن بود در صورت رفتن چه اتفاقی برایش بیفتد و برعکس. بنابراین برای هر شخص دو متغیر پتانسیل سلامت وجود دارد:

$$\text{خروجی بالقوه} = \begin{cases} Y_{1i} & \text{if } D_i = 1 \\ Y_{0i} & \text{if } D_i = 0 \end{cases}$$

به عبارت دیگر، $Y0i$ وضعیت سلامت یک شخص است؛ اگر به بیمارستان نرفته بود، صرفنظر از اینکه او در حقیقت رفت یا خیر، در حالی که $Y1i$ وضعیت سلامتی شخص است؛ اگر او به بیمارستان برود. ما می‌خواهیم تفاضل بین $Y0i$ و $Y1i$ را که اثر علّیِ[1] رفتن به بیمارستان را برای شخص i است، بدانیم. اگر ما می‌توانستیم زمان را برگردانیم و شرایط درمانی شخص را تغییر دهیم، این چیزی است که اندازه می‌گرفتیم.

برآمد[2] مشاهده شده، Yi، می‌تواند به شکل به صورت خروجی‌های بالقوه[3] نوشته شود:

1. Casual Effect
2. Outcome
3. Potential Outcomes

$$\begin{aligned} Y_i &= \begin{cases} Y_{1i} & \text{if } D_i = 1 \\ Y_{0i} & \text{if } D_i = 0 \end{cases} \\ &= Y_{0i} + (Y_{1i} - Y_{0i})D_i. \end{aligned} \tag{2.1.1}$$

این نمادگذاری مفید است، زیرا $Y_{1i} - Y_{0i}$ اثر علّی بستری شدن برای یک فرد است. در کل، احتمال توزیع Y_{1i} و Y_{0i} در جامعه وجود دارد، بنابراین اثر درمان می‌تواند برای افراد مختلف، متفاوت باشد، اما به این دلیل که ما هیچگاه خروجی بالقوه را برای هرکسی مشاهده نمی‌کنیم، باید با مقایسه میانگین سلامت کسانی که بستری شدند و آنهایی که نشدند، اثر بستری شدن را پیدا کنیم.

مقایسه ناپخته[1] میانگین‌ها توسط وضعیت بستری شدن، به ما اطلاعاتی درباره خروجی بالقوه خواهد داد، گرچه نه لزوماً آنچه می‌خواهیم بدانیم. مقایسه میانگین عبارت شرطی سلامت در شرایط بستری شدن به شکلی جدی توسط معادله زیر به میانگین اثر عِلّی بستگی دارد:

$$\underbrace{E[Y_i|D_i=1] - E[Y_i|D_i=0]}_{\text{تفاضل دیده شده در میانگین سلامت}} = \underbrace{E[Y_{1i}|D_i=1] - E[Y_{0i}|D_i=1]}_{\text{میانگین اثر تیمار بر کسانی که درمان شدند}} + \underbrace{E[Y_{0i}|D_i=1] - E[Y_{0i}|D_i=0]}_{\text{اریبی گزینش}}$$

عبارت:

$$E[Y_{1i}|D_i=1] - E[Y_{0i}|D_i=1] = E[Y_{1i} - Y_{0i}|D_i=1]$$

میانگین اثر عِلّی برای کسانی است که بستری شدند. این عبارت تفاضل میانگین بین سلامت شخص بستری شده، یعنی $E[Y_{1i}|D_i = 1]$ و اینکه اگر شخص بستری نشده، چه اتفاقی برایشان می‌افتاد، یعنی $E[Y_{0i}|D_i = 1]$، را به دست می‌آورد. اما، تفاضل مشاهده شده در شرایط سلامت، عبارتی را به نام *اریبی گزینش*[2] به اثر عِلّی اضافه می‌کند. این عبارت تفاضل در میانگین Y_{0i} است بین کسانی که بستری شده و نشده‌اند. زیرا بیماران با احتمال زیادتری نسبت به افراد سالم جویای درمان هستند، کسانی که بستری شدند، Y_{0i} بدتری دارند که اریبی گزینش را در این مثال منفی می‌کند. قدر مطلق اریبی گزینش ممکن است بزرگ باشد که اثری درمانی مثبت را نامحسوس می‌کند. هدف بیشتر پژوهش‌های اقتصاد تجربی غلبه بر اریبی گزینش، و بنابراین گرفتن نتیجه درباره اثر عِلّی متغیری مثل D_i است.

1. Naive Comparisons
2. Selection Bias

۲-۲. تخصیص تصادفی مشکل گزینش را حل می‌کند

تخصیص تصادفی Di مشکل گزینش را حل می‌کند، زیرا تخصیص تصادفی، Di را از خروجی بالقوه مستقل می‌کند. برای مشاهده این نکته توجه کنید که:

$$\begin{aligned} E[Y_i|D_i=1]-E[Y_i|D_i=0] &= E[Y_{1i}|D_i=1]-E[Y_{0i}|D_i=0] \\ &= E[Y_{1i}|D_i=1]-E[Y_{0i}|D_i=1], \end{aligned}$$

وقتی که استقلال Y_{0i} و D_i به ما اجازه جابه‌جایی $E[Y_{0i}|D_i=1]$ را با $E[Y_{0i}|D_i=0]$ می‌دهد. در حقیقت، با داشتن تخصیص تصادفی، معادله بالا به این گونه ساده می‌شود:

$$\begin{aligned} E\left[Y_{1i}|D_i=1\right]-E\left[Y_{0i}|D_i=1\right] &= E\left[Y_{1i}-Y_{0i}|D_i=1\right] \\ &= E\left[Y_{1i}-Y_{0i}\right]. \end{aligned}$$

اثر بستری شدن به صورت تخصیص تصادفی بر بیماران بستری شده با اثر بستری شدن بر بیماران به صورت تصادفی انتخاب شده مساوی است. مسئله اصلی آن است که تخصیص تصادفیِ Di اریبی گزینش را حذف می‌کند. این به این معنا نیست که آزمایشات تصادفی بدون هیچ مشکلی هستند، اما آنها در اصل مهم‌ترین مشکلی که در تحقیقات تجربی اتفاق می‌افتد را حل می‌کنند.

اساساً این داستان بستری شدن در بیمارستان چقدر در رابطه با این مسئله مرتبط است؟ آزمایش‌ها، اغلب چیزهایی را نشان می‌دهند که در نتایج مقایسه‌ای ناپخته قابل مشاهده نیست. مثال اخیر در پزشکی ارزیابی درمان جایگزینی هورمون[1] است. این نوع درمان، مداخله پزشکی است که برای زنان میانسال برای کاهش علائم یائسگی تجویز می‌شود. اسناد به دست آمده از مطالعه سلامت پرستاران،[2] یک مطالعه غیرتجربی بزرگ و مؤثر در مورد پرستاران، وضعیت سلامت بهتری را در میان استفاده‌کنندگان از HRT نشان داد. در مقابل، نتایج آزمایه تصادفیده[3] که اخیراً به انجام رسیده است، اثر مثبت ناچیزی را از HRT نشان می‌دهد. علاوه بر این، این آزمایه تصادفیده، اثرات جانبی جدی را که در داده‌های غیرآزمایشی ظاهر نشده بود، آشکار کرد (برای مثال طرح بهداشت زنان[4] را ببینید، سیا[5] و همکاران، ۲۰۰۶).

مثالی نمادین از زمینه کاری خودمان در اقتصاد، کار بررسی برنامه آموزشی یارانه‌ای دولت است. این برنامه‌ها ترکیبی از تدریس کلاسی و آموزش در حین انجام کار برای کارگران محروم، کسانی که به مدت زیادی بیکار بوده‌اند، معتادان به دارو، و تبهکاران جنسی است و هدف آن افزایش فرصت شغلی و درآمد است. به طور متضاد، مطالعات بر اساس مقایسه‌های غیرتجربی بین شرکت‌کنندگان و کسانی که شرکت نکردند

1. Hormone Replacement Therapy
2. Nurses Health Study
3. Randomized trial
4. Women’s Health Initiative (WHI)
5. Hsia

نشان می‌دهد که بعد از آموزش، شرکت‌کنندگان آموزش‌دیده نسبت به گروه‌های مقایسه‌ای ممکن درآمد کمتری دارند (به طور مثال آشنفلتر،[1] ۱۹۷۸؛ آشنفلتر و کارد،[2] ۱۹۸۵؛ لالوند،[3] ۱۹۹۵). از آنجایی که هدف برنامه آموزشی، خدمت‌رسانی به مردان و زنان کم‌درآمد بود، در اینجا نیز اریبی گزینش موجب نگرانی است. بنابراین، مقایسه ناپخته بین شرکت‌کنندگان در این برنامه با کسانی که شرکت نکردند، درآمد کمتر را برای شرکت‌کنندگان نشان می‌دهد. در مقابل، اطلاعات به دست آمده از ارزیابی تصادفیده[4] نشان می‌دهد که برنامه آموزشی، عمدتاً اثرات مثبت به دست می‌دهد (برای مثال لالوند، ۱۹۸۶؛ اور و همکاران، ۱۹۹۶ را ببینید).

آزمایه‌های تصادفی هنوز در علوم اجتماعی مرسوم نیست. یکی از زمینه‌هایی که اهمیت تخصیص تصادفی به سرعت در آن در حال گسترش است، تحقیقات آموزشی است (آنگریست،[5] ۲۰۰۴). قانون اصلاحات علوم تربیتی در سال ۲۰۰۲ که توسط کنگره آمریکا تصویب شد، استفاده از طرح‌های پژوهشی شبه‌آزمایشی و کاملاً آزمایشی را برای تمام پژوهش‌های آموزشی سرمایه‌گذاری توسط دولت اجباری کرد. بنابراین ما می‌توانیم انتظار داشته باشیم که در سال‌های پیش رو، آزمایه‌های تصادفی بسیار بیشتری را در تحقیقات آموزشی ببینیم.

پژوهش تصادفیده پیشگام در زمینه آموزش، آزمایش تِنِسی استار[6] است که برای ارزیابی اثر کلاس‌های کوچک‌تر در دبستان طراحی شد. اقتصاددانان کار و سایر محققان به طور سنتی در تلاش برای برقراری ارتباط بین ویژگی‌های محیط کلاس و یادگیری دانش‌آموزان بوده‌اند، یعنی زمینه‌ای از پژوهش که ما «تولید آموزش»[7] می‌نامیم. این عبارت منعکس‌کننده این حقیقت است که ما ویژگی‌های محیط آموزشی را به عنوان ورودی‌هایی در نظر می‌گیریم که هزینه‌بر است، در حالی که خروجی‌ای که مدارس تولید می‌کنند، یادگیری دانش‌آموز است. سؤال کلیدی در پروهش‌های تولید آموزشی این است که کدام ورودی بیشترین خروجی را تولید می‌کند. یکی از گران‌ترین ورودی‌ها، اندازه کلاس است ـ زیرا کلاس‌های کوچک‌تر استخدام معلم‌های بیشتری می‌طلبد. بنابراین دانستن اینکه آیا هزینه کلاس‌های کوچک‌تر در صورت دستاورد بالاتر دانش‌آموزان جبران شده است یا خیر، اهمیت دارد. هدف آزمایش استار پاسخ این سؤال بود.

بسیاری از پژوهش‌های تولید آموزش که از داده‌های غیرآزمایشی استفاده می‌کنند نشان می‌دهند که ارتباط بسیار کمی بین اندازه کلاس و یادگیری دانش‌آموزان وجود دارد یا اصلاً ارتباطی وجود ندارد. ممکن است سیستم‌های مدرسه بتوانند با استخدام تعداد بسیار کمی معلم بدون هیچ کاهشی در دستاوردهایشان، پول بیشتری ذخیره کنند. اما، از آنجایی که دانش‌آموزان ضعیف‌تر اغلب عمدی به کلاس‌های کوچک‌تر گروه‌بندی می‌شوند، ارتباط مشاهده شده بین اندازه کلاس و دستاورد دانش‌آموزان

1. Ashenfelter
2. Card
3. Lalonde
4. Randomized Evaluations
5. Angrist
6. Tennessee Star
7. Education Production

نباید به راحتی پذیرفته شود. آزمایه تصادفی با مطمئن شدن از اینکه مقایسه بین دو مورد همگون انجام می‌شود، به این مشکل فائق می‌آید و اطمینان حاصل می‌کند که سیب با سیب مقایسه می‌شود، به عبارت دیگر، دانش‌آموزانی که در کلاس‌هایی با اندازه‌های مختلف قرار می‌گیرند، قابل مقایسه هستند. نتایج از آزمایش تنسی استار به کارایی قوی و ماندگار کلاس‌های کوچک‌تر اشاره کرد (فین و آچیل،[1] ۱۹۹۰؛ کروگر،[2] ۱۹۹۹؛ برای تحلیل اقتصادسنجی داده‌های استار).

آزمایش استار به شکل غیرطبیعی بلندپروازانه و قدرتمند بود، بنابراین ارزش بیان بعضی از جزئیات را دارد. هزینه پروژه معادل ۱۲ میلیون بود و در سال ۱۹۸۶- ۱۹۸۵ روی گروهی از بچه‌ها در کودکستان‌ها اجرا شد. این پژوهش برای چهار سال ادامه داشت، به عبارت دیگر تا زمانی که آن بچه‌ها به سال سوم دبستان راه یافتند و تعداد بچه‌های شرکت‌کننده ۱۱۶۰۰ نفر بود، میانگین اندازه کلاس برای یک کلاس عادی تنسی در سال ۱۹۸۶-۱۹۸۵ حدود ۲۲/۳ بود. این آزمایش، بچه‌ها را در یکی از سه گروه تیمار[3] قرار داد. کلاس‌های کوچک با ۱۷-۱۳ کودک، کلاس عادی ۲۲ تا ۲۵ کودک و یک دستیار معلم پاره‌وقت و یا یک کلاس عادی با دستیار تمام‌وقت. مدارسی با حداقل سه کلاس در هر مقطع تحصیلی می‌توانستند که به این آزمایش بپیوندند. اولین سؤال درباره آزمایش تصادفی این است که آیا تصافیده شدن به طور موفقیت‌آمیز می‌تواند ویژگی‌های شخص تحت آزمایش را در کل گروه‌های تیمار به تعادل رساند یا خیر؟ برای ارزیابی این مطلب، به طور معمول باید نتایج قبل از آزمایش یا سایر متغیرهای کمکی در گروه‌ها مشخص شوند. متأسفانه، داده‌های استار نمره‌های پیش‌آزمون را وارد نکرد، گرچه ممکن است که به خصوصیات بچه‌ها مثل نژاد و سن توجه کرده باشد.

جدول ۱-۲-۲ که از آزمایش کروگر (۱۹۹۹) به دست آمد، میانگین این متغیرها را نشان می‌دهد. خصوصیات دانش‌آموزان در جدول شامل متغیرهای ناهار رایگان، نژاد و سن دانش‌آموز بود. از آنجایی که کودکان فقیر ناهار رایگان دریافت می‌کردند، وضعیت ناهار رایگان ارزیابی خوبی از درآمد خانواده است. تفاوت در این ویژگی‌ها در این سه نوع کلاس کم است و هیچ‌کدام فاصله چندانی با صفر ندارند. این نشان می‌دهد که تخصیص تصادفی همان طور که مورد نظر بود، عمل کرد.

---

1. Finn and Achilles
2. Krueger
3. Treatment Group

جدول ۱-۲-۲. مقایسه مشخصات تیمار و کنترل در آزمایش تنسی استار

	Variable	Small	Regular	Regular/Aide	Joint P-value
1.	Free lunch	.47	.48	.50	.09
2.	White/Asian	.68	.67	.66	.26
3.	Age in 1985	5.44	5.43	5.42	.32
4.	Attrition rate	.49	.52	.53	.02
5.	Class size in kindergarten	15.10	22.40	22.80	.00
6.	Percentile score in kindergarten	54.70	48.90	50.00	.00

نکته: این جدول، از جدول ۱ کروگر (۱۹۹۹) اخذ شده است. این جدول میانگین متغیرها را توسط وضعیت تیمار[1] نشان می‌دهد. مقدار پی (P) در ستون آخر برای آزمون اف (F) برابری میانگین متغیرها در کل سه گروه است. تمام متغیرها بجز نرخ کاهش برای اولین سالی است که دانش‌آموزی تحت بررسی قرار گرفته است. متغیر ناهار رایگان برای بچه‌هایی است که ناهار رایگان دریافت می‌کنند. مقدار صدک، مقدار صدک میانگین در سه آزمون پیشرفت استنفورد[2] است. نرخ کاهش، دانش‌آموزانی هستند که تا پایان سال سوم، به طور کامل تحت آزمایش نبودند.

جدول ۱-۲-۲ همچنین اطلاعات کلاس‌های با اندازه متوسط، نرخ کاهش و نمره آزمون را نشان می‌دهد **که** در اینجا به شکل صدک اندازه‌گیری شده است. نرخ کاهش در کلاس‌های کودکستان کمتر بود. این نکته حداقل از لحاظ اصولی، یک مشکل بالقوه است.[3] در کلاس‌هایی که تعیین شده است تا کوچک باشند،[4] اندازه کلاس به طور قابل ملاحظه‌ای کمتر است، یعنی آزمایش در ایجاد تغییرات دلخواه موفق بوده است. اگر بسیاری از اولیای دانش‌آموزانی که در کلاس‌های عادی قرار داده شده بودند، مدیران و معلمان را مجبور می‌کردند که بچه‌هایشان را در کلاس‌های کوچک قرار دهند، شکاف در اندازه کلاس، بین گروه‌ها بسیار کمتر می‌شد.

به این دلیل که تصادفی‌سازی، اریبی گزینش را حذف می‌کند، تفاضل خروجی بین گروه‌های تیمار، میانگین اثر عِلّی[5] اندازه کلاس را به دست می‌آورد (که مربوط به کلاس‌های عادی با دستیار پاره‌وقت است). در عمل، تفاضل میانگین‌ها بین گروه تیمار و کنترل می‌تواند با رگرسیون نمره آزمون بر متغیرهای موهومی هر گروه تیمار به دست آید که در زیر آن را شرح می‌دهیم. تفاضل گروه کنترل از آزمایش برای بچه‌های کودکستان که در جدول ۲-۲-۲ آمده است (استخراج شده از کروگر، ۱۹۹۹، جدول ۵)، اثر کلاس را در حدود ۵ تا ۶ واحد درصد نشان می‌دهد. اندازه اثر[6] حدود ۲σ است که در آن σ انحراف استاندارد به واحد درصد در کودکستان است. اثر کلاس کوچک فاصله زیادی از صفر دارد، در حالی که اثر کلاس عادی/ دستیار قابل ملاحظه نیست.

1. Treatment Status
2. Stanford Achievement Test
۳. کروگر توجه زیادی به نرخ کاهش داشت. تفاضل نرخ‌های کاهش در میان گروه‌ها ممکن است باعث شود که نمونه‌ای از دانش‌آموزان به مقاطع بالاتر بروند در حالی که این نمونه به صورت تصادفی در انواع کلاس‌ها توزیع نشده‌اند. نتایجی از کودکستان که تحت تأثیر نرخ کاهش نبودند، قابل اعتمادتر بودند.
4. Assigned-to-be-small
5. Casual Effect
6. Effect Size

جدول ۲-۲-۲. برآورد تجربی اثر تخصیص برحسب اندازه کلاس بر نمره آزمون

Explanatory variable	(1)	(2)	(3)	(4)
Small class	4.82	5.37	5.36	5.37
	(2.19)	(1.26)	(1.21)	(1.19)
Regular/aide class	.12	.29	.53	.31
	(2.23)	(1.13)	(1.09)	(1.07)
White/Asian (1 = yes)	–	–	8.35	8.44
			(1.35)	(1.36)
Girl (1 = yes)	–	–	4.48	4.39
			(.63)	(.63)
Free lunch (1 = yes)	–	–	-13.15	-13.07
			(.77)	(.77)
White teacher	–	–	–	-.57
				(2.10)
Teacher experience	–	–	–	.26
				(.10)
Master's degree	–	–	–	-0.51
				(1.06)
School fixed effects	No	Yes	Yes	Yes
R^2	.01	.25	.31	.31

نکته: این جدول از جدول ۵ کروگر (۱۹۹۹) گرفته شده است. متغیر مستقل نمره آزمون پیشرفت استنفورد به واحد درصد است. خطای استاندارد استوار[1] که مانده‌های همبسته در کلاس‌ها را ممکن می‌سازد در پرانتزها نشان داده شده است. اندازه نمونه ۵۶۶۸۱ است.

پژوهش استار نمونه‌ای خوب از آزمایه تصادفی در تاریخچه علوم اجتماعی است. همین طور مشکلات لجستیکی، زمان طولانی و هزینه زیاد آزمایه‌های تصادفی را نشان می‌دهد. در بسیاری از موارد، چنین آزمایه‌هایی غیرعملی است.[2] در بقیه موارد، ما دوست داریم هرچه سریع‌تر پاسخی داشته باشیم.

بنابراین، ما تلاش می‌کنیم که از منابع تغییر آماده‌تر و آسان‌تر استفاده کنیم. ما امیدواریم که شبه‌آزمایش[3] یا آزمایش‌های طبیعی‌ای را بیابیم که عوامل دیگر را متعادل نگه دارد، از آزمایه تصادفی با تغییر متغیرهای مورد نظر تقلید کند. آیا می‌توانیم آزمایش طبیعی قانع‌کننده‌ای بیابیم؟ البته که نه، اما عقیده داریم که آزمایه تصادفی عمومی شاخص ماست. بسیاری از پژوهشگران، البته نه همه آنها، با ما هم‌عقیده هستند. ما این نظر را نخستین بار از استاد و مشاور پایان‌نامه‌مان، اورلی آشنفلتر،[4] که مدافع پیشگام طرح‌های پژوهشی آزمایشی و شبه‌آزمایشی در علوم اجتماعی بود، شنیدیم. در اینجا نظر آشفلر

1. Robust Standard Errors

۲. آزمایه‌های تصادفی هیچگاه بی‌نقص نبوده‌اند و استار هم از این قائده مستثنا نیست. دانش‌آموزانی بودند که یک سال تحصیلی را تکرار کردند یا جهشی رفتند و آن سال را نگذراندند. دانش‌آموزانی که به مدرسه یا سال تحصیلی تحت آزمایش اضافه شدند، به صورت تصادفی در یکی از کلاس‌ها قرار گرفتند. یکی از جنبه‌های بدیمن آزمایش این بود که دانش‌آموزانی که در کلاس‌های عادی و عادی/ دستیار بودند، شاید به دلیل اعتراض والدین به حضور بچه‌هایشان در کلاس عادی، بعد از کودکستان دوباره کلاس آنها عوض شد. همچنین فرایند تعویض بچه‌ها بعد از کودکستان نیز وجود داشت. علی‌رغم این مشکلات، آزمایش استار، آزمایه تصادفی بود که به خوبی اجرا شد. تحلیل کروگر (۱۹۹۹) نشان می‌دهد که هیچ‌کدام از این مشکلات، بر نتیجه اصلی این پژوهش اثرگذار نبود.

3. Quasi-experiment

4. Orley Ashenfelter

(۱۹۹۱) آمده است که اعتبار مطالعاتی را که تحصیل و درآمد را به هم مرتبط می‌کرد، ارزیابی می‌کند:

شواهدی که تحصیل و آموزش را به هم مرتبط می‌کنند تا چه میزان معتبرند؟ این پاسخ من است: بسیار متقاعدکننده. اگر باید شرط می‌بستم که آزمایش ایدئال چه چیزی را نشان می‌دهد، روی این شرط می‌بستم که آن آزمایش نشان می‌داد که کارگران با تحصیل بیشتر، درآمد بیشتری دارند.

پژوهش شبه‌آزمایشی[1] در مورد اندازه کلاس که توسط آنگریست و لاوی[2] انجام شد، نشان داد که داده‌های غیرآزمایشی می‌توانند به شکل داده‌های آزمایشی تحلیل شوند. پژوهش آنگریست و لاوی بر اساس این حقیقت است که اندازه کلاس‌ها به ۴۰ محدود شدند. بنابراین دانش‌آموزی در کلاس پنجم با ۴۰ دانش‌آموز به کلاس ۴۰ نفره می‌رود، در حالی که اگر در گروه کلاس پنجمی‌های ۴۱ نفره است، به کلاسی که نصف این اندازه است می‌رود، به این دلیل که گروه نصف می‌شود. از آنجایی که دانش‌آموزان در گروه ۴۰ نفره و ۴۱ نفره احتمال دارد که در ابعاد دیگر مثل توانایی یا پیشینه خانوادگی، شبیه هم باشند، ما می‌توانیم تفاضل بین ۴۰ و ۴۱ دانش‌آموز ثبت نام شده را به عنوان تخصیص تصادفی در نظر بگیریم.

پژوهش آنگریس ـ لاوی دانش‌آموزان را در مقاطع مختلف با ثبت نام بالاتر و پایین‌تر از حد مشخص اندازه کلاس مقایسه می‌کند تا بدون داشتن فرصتی برای انجام یک آزمایش واقعی، ارزیابی به خوبی کنترل‌شده‌ای از اثر تغییر سریع در اندازه کلاس انجام دهد. همانند آزمایش تنسی ـ استار، نتایج آنگریس و لاوی (۱۹۹۹) به ارتباط قوی بین اندازه کلاس و دستاورد دانش‌آموز اشاره می‌کند. همچنین آنگریس و لاوی بر اساس مقایسه‌ای ساده بین دانش‌آموزانی که در کلاس‌های کم‌جمعیت‌تر و پُرجمعیت ثبت نام کردند، گزارش دادند که، نتیجه بسیار با تحلیل‌های ناپخته متفاوت بود. این مقایسه‌ها نشان می‌دهد که دانش‌آموزان در کلاس‌های کم‌جمعیت‌تر در آزمون‌های استاندارد شده بدتر عمل می‌کنند. بنابراین، اریبیِ گزینش مثال بیمارستان در سؤالات پژوهش اندازه کلاس هم کاربرد دارد.[3]

۳-۲. تحلیل رگرسیون آزمایش‌ها[4]

رگرسیون، ابزاری کاربردی برای مطالعه سؤالات عِلّی از جمله تحلیل داده نتایج آزمایش‌هاست. اکنون فرض کنید که اثر تیمار برای همه برابر است و $Y_{1i} - Y_{0i} = \rho$، ثابت است. با اثر تیمار ثابت می‌توانیم معادله ۱-۱-۲ را دوباره بنویسیم.

1. Quasi-experimental Study
2. Angrist and Lavy
۳. نتایج آنگریست ـ لاوی (۱۹۹۹) در بخش ۶ به عنوان نمایش طرح پژوهشی رگرسیون ـ ناپیوستگی شبه‌آزمایشی دوباره نشان داده می‌شود.
4. Regression Analysis of Experiments

$$\begin{array}{ccccccc} Y_i = & \alpha & + & \rho & D_i & + & \eta_i, \\ & \shortparallel & & \shortparallel & & & \shortparallel \\ & E(Y_{0i}) & & (Y_{1i}-Y_{0i}) & & & Y_{0i}-E(Y_{0i}) \end{array} \quad (2.3.1)$$

η_i قسمت تصادفی Y_{0i} است. که امید ریاضی شرطی این عبارت با وضعیت تیمار که فعال و غیرفعال می‌شود، **عبارت** زیر را می‌دهد:

$$\begin{aligned} E[Y_i|D_i=1] &= \alpha+\rho+E[\eta_i|D_i=1] \\ E[Y_i|D_i=0] &= \alpha+E[\eta_i|D_i=0], \end{aligned}$$

بنابراین:

$$E[Y_i|D_i=1]-E[Y_i|D_i=0] = \underbrace{\rho}_{\text{اثر تیمار}} + \underbrace{E[\eta_i|D_i=1]-E[\eta_i|D_i=0]}_{\text{اریب گزینش}}.$$

بنابراین اریبیِ گزینش با مجموع همبستگی بین جمله خطای رگرسیون، یعنی η_i، و رگرسور، یعنی D_i برابر است. از آنجایی که:

$$E[\eta_i|D_i=1]-E[\eta_i|D_i=0]=E[Y_{0i}|D_i=1]-E[Y_{0i}|D_i=0],$$

این همبستگی، تفاضل خروجی بالقوه (غیرتیمار) را بین آنهایی که تحت تیمار قرار گرفتند و آنهایی که قرار نگرفتند نشان می‌دهد. در مثال بیمارستان، آنهایی که درمان شدند، در شرایط غیردرمانی خروجی سلامت ضعیف‌تری داشتند، در حالی که در پژوهش آنگریست و لاوی (۱۹۹۹)، دانش‌آموزان در کلاس‌های کم‌جمعیت‌تر در اصل نمره آزمونشان کمتر بود.

در آزمایش استار، که در آن D_i به صورت تصادفی معین شد، جمله گزینش[۱] وجود ندارد و رگرسیون Y_iبر D_i. اثر علّی مورد نظر، یعنی ρ را ارزیابی می‌کند. پسماند جدول ۲-۲-۲ مشخصات رگرسیون[۲] مختلف را نشان می‌دهد، که بعضی از آنها علاوه بر شاخص تخصیص تصادفی،[۳] یعنی D_i، شامل متغیرهای کمکی[۴] است. متغیرهای کمکی در تحلیل رگرسیون داده‌های آزمایشی دو نقش دارند. ابتدا، طرح آزمایشی استار از تخصیص تصادفی شرطی[۵] استفاده کرد. به ویژه اینکه، تخصیص به کلاس‌های مختلف در مدارس تصادفی بود. دانش‌آموزانی که در مدارس مختلف شرکت می‌کردند (مثلاً

1. Selection Term
2. Regression Specifications
3. Random Assignment Indicator
4. Covariate
5. Conditional Random Assignment

شهری، در برابر روستایی) کم و بیش احتمال داشت که در کلاس‌های کوچک‌تر قرار داده شوند. مقایسه در ستون ۱ جدول ۲-۲-۲ که، تعدیلی[1] برای آن ایجاد نکرد، ممکن است به دلیل وجود نتایج مختلف از مدارس مختلف باشد. برای تعدیل نتایج، چندین مدل رگرسیون کروگر شامل اثر ثابت مدرسه می‌شود، به عبارت دیگر، شامل عرض از مبدأ[2]های جداگانه برای هر مدرسه در داده‌های استار می‌شود. در عمل، نتیجه تعدیل برای اثر ثابت مدرسه نسبتاً کم است، اما ما بدون مشاهده، متوجه نمی‌شدیم. درباره مدل رگرسیون با اثر ثابت در بخش ۵ خواهیم گفت.

کنترل‌های دیگر در جدول کروگر مشخصات دانش‌آموزان مثل نژاد، سن و وضعیت ناهار رایگان را توصیف می‌کند. ما سابقاً مشاهده کردیم که این ویژگی‌ها در انواع کلاس‌ها به تعادل رسیدند، به عبارت دیگر، آنها به طور سیستماتیک به تخصیص دانش‌آموزان بر حسب اندازه کلاس مرتبط نیستند. اگر این کنترل‌ها، یا X_i با تیمار D_i همبستگی نداشته باشد، بر تخمین ρ اثر نمی‌گذارد. به عبارت دیگر، برآورد $\boldsymbol{\rho}$ در رگرسیون بلند (طولانی)[3]

$$Y_i = \alpha + \rho D_i + X_i'\gamma + \eta_i \qquad (2.3.2)$$

نزدیک به برآورد ρ در رگرسیون کوتاه[4] می‌شود (۲-۳-۱). این موضوعی است که در بخش ۳ به آن می‌پردازیم. با وجود این، شمول متغیرهای X_i، برآوردهای دقیق‌تری از اثر علیت مورد نظر تولید می‌کند. توجه کنید که خطای استاندارد اثر تیمار برآورد شده در ستون ۳ کوچک‌تر از خطای استاندارد متناظر آن در ستون ۲ است و اگرچه متغیرهای کنترل، یعنی X_i، با D_i مطابق نیست، اما آنها قدرت توضیح‌دهندگی زیادی برای Y_i دارند. بنابراین، شمول این متغیرهای کنترل، واریانس پسماند[5] را کاهش می‌دهد، که به نوبت خطای استانداردِ برآوردِ رگرسیون را کاهش می‌دهد. به طور مشابه، خطای مشابه برآورد ρ با شمول اثر ثابت مدرسه کاهش می‌یابد، زیرا اینها نیز بخش مهمی از واریانس عملکرد دانش‌آموز را شرح می‌دهد. ستون آخر ویژگی‌های معلم‌ها را اضافه می‌کند. به این دلیل که معلم‌ها به صورت تصادفی در کلاس‌ها قرار گرفتند و در این داده‌ها، ویژگی معلم‌ها رابطه کمی با دستاورد دانش‌آموز دارند، اثر برآورد شده کلاس‌های کوچک و خطاهای استانداردش با اضافه شدن متغیرهای معلم‌ها ثابت می‌ماند.

رگرسیون نقش مهمی در پژوهش‌های اقتصادی تجربی بازی می‌کند. بعضی از رگرسیون‌ها، همانند بیشتر پژوهش‌ها درباره نابرابری حقوق، ابزار توصیفی ساده‌ای هستند. همان طور که ما در این بخش دیدیم، رگرسیون مناسب تجزیه و تحلیل داده‌های تجربی است. در بعضی از موارد، رگرسیون می‌تواند

1. Adjustment
2. Intercept
3. Long Regression
4. Short Regression
5. Residual Variance

در غیاب تخصیص تصادفی برای تقریب زدن آزمایش‌ها استفاده شود، اما پیش از اینکه به این سؤال مهم بپردازیم که چه زمانی رگرسیون احتمال دارد تفسیر عِلّی داشته باشد، مرور تعدادی از نکات و ویژگی‌های اساسی رگرسیون سودمند است. بدون توجه به هدف استفاده از رگرسیون، این نکات و ویژگی‌ها به طور قابل اعتمادی برای هر رگرسیون درست است.

ب) هسته

۳. معنادار کردن رگرسیون

«بیایید به چیزی که غیرقابل فکر کردن است فکر کنیم، بیایید کاری را که غیرقابل انجام دادن است انجام دهیم.

بیایید آماده باشیم تا با چیزهای غیرقابل بیان درگیر شویم،

و بفهمیم که شاید در نهایت نمی‌توانیم آنها را بیان کنیم».

داگلاس آدامز،[۱] آژانس کارآگاهی جامع درک جنتلی[۲] (سال ۱۹۹۰)

آنگریست[۳] چنین به یاد می‌آورد:

من اولین رگرسیونم را در تابستان سال ۱۹۷۹ به عنوان دانشجوی کالج ابرلین[۴] و بین سال اول و دوم تحصیلیم اجرا کردم. در آن زمان به عنوان دستیار تحقیق برای آلن ملتزر[۵] و اسکات ریچارد،[۶] اعضای هیئت علمی دانشگاه کارنگی ملون[۷] که نزدیک خانه‌ام در پیتزبورگ[۸] بود کار می‌کردم. بیشتر علاقه‌مند به شغلی در زمینه آموزش افراد با شرایط خاص بودم و قصد داشتم که به عنوان کارمند برای کار به بیمارستان روانی دولتی که کار تابستانی قبلی‌ام بود بازگردم، اما کلاس Econ 101 (اقتصاد مقدماتی) مرا به فکر فرو برد و همچنین می‌توانستم ببینم که با همان نرخ درآمد، ساعت و شرایط کاری دستیار تحقیق بهتر از کارمند بیمارستان است. وظایف من به عنوان دستیار تحقیق شامل جمع‌آوری داده‌ها و تجزیه و تحلیل رگرسیون بود، هرچند که من در آن زمان رگرسیون یا حتی آمار را درک نمی‌کردم.

مقاله‌ای که در آن تابستان روی آن کار می‌کردم (ملتزر و ریچارد، سال ۱۹۸۳) تلاشی برای ارتباط دادن اندازه دولت‌ها در نظام‌های مردم‌سالار و نابرابری درآمد است که با اندازه‌گیری هزینه‌های دولت نسبت به تولید ناخالص داخلی (GDP) سنجیده می‌شود. بیشترین توزیع‌های درآمد دارای دنباله راست

1. Douglas Adams
2. Dirk Gently's Holistic Detective Agency
3. Angrist
4. Oberlin College
5. Allan Meltzer
6. Allan Meltzer
7. Carnegie-mellon University
8. Pittsburgh

طولانی هستند به این معناست که به درآمد متوسط گرایش دارد که بسیار بالاتر از میانه قرار گیرد. با رشد نابرابری، بیشتر رأی‌دهندگان، درآمدی زیر میانگین پیدا می‌کنند. کسانی که درآمدی بین میانه و میانگین دارند، ممکن است به علت خشم نسبت به نابرابری درآمد، برای رأی دادن به سیاست‌های مالی که ـ مانند رابین هود ـ از ثروتمندان می‌گیرند و به فقرا می‌دهند، به کسانی که درآمد پایین‌تر از میانه دارند بپیوندند و در نتیجه حجم دولت افزایش می‌یابد.

من تئوری پایه پشت پروژه ملتزر و ریچارد[1] را فهمیدم، هرچند که آن را کاملاً باور نکردم، زیرا احتمال به دست آمدن نتیجه مورد نظر رأی‌دهندگان فقیر کم است. همچنین به یاد دارم که با آلن ملتزر در مورد اینکه آیا هزینه‌های دولت برای آموزش باید در دسته کالا و خدمات عمومی (چیزی که علاوه بر افرادی که مستقیماً تحت تأثیر قرار می‌دهد، به نفع همه است) یا کالا و خدمات خصوصی که به عموم عرضه می‌شود و در نتیجه شکلی از توزیع مجدد مانند سیاست‌های رفاه است، تقسیم‌بندی شود بحث می‌کردم. ممکن است بگویید که این پروژه نقطه شروع علاقه من به رویکرد بازدهی اجتماعی به آموزش بود، موضوعی که من با اشتیاق و درک بیشتر در عاصم‌اوغلو[2] و آنگریست (۲۰۰۰) به آن بازگشتم.

امروزه من مطالعه ملتزر و ریچارد (سال ۱۹۸۳) را تلاشی برای استفاده از رگرسیون برای کشف و اندازه‌گیری یک رابطه علّی جالب، می‌دانم. هرچند در آن زمان من فقط یک استفاده‌کننده از رگرسیون بودم. گاهی اوقات تحلیل رگرسیون را ناامیدکننده می‌دانستم. روزها می‌گذشت که من با کسی غیر از کارفرمایانم و گاهی دانشجویان دکترای کارنگی ملون که اکثر آنها هم خیلی کم انگلیسی صحبت می‌کردند صحبت نمی‌کردم. بهترین بخش کار، ناهار با آلن ملتزر بود، محقق برجسته و سرپرستی صبور و خوش‌قلب که از گفتگو در مدتی که محتوای کیسه‌های قهوه‌ای را می‌خوردیم خوشحال بود (این مدت زیاد طول نمی‌کشید چون آلن کم می‌خورد و من سریع می‌خوردم). به یاد می‌آورم که از آلن پرسیدم که آیا از اینکه وقت خود را صرف بررسی خروجی رگرسیونی می‌کند که پس از آن بر روی مقادیر زیادی از کاغذ نواری سبز عریض چاپ می‌شود رضایت دارد. ملتزر خندید و گفت که هیچ کار دیگری را به این ترجیح نمی‌دهد.

در حال حاضر، ما نیز روزهای خود (حداقل بهترین روزهایمان) را مانند اساتید و مشاورانمان در دانشگاه با خوشحالی به بررسی خروجی رگرسیون می‌گذرانیم. این فصل توضیح می‌دهد که علت آن چیست.

۱-۳. مبانی رگرسیون

پایان فصل قبلی، مدل‌های رگرسیونی را به عنوان ابزار محاسباتی برای برآورد تفاوت‌های تیمار ـ کنترل

1. Meltzer and Richards
2. Acemoglu

در یک آزمایش با و بدون متغیرهای کمکی معرفی کرد. از آنجا که رگرسور مورد نظر در مطالعه اندازه کلاس که در بخش ۳-۲ مورد بحث قرار گرفت به طور تصادفی تخصیص داده شده است، برآوردهای نتیجه شده از آن تفسیر علیتی دارند. با این حال در اغلب موارد، رگرسیون با داده‌های مشاهده شده[1] مورد استفاده قرار می‌گیرد. بدون استفاده از تخصیص تصادفی، برآورد رگرسیون ممکن است یا تفسیر علیتی داشته باشد یا نداشته باشد. بعداً در این فصل به این سؤال اصلی که چه چیزی باعث علّی شدن رگرسیون می‌شود باز خواهیم گشت.

برای لحظاتی، این بخش را با کنار گذاشتن مسئله نسبتاً انتزاعی علیت، با خواص مکانیکی برآوردهای رگرسیون آغاز می‌کنیم. اینها ویژگی‌های کلی بردار رگرسیون جامعه و قیاس نمونه‌ای آن است که هیچ ارتباطی با تفسیر محققان از خروجی آن ندارد. این فصل با بررسی این ویژگی‌ها آغاز می‌شود که شامل موارد زیر است:

۱. ارتباط نزدیک بین تابع رگرسیون جامعه[2] و تابع امید ریاضی شرطی،[3]

۲. ضرایب رگرسیون با اضافه یا حذف کردن متغیرهای کمکی چگونه و چرا تغییر می‌کنند،

۳. ارتباط نزدیک بین رگرسیون و سایر «استراتژی‌های کنترل» مانند جورسازی،[4]

۴. توزیع نمونه‌گیری برآوردهای رگرسیون.

۳-۱-۱. روابط اقتصادی و تابع امید ریاضی شرطی

تحقیقات تجربی اقتصادی ما در زمینه اقتصاد کار، به طور معمول مربوط به تحلیل آماری شرایط اقتصادی فردی و به ویژه تفاوت‌های بین افراد است که ممکن است علت تفاوت‌های موجود در ثروت اقتصادی آنها باشد. چنین تفاوت‌هایی در ثروت اقتصادی بسیار سخت توضیح داده می‌شود؛ در یک کلمه می‌شود گفت آنها تصادفی هستند. با این حال، همانند اقتصادسنجی کارهای کاربردی ما هم اعتقاد داریم که می‌توانیم تصادفی بودن‌ها را به صورت مفید خلاصه و تفسیر کنیم. مثالی از «تصادفی بودن سیستماتیک» که در مقدمه ذکر شد ارتباط بین آموزش و درآمد است. به طور متوسط، افرادی که تحصیلات بیشتری دارند، درآمد بیشتری نسبت به افرادی که تحصیلات کمتری دارند کسب می‌کنند. ارتباط بین تحصیلات و متوسط درآمد، با وجود تنوع بسیار زیاد در شرایط افراد که گاهی اوقات رابطه را تحت تأثیر قرار می‌دهد، قدرت پیش‌بینی قابل توجهی دارد. البته، این واقعیت که افراد تحصیلکرده درآمد بیشتری نسبت به افراد کمتر تحصیلکرده کسب می‌کنند به این معنا نیست که تحصیل *باعث* افزایش درآمد می‌شود. این سؤال که آیا رابطه درآمد ـ تحصیلات علّی است، اهمیت زیادی دارد و ما چندین بار به آن باز خواهیم گشت. با وجود این، حتی بدون حل سؤال دشوار علیت، واضح است که آموزش، درآمد را در جهت آماری

1. Observational Data
2. Population Regression Function (PRF)
3. Conditional Expectation Function (CEF)
4. Matching

مشخصی پیش‌بینی می‌کند. این قدرت پیش‌بینی به صورت قانع‌کننده‌ای توسط تابع امید ریاضی شرطی (**CEF**) خلاصه می‌شود.

CEF برای یک متغیر وابسته، Y_i که یک بردار $K \times 1$ از متغیرهای کمکی است، X_i (با عناصر X_{ki}) امید ریاضی، یا میانگین جامعه آماری Y_i با ثابت نگه داشتن X_i است. میانگین جامعه آماری را می‌توان به عنوان میانگین در یک نمونه بزرگ بی‌نهایت یا میانگین در یک جامعه آماری متناهی که به طور کامل شمارش شده است در نظر گرفت. CEF به صورت $E[Y_i|X_i]$ نوشته می‌شود و تابعی از X_i است. از آنجا که X_i تصادفی است، CEF نیز تصادفی است، هرچند گاهی اوقات با یک مقدار خاص از CEF کار می‌کنیم و می‌گوییم $E[Y_i|X_i = 42]$ و فرض می‌کنیم که ۴۲ یک مقدار ممکن برای X_i است. در فصل ۲ برای مدت کوتاهی CEF را $E[Y_i|D_i]$ در نظر گرفتیم که در آن $\mathbf{D_i}$ یک متغیر صفر ـ یک است. CEF دو مقدار $E[Y_i|D_i = 1]$ و $E[Y_i|D_i = 0]$ را می‌گیرد. اگرچه این مورد خاص نیز مهم است، اما اغلب علاقمند به CEFهایی هستیم که توابعی از متغیرهای بسیار هستند که به راحتی در بردار X_i وارد می‌شوند. برای یک مقدار خاص X_i، مثلاً $X_i = x$ می‌نویسیم $E[Y_i|X_i = x]$. برای Y_i پیوسته با چگالی شرطی $f_y(\cdot\,|X_i = x)$، CEF به صورت زیر است:

$$E[Y_i|X_i = x] = \int t f_y(t|X_i = x)dt.$$

اگر Y_i گسسته باشد، $E[Y_i|X_i = x]$ برابر با مجموع $\sum_t t f_y(t|X_i = x)$ می‌شود.

امید ریاضی یک مفهوم مرتبط با جامعه آماری است. در عمل، داده‌ها معمولاً به شکل نمونه می‌آیند و به ندرت شامل کل جامعه آماری می‌شوند. بنابراین ما از نمونه‌ها برای استنباط در مورد جامعه آماری استفاده می‌کنیم. برای مثال، **CEF** نمونه برای یادگیری در مورد **CEF** جامعه آماری استفاده می‌شود. اگرچه این مسئله همیشه ضروری است، اما بحث در مورد گام رسمی استنباط که ما را از نمونه به جامعه آماری می‌رساند تا بخش ۳-۱-۳ به تعویق می‌اندازیم. رویکرد «اول، جامعه آماری»[1] به اقتصادسنجی نتیجه این واقعیت است که ما باید اهداف مورد نظر را قبل از این که بتوانیم از داده‌ها برای مطالعه آنها استفاده کنیم، تعریف کنیم.[2]

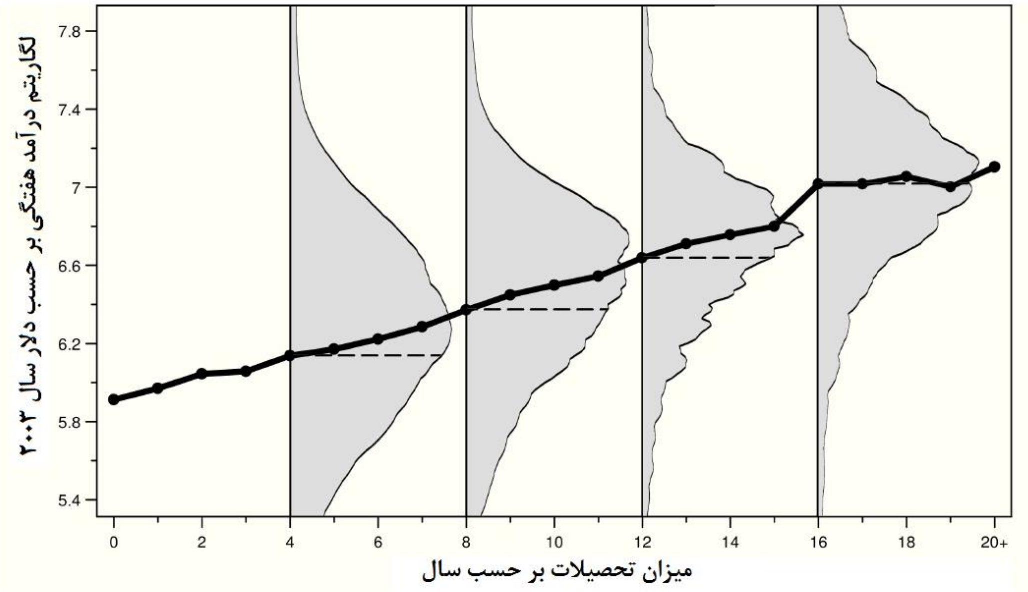
شکل ۱-۱-۳، نمودار CEF لگاریتم درآمد هفتگی با توجه به تحصیلات را برای نمونه‌ای از مردان سفیدپوست میانسال در سرشماری سال ۱۹۸۰ نشان می‌دهد. نمودار توزیع درآمد نیز برای چندین مقدار کلیدی مشخص شده است: ۴، ۸، ۱۲ و ۱۶ سال تحصیل. CEF در این تصویر این واقعیت را نشان می‌دهد که ـ علی‌رغم تفاوت بسیار در شرایط فردی ـ افرادی که تحصیلات بیشتری دارند معمولاً به

1. Population First Approach

۲. نمونه‌هایی از نوشتار مربوط به آموزش با استفاده از رویکرد «اول، جامعه آماری» در اقتصادسنجی شامل کارهای چمبرلین (Chamberlain) (۱۹۸۴)، گلدبرگر (Goldberger) (۱۹۹۱) و مانسکی (Manski) (سال ۱۹۹۱) است.

طور متوسط درآمد بیشتری کسب می‌کنند. متوسط درآمد ناشی از یک سال تحصیل در حدود ۱۰ درصد است.

شکل ۱-۱-۳. داده‌های خام و CEF لگاریتم میانگین درآمد هفتگی با توجه به تحصیلات



نمونه شامل مردان سفیدپوست ۴۰ تا ۴۹ساله در فایل نمونه ۵ درصد از سرشماری‌های موجود در پایگاه IPUMS سال ۱۹۸۰ است.

یک مکمل مهم در CEF قانون امید ریاضی مکرر[1] است. این قانون می‌گوید که امید ریاضی غیرشرطی را می‌توان به عنوان میانگین جامعه آماری CEF نوشت. به عبارت دیگر:

$$E[Y_i] = E\{E[Y_i|X_i]\}, \qquad (3.1.1)$$

که در آن امید ریاضی بیرونی از توزیع X_i استفاده می‌کند. در اینجا اثبات قانون امید ریاضی مکرر برای توزیع مستمر (X_i, Y_i) با چگالی توأم $f_{xy}(u,t)$، ارائه می‌شود که در آن $f_y(t|X_i = x)$ توزیع شرطی Y_i با $X_i = x$ است و $g_y(t)$ و $g_x(u)$ چگالی حاشیه‌ای هستند:

$$E\{E[Y_i|X_i]\} = \int E[Y_i|X_i = u] g_x(u) du$$

$$= \int \left[\int t f_y(t|X_i = u) dt\right] g_x(u) du$$

1. Law of Iterated Expectations.

$$= \int \int t f_y(t|X_i = u)\, g_x(u) du dt$$

$$= \int t \left[\iint f_y(t|X_i = u) g_x(u) du \right] dt = \int t \left[\iint f_{xy}(u,t) du \right] dt$$

$$= \int t g_y(t) dt.$$

انتگرال‌های این مشتق بیش از مقادیر ممکن X_i و Y_i (مشخص شده با u و t) هستند. ما این مراحل را مشخص کرده‌ایم، زیرا CEF و خواص آن در ادامه این فصل نقش محوری دارند.

قدرت قانون امید ریاضی مکرر ناشی از شکسته شدن متغیر تصادفی به دو بخش توسط آن است.

قضیه ۱-۱-۳. ویژگی تجزیه CEF[1]

$$Y_i = E[Y_i|X_i] + \varepsilon_i$$

که در آن (**i**) ε_i مستقل از میانگین X_i است، به عنوان مثال، $E\left[\varepsilon_{i_i}|X_i\right] = 0$؛ بنابراین، (**ii**) ε_i با هیچ تابعی از X_i همبستگی ندارد.

اثبات:

(**i**) $E[\varepsilon_i|X_i] = E[Y_i - E[Y_i|X_i]|X_i] = E[Y_i|X_i] - E[Y_i|X_i] = 0$؛

(**ii**) از (**i**) حاصل می‌شود: اگر $h(X_i)$ هر تابعی از X_i باشد. با قانون امید ریاضی مکرر، $E[h(X_i)\varepsilon_i] = E\{h(X_i)E[\varepsilon_i|X_i]\}$ و با استقلال از میانگین، $E[\varepsilon_i|X_i] = 0$. ■

این قضیه می‌گوید هر متغیر تصادفی Y_i را می‌توان به قطعه‌ای که «توسط X_i توضیح داده می‌شود» تجزیه کرد، یعنی **CEF**، و قطعه پسماند که متعامد (یعنی ناهمبسته) با هر تابعی از X_i است.

CEF به دلایل مختلف خلاصه خوبی از رابطه بین Y_i و X_i است. اول اینکه ما به در نظر گرفتن میانگین به عنوان ارائه‌دهنده مقدار نمایندگی‌کننده یک متغیر تصادفی عادت داریم. به طور رسمی‌تر، CEF بهترین پیش‌بینی‌کننده Y_i با در نظر گرفتن X_i به صورتی است که یک مسئله پیش‌بینی حداقل میانگین مربعات خطا (MMSE) را حل کند. این ویژگی پیش‌بینی CEF، نتیجه‌ای از ویژگی تجزیه CEF است:

قضیه ۲-۱-۳. ویژگی پیش‌بینی CEF[2]

اگر $m(X_i)$ هر تابعی از X_i باشد. **CEF** رابطه زیر را حل می‌کند:

1. CEF-decomposition Property
2. The CEF-prediction Property

$$E[Y_i|X_i] = arg \min_{m(X_i)} E[(Y_i - m(X_i))^2]$$

بنابراین، این عبارت پیش‌بینی‌کننده **MMSE** از Y_i با در نظر گرفتن X_i است.

اثبات. می‌نویسیم:

$$(Y_i - m(X_i))^2 = \left((Y_i - E[Y_i|X_i]) + (E[Y_i|X_i] - m(X_i))\right)^2$$
$$= (Y_i - E[Y_i|X_i])^2 + 2(E[Y_i|X_i] - m(X_i))(Y_i - E[Y_i|X_i])$$
$$+(E[Y_i|X_i] - m(X_i))^2$$

عبارت اول مهم نیست، زیرا شامل $m(X_i)$ نیست. عبارت دوم را می‌توان به صورت $h(X_i)\varepsilon_i$ نوشت که در آن $h(X_i) \equiv 2(E[Y_i|X_i] - m(X_i))$ است و در نتیجه در ویژگی تجزیه **CEF** امید ریاضی صفر دارد. عبارت آخر در صفر حداقل می‌شود که در آن **CEF** برابر با $m(X_i)$ است. ■

یک ویژگی نهایی از **CEF** که وابستگی نزدیک به هر دو ویژگی تجزیه و پیش‌بینی **CEF** دارد، قضیه آنالیز واریانس (**ANOVA**) است:

قضیه ۳-۱-۳. قضیه ANOVA

$$V(Y_i) = V(E[Y_i|X_i]) + E[V(Y_i|X_i)]$$

که در آن $V(\cdot)$ نشان‌دهنده واریانس است و $V(Y_i|X_i)$ واریانس شرطی Y_i با در نظر گرفتن X_i است.

اثبات. ویژگی تجزیه **CEF** نشان می‌دهد که واریانس Y_i، واریانس **CEF** به علاوه واریانس پسماند است، $\varepsilon_i \equiv Y_i - E[Y_i|X_i]$ زیرا ε_i و $E[Y_i|X_i]$ همبستگی ندارند. واریانس ε_i به صورت زیر است:

$$E\left[\varepsilon_i^2\right] = E\left[E\left[\varepsilon_i^2|X_i\right]\right] = E\left[V[Y_i|X_i]\right]$$

که در آن $E\left[\varepsilon_i^2|X_i\right] = V[Y_i|X_i]$ است زیرا $\varepsilon_i \equiv Y_i - E[Y_i|X_i]$. ■

دو ویژگی **CEF** و قضیه **ANOVA** ممکن است یک حلقه آشنا داشته باشند. برای مثال ممکن است شما عادت داشته باشید که یک جدول **ANOVA** در خروجی رگرسیون خود ببینید. **ANOVA** در پژوهش‌های حوزه نابرابری نیز مهم است که در آن اقتصاددانان کار، تغییرات توزیع درآمد را به بخش‌هایی تجزیه می‌کنند که می‌تواند توسط تغییرات ویژگی‌های کارگر و تغییرات آنچه پس از محاسبه این عوامل باقی مانده است، محاسبه شود (برای مثال، **Autor**، **Katz** و **Kearney**، ۲۰۰۵ را ببینید). آنچه ممکن است ناآشنا باشد، این واقعیت است که ویژگی‌های **CEF** و تحلیل واریانس **ANOVA**

در جامعه آماری نیز، علاوه بر نمونه‌ها، عمل می‌کند و فرض **CEF** خطی را رد نمی‌کند. در واقع، اعتبار رگرسیون خطی به عنوان یک ابزار تجربی هم خطی بودن را رد نمی‌کند.

۲-۱-۳. رگرسیون خطی و CEF

رگرسیونی که می‌خواهید اجرا کنید چیست؟

در دنیای ما، این سؤال یا مشابه آن تقریباً هر روز شنیده می‌شود. برآورد رگرسیون تقریباً برای تمام تحقیقات تجربی مبنای ارزشمندی فراهم می‌کند، زیرا رگرسیون به صورت تنگاتنگی با **CEF** ارتباط دارد و **CEF** خلاصه‌ای طبیعی از روابط تجربی ارائه می‌دهد. ارتباط بین توابع رگرسیون ـ به عبارت دیگر بهترین خط برازش شده ناشی از به حداقل رساندن مربعات خطای مورد انتظار ـ و **CEF** را می‌توان حداقل با سه روش توضیح داد. برای تدوین دقیق این توضیحات، بهتر است در مورد تابع رگرسیونی که در ذهن داریم دقیق باشیم. این فصل مربوط به بردار ضرایب رگرسیون *جامعه آماری* است، که به عنوان راه‌حلی برای مسئله کمترین مربعات جامعه آماری تعریف شده است. ما در این مرحله نگران علیت نیستیم. در عوض اجازه می‌دهیم که β بردار ضریب رگرسیون $K \times 1$ با حل کردن رابطه زیر تعریف شود:

$$\beta = \arg\min_{b} E\left[\left(Y_i - X_i'b\right)^2\right]. \qquad (3.1.2)$$

با استفاده از شرط مرتبه اول داریم:

$$E[X_i(Y_i - X_i'b)] = 0$$

راه‌حل برای **b** می‌تواند به این صورت نوشته شود $\beta = E[X_iX_i']^{-1}E[X_iY_i]$. توجه داشته باشید که با اجرا کردن، $E[X_i(Y_i - X_i'\beta)] = 0$ می‌شود. به عبارت دیگر، باقی‌مانده جامعه آماری که آن را به عنوان $Y_i - X_i'\beta = e_i$ تعریف می‌کنیم، با رگرسور X_i همبستگی ندارد. این امر تأکید دارد که این عبارت خطا به تنهایی موجودیت ندارد. موجودیت و به معنی خود را به β وابسته است.

در موارد ساده دومتغیره که در آن بردار رگرسیون فقط شامل رگرسیون تک x_i و یک ثابت است، ضریب شیب $\beta_1 = \frac{Cov(Y_i,x_i)}{V(x_i)}$ و عرض از مبدأ $\alpha = E[Y_i] - \beta_1 E[X_i]$ است. در موارد چندمتغیره که بیش از یک رگرسور غیرثابت دارد، ضریب شیب برای رگرسور **k-th** در زیر آمده است:

آناتومی رگرسیون

$$\beta_k = \frac{Cov\left(Y_i, \tilde{x}_{ki}\right)}{V\left(\tilde{x}_{ki}\right)}, \qquad (3.1.3)$$

که در آن $\tilde{x}_{ki}$ پسماند رگرسیون x_{ki} در تمام متغیرهای کمکی دیگر است.

به عبارت دیگر، $E[X_i X_i']^{-1} E[X_i Y_i]$ بردار $k \times 1$ با عنصر **k-th** $\frac{Cov(Y_i, \tilde{x}_{ki})}{V(\tilde{x}_{ki})}$ است. گفته می‌شود که این فرمول مهم توصیف «آناتومی ضریب رگرسیون چندمتغیره» است، زیرا موارد بیشتری از فرمول ماتریسی $\boldsymbol{\beta} = E[X_i X_i']^{-1} E[X_i Y_i]$ را آشکار می‌کند. این فرمول نشان می‌دهد که هر ضریب در رگرسیون چندمتغیره، ضریب شیب دومتغیره برای رگرسور مربوطه پس از «تقسیم کردن» تمام متغیرهای دیگر در مدل است.

برای تأیید فرمول آناتومی رگرسیون، رابطه زیر را رابطه (۳-۱-۳) جایگزین کنید:

$$\boldsymbol{Y_i = \beta_0 + \beta_1 x_{1i} + \cdots + \beta_k x_{ki} + \cdots + \beta_k x_{ki} + e_i}$$

از آنجا که $\tilde{x}_{ki}$ یک ترکیب خطی از رگرسورها است، با e_i همبستگی ندارد. همچنین از آنجا که $\tilde{x}_{ki}$ یک پسماند از رگرسیون بر روی همه متغیرهای کمکی دیگر در مدل است، باید با این متغیرهای کمکی همبستگی نداشته باشد. در نهایت، به همین دلیل، کوواریانس $\tilde{x}_{ki}$ با x_{ki} فقط واریانس $\tilde{x}_{ki}$ است. بنابراین $Cov(Y_i, \tilde{x}_{ki}) = \beta_k V(\tilde{x}_{ki})$ می‌شود.[1]

فرمول آناتومی رگرسیون احتمالاً به علت گذراندن دوره رگرسیون یا آمار برای شما آشناست، البته شاید با یک پیچیدگی: ضرایب رگرسیون تعریف شده در این بخش برآوردگر[2] نیستند، بلکه اجزایی غیرتصادفی از توزیع توأم[3] متغیرهای وابسته و مستقل هستند. توزیع توأم چیزی است که در صورت شمارش کامل جامعه آماری مورد نظر (یا دانستن فرایند تصادفی تولید داده‌ها) می‌توانستید مشاهده کنید. شما احتمالاً چنین اطلاعاتی ندارید. با این حال، مجاز ـ حتی مطلوب ـ است که در مورد میانگین ممکن مجموعه‌ای از پارامترهای جامعه آماری بدون نگرانی ابتدایی در مورد چگونگی برآورد آنها فکر کنیم.

در زیر در مورد سه دلیلی که چرا بردار ضرایب رگرسیون جامعه آماری ممکن است مورد توجه قرار گیرد بحث خواهیم کرد. این دلایل را می‌توان به این صورت خلاصه کرد که اگر علاقمند به CEF باشید، به پارامترهای رگرسیون هم علاقمند هستید.

---

۱. فرمول آناتومي رگرسیون معمولاً به فریش (Frisch) و وو (Waugh) (۱۹۳۳) نسبت داده می‌شود. همچنین می‌توانید آناتومی رگرسیون را این گونه انجام دهید:

$$\beta_k = \frac{Cov(\tilde{Y}_{ki}, \tilde{x}_{ki})}{V(\tilde{x}_{ki})},$$

که در آن $\tilde{Y}_{ki}$ باقی‌مانده رگرسیون Y_i در هر متغییر کمکی بجز x_{ki} است. این رابطه به این دلیل درست است که مقادیر متناسب که از $\tilde{Y}_{ki}$ حذف شده است با $\tilde{x}_{ki}$ همبستگی ندارد. اغلب مفید است که $\tilde{Y}_{ki}$ را در نمودار در مقابل $\tilde{x}_{ki}$ قرار دهیم و در این نمودار، حتی اگر نمودار دو بُعدی باشد، شیب حداقل مربعات متناسب با این نمودار پراکنش، برآورد شما از β_k چندمتغیره است. توجه داشته باشید که با این حال کافی نیست که سایر متغیرهای کمکی را فقط از Y_i جدا کنیم. به این معنا که:

$$\frac{Cov(\tilde{Y}_{ki}, x_{ki})}{V(x_{ki})} = \left[\frac{Cov(\tilde{Y}_{ki}, \tilde{x}_{ki})}{V(\tilde{x}_{ki})}\right]\left[\frac{V(\tilde{x}_{ki})}{V(x_{ki})}\right] \neq \beta_k,$$

مگر اینکه x_{ki} با سایر متغیرهای کمکی همبستگی نداشته باشد.

2. Estimator
3. Joint Distribution

قضیه ۴-۱-۳. قضیه CEF خطی (توجیه اول رگرسیون)[۱]

فرض کنید **CEF** خطی است. سپس تابع رگرسیون جامعه آماری برابر با آن است.

اثبات. فرض کنید برای یک بردار $\boldsymbol{k \times 1}$ از ضرایب $\boldsymbol{\beta^*}$، $\boldsymbol{E[Y_i|X_i] = X_i'\beta^*}$ باشد. به یاد آورید که توسط ویژگی تجزیه **CEF** $\boldsymbol{E[X_i(Y_i - E[Y_i|X_i])] = 0}$ است. $\boldsymbol{E[Y_i|X_i] = X_i'\beta^*}$ را جایگزین کنید تا $\boldsymbol{\beta^* = E[X_iX_i']^{-1}E[X_iY_i] = \beta}$ به دست آید.■

قضیه **CEF** خطی این سؤال را ایجاد می‌کند که در چه شرایطی **CEF** خطی است. سناریوی کلاسیک، نرمال بودن مشترک است، به عبارت دیگر بردار $\boldsymbol{(Y_i, x_i')'}$ توزیع نرمال چندمتغیره دارد. این سناریو توسط گالتون[۲] (سال ۱۸۸۶ (۱۲۶۵))، پدر علم رگرسیون که علاقمند به ارتباط بین نسلی بین صفت‌های توزیع نرمال مانند قد و هوش بود در نظر گرفته شده است. مورد عادی به وضوح ارتباط تجربی محدودی دارد، زیرا رگرسورها و متغیرهای وابسته اغلب گسسته هستند، در حالی که توزیع‌های نرمال پیوسته هستند. یکی دیگر از سناریوهای خطی بودن زمانی پیش می‌آید که مدل‌های رگرسیون اشباع می‌شوند. همان طور که در بخش ۴-۱-۳ بررسی شده است، مدل رگرسیون اشباع دارای یک پارامتر مجزا برای هر ترکیب ممکن از مقادیری است که به مجموعه رگرسورها اعمال می‌شود. به عنوان مثال یک مدل رگرسیون اشباع با دو متغیر کمکی موهومی شامل هر دو متغیر کمکی (با ضرایب شناخته شده به عنوان اثرات اصلی) و محصولات آنها (که به عنوان عبارت تعامل شناخته می‌شود) است. چنین مدل‌هایی به طور ذاتی خطی هستند و این نکته‌ای است که در بخش ۴-۱-۳ نیز در مورد آن بحث می‌کنیم.

دو دلیل بعدی برای تمرکز بر رگرسیون مربوط به زمانی است که قضیه **CEF** خطی صادق نیست.

قضیه ۵-۱-۳. بهترین قضیه پیش‌بینی‌کننده خطی (توجیه دوم رگرسیون)

تابع $\boldsymbol{X_i'\beta}$ بهترین پیش‌بینی‌کننده خطی $\boldsymbol{Y_i}$ با در نظر گرفتن $\boldsymbol{X_i}$ در مفهوم **MMSE** است.

اثبات. $\boldsymbol{\beta = E[X_iX_i']^{-1}E[X_iY_i]}$ مشکل حداقل مربعات جامعه آماری را حل می‌کند (۳-۱-۲). ■به عبارت دیگر درست همان طور که در **CEF**، $\boldsymbol{E[Y_i|X_i]}$ بهترین پیش‌بینی‌کننده $\boldsymbol{Y_i}$ (یعنی **MMSE**) با در نظر گرفتن $\boldsymbol{X_i}$ در کلاس *تمام* توابع $\boldsymbol{X_i}$ است، در کلاس توابع *خطی* تابع رگرسیون جامعه آماری بهترین اقدامی است که می‌توانیم انجام دهیم.

قضیه ۶-۱-۳. قضیه رگرسیون CEF (توجیه سوم رگرسیون)

تابع $\boldsymbol{X_i'\beta}$ تقریب خطی **MMSE** به $\boldsymbol{E[Y_i|X_i]}$ را ارائه می‌دهد، یعنی:

$$\beta = \arg\min_b E\{(E[Y_i|X_i] - X_i'b)^2\}. \qquad (3.1.4)$$

اثبات. می‌نویسیم:

---

1. Regression Justification I
2. Galton

$$(Y_i - X_i'b)^2 = \{(Y_i - E[Y_i|X_i]) - (E[Y_i|X_i] - X_i'b)\}^2$$
$$= (Y_i - E[Y_i|X_i])^2 - (E[Y_i|X_i] - X_i'b)^2$$
$$+2(Y_i - E[Y_i|X_i])(E[Y_i|X_i] - X_i'b)$$

عبارت اول شامل b نمی‌شود و عبارت آخر با ویژگی تجزیه CEF امید ریاضی صفر دارد. در نتیجه مسئله تقریب CEF (۴-۱-۳) راه‌حلی یکسان با مسئله حداقل مربعات جامعه آماری (۲-۱-۳) دارد. ■این دو قضیه دو روش دیگر برای مشاهده رگرسیون را به ما نشان می‌دهد. رگرسیون بهترین پیش‌بینی‌کننده خطی برای متغیر وابسته را با همان روشی ارائه می‌دهد که CEF بهترین پیش‌بینی‌کننده نامحدود متغیر وابسته است. از سوی دیگر، اگر ترجیح دهیم که تقریب $E[Y_i|X_i]$ را متضاد پیش‌بینی Y_i در نظر بگیریم، قضیه رگرسیون CEF می‌گوید که حتی اگر CEF غیرخطی باشد، رگرسیون بهترین تقریب خطی به آن را ارائه می‌دهد.

قضیه رگرسیون CEF روش مورد علاقه ما برای تحریک انگیزه‌ها در مورد رگرسیون است. عبارتی که در آن رگرسیون CEF را برآورد می‌کند با دیدگاه ما از کار تجربی به عنوان تلاشی برای توصیف ویژگی‌های اساسی روابط آماری، بدون اینکه لزوماً تلاش کند جزئیات دقیق آن را آشکار کند، هم‌جهت است. قضیه خطی CEF تنها برای موارد خاص است. بهترین قضیه پیش‌بینی‌کننده خطی به طور رضایت‌بخشی کلی است، اما قضیه دید بیش از حد تئوری به تحقیقات تجربی را تشویق می‌کند. ما در واقع علاقمند به پیش‌بینی Y_i تکی نیستیم؛ چیزی که مورد نظر ماست توزیع Y_i است.

شکل ۲-۱-۳ ویژگی تقریب CEF را برای همان نمودار تحصیلات CEF که در شکل ۱-۱-۳ ترسیم شده است نشان می‌دهد. خط رگرسیون با CEF که تا حدودی ناپیوسته و غیرخطی است متناسب است به طوری که این گونه به نظر می‌رسد که مدلی از $E[Y_i|X_i]$ را به جای مدلی از Y_i برآورد کرده‌ایم. در حقیقت این دقیقاً همان چیزی است که در حال وقوع است. مفهومی از قضیه رگرسیون CEF این است که ضرایب رگرسیون را می‌توان با استفاده از $E[Y_i|X_i]$ به عنوان یک متغیر وابسته به جای خود Y_i به دست آورد. برای درک این مطلب فرض کنید که X_i یک متغیر تصادفی گسسته با تابع جرم احتمال $g_x(u)$ با $X_i = u$ است. بنابراین:

$$E\{(E[Y_i|X_i] - X_i'b)^2\} = \sum_u (E[Y_i|X_i = u] - u'b)^2 g_x(u)$$

این به این معناست که β می‌تواند از رگرسیون حداقل مربعات وزنی $E[Y_i|X_i = u]$ برای u ساخته شود، که در آن u بر روی مقادیر گرفته شده توسط X_i اجرا می‌شود. وزن‌ها توسط توزیع X_i، یعنی $g_x(u)$، وقتی که $X_i = u$ است داده می‌شود. روش دیگر این محاسبه این است که امید ریاضی را در فرمول β تکرار کنیم:

$$\beta = E[X_iX_i']^{-1}E[X_iY_i] = E[X_iX_i']^{-1}E[X_iE(Y_i|X_i)]. \qquad (3.1.5)$$

CEF یا نسخه‌ای با داده‌های گروه‌بندی شده از فرمول رگرسیون، هنگام کار روی پروژه‌ای که شامل تجزیه و تحلیل داده‌های خُرد نمی‌شود، کاربردی است. به عنوان مثال، آنگریست (۱۹۹۸)، تأثیر خدمات داوطلبانه نظامی را بر درآمد در ادامه زندگی مطالعه کرد. یکی از استراتژی‌های برآورد مورد استفاده در این پروژه، رگرسیون درآمد شهروندان بر یک متغیر موهومی نشان‌دهنده وضعیت سربازی، همراه با ویژگی‌های شخصی و متغیرهای استفاده شده توسط ارتش برای غربال سربازان است. داده‌های درآمد از سیستم تأمین اجتماعی ایالات متحده آمریکا به دست آمده است، اما سوابق درآمد تأمین اجتماعی انتشار برای عموم ندارد. آنگریست به جای درآمد فردی از درآمد متوسط مقید به نژاد، جنس، نمرات آزمون، تحصیلات و وضعیت سربازی استفاده کرد.

تصویری از رویکرد داده‌های گروه‌بندی شده برای رگرسیون در ادامه نشان داده شده است. ما ضریب تحصیل را در یک معادله درآمد با استفاده از ۲۱ معیار شرطی و نمونه **CEF** درآمد با در نظر گرفتن تحصیلات برآورد کردیم. همان طور که خروجی **Stata** گزارش شده در ادامه نشان می‌دهد، رگرسیون داده‌های گروه‌بندی شده که توسط تعدادی از افراد در هر سطح تحصیلات در نمونه وزن‌دهی شده است، ضرایب یکسانی تولید می‌کند که در برابر چیزی قرار دارد که می‌توانست با استفاده از نمونه داده‌های خُرد زمینه‌ای با صدها هزار مشاهده به دست آید. با وجود این، توجه داشته باشید که خطاهای استاندارد رگرسیون با داده‌های گروه‌بندی به درستی نشان‌دهنده واریانس نمونه‌گیری مجانبی از برآورد شیب در نمونه‌های تکرار شده *داده‌های خُرد* نیست و برای آن نیاز به برآورد واریانس $\boldsymbol{Y_i - X_i'\beta}$ دارید. این واریانس به داده‌های خُرد و به ویژه گشتاورهای دوم $\boldsymbol{W_i \equiv [Y_i;\ X_i']'}$ بستگی دارد که مسئله‌ای است که در بخش بعد شرح داده می‌شود.

۳-۱-۳. استنباط OLS مجانبی[1]

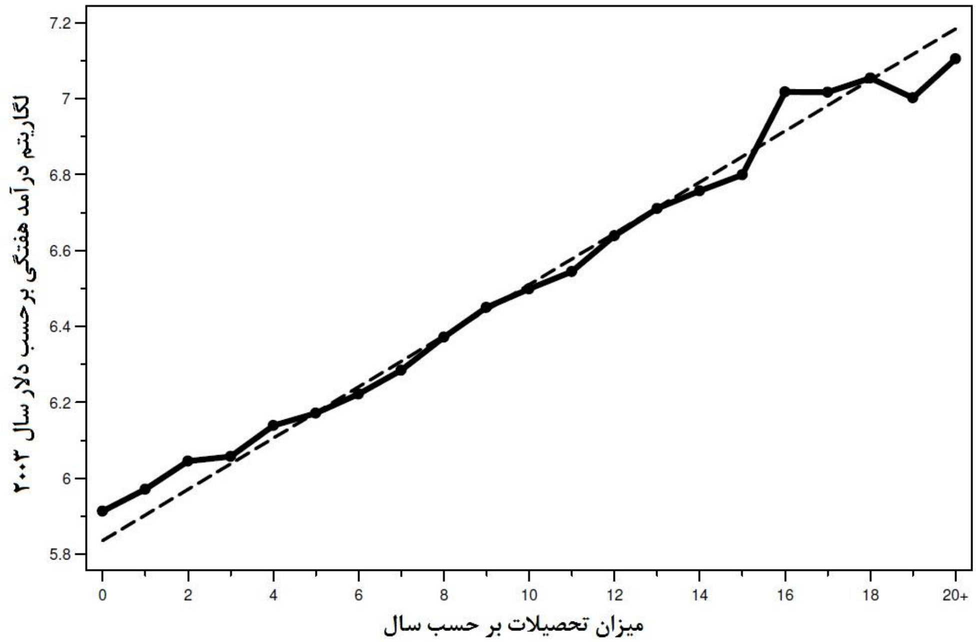
در عمل ما معمولاً نمی‌دانیم **CEF** یا بردار رگرسیون جامعه آماری چیست. بنابراین استنباط‌های آماری در مورد این مقادیر را با استفاده از نمونه‌ها به دست می‌آوریم. استنباط آماری بیشترین بخش اقتصادسنجی سنتی است. این موضوع در هر متن اقتصادسنجی پوشش داده شده است، با وجود این ما نمی‌خواهیم استنباط را به طور کامل از متن خودمان حذف کنیم. بازبینی اساسی تئوری مجانبی به ما اجازه می‌دهد که این واقعیت مهم را برجسته کنیم که فرایند استنباط آماری کاملاً متمایز از سؤال چگونگی تفسیر مجموعه مشخصی از برآوردهای رگرسیون است. معنای ضریب رگرسیون هرچه که باشد، توزیع نمونه‌ای دارد که توصیف و استفاده از آن برای استنباط آماری آسان است.[2]

---

1. Asymptotic OLS Inference

۲. بحث در مورد استنباط OLS مجانبی در این بخش عمدتاً چکیده مطالب چمبرلین (Chamberlain) (۱۹۸۴) است. موانع

شکل ۲-۱-۳. رگرسیون CEF میانگین درآمد هفتگی با توجه به تحصیلات



نمونه محدود به مردان سفیدپوست ۴۹-۴۰ ساله است. داده‌ها از سرشماری IPUMS در سال ۱۹۸۰ و نمونه‌ی ۵٪ است.

ما به توزیع مشابه نمونه

$$\boldsymbol{\beta} = E[X_i X_i']^{-1} E[X_i Y_i]$$

در نمونه‌های تکرار شده علاقمند هستیم. فرض کنید بردار $\boldsymbol{W_i} \equiv [\boldsymbol{Y_i};\ \boldsymbol{X_i'}]'$ به طور مستقل و یکسان در نمونه‌ای با اندازه $\boldsymbol{N}$ توزیع شده است. برآوردگر طبیعی گشتاور جامعه آماری اول $\boldsymbol{E[W_i]}$، مجموع $\frac{1}{N}\sum_{i=1}^{N} \boldsymbol{W_i}$ است. با توجه به قانون اعداد بزرگ، با رشد اندازه نمونه، این گشتاور نمونه به طور دلخواه به گشتاور جامعه آماری مربوطه نزدیک می‌شود. ما ممکن است به طور مشابه لحظات گشتاورهایی با مرتبه بالاتر را از عناصر $\boldsymbol{W_i}$ یا به عبارت دیگر ماتریس گشتاورهای دوم $\boldsymbol{E[W_i W_i']}$ با مشابه نمونه $\frac{1}{N}\sum_{i=1}^{N} \boldsymbol{W_i W_i'}$ در نظر بگیریم. به دنبال این اصل، روش برآوردگر گشتاورهای $\boldsymbol{\beta}$ هر امید ریاضی را با یک مجموع جایگزین می‌کند. این منطق منجر به برآوردگر حداقل مربعات معمولی (**OLS**) می‌شود:

$$\widehat{\boldsymbol{\beta}} = \left[\sum_i X_i X_i'\right]^{-1} \sum_i X_i Y_i$$

اگرچه ما $\widehat{\boldsymbol{\beta}}$ را به عنوان روشی برای برآوردگر روش گشتاورها[1] استخراج کردیم، اما $\widehat{\boldsymbol{\beta}}$ برآوردگر **OLS** برای $\boldsymbol{\beta}$ نامیده می‌شود، زیرا مشابه نمونه مسائل حداقل مربعات توضیح داده در آغاز بخش ۲-

و مشکلات مهم این تئوری مجانبی در فصل آخر مورد بررسی قرار گرفته است.

1. Method of Moments Estimator

۳-۱ را حل می‌کند.[1]

الف) داده‌های سطح فرد

```
. regress earnings school, robust

      Source |       SS       df       MS              Number of obs =  409435
-------------+------------------------------           F(  1,409433) =49118.25
       Model | 22631.4793      1  22631.4793           Prob > F      =  0.0000
    Residual |  188648.31 409433  .460755019           R-squared     =  0.1071
-------------+------------------------------           Adj R-squared =  0.1071
       Total | 211279.789 409434  .51602893            Root MSE      =  .67879
------------------------------------------------------------------------------
             |              Robust                        Old Fashioned
    earnings |     Coef.   Std. Err.       t               Std. Err.       t
-------------+----------------------------------------------------------------
      school |   .0674387   .0003447    195.63             .0003043    221.63
      const. |   5.835761   .0045507   1282.39             .0040043   1457.38
------------------------------------------------------------------------------
```

ب) میانگین‌ها بر اساس سال‌های تحصیل

```
. regress average_earnings school [aweight=count], robust
(sum of wgt is   4.0944e+05)

      Source |       SS       df       MS              Number of obs =      21
-------------+------------------------------           F(  1,     19) =  540.31
       Model |  1.16077332     1  1.16077332           Prob > F      =  0.0000
    Residual |  .040818796    19  .002148358           R-squared     =  0.9660
-------------+------------------------------           Adj R-squared =  0.9642
       Total |  1.20159212    20  .060079606           Root MSE      =  .04635
------------------------------------------------------------------------------
     average |              Robust                        Old Fashioned
   _earnings |     Coef.   Std. Err.       t               Std. Err.       t
-------------+----------------------------------------------------------------
      school |   .0674387   .0040352     16.71             .0029013     23.24
      const. |   5.835761   .0399452    146.09             .0381792    152.85
------------------------------------------------------------------------------
```

شکل ۳-۱-۳: برآوردهای داده‌های خُرد و داده‌های گروه‌بندی شده از بازده درآمدی تحصیلات. منبع: ۱۹۸۰ سرشماری ـ **IPUMS** سال ۱۹۸۰ (۱۳۷۹)، نمونه ۵ درصد. نمونه محدود به مردان سفیدپوست ۴۰-۴۹ساله است. حاصل از خروجی رگرسیون **Stata** است. خطاهای استاندارد سنتی به طور پیش‌فرض گزارش شده است. اشتباهات استاندارد استوار ناهم‌واریانسی ثابت دارد. پنل **A** از داده‌های سطح فردی استفاده می‌کند. پنل **B** از درآمد میانگین بر اساس سال‌های تحصیل استفاده می‌کند.

توزیع نمونه‌گیری مجانبی $\hat{\boldsymbol{\beta}}$ فقط به تعریف مورد برآورد (یعنی ماهیت $\boldsymbol{\beta}$ که ما قصد برآورد کردن

---

۱. متخصصان اقتصادسنجی مایل به استفاده از ماتریس‌ها هستند، زیرا نمادگذاری در آن بسیار فشرده است. گاهی اوقات (نه اغلب) ما هم همین کار را می‌کنیم. فرض کنید X ماتریسی است که ردیف‌های آن با X_i' مشخص شده است و y برداری با عناصر y_i، برای $i = 1, \ldots, N$ است. گشتاور نمونه $\frac{1}{N}\sum X_i X_i'$ برابر با $X'X/N$ است و گشتاور نمونه $\frac{1}{N}\sum X_i y_i$ برابر با $X'y/N$ است. سپس می‌توانیم بنویسیم $\hat{\beta} = (X'X)^{-1}X'y$، که یک فرمول ماتریس آشناست.

آن را داریم) و فرض اینکه داده‌ها یک نمونه تصادفی را تشکیل می‌دهند بستگی دارد. قبل از به دست آوردن این توزیع، بهتر است که تئوری توزیع مجانبی عمومی[1] که نیازهای ما را در این مورد پوشش می‌دهد، بیان کنیم. این تئوری پایه را می‌توان بیشتر در قالب کلمات بیان کرد. برای اهداف این بیانیه فرض می‌کنیم که خواننده با اصطلاحات و مفاهیم اصلی نظریه آماری آشنایی دارد (به عبارت دیگر گشتاورها، امید ریاضی، محدودیت‌های احتمال و توزیع‌های مجانبی). برای تعاریف این اصطلاحات و بیان ریاضی رسمی گزاره‌های تئوری که در زیر آمده است، نایت[2] (سال ۲۰۰۰) را ببینید.

قانون اعداد بزرگ[3] گشتاورهای نمونه در احتمال گشتاورهای جامعه آماری مربوطه همگرا می‌شوند. به عبارت دیگر، احتمال اینکه میانگین نمونه نزدیک به میانگین جامعه آماری باشد می‌تواند با جمع‌آوری نمونه‌ای که به اندازه کافی بزرگ باشد تا حدی که می‌خواهید بالا رود.

تئوری حد مرکزی[4] گشتاورهای نمونه به صورت نرمال مجانبی توزیع می‌شوند (بعد از تفریق گشتاور جامعه آماری مربوطه و ضرب توسط ریشه مربع اندازه نمونه). ماتریس کوواریانس توسط واریانس متغیر تصادفی پایه به دست می‌آید. به عبارت دیگر، در نمونه‌هایی که به اندازه کافی بزرگ است، گشتاورهای نمونه که به طور مناسبی نرمال شده‌اند تقریباً به طور نرمال توزیع می‌شوند.

نظریه اسلاتسکی[5]

(a) مجموع دو متغیر تصادفی را در نظر بگیرید که یکی از آنها در توزیع و دیگری در احتمال به یک مقدار ثابت همگرا می‌شود: توزیع مجانبی این مجموع با جایگزینی متغیری که به مقدار ثابت همگرا می‌شود با آن مقدار ثابت، تحت تأثیر قرار نمی‌گیرد. اگر a_N را یک آماره با توزیع محدود و b_N را یک آماره با محدودیت احتمال b در نظر بگیریم، نتیجه حاصل خواهد شد که $a_N + b_N$ و $a_N + b$ توزیع محدود یکسانی دارند.

(b) ضرب دو متغیر تصادفی را در نظر بگیرید که یکی از آنها در توزیع و دیگری در احتمال به یک ثابت همگرا می‌شود: توزیع مجانبی این محصول با جایگزینی متغیری که به ثابت همگرا می‌شود با آن ثابت، تحت تأثیر قرار نمی‌گیرد. این نکته به ما اجازه می‌دهد که برخی از گشتاورهای نمونه را با گشتاورهای جامعه آماری (به عبارت دیگر با محدودیت‌های احتمال آنها) برای به دست آوردن توزیع جایگزین کنیم. اگر a_N را یک آماره با توزیع محدود و b_N را یک آماره با محدودیت احتمال b در نظر بگیریم، نتیجه گرفته می‌شود که $a_N b_N$ و $a_N b$ توزیع مجانبی یکسانی دارند.

تئوری نگاشت پیوسته[6] محدودیت‌های احتمال از توابع پیوسته می‌گذرد. به عنوان مثال، محدودیت

1. General Asymptotic Distribution Theory
2. Knight
3. The Law Of Large Numbers
4. The Central Limit Theorem
5. Slutsky's Theorem
6. The Continuous Mapping Theorem

احتمال هر تابع پیوسته از یک گشتاور نمونه، تابعی است که در گشتاور جامعه آماری مربوطه ارزیابی می‌شود. به عبارت دیگر، محدودیت احتمال $h(b_N)$ برابر با $h(b)$ است که در آن $\text{plim } b_N = b$ و $h(\cdot)$ در b پیوسته است.

روش دلتا[1] یک متغیر تصادفی با بردار ارزشگذاری شده[2] را که به طور نرمال توزیع مجانبی شده است را در نظر بگیرید. بیشتر توابع اسکالر این متغیر تصادفی نیز با ماتریس کوواریانس داده شده در قالب درجه دو با ماتریس کوواریانس متغیر تصادفی در داخل و گرادیان تابع ارزیابی شده در محدودیت احتمال متغیر تصادفی در بیرون به طور نرمال توزیع مجانبی شده است. به عبارت دیگر، توزیع مجانبی $h(b_N)$ با ماتریس کوواریانس $\nabla h(b)'\Omega\nabla h(b)$ نرمال است که در آن $\text{plim } b_N = b$ و $h(\cdot)$ به طور پیوسته در b با گرادیان $\nabla h(b)$ دیفرانسیل‌پذیر است و b_N کوواریانس مجانبی ماتریس Ω را دارد.[3]

ما می‌توانیم از این نتایج برای به دست آوردن توزیع مجانبی $\boldsymbol{\hat{\beta}}$ به دو روش استفاده کنیم. یک رویکرد که از نظر مفهومی ساده، اما نه چندان مطلوب است استفاده از روش دلتاست: $\boldsymbol{\hat{\beta}}$ تابعی از گشتاورهای نمونه است. بنابراین به طور نرمال توزیع مجانبی شده است. در اینجا، تنها نیاز است ماتریس کوواریانس توزیع مجانبی از گرادیان این تابع پیدا شود. (توجه داشته باشید که ثابت بودن $\boldsymbol{\hat{\beta}}$ مستقیماً از قضیه نگاشت پیوسته حاصل می‌شود). یک روش استخراج ساده‌تر و سازنده‌تر، استفاده از قضیه‌های حد مرکزی و اسلاتسکی است. ابتدا توجه کنید که می‌توانیم بنویسیم:

$$\boldsymbol{Y_i = X_i'\beta + [Y_i - X_i'\beta] \equiv X_i'\beta + e_i} \qquad (۳-۱-۶)$$

که در آن $\boldsymbol{e_i}$ پسماند، مانند قبل به عنوان تفاوت بین متغیر وابسته و تابع رگرسیون جامعه آماری *تعریف* می‌شود. موقعیت خوبی است که اشاره کنیم که این پسماندها با رگرسورهای تعریف $\boldsymbol{\beta}$ همبستگی ندارند. به عبارت دیگر $\boldsymbol{E[X_i e_i] = 0}$ نتیجه $\boldsymbol{\beta = E[X_i X_i']^{-1} E[X_i Y_i]}$ و $\boldsymbol{e_i = Y_i - X_i'\beta}$ و نه فرضی درباره یک رابطه اقتصادی پایه‌ای است. در بحث مدل‌های رگرسیون علیتی در بخش ۲-۳ به این نکته مهم باز می‌گردیم.[4]

با جایگزینی شاخص ۶-۱-۳ با $\boldsymbol{Y_i}$ در فرمول $\boldsymbol{\hat{\beta}}$ داریم:

$$\boldsymbol{\hat{\beta} = \beta + \left[\sum X_i X_i'\right]^{-1} \sum X_i e_i}$$

توزیع مجانبی $\boldsymbol{\hat{\beta}}$ توزیع مجانبی $\boldsymbol{\sqrt{N}(\hat{\beta} - \beta) = N[\sum X_i X_i']^{-1} \frac{1}{\sqrt{N}} \sum X_i e_i}$ است. با قضیه

1. The Delta Method
2. Vector-valued Random Variable
۳. برای استخراج از فرمول روش دلتا با استفاده از اسلاتسکي و قضیه‌های نگاشت پیوسته، به نایت (Knight)، سال ۲۰۰۰، صفحات ۱۲۱-۱۲۰ مراجعه کنید.
٤. پسماندهاي تعریف شده در این روش لزوماً مستقل از میانگین X_i نیستند؛ برای استقلال از میانگین، نیاز به CEF خطی داریم.

اسلاتسکی، این توزیع مجانبی با توزیع مجانبی $E[X_iX_i']^{-1}\frac{1}{\sqrt{N}}\sum X_ie_i$ برابر است. از آنجا که $E[X_ie_i]=0$ است، $\frac{1}{\sqrt{N}}\sum X_ie_i$ یک گشتاور نمونه نرمال و مرکزی شده ریشه N است. با توجه به قضیه حد مرکزی، از آنجا که این گشتاور چهارم ماتریس کوواریانس X_ie_i است، این عبارت به طور نرمال با میانگین صفر و ماتریس کوواریانس $E[X_iX_i'e_i^2]$ توزیع مجانبی می‌شود. بنابراین $\widehat{\beta}$ دارای توزیع نرمال مجانبی با محدودیت احتمال β و ماتریس کوواریانس زیر است:

$$E[X_iX_i']^{-1}E[X_iX_i'e_i^2]E[X_iX_i']^{-1}. \qquad (3.1.7)$$

خطاهای استاندارد استفاده شده برای ایجاد آمار-t، مربع ریشه‌های عناصر قطری این ماتریس است. در عمل، این خطاهای استاندارد با جایگزینی مجموع با امید ریاضی و استفاده از پسماندهای برآورد شده $\hat{e}_i=Y_i-X_i'\widehat{\beta}$ برای تشکیل گشتاور چهارم تجربی $\sum[X_iX_i\hat{e}_i^2]/N$ برآورد می‌شود.

خطاهای استاندارد مجانبی محاسبه شده با این روش به عنوان خطاهای استاندارد سازگار با ناهمسانی واریانس (ناهم‌واریانس)،[1] خطاهای استاندارد وایت (**a**۱۹۸۰) یا خطاهای استاندارد ایکر وایت[2] با اشاره به کار ایکر (۱۹۶۷) شناخته می‌شود. آنها همچنین به عنوان خطاهای استاندارد «استوار»[3] (مثلاً در **Stata**) شناخته می‌شوند. از آنجایی به این خطاهای استاندارد، استوار گفته می‌شود که در نمونه‌هایی که به اندازه کافی بزرگ هستند، آزمون‌های فرضیه دقیق و فواصل اطمینان را با حداقل فرضیات در مورد داده‌ها و مدل‌ها ارائه می‌دهند. به طور خاص، شیوه استخراج ما از توزیع محدودکننده،[4] هیچ فرضی بجز آنچه برای اطمینان حاصل کردن از صحت نتایج آماری پایه مانند قضیه حد مرکزی مورد نیاز است ندارد. با وجود این، اینها خطاهای استانداردی نیستند که به طور پیش‌فرض از بسته‌های نرم‌افزاری دریافت می‌کنید. خطاهای استاندارد پیش‌فرض بر اساس فرض ناهمسان واریانس و به طور خاص $E[e_i^2|X_i]=\sigma^2$ که یک مقدار ثابت است استخراج می‌شوند. با توجه به این فرض، با امید ریاضی مکرر داریم:

$$E[X_iX_i'e_i^2]=E(X_iX_i'E[e_i^2|X_i])=\sigma^2E[X_iX_i']$$

سپس ماتریس کوواریانس مجانبی $\widehat{\beta}$ به رابطه زیر ساده می‌شود:

$$\begin{aligned} E[X_iX_i']^{-1}E[X_iX_i'e_i^2]E[X_iX_i']^{-1} &= E[X_iX_i']^{-1}\sigma^2E[X_iX_i']E[X_iX_i]^{-1} \\ &= E[X_iX_i']^{-1}\sigma^2. \qquad (3.1.8) \end{aligned}$$

عناصر قطری (۸-۱-۳) مطابق با خروجی **SAS** یا **Stata** هستند مگر اینکه شما درخواست دیگری داشته باشید.

دیدگاه ما نسبت به رگرسیون به عنوان تقریبی از **CEF** باعث می‌شود که ناهمسانی واریانس

1. Heteroskedasticity-Consistent Standard Errors
2. Eicker-White Standard Errors
3. Robust Standard Errors
4. Limiting Distribution

طبیعی به نظر برسد. اگر **CEF** غیرخطی باشد و از یک مدل خطی برای تقریب آن استفاده کنید، کیفیت تناسب بین خط رگرسیون و **CEF** با توجه به $\boldsymbol{X_i}$ متفاوت خواهد بود. بنابراین پسماندها به طور میانگین در مقادیری از $\boldsymbol{X_i}$ که برازش ضعیف‌تر است، بزرگ‌تر خواهد بود. حتی اگر شما می‌خواهید بپذیرید که واریانس شرطی $\boldsymbol{Y_i}$ با توجه به $\boldsymbol{X_i}$ ثابت است، این واقعیت که **CEF** غیرخطی است، به این معناست که $\boldsymbol{E[(Y_i - X_i'\beta)^2|X_i]}$ با توجه به $\boldsymbol{X_i}$ متفاوت خواهد بود. برای درک این مطلب توجه داشته باشید که به عنوان یک قاعده داریم:

$$\begin{aligned} E[(Y_i - X_i'\beta)^2|X_i] &= \\ & E\{[(Y_i - E[Y_i|X_i]) + (E[Y_i|X_i] - X_i'\beta)]^2|X_i\} \\ &= V[Y_i|X_i] + (E[Y_i|X_i] - X_i'\beta)^2. \end{aligned} \tag{3.1.9}$$

بنابراین، حتی اگر $\boldsymbol{V[Y_i|X_i]}$ ثابت باشد، واریانس پسماند با مربع فاصله بین خط رگرسیون و **CEF** افزایش می‌یابد که واقعیتی است که در کار وایت (**b**۱۹۸۰) به آن اشاره شده است.[۱]

به طور مشابه شایان ذکر است که با اینکه **CEF** خطی امکان همسانی واریانس (هم‌واریانسی)[۲] را فراهم می‌کند، اما شرط کافی برای هم‌واریانسی نیست. مثال مورد علاقه ما در این زمینه، مدل احتمال خطی (**LPM**[۳]) است. مدل احتمالی خطی، هر رگرسیونی است که در آن متغیر وابسته به صورت صفر ـ یک است، یعنی یک متغیر موهومی مانند شاخصی برای مشارکت نیروی کار. فرض کنید که مدل رگرسیون اشباع شده است، بنابراین **CEF** خطی است. از آنجا که **CEF** خطی است، واریانس پسماند نیز واریانس شرطی $\boldsymbol{V[Y_i|X_i]}$ است، اما متغیر وابسته یک آزمون برنولی است و واریانس یک آزمون برنولی $\boldsymbol{P[Y_i|X_i](1 - P[Y_i|X_i])}$ است. نتیجه می‌گیریم که پسماندهای **LPM** لزوماً ناهم‌واریانس هستند مگر اینکه تنها رگرسور ثابت باشد.

با وجود این نکات اصلی به عنوان یک مطلب تجربی، ناهم‌واریانسی ممکن است کم‌اهمیت باشد. در رگرسیون داده‌های خُرد تحصیلات که در شکل ۳-۱-۳ نشان داده شده است، خطای استاندارد استوار ۰/۰۰۰۳۴۴۷ است، در حالی که خطای استاندارد سنتی ۰/۰۰۰۳۰۴۳ است که فقط کمی کوچک‌تر است. خطای استاندارد رگرسیون داده‌های گروه‌بندی شده که در صورت تفاوت اندازه گروه‌ها لزوماً ناهم‌واریانس است کمی بیشتر تغییر می‌کند؛ استاندارد استوار ۰/۰۰۴ را با خطای استاندارد معمولی ۰/۰۰۲۹ مقایسه کنید. بر اساس تجربه ما، این اختلافات معمول هستند. اگر ناهم‌واریانسی بیش از حد مهم باشد، مثلاً بیش از ۳۰ درصد افزایش یا هر کاهش مشخصی در خطاهای استاندارد، شما باید در

---

۱. عبارت حاصلضرب که نتیجه بسط درجه دوم در اواسط ۹-۱-۳ است صفر است، زیرا $Y_i - E[Y_i|X_i]$ میانگین مستقل X_i است.

2. Homoskedasticity

3. Linear Probability Model

مورد اشتباهات برنامه‌نویسی ممکن یا مشکلات دیگر نگران شوید (به عنوان مثال، خطاهای استاندارد استوار کمتر از مقدار متعارف ممکن است نشانه‌ای از اریب نمونه متناهی[1] در محاسبه استواری است؛ فصل ۸ را در زیر ببینید).

۴-۱-۳. مدل‌های اشباع شده، اثرات اصلی و بحث‌های دیگر در مورد رگرسیون

ما اغلب در بحث در مورد مدل‌های رگرسیون از اصطلاحاتی مانند *اشباع شده*[2] و *اثرات اصلی*[3] استفاده می‌کنیم. این اصطلاحات از یک سنت تجربی که از رگرسیون برای مدل‌سازی متغیرهای نوع تیمار گسسته[4] استفاده می‌کند، سرچشمه می‌گیرد. با این حال این زبان در حال حاضر به طور گسترده‌تری در بسیاری از زمینه‌ها از جمله اقتصادسنجی کاربردی استفاده می‌شود. این بخش بررسی مختصری برای خوانندگان ناآشنا با این اصطلاحات ارائه می‌دهد.

مدل‌های رگرسیون اشباع شده، مدل‌های رگرسیونی با متغیرهای توضیحی گسسته هستند که در آنها مدل شامل یک پارامتر جداگانه برای تمام مقادیر ممکن اختیار شده توسط متغیرهای توضیحی است. برای مثال، هنگام کار با یک متغیر توضیحی که نشان می‌دهد آیا یک کارگر فارغ‌التحصیل کالج است یا نه، مدل با شامل کردن یک متغیر موهومی تکی برای فارغ‌التحصیلان کالج و یک ثابت، اشباع شده است. اشباع ممکن است هنگامی که رگرسور مقادیر بسیاری می‌گیرد نیز رخ دهد. برای مثال فرض کنید $\boldsymbol{s_i = 0, 1, 2, \dots, \tau}$ است. یک مدل رگرسیون اشباع شده برای $\boldsymbol{s_i}$ به صورت زیر است:

$$\boldsymbol{Y_i = \beta_0 + \beta_1 d_{1i} + \beta_2 d_{2i} + \cdots + \beta_\tau d_{\tau i} + \varepsilon_i}$$

که در آن $\boldsymbol{d_{ji} = 1[s_i = j]}$ یک متغیر موهومی است که نشان‌دهنده سطح تحصیلات $\boldsymbol{j}$ است و $\boldsymbol{\beta_j}$ *اثر* تحصیلات سطح $\boldsymbol{j}$ است. توجه داشته باشید که:

$$\boldsymbol{\beta_j = E[Y_i|s_i = j] - E[Y_i|s_i = 0]}$$

در حالی که $\boldsymbol{\beta_0 = E[Y_i|s_i = 0]}$ است. در عمل می‌توانید هر مقداری از $\boldsymbol{s_i}$ را برای گروه مرجع انتخاب کنید؛ یک مدل رگرسیون تا زمانی اشباع شده است که یک پارامتر برای هر احتمال ممکن $\boldsymbol{j}$ در $\boldsymbol{E[Y_i|s_i = j]}$ داشته باشد. مدل‌های اشباع شده به صورت کامل با **CEF** برازش دارند، زیرا **CEF** در رگرسورهای موهومی که برای اشباع استفاده می‌شوند خطی است. این یک مورد خاص مهم از قضیه رگرسیون **CEF** است.

اگر دو متغیر توضیحی وجود داشته باشد، مثلاً یک متغیر موهومی نشان‌دهنده فارغ‌التحصیلان کالج و یک متغیر موهومی نشان‌دهنده جنسیت باشد، مدل با شامل کردن این دو متغیر موهومی، ضرب آنها و یک ثابت اشباع می‌شود. ضرایب روی متغیرهای موهومی به عنوان اثرات اصلی شناخته می‌شوند، در

1. Finite-sample Bias
2. Saturated
3. Main Effects
4. Discrete Treatment-type Variables

حالی که ضرب به عنوان عبارت اثر متقابل شناخته می‌شود. این تنها پارامتر اشباع شده نیست؛ هر مجموعه‌ای از شاخص‌ها (متغیرهای موهومی) که می‌تواند برای شناسایی هر مقدار گرفته شده توسط متغیرهای کمکی استفاده شود، یک مدل اشباع شده را تولید می‌کند. به عنوان مثال، یک مدل اشباع شده جایگزین متغیرهای موهومی برای فارغ‌التحصیلان مرد کالج، انصراف‌دهندگان مرد، فارغ‌التحصیلان زن کالج و انصراف‌دهندگان زن را بدون هیچ وقفه‌ای شامل می‌کند.

در اینجا برخی از نشانه‌هایی که این مطلب را قطعی‌تر می‌کند ارائه شده است. اگر x_{1i} نشان‌دهنده فارغ‌التحصیلان کالج و x_{2i} نشان‌دهنده زنان باشد، **CEF** با توجه به x_{1i} و x_{2i} چهار مقدار می‌گیرد:

$$E[Y_i|x_{1i}=0, x_{2i}=0],$$
$$E[Y_i|x_{1i}=1, x_{2i}=0],$$
$$E[Y_i|x_{1i}=0, x_{2i}=1],$$
$$E[Y_i|x_{1i}=1, x_{2i}=1],$$

می‌توانیم این را با استفاده از طرح زیر برچسب‌گذاری کنیم:

$$E[Y_i|x_{1i}=0, x_{2i}=0]=\alpha$$
$$E[Y_i|x_{1i}=1, x_{2i}=0]=\alpha+\beta$$
$$E[Y_i|x_{1i}=0, x_{2i}=1]=\alpha+\gamma$$
$$E[Y_i|x_{1i}=1, x_{2i}=1]=\alpha+\beta+\gamma+\delta$$

از آنجا که چهار حرف یونانی وجود دارد و **CEF** چهار مقدار می‌گیرد، این پارامترسازی **CEF** را محدود نمی‌کند. این عبارت می‌تواند با حروف یونانی به صورت زیر نوشته شود:

$$E[Y_i|x_{1i}, x_{2i}]=\alpha+\beta x_{1i}+\gamma x_{2i}+\delta(x_{1i}, x_{2i})$$

که پارامترسازی با دو اثر اصلی و یک عبارت اثر متقابل است.[۱] در نتیجه معادله رگرسیون اشباع شده به صورت زیر نوشته می‌شود:

$$Y_i=\alpha+\beta x_{1i}+\gamma x_{2i}+\delta(x_{1i}, x_{2i})+\varepsilon_i$$

در نهایت، می‌توانیم متغیر تحصیلات چند مقداری را برای تولید یک مدل اشباع شده که اثرات اصلی تحصیلات $\boldsymbol{\tau}$، یک اثر اصلی برای جنسیت و اثر متقابل جنسیت ـ تحصیلات $\boldsymbol{\tau}$ را دارد، با جنسیت ترکیب کنیم:

$$\mathrm{Y}_i=\beta_0+\sum_{j=1}^{\tau}\beta_j d_{ji}+\gamma x_{2i}+\sum_{j=1}^{\tau}\delta_j(d_{ji}x_{2i})+\varepsilon_i. \tag{3.1.10}$$

عبارات اثر متقابل $\boldsymbol{\delta_j}$ نشان می‌دهد که چگونه هریک از اثرات تحصیلات با توجه به جنسیت متفاوت می‌شود. **CEF** در این مورد $\boldsymbol{2(\tau+1)}$ مقدار می‌گیرد در حالی که رگرسیون این تعداد پارامتر دارد.

---

۱. با متغیر موهومی سوم در مدل که با x_{3i} نشان داده می‌شود، مدل اشباع شده شامل سه اثر اصلی، سه عبارت اثر متقابل مرتبه دوم $\{x_{1i}x_{2i}, x_{2i}x_{3i}, x_{1i}x_{3i}\}$ و یک عبارت مرتبه سوم $x_{1i}x_{2i}x_{3i}$ می‌شود.

توجه داشته باشید که یک سلسله‌مراتب طبیعی از استراتژی‌های مدل‌سازی با مدل‌های اشباع شده در بالا وجود دارد. طبیعی است که با یک مدل اشباع شده شروع کنیم، زیرا با **CEF** متناسب است. از سوی دیگر، مدل‌های اشباع شده، عبارات اثر متقابل زیادی ایجاد می‌کنند که بسیاری از آنها ممکن است مورد توجه یا دقیق نباشد. بنابراین ممکن است به صلاحدید خود انتخاب کنید که برخی یا همه آنها را حذف کنید. معادله (۱۰-۱-۳) بدون عبارات اثر متقابل، تقریب **CEF** را با یک مدل صرفاً جمعی برای تحصیلات و جنسیت انجام می‌دهد. این تقریب خوبی برای وضعیتی است که بازده کالج برای مردان و زنان مشابه باشد و در هر صورت، ضرایب تحصیلات در مشخصات جمعی، همان طور که در بخش ۱-۳-۳ در زیر بحث شده است، بازده متوسط (وزنی) را در هر دو جنس به دست می‌دهد. از سوی دیگر، برآورد مدلی که عبارات اثر متقابل را شامل کرده است، اما اثرات اصلی مربوطه را حذف کرده است، عجیب خواهد بود. این مسئله در مورد تحصیلات، چیزی شبیه به این خواهد بود:

$$Y_i = \beta_0 + \gamma x_{2i} + \sum_{j=1}^{\tau} \delta_j (d_{ji} x_{2i}) + \varepsilon_i. \qquad (3.1.11)$$

این مدل به تحصیلات اجازه می‌دهد که فقط دستمزد زنان را تغییر دهد که چیزی بسیار دور از حقیقت است. در نتیجه، نتایج برآورد (۱۱-۱-۳) به احتمال زیاد به سختی قابل تفسیر است.

در نهایت، مهم است تشخیص دهیم که یک مدل اشباع شده، بدون در نظر گرفتن توزیع $\boldsymbol{Y_i}$، به طور عالی با **CEF** متناسب است. برای مثال، این مطلب برای مدل‌های احتمال خطی و دیگر مدل‌های متغیر وابسته محدود (به عبارت دیگر $\boldsymbol{Y_i}$ غیرمنفی) درست است و این نکته‌ای است که در انتهای این فصل به آن باز می‌گردیم.

۲-۳. رگرسیون و علیت

بخش ۲-۱-۳ نشان می‌دهد؛ چگونه رگرسیون، بهترین (**MMSE**) تقریب خطی **CEF** را به دست می‌دهد. با وجود این، این درک به ما در مورد پرسش عمیق‌تری که چه زمانی رگرسیون دارای تفسیر علیت است، کمکی نمی‌کند. چه زمانی می‌توانیم یک ضریب رگرسیون را به عنوان تقریبی از اثر علیتی که ممکن است در یک آزمایش نشان داده شود، در نظر بگیریم؟

۱-۲-۳. فرض استقلال شرطی[1]

رگرسیون زمانی علّی است که **CEF**ای که به آن تقریب داده می‌شود علّی باشد. البته این به سؤال پاسخ نمی‌دهد. این فقط یک سطح آماده‌سازی را پشت سر می‌گذارد، زیرا همان طور که مشاهده کردیم، رگرسیون درستی خود را از **CEF** به دست می‌آورد. علیت برای افراد مختلف معانی متفاوتی دارد، اما

1. The Conditional Independence Assumption (CIA)

به نظر محققان مشغول به کار در بسیاری از رشته‌ها، در نظر گرفتن روابط علّی به عنوان نشانه‌های بالقوه نتایج استفاده شده در فصل ۲ برای توصیف آنچه برای یک فرد در مقایسه فرضی سناریوهای بستری شدن جایگزین اتفاق می‌افتد، مفید است. گفته شده است که تفاوت در این نتایج بالقوه اثر علّی بستری شدن است. **CEF** وقتی علّی است که تفاوت‌ها را در نتایج بالقوه میانگین برای یک جامعه آماری مرجع ثابت توصیف کند.

ساده‌ترین راه این است که مفهوم تا حدودی مبهم یک **CEF** علّی را در متن یک سؤال خاص تعمیم بدهیم. بنابراین با مثال تحصیلات ادامه می‌دهیم. ارتباط علّی بین تحصیلات و درآمد می‌تواند به عنوان رابطه کارکردی تعریف شود که توصیف می‌کند فرد بر اساس سطح تحصیلات خود به چه میزان می‌تواند کسب درآمد کند. به طور خاص، ما ممکن است تصور کنیم که تصمیمات تحصیلی در مجموعه‌ای از قسمت‌ها صورت می‌گیرد که تصمیم‌گیرنده ممکن است به طور واقعی به یک راه یا راه دیگر برود، حتی اگر انتخاب‌های خاصی احتمال بیشتری نسبت به دیگران داشته باشد. به عنوان مثال، آنگریست برای اواسط سال سوم دبیرستان و وضعیت بی‌طاقتی و ناراحتی همراه با ترس گزینه‌های خود را این چنین در نظر می‌گیرد: رها کردن دبیرستان و امیدوار برای پیدا کردن شغل، ماندن در مدرسه، اما انتخاب کلاس‌های آسان که منجر به گرفتن دیپلمی سریع و شلخته از دبیرستان می‌شود و یا تلاش در یک مسیر علمی که منجر به کالج می‌شود. اگرچه عواقب چنین انتخاب‌هایی معمولاً از پیش ناشناخته است، اما ایده مسیرهای جایگزینی که منجر به نتایج جایگزین برای یک فرد می‌شود به نظر غیرقابل انکار می‌رسد. فیلسوفان بر سر اینکه آیا این مفهوم شخصی از نتایج بالقوه به اندازه کافی دقیق است که از لحاظ علمی مفید باشد یا نه بحث کرده‌اند، اما به نظر می‌رسد تصمیم‌گیرندگان فردی مشکلی در تفکر در مورد زندگی و انتخاب‌های خود به این شیوه نداشته باشند (همان طور که در کتاب مشهور رابرت فراست[1] *راه انتخاب نشده*[2] آمده است: مسافر ـ راوی خود را می‌بیند که به لحظه تصمیم‌گیری در گذشته نگاه می‌کند. او معتقد است که تصمیم به رفتن در جاده‌ای که کمتر در آن سفر رفته‌اند «دلیل تمام تفاوت‌هاست»، هرچند او همچنین می‌داند که نتایج پیش نیامده ناشناخته هستند).

در کار تجربی، رابطه علّی بین تحصیلات و درآمد به ما می‌گوید که اگر بتوانیم تحصیلات افراد را در یک محیط کاملاً کنترل شده تغییر دهیم یا تحصیلاتشان را به صورت تصادفی تغییر دهیم تا افراد با سطوح مختلف تحصیلات قابل مقایسه باشند، افراد به طور متوسط به چه میزان می‌توانند درآمد کسب کنند. همان طور که در فصل ۲ بحث کردیم، آزمایش‌ها اطمینان می‌دهند که متغیر علّی مورد نظر مستقل از نتایج بالقوه است. بنابراین گروه‌های مقایسه شده واقعاً قابل مقایسه هستند. در اینجا، ما

1. Robert Frost
2. The Road Not Taken

می‌خواهیم این مفهوم را به متغیرهای علّی که بیش از دو مقدار می‌گیرند و به شرایط پیچیده‌تری که باید انواع «متغیرهای کنترلی» را برای درست بودن استنباط‌های علیتی ثابت نگه داریم، تعمیم دهیم. این منجر به *فرض استقلال شرطی* (**CIA**) می‌شود، یک فرضیه اصلی که توجیه (گاهی ضمنی) برای تفسیر علت رگرسیون را فراهم می‌کند. این فرض گاهی اوقات انتخاب بر اساس مشاهده نامیده می‌شود، به این دلیل که متغیرهای ثابت، شناخته شده و مشاهده می‌شوند (برای توضیح بیشتر در گلدبرگر، ۱۹۷۲؛ بارنو،[۱] کین[۲] و گلدبرگر، ۱۹۸۱ را ببینید). بنابراین سؤال بزرگ این است که این متغیرهای کنترل چه هستند یا چه باید باشند. به زودی در مورد آن صحبت خواهیم کرد. در حال حاضر، صرفاً اقتصادسنجی انجام می‌دهیم و متغیر کمکی را «X_i» می‌نامیم. با توجه به مشکل تحصیلات، طبیعی است که تصور کنید که X_i برداری شامل معیارهای توانایی و زمینه خانوادگی است.

برای شروع، تحصیل را به عنوان یک تصمیم دوگانه در نظر بگیرید، مثلاً اینکه آنگریست به کالج می‌رود یا نه. این را با یک متغیر موهومی C_i تعریف کنید. رابطه علیتی بین حضور کالج و یک نتیجه در آینده مانند درآمد را می‌توان با استفاده از نشانه‌های بالقوه نتایجی که برای توصیف آزمایشات در فصل دوم استفاده کردیم، توصیف کرد. برای پاسخگویی به این سؤال، دو متغیر درآمد بالقوه را تصور می‌کنیم:

$$\text{نتیجه بالقوه} = \begin{cases} Y_{1i} & \text{اگر} \quad C_i = 1 \\ Y_{0i} & \text{اگر} \quad C_i = 0 \end{cases}$$

در این مورد، Y_{0i} درآمد i بدون کالج است، در حالی که Y_{1i} درآمد i با رفتن او به کالج است. ما می‌خواهیم تفاوت بین Y_{1i} و Y_{0i} را بدانیم، که اثر علّی کالج رفتن برای فرد i است. این چیزی است که اگر می‌توانستیم در زمان به عقب برگردیم و i را به جاده انتخاب نشده‌اش، هل دهیم می‌توانستیم اندازه‌گیری کنیم. نتیجه مشاهده شده Y_i می‌تواند به صورت زیر به عنوان نتایج بالقوه نوشته شود:

$$Y_i = Y_{0i} + (Y_{1i} - Y_{0i})C_i$$

ما می‌توانیم یکی از Y_{1i} یا Y_{0i} را ببینیم، اما دیدن هر دو با هم هرگز ممکن نیست. بنابراین امیدواریم میانگین $Y_{1i} - Y_{0i}$ یا میانگین بعضی از گروه‌ها، مانند کسانی که به کالج رفته‌اند، را اندازه‌گیری کنیم. این $E[Y_{1i} - Y_{0i} | C_i = 1]$ است.

به طور کلی، مقایسه افرادی که به کالج می‌روند و نمی‌روند، احتمالاً معیار ضعیفی از اثر علیتی کالج رفتن است. به دنبال منطق فصل ۲، داریم:

1. Barnow
2. Cain

$$\underbrace{E\left[Y_i|C_i=1\right]-E[Y_i|C_i=0]}_{\text{تفاوت مشاهده شده در درآمد}} = \underbrace{E[Y_{1i}-Y_{0i}|C_i=1]}_{\text{اثر تیمار میانگین بر تیمار شده}} + \underbrace{E\left[Y_{0i}|C_i=1\right]-E\left[Y_{0i}|C_i=0\right]}_{\text{انتخاب اریب}}. \quad (3.2.1)$$

به نظر احتمال دارد کسانی که به کالج می‌روند، به هر حال درآمد بیشتری داشته باشند. اگر چنین است، انتخاب اریب مثبت است و مقایسه ساده $E[Y_i|C_i=1]-E[Y_i|C_i=0]$ فواید حضور در کالج را بزرگنمایی می‌کند.

CEF ادعا می‌کند که مشروط بر مشخصات مشاهده شده X_i انتخاب اریب ناپدید می‌شود. در این مثال، **CIA** می‌گوید:

$$\{Y_{0i},Y_{1i}\} \amalg C_i|X_i. \quad (3.2.2)$$

با توجه به **CIA**، مقایسه مشروط بر X_i درآمد میانگین معرض از مبدأ با سطح تحصیلات، یک تفسیر علیت دارد. به عبارت دیگر:

$$E\left[Y_i|X_i,C_i=1\right]-E\left[Y_i|X_i,C_i=0\right]=E[Y_{1i}-Y_{0i}|X_i].$$

اکنون می‌خواهیم فرض استقلال شرطی را به روابط علیتی که شامل متغیرهایی است که می‌توانند بیش از دو مقدار مانند سال‌های تحصیل S_i را بگیرند تعمیم دهیم. رابطه علّی بین تحصیل و درآمد احتمالاً برای هر فرد متفاوت است. ازاین‌رو از نشانه‌های خاص فردی استفاده می‌کنیم:

$$Y_{si}\equiv f_i(s)$$

تا درآمد بالقوه‌ای که فرد i پس از s سال تحصیل دریافت می‌کند نشان دهیم. اگر s فقط دو مقدار ۱۲ و ۱۶ را بگیرد، پس به مثال با کالج/ بدون کالج باز می‌گردیم:

$$Y_{0i}=f_i(12);\, Y_{1i}=f_i(16).$$

به طور کلی، تابع $f_i(s)$ به ما می‌گوید که i برای هر مقدار از تحصیلات s چه درآمدی کسب می‌کند. به عبارت دیگر $f_i(s)$ به سؤالات علّی «چه می‌شد اگر» پاسخ می‌دهد. در زمینه مدل‌های نظری رابطه سرمایه انسانی و درآمد، شکل $f_i(s)$ می‌تواند توسط جنبه‌های رفتار فردی و/یا نیروهای بازار تعیین شود.

CIA در این تنظیم بیشتر عمومی تبدیل به حالت زیر می‌شود:

$$Y_{si} \amalg s_i|X_i \quad \text{(CIA)}$$

در بسیاری از آزمایشات تصادفی، **CIA** به دلیل اینکه S_i به صورت تصادفی مشروط بر X_i تعیین می‌گردد، بالا می‌رود (به عنوان مثال در آزمایش تنسی استار،[1] کلاس‌های کوچک به طور تصادفی در

1. Tennessee Star

مدارس تعیین شده‌اند). در یک مطالعه مبتنی بر مشاهده، **CIA** به این معناست که می‌توان گفت که S_i می‌تواند مشروط به X_i «تا حدی خوب باشد که در تعیین تصادفی ممکن است».

مشروط بر X_i، *اثر علیت میانگین* یک سال افزایش در تحصیلات $E[f_i(s) - f_i(s-1)|X_i]$ است، در حالی که اثر علیت میانگین افزایش چهارساله در تحصیلات $E[f_i(s) - E[f_i(s-4)]|X_i]$ است. داده‌ها فقط $Y_i = f_i(S_i)$ را نشان می‌دهند، اگرچه $f_i(s)$ برای $s = S_i$ است، اما با توجه به **CIA**، مقایسه مشروط بر X_i درآمد متوسط معرض از مبدأ بر سطح تحصیلات، یک تفسیر علّی دارد. به عبارت دیگر:

$$E[Y_i|X_i, S_i = s] - E[Y_i|X_i, S_i = s-1]$$
$$= E[f_i(s) - f_i(s-1)|X_i]$$

برای هر مقدار از s. به عنوان مثال، ما می‌توانیم درآمد افرادی با سابقه ۱۱ تا ۱۲ سال تحصیل را مقایسه کنیم تا بتوانیم در مورد اثر علیتی میانگین فارغ‌التحصیلی از دبیرستان بدانیم:

$$E[Y_i|X_i, S_i = 12] - E[Y_i|X_i, S_i = 11] = E[f_i(12)|X_i, S_i = 12] - E[f_i(11)|X_i, S_i = 11]$$

این مقایسه یک تفسیر علیتی دارد، زیرا با توجه به **CIA** داریم:

$$E[f_i(12)|X_i, S_i = 12] - E[f_i(11)|X_i, S_i = 11] = E[f_i(12) - f_i(11)|X_i, S_i = 12]$$

در اینجا، انتخاب اریب، میانگین تفاوت درآمد بالقوه فارغ‌التحصیلان دبیرستان و ترک تحصیل‌کنندگان است. با این حال، با توجه به **CIA**، فارغ‌التحصیلی از دبیرستان مستقل از درآمد بالقوه مشروط بر X_i است، بنابراین انتخاب اریب ناپدید می‌شود. همچنین توجه داشته باشید که در این مورد، اثر علیت فارغ‌التحصیلی از دبیرستان، اثر میانگین جامعه آماری فارغ‌التحصیلان دبیرستان است:

$$E[f_i(12) - f_i(11)|X_i, S_i = 12] = E[f_i(12) - f_i(11)|X_i].$$

این نکته مهم است، اما کمتر از حذف انتخاب اریب در (۱-۲-۳) اهمیت دارد.

تاکنون، ما اثرات علّی جداگانه را برای هر مقدار گرفته شده توسط متغیر شرطی X_i ایجاد کرده‌ایم که این تعداد اثرات علّی به برابری با تعداد مقادیر X_i منجر می‌شود که مایه خجالت ثروتمندان است. تجربیات، تقریباً همیشه نشان داده‌اند؛ تبدیل مجموعه‌ای از تخمین‌ها به یک معیار خلاصه واحد، مانند اثر علیت میانگین جامعه آماری، مفید است. طبق قانون امید ریاضی مکرر، اثر علیت میانگین جامعه آماری فارغ‌التحصیلان دبیرستان به صورت زیر است:

$$E\{E[Y_i|X_i, S_i = 12] - E[Y_i|X_i, S_i = 11]\} \quad (3.2.3)$$
$$= E\{E[f_i(12) - f_i(11)|X_i]\}$$
$$= E[f_i(12) - f_i(11)] \quad (3.2.4)$$

به همان صورت، ممکن است به اثر علیت میانگین فارغ‌التحصیلی از دبیرستان بر فارغ‌التحصیلان دبیرستان علاقمند باشیم:

$$E\{E[Y_i|X_i, S_i = 12] - E[Y_i|X_i, S_i = 11]|S_i = 12\} \quad (3.2.5)$$
$$= E\{E[f_i(12) - f_i(11)|X_i]|S_i = 12\}$$
$$= E[f_i(12) - f_i(11)|S_i = 12]. \quad (3.2.6)$$

این پارامتر به ما می‌گوید که فارغ‌التحصیلان دبیرستان پس از فارغ‌التحصیلی چه میزانی از درآمد را کسب کرده‌اند. به همین ترتیب، برای اثرات فارغ‌التحصیلی از کالج، تفاوتی بین اثر علیت میانگین فارغ‌التحصیلان کالج $E[f_i(16) - f_i(12)|S_i = 16]$ و اثر میانگین جامعه آماری $E[f_i(16) - f_i(12)]$ وجود دارد.

اثر میانگین جامعه آماری (۳-۲-۳) را می‌توان با میانگین گرفتن از همه اثرات خاص $\boldsymbol{X}$ با استفاده از توزیع نهایی X_i محاسبه کرد؛ در حالی که اثر میانگین در فارغ‌التحصیلان دبیرستان یا کالج، میانگین اثرات خاص $\boldsymbol{X}$ را با استفاده از توزیع X_i در این گروه‌ها تعیین می‌کند. در هر دو مورد، همتای تجربی، برآوردگر جورسازی[1] است: ما در گروه‌های تحصیلات فارغ‌التحصیلان برای افرادی با مقادیر یکسان از متغیر کمکی مقایسه انجام می‌دهیم، تفاوت درآمدشان را محاسبه می‌کنیم و سپس این تفاوت‌ها را به روشی میانگین می‌گیریم.

در عمل، جزئیات زیادی برای نگرانی در مورد اجرای یک استراتژی جورسازی وجود دارد. برخی از جزئیات فنی مکانیک جورسازی در بخش ۱-۳-۳ را در زیر ذکر می‌کنیم. در اینجا اشاره می‌کنیم که یک نقص جهانی در رویکرد جورسازی این است که این فرایند «به صورت خودکار» نیست، بلکه نیاز به دو مرحله دارد، جورسازی و میانگین گرفتن. برآورد خطاهای استاندارد برآوردهای حاصل نیز ممکن است واضح نباشد. ملاحظه دیگر این است که کنتراست دوطرفه در قلب این زیربخش (فارغ‌التحصیلان دبیرستان یا کالج در مقابل ترک تحصیل‌کنندگان) با مشکل مواجه است. از آنجا که S_i مقادیر بسیاری می‌گیرد، اثرات علّی میانگین جداگانه‌ای برای هر افزایش ممکن در S_i وجود دارد که باید به نحوی خلاصه شود. این ملاحظات ما را به رگرسیون هدایت می‌کند.[2]

---

1. Matching Estimator

۲. به عنوان مثال، می‌توانیم اثر میانگین بر s را با استفاده از توزیع S_i ایجاد کنیم. به عبارت دیگر، $E[f_i(s) - f_i(s-1)]$

رگرسیون یک راهبرد تجربی آسان برای استفاده است که به طور خودکار **CIA** را به اثرات علّی تبدیل می‌کند. از **CIA** تا رگرسیون دو مسیر را می‌توان بررسی کرد. یکی از مسیرها فرض می‌کند که $f_i(s)$ هم در s خطی است و هم برای همه بجز یک عبارت خطای جمعی یکسان است که در آن رگرسیون خطی یک ابزار طبیعی برای برآورد ویژگی‌های $f_i(s)$ است. یک مسیر کلی‌تر، اما تا حدودی طولانی‌تر درمی‌یابد که $f_i(s)$ تقریباً به طور قطعی برای افراد مختلف متفاوت است و علاوه بر این نیازی نیست در $f_i(s)$ خطی باشد. با این حال، با در نظر گرفتن تغییرات تصادفی در $f_i(s)$ برای افراد و برای غیرخطی برای فرد مورد نظر، می‌توان رگرسیون را به عنوان استراتژی برآورد میانگین وزنی تفاوت خاص فردی $f_i(s) - f_i(s-1)$ در نظر گرفت. در واقع، رگرسیون را می‌توان به عنوان نوع خاصی از برآوردگر جورسازی در نظر گرفت، که اثر علیت میانگین را بسیار شبیه به ۳-۲-۳ یا ۵-۲-۳ ثبت می‌کند.

در این مرحله می‌خواهیم بر شرایط مورد نیاز برای تفسیر علیتی داشتن رگرسیون تمرکز کنیم، نه جزئیات مشابه جورسازی[1] با رگرسیون. بنابراین با اولین مسیر که یک مدل علّی خطی اثرات ثابت است شروع می‌کنیم. فرض کنید که:

$$f_i(s) = \alpha + \rho s + \eta_i. \qquad (3.2.7)$$

این معادله علاوه بر خطی بودن می‌گوید که رابطه عملکردی مورد نظر برای همه یکسان است. باز هم، s بدون اندیس i برای نشان دادن افراد نوشته می‌شود، زیرا معادله (۷-۲-۳) به ما می‌گوید که شخص i برای هر مقدار از s و نه صرفاً مقدار تحقق‌یافته s_i کسب درآمد می‌کند. با این حال در این مورد تنها بخش خاص فردی و تصادفی $f_i(s)$، یک جزء خطای میانگین ـ صفر η_i است که فاکتورهای مشاهده نشده‌ای را که درآمد بالقوه را تعیین می‌کند ثبت می‌کند.

با جایگزین کردن مقدار مشاهده شده s_i با s در معادله (۷-۲-۳) داریم:

$$\mathrm{Y}_i = \alpha + \rho \mathrm{s}_i + \eta_i. \qquad (3.2.8)$$

معادله (۸-۲-۳) یک مدل رگرسیون دومتغیره به نظر می‌رسد، بجز اینکه معادله (۷-۲-۳) به طور صریح ضرایب (۸-۲-۳) را با یک رابطه علیتی مرتبط می‌کند. مهم اینکه چون معادله (۷-۲-۳) یک مدل علّی است، ممکن است s_i با نتایج بالقوه $f_i(s)$ یا در این مورد با عبارت پسماند (۸-۲-۳) η_i همبستگی داشته باشد.

اکنون فرض کنید که **CIA** دارای یک بردار از متغیرهای کمکی مشاهده شده X_i است. علاوه بر

را برای هر s با تطبیق و سپس محاسبه تفاوت میانگین برآورد کنید:

$$\sum E[f_i(s) - f_i(s-1)]P(s).$$

که در آن $P(s)$ تابع احتمال جرم برای s_i است. این یک تقریب گسسته به مشتق میانگین $E[f_i'(\mathrm{s}_i)]$ است.

1. Regression-matching Analog

فرض فرم عملکردی برای نتایج بالقوه مندرج در (۸-۲-۳)، بخش تصادفی درآمد بالقوه η_i را به یک تابع خطی با مشخصات قابل مشاهده X_i و یک عبارت خطای v_i تجزیه می‌کنیم:

$$\eta_i = X_i'\gamma + v_i,$$

که در آن γ یک بردار ضرایب رگرسیون جامعه آماری است که فرض بر این است که $E[\eta_i|X_i] = X_i'\gamma$ را برآورده می‌کند. از آنجا که γ توسط رگرسیون η_i در X_i تعریف شده است؛ پسماند v_i و X_i به خاطر *ساختشان*،[1] همبستگی ندارند. علاوه بر این به موجب **CIA** داریم:

$$E[f_i(s)|X_i, s_i] = E[f_i(s)|X_i] = \alpha + \rho s + E[\eta_i|X] = \alpha + \rho s + X_i'\gamma$$

از آنجا که استقلال میانگین مستلزم متعامد بودن است، پسماند در مدل خطی علّی زیر:

$$Y_i = \alpha + \rho s_i + X_i'\gamma + v_i \qquad (3.2.9)$$

با رگرسورهای s_i و X_i همبستگی ندارد و ضریب رگرسیون ρ اثر علّی مورد نظر ماست. جا دارد بار دیگر تأکید کنیم که فرض کلیدی در اینجا این است که مشخصات قابل مشاهده X_i تنها همبستگی η_i و s_i (معادل $f_i(s)$ و s_i) هستند. این فرض انتخاب از قابل مشاهده‌ها[2] برای مدل‌های رگرسیون بیش از ۲۵ سال پیش توسط باروانو، کین و گلدبرگر (سال ۱۹۸۱) مورد بحث قرار گرفته و همچنان اساس بیشتر کارهای تجربی در اقتصاد است.

۲-۲-۳. فرمول اریب متغیرهای حذف شده

فرمول اریب متغیرهای حذف شده (**OVB**[3]) رابطه بین تخمین رگرسیون در مدل‌های با مجموعه‌های مختلف متغیرهای کنترل را بیان می‌کند. این فرمول مهم اغلب با این تصور شکل می‌گیرد که رگرسیون طولانی‌تر (بلندتر)، که دارای متغیرهای کنترل بیشتری مثل معادله ۹-۲-۳ است، تفسیری علّی دارد، در حالی که رگرسیون کوتاه‌تر چنین نیست. بنابراین، این طور گفته می‌شود که ضرایب متغیرها در رگرسیون کوتاه‌تر دارای اریب هستند. در حقیقت، فرمول **OVB** یک ارتباط مکانیکی بین بردارهای ضرایب است که بر رگرسیون کوتاه و بلند اعمال می‌شود، چه این رگرسیون بلندتر علّی باشد یا نباشد. با این حال، ما به طور سنتی، به تفاوت بین ضرایب موجود در رگرسیون بلند و رگرسیون کوتاه معین شده توسط فرمول **OVB** اشاره می‌کنیم.

به منظور انسجام بخشیدن به این بحث، فرض کنید مجموعه متغیرهای کنترل مربوطه در رگرسیون تحصیل بتواند ترکیبی از سابقه خانوادگی، هوشمندی و انگیزه باشد. فرض کنید این عوامل خاص توسط

1. By Construction
2. Selection-on-observables Assumption
3. Omitted Variables Bias

برداری مشخص شود، A_i که ما آن را «توانایی»[1] می‌نامیم. رگرسیون دستمزدها بر تحصیل، S_i، کنترل‌کننده این «توانایی»، می‌تواند به صورت زیر نوشته شود:

$$Y_i = \alpha + \rho S_i + A_i'\gamma + \varepsilon_i, \qquad (3.2.10)$$

در این رابطه α، ρ، و γ ضرایب رگرسیونی جامعه و ε_i پسماند رگرسیون است که بنابر تعریف، با هیچکدام از رگرسورها همبستگی ندارد. اگر فرض استقلال شرطی (CIA) با وجود A_i مفروض اعمال شود، ρ می‌تواند با ضریب مدل علّی خطی ۷-۲-۳ مساوی باشد، در حالی که پسماند ε_i بخش تصادفی درآمد بالقوه است که پس از کنترل A_i پسماند می‌شود.

در عمل، «توانایی» به سختی اندازه‌گیری می‌شود. به عنوان مثال، پیمایش جاری جامعه آمریکا (CPS[2])، مجموعه وسیعی از داده‌های به کار رفته در اقتصاد خُرد کاربردی (و منبع داده‌های دولت آمریکا درباره نرخ بیکاری)، به ما درباره سابقه خانوادگی پاسخگویان بزرگسال، هوشمندی، یا انگیزش هیچ چیز نمی‌گوید. نتایج باقی گذاشتن توانایی در خارج از دسترس رگرسیون چیست؟ (۱۰-۲-۳). ضریب منتج رگرسیون کوتاه مرتبط با ضریب رگرسیون بلند در معادله ۱۰-۲-۳ به صورت زیر است:

$$\frac{Cov(Y_i, S_i)}{V(S_i)} = \rho + \gamma'\delta_{As}, \qquad (3.2.11)$$

که در آن δ_{A_s} بردار ضرایب حاصل از رگرسیون اجزای A_i روی S_i است. فرمول OVB به این معناست که:

رگرسیون کوتاه مساوی است با رگرسیون بلند به علاوه تأثیر متغیر حذف شده، ضرب در رگرسیون حذف شده روی متغیرهای در نظر گرفته شده (شامل شده) است.

این فرمول به سهولت استخراج می‌شود: رگرسیون بلند را در فرمول رگرسیون کوتاه قرار دهید: $\frac{Cov(Y_i,S_i)}{V(S_i)}$. تعجب‌برانگیز نیست که، فرمول OVB به طور نزدیکی با فرمول آناتومی رگرسیون، ۳-۱-۳، از فصل ۲-۱-۳ مرتبط است. هر دو فرمول OVB و آناتومی رگرسیون به ما می‌گویند که ضرایب رگرسیون کوتاه و بلند هروقت متغیرهای حذف شده و شامل شده ناهمبسته باشند، یکسان هستند.[3]

ما می‌توانیم فرمول OVB را برای به دست آوردن حسی از پیامدهای احتمالی حذف توانایی برای ضرایب تحصیلی به کار ببریم. متغیرهای توانایی، دارای تأثیر مثبت بر دستمزدها هستند و این متغیرها احتمالاً رابطه مثبتی با تحصیل دارند. ضریب رگرسیون کوتاه ممکن است در رابطه با آنچه می‌خواهیم خیلی بزرگ باشد. از طرف دیگر، به عنوان یک نظریه اقتصادی، رابطه بین تحصیل و توانایی به طور

1. Ability
2. American Current Population Survey
۳. يك تعميم چندمتغيره از OVB به اين صورت است: فرض كنيد كه β_1^s بردار ضريبي از يك بردار متغيرهاي $K_1 \times 1$ باشد، X_{1i} در يك رگرسيون كوتاه كه هيچ متغير ديگري ندارد و β_1^l بردار ضريب اين متغيرها در يك رگرسيون بلند كه شامل بردار $K_2 \times 1$ از متغيرهاي كنترل بوده و X_{2i} بردار ضريب β_2^l باشد. خواهيم داشت: $\beta_1^s = \beta_1^l + E[X_{1i}X_{1i}']^{-1}E[X_{1i}X_{2i}']\beta_2^l$.

کامل روشن نیست. برخی متغیرهای حذف شده ممکن است با تحصیل رابطه منفی داشته باشند، که در آن صورت ضریب رگرسیون کوتاه خیلی کوچک خواهد بود.[۱]

جدول ۱-۲-۳ این نقاط را با استفاده از داده‌های پیمایش طولی[۲] ملی جوانان (NLSY[۳]) نشان می‌دهد. سه ورودی ابتدایی در این جدول نشان می‌دهند که ضریب تحصیل از ۰/۱۳۲ تا ۰/۱۱۴ وقتی متغیرهای سابقه خانوادگی ـ در این مورد، تحصیلات والدین ـ همانند برخی از خصوصیات جامعه (مثل سن، نژاد، ناحیه سرشماری ساکنان) به عنوان متغیر کنترل در نظر گرفته شده‌اند، کاهش می‌یابد. کنترل بیشتر در توانایی فردی، همانند استفاده از پروکسیِ آزمون کیفیت نیروهای نظامی (AFQT[۴])، ضریب تحصیل را تا ۰/۰۸۷ کاهش می‌دهد (AFQT توسط ارتش جهت انتخاب سربازان استفاده می‌شود). فرمول اریب متغیرهای حذف شده به ما می‌گوید که این کاهش‌ها در نتیجه این حقیقت است که کنترل اضافی رابطه مثبتی با دستمزد و تحصیل دارد.[۵]

جدول ۱-۲-۳. تخمین رگرسیون بازده درآمدی به تحصیل برای مردان در NLSY

	(1)	(2)	(3)	(4)	(5)
Controls:	None	Age dummies	Col. (2) and additional controls*	Col. (3) and AFQT score	Col. (4), with occupation dummies
	0.132 (0.007)	0.131 (0.007)	0.114 (0.007)	0.087 (0.009)	0.066 (0.010)

نکته: داده‌ها برگرفته از سرشماری ملی طولی جوانان (کوهورت ۱۹۷۹، سرشماری ۲۰۰۲) است. این جدول ضریب سال‌های تحصیل در رگرسیون دستمزدها در سال تحصیل و کنترل‌های مشهود را گزارش می‌دهد. خطاهای استاندارد در پرانتز نشان داده شده‌اند. نمونه محدود به مردان و توسط نمونه‌گیری **NLSY** وزن‌بندی شده است. اندازه نمونه ۲۴۳۴ نفر است.

*کنترل‌های اضافی سال‌های تحصیل مادران و پدران و متغیرهای موهومی در ناحیه نژادی و سرشماری است.

با وجود سادگی، فرمول **OVB** یکی از مهم‌ترین نکاتی است که باید درباره رگرسیون بدانیم. اهمیت فرمول **OVB** از این حقیقت ناشی شده که اگر شما مدعی شوید که اریب متغیرهای حذف شده وجود ندارد، پس به طور طبیعی مدعی این گزاره هستید که رگرسیون حاصل، آن چیزی است که به

۱. به عنوان افرادي با تحصیلات عالیه، ما دوست داریم تصور کنیم که توانایي و تحصیل با همدیگر همبستگي مثبت دارند. هرچند، این یك پیش‌بیني مسلم نیست، چراكه مایك جاگر از مدرسه اقتصاد لندن و بیل گیتس از هاروارد خارج شدند، شاید به این خاطر که هزینه فرصت ادامه تحصیل براي این افراد با توانایي زیاد، بالا بود (هرچند، آنها ممکن است در زمره خوش‌شانس‌ترین خارج‌شدگان از تحصیل باشند).

۲. توضیح مترجم: مطالعه طولی به دسته‌ای از مطالعات مشاهده‌ای گفته می‌شود که در آن گروه مورد مطالعه، در طول زمان، مورد مطالعه قرار می‌گیرد. مطالعات طولی در مقابل مطالعات مقطعی مطرح می‌شوند.

3. National Longitudinal Survey of Youth

4. Armed Forces Qualification Test

۵. ادبیات وسیعي، موضوع نتایج حذف متغیرهاي توانایي از معادلات تحصیل را بررسي مي‌کنند. ارجاعات اساسي عبارتند از: گریلیچز و میسون (۱۹۷۲)، توبمن (۱۹۷٦)، گریلیچز (۱۹۷۷) و چمبرلین (۱۹۷۸).

دنبالش هستید و البته رگرسیونی که به دنبال آن هستید اغلب دارای تفسیر علّی است. به عبارت دیگر، برای تفسیر علّی تخمین رگرسیون بلند به فرض استقلال شرطی (**CIA**) اتکا می‌کنید.

شایان توجه است که **CIA** به احتمال قوی مبنای پذیرفتنی برای کار تجربی ارائه می‌کند. بهترین سناریو، اختصاص تصادفی s_i، مشروط به X_i، در بعضی انواع تجربیات (احتمالاً تجربیات طبیعی) است. یک نمونه از بررسی برنامه بازآموزی اجباری برای کارگران بیکار توسط بلاک و همکاران (۲۰۰۳) انجام شده است. پژوهشگران این تحقیق علاقمند بودند که بدانند آیا برنامه بازآموزی در افزایش درآمدها پس از آن موفق بوده است یا خیر. آنان از این حقیقت استفاده کردند که صلاحیت پذیرش در این برنامه تربیتی، بر اساس خصوصیات فردی و سابقه بیکاری و اشتغال مشخص شده است. کارگران به گروه‌هایی بر اساس خصوصیات خود تقسیم‌بندی شدند. در حالی که برخی از این گروه‌های کارگری برای آموزش صلاحیت نداشتند، کسانی که در گروه‌های دیگر بودند اگر شغلی به دست نمی‌آوردند باید این آموزش‌ها را دریافت می‌کردند. وقتی بعضی از گروه‌های آموزشی اجباری، کارگران بیشتری نسبت به بازه زمان آموزشی اختصاص داده شده داشتند، فرصت‌های آموزشی توسط قرعه‌کشی توزیع شده است. بنابراین، الزامات آموزشی به طور تصادفی اختصاص داده می‌شد که مشروط به متغیرهای مشترک به کار رفته جهت تخصیص کارگران به گروه‌های آموزشی بوده است. رگرسیون بر روی یک متغیر موهومی برای آموزش، به علاوه خصوصیات فردی، متغیرهای سابقه بیکاری و متغیرهای گذشته شغلی برای طبقه‌بندی کارگران، به احتمال قوی تخمین قابل اعتمادی از تأثیر علّی آموزش به دست می‌دهد.[۱]

در زمینه تحصیلی، اغلب هیچ قرعه‌کشی که مستقیماً معین کند آیا فردی به دانشگاه می‌رود یا دبیرستان را تمام می‌کند، وجود ندارد.[۲] با این حال، ما ممکن است تصور کنیم که افراد با قابلیت‌های یکسان و سوابق خانوادگی مشابه در برابر تجربه‌ای که مشوق افراد به تحصیل در مدرسه است قرار گیرند. کمک مالی تحصیلی، که در برخی مناطق انگلیس به دانش‌آموزان مدارس داده می‌شود، یکی از تجربیات در این خط‌مشی است. (دیردن و همکاران، ۲۰۰۴)

نوع دوم بررسی که به نفع **CIA** است از اطلاعات جزئیات نهادی، درباره روندی استفاده می‌کند که تعیین‌کننده s_i است. یک نمونه از آن بررسی آنگریست (۱۹۹۸) درباره تأثیر خدمات نظامی داوطلبانه بر درآمد بعدی سربازان است. این تحقیق می‌خواهد بداند که آیا مردانی که داوطلب خدمت در نیروهای نظامی آمریکا هستند از نظر اقتصادی در بلندمدت شرایط بهتری داشته‌اند یا خیر؟ از آنجا که خدمت داوطلبانه سربازی به طور تصادفی تخصیص داده نمی‌شود، ما هرگز نمی‌توانیم از این مسئله مطمئن شویم. آنگریست سپس از تکنیک‌های جورسازی و رگرسیون برای کنترل تفاوت مشهود بین کهنه‌سربازان و غیرکهنه‌سربازان که برای ورود به نیروهای نظامی بین سال‌های ۱۹۷۹ و ۱۹۸۲ اقدام کردند، استفاده

---

۱. ظاهراً این برنامه درآمدها را افزایش می‌دهد، چراکه افرادی که در گروه‌های آموزشی بودند، سریع‌تر به کار برمی‌گشتند.
۲. از قرعه‌کشی ـ لوتاری ـ برای توزیع یارانه‌های شهریه مدارس استفاده می‌شده است، به عنوان مثال نگاه کنید به انگریست و همکاران (۲۰۰۲).

کرده است. انگیزه راهبرد کنترل در این مورد حقیقتی است که ارتش در ابتدا تقاضای سربازان را بر اساس متغیرهای مشترک مثل سن، تحصیلات و امتیاز آزمون غربالگری می‌کند.

CIA در مورد آنگریست (۱۹۹۸) در حدی است که ادعا می‌شود پس از مقید کردن کلیه این خصوصیات مشهود، کهنه‌سربازان و غیرکهنه‌سربازان قابل مقایسه است. این فرض به نظر می‌رسد ارزش پذیرش داشته باشد، زیرا شرایط X_i تغییرات در وضعیت سربازان در بررسی آنگریست (۱۹۹۸) فقط از این حقیقت سرچشمه می‌گیرد که بعضی متقاضیان صلاحیت‌دار در دقیقه آخر در فراخوان خدمت سربازی قرار نمی‌گیرند. البته، ملاحظاتی که منجر می‌شود متقاضیان صلاحیت‌دار در روند فراخوان نادیده گرفته شوند، می‌تواند مربوط به پتانسیل درآمدزایی باشد، بنابراین **CIA** به روشنی حتی در این مورد تضمین نمی‌شود.

۳-۲-۳. کنترل بد

ما به این نکته پرداختیم که کنترل متغیرهای مشترک می‌تواند احتمال **CIA** را بیشتر کند، اما کنترل بیشتر همیشه بهتر نیست. بعضی متغیرها کنترل بد هستند و حتی وقتی که شمول آنها احتمالاً تغییری در ضرایب رگرسیون کوتاه می‌دهند و باید در مدل رگرسیون قرار گیرند. کنترل‌های بد متغیرهایی هستند که به خودی خود متغیرهای خروجی تجربه مفهومی آماده هستند. یعنی اینکه، کنترل‌های بد ممکن است متغیرهای وابسته نیز باشند. کنترل‌های خوب متغیرهایی هستند که می‌توان به عنوان ثابت در زمان تعیین رگرسیون مورد نظر تصور کرد.

مفهوم مشکل کنترل بد نسخه‌ای از اریب گزینش است، البته چیزی نامحسوس‌تر از اریب گزینش که در فصل ۲ و بخش ۳-۲ بحث شده است. برای نشان دادن این امر، فرض کنید ما به تأثیر درجه دانشگاهی در درآمدزایی علاقمندیم و اینکه افراد می‌توانند در یکی از دو شغل فعالیت کنند، یقه‌سفید و یقه‌آبی. درجه دانشگاهی به روشنی در را برای شغل‌های با پرداخت بالاتر (یقه‌سفیدی) باز می‌کند. آیا اشتغال به عنوان متغیر حذف شده در رگرسیون دستمزدها در تحصیلات باید دیده شود؟ به طور کلی، اشتغال با تحصیلات و پرداخت، رابطه مستقیم دارد. شاید بهتر است تأثیر دانشگاه بر دستمزدهای کسانی با یک شغل، مثل یقه‌سفیدها را ملاحظه کرد. مشکل این استدلال این است که وقتی ما حقیقتی را می‌پذیریم که دانشگاه بر اشتغال تأثیر دارد، حتی اگر تکمیل درجه دانشگاهی به طور تصادفی در نظر گرفته شود، مقایسه دستمزدها به واسطه درجه دانشگاهی در اشتغال دیگر یک به یک نیست.

در اینجا یک نمایش از مسئله کنترل بد در مثال دانشگاه/ اشتغال آورده می‌شود.[1] فرض کنید W_i یک متغیر موهومی است که کارکنان یقه‌سفید را نشان می‌دهد و Y_i نشانه درآمد باشد. تحقق این متغیرها توسط وضعیت فارغ‌التحصیلی از دانشگاه و دستاوردهای بالقوه که در برابر C_i نمایه شدند، تعیین می‌شود. خواهیم داشت:

۱. همین مشکل در مقایسه‌هاي «مشروط به مثبت» كه در بخش ۲-۴-۳ در ادامه بررسي مي‌شود وجود دارد.

$$\begin{aligned} Y_i &= C_i Y_{1i} + (1 - C_i) Y_{0i} \\ W_i &= C_i W_{1i} + (1 - C_i) W_{0i} \end{aligned}$$

که در آنها $C_i = 1$ برای فارغ‌التحصیلان دانشگاه صفر است در غیر این صورت، $\{Y_{1i}, Y_{0i}\}$ نشانه درآمد بالقوه است، و $\{W_{1i}, W_{0i}\}$ نشانه وضعیت بالقوه یقه‌سفیدهاست. فرض می‌کنیم که C_i به طور تصادفی تخصیص یافته و بنابراین مستقل از کلیه درآمدهای بالقوه است. مشکلی با تأثیر علّی بر C_i روی Y_i یا W_i نداریم، زیرا عدم وابستگی به ما شرایط زیر را می‌دهد:

$$\begin{aligned} E[Y_i|C_i = 1] - E[Y_i|C_i = 0] &= E[Y_{1i} - Y_{0i}], \\ E[W_i|C_i = 1] - E[W_i|C_i = 0] &= E[W_{1i} - W_{0i}]. \end{aligned}$$

در عمل، ممکن است این تأثیرات متوسط را توسط رگرسیون Y_i و W_i و روی C_i تخمین بزنیم.

کنترل بد به معنای مقایسه درآمدها به شرط آن است که W_i هیچ تفسیر علّی نداشته باشد. تفاوت درآمدهای متوسط بین فارغ‌التحصیلان دانشگاه و سایرین را مشروط به کار در یک شغل یقه‌سفید را ملاحظه کنید. می‌توانیم آن را با یک مدل رگرسیونی که شامل W_i یا با رگرسیون Y_i روی C_i در یک نمونه که $W_i = 1$ است محاسبه کنیم. تخمین‌ها در مورد دوم برابر تفاوت در متوسط‌ها با صفر و یک کردن C_i، مشروط به $W_i = 1$ است:

$$E[Y_i|W_i = 1, C_i = 1] - E[Y_i|W_i = 1, C_i = 0] = E[Y_{1i}|W_{1i} = 1, C_i = 1] - E[Y_{0i}|W_{0i} = 1, C_i = 0] \quad (3.2.12)$$

با عدم وابستگی در $\{Y_{1i}, W_{1i}, Y_{0i}, W_{0i}\}$ و C_i ، خواهیم داشت:

$$E[Y_{1i}|W_{1i} = 1, C_i = 1] - E[Y_{0i}|W_{0i} = 1, C_i = 0] = E[Y_{1i}|W_{1i} = 1] - E[Y_{0i}|W_{0i} = 1].$$

این عبارت، طبیعت ناهمگون مسئله کنترل بد را نشان می‌دهد:

$$\begin{aligned} & E[Y_{1i}|W_{1i} = 1] - E[Y_{0i}|W_{0i} = 1] \\ &= \underbrace{E[Y_{1i} - Y_{0i}|W_{1i} = 1]}_{\text{تاثیر علّی روی درجات دانشگاهی}} + \underbrace{\{E[Y_{0i}|W_{1i} = 1] - E[Y_{0i}|W_{0i} = 1]\}}_{\text{اریب گزینش}}. \end{aligned}$$

به عبارت دیگر، تفاوت دستمزدها بین کسانی با/و بدون درجه دانشگاهی مشروط به کار در شغل یقه‌سفیدی معادل تأثیر علّی دانشگاه بر کسانی است که $W_{1i} = 1$ (افرادی که وقتی دارای درجه دانشگاهی هستند در شغل یقه‌سفیدی کار می‌کنند) و یک عبارت اریب گزینش که نشانه حقیقتی است که دانشگاه ترکیب کارکنان یقه‌سفید را عوض می‌کند.

اریب گزینش در این زمینه می‌تواند مثبت یا منفی باشد و بستگی به رابطه بین انتخاب شغل، حضور در دانشگاه و درآمد بالقوه دارد. نکته اصلی این است که حتی اگر $Y_{1i} = Y_{0i}$، به طوری که هیچ تأثیر علّی دانشگاه بر دستمزدها وجود نداشته باشد، مقایسه مشروط در ۱۲-۲-۳ به ما این را نمی‌گوید (رگرسیون Y_i روی W_i و C_i کاملاً مشکل مشابهی دارد). همچنین درست نیست که گفته

شود مقایسه مشروط بخشی از تأثیر دانشگاه را دربرمی‌گیرد که توسط اشتغال بیان نمی‌شود. در حقیقت، مقایسه مشروط به ما چیز زیادی نمی‌گوید که بدون مدل ارتباط بین دانشگاه، اشتغال و درآمد بتواند مفید باشد.[۱]

به عنوان یک نمایش تجربی، مشاهده می‌کنیم که افزودن اشتغال دورقمی موهومی در واقع ضریب تحصیلات را در مدل‌های **NLSY** که در جدول ۱-۲-۳ گزارش شده، کاهش می‌دهد، که در این مورد از ۰/۰۸۷ تا ۰/۰۶۶ است. در هر حال، به سختی می‌توان گفت از این تنزل چه استفاده‌ای می‌بریم. وقتی اشتغال موهومی را افزایش می‌دهیم، تغییر در ضرایب تحصیلات، ممکن است به طور ساده یک اریب گزینشی موهومی باشد. بنابراین بهتر است فقط متغیرهایی کنترل شوند که به خودی خود به واسطه تحصیلات ایجاد نشده باشند.

نسخه دوم از سناریوی کنترل بد شامل کنترل پروکسی است، یعنی شمول متغیرهایی که ممکن است به طور نسبی عوامل حذف شده را کنترل کنند، اما به خودی خود تحت تأثیر متغیرهای مطلوب قرار دارند. یک نسخه ساده از سناریوی کنترل پروکسی به شکل زیر تداوم دارد: فرض کنید مطلوب شما رگرسیون بلند مشابه معادله ۱۰-۲-۳ باشد:

$$\mathrm{Y}_i = \alpha + \rho \mathrm{S}_i + \gamma a_i + \varepsilon_i, \qquad (3.2.13)$$

به این منظور، ما بردار کنترل‌های $\boldsymbol{A_i}$ را با قابلیت اسکالر ارزیابی $\boldsymbol{a_i}$ جایگزین کردیم. این کمیت را به عنوان امتیاز **IQ** تصور کنید که قابلیت ذاتی را در مقطع تحصیلی هشتم قبل از هر انتخاب تحصیلی مربوطه ارزیابی می‌کند (با فرض اینکه هر کس مقطع تحصیلی هشتم را تکمیل می‌کند). جمله خطا در این معادله با رابطه $E[\mathrm{S}_i\varepsilon_i] = E[a_i\varepsilon_i] = 0$ مطابقت دارد. از آنجا که $\boldsymbol{a_i}$ قبل از تعیین $\boldsymbol{S_i}$ اندازه‌گیری می‌شود، این کنترل خوبی است.

معادله ۱۳-۲-۳ رگرسیون مطلوب است، اما متأسفانه، داده‌های $\boldsymbol{a_i}$ در دست نیستند. در هر حال، شما دارای قابلیت دوم در ارزیابی جمع‌آوری و پس از تکمیل تحصیلات هستید (می‌توان گفت، امتیاز در تست به کار رفته در غربالگری متقاضیان شغل). این متغیر را قابلیت تأخیری[۲] بنامید، $\boldsymbol{a_{li}}$. به طور کلی، تحصیلات افزاینده قابلیت تأخیری نسبت به قابلیت ذاتی است. برای مشخص کردن آن، فرض کنید:

$$a_{li} = \pi_0 + \pi_1 \mathrm{S}_i + \pi_2 a_i. \qquad (3.2.14)$$

توسط آن، مقصود ما این است که بگوییم هم تحصیلات و هم قابلیت ذاتی افزاینده قابلیت تأخیری

۱. در این مثال، اریب گزینش احتمالاً منفي است به این صورت که $E[\mathrm{Y}_{0i}|\mathrm{W}_{1i} = 1] < E[\mathrm{Y}_{0i}|\mathrm{W}_{0i} = 1]$. منطقي به نظر مي‌رسد که هر فارغ‌التحصیل دانشگاه بتواند يك شغل یقه‌سفیدي بگیرید، به این ترتیب، $E[\mathrm{Y}_{0i}|\mathrm{W}_{1i} = 1]$ خیلي از $E[\mathrm{Y}_{0i}]$ دور نیست. اما اگر فردي بدون داشتن مدرك دانشگاهي، یعني $\mathrm{W}_{0i} = 1$ شغل یقه‌سفید مي‌گیرد، این بدان معناست که Y_{0i} بهتر از متوسط دارد.

2. Late Ability

یا ارزیابی شده است. تقریباً حتمی است که مقداری تصادفی بودن در قابلیت اندازه‌گیری شده وجود دارد، اما با این حال ما می‌توانیم نظر خود را از طریق ارتباط قطعی[1] ساده‌تر کنیم (۱۴-۲-۳).

از **OVB** در رگرسیون $\boldsymbol{Y_i}$ روی $\boldsymbol{S_i}$ تعجب می‌کنید، پس فرض می‌کنید رگرسیون $\boldsymbol{Y_i}$ روی $\boldsymbol{S_i}$ و قابلیت تأخیری، $\boldsymbol{a_{li}}$ باشد، زیرا کنترل مطلوب، $\boldsymbol{a_i}$ موجود نیست. با استفاده از ۱۴-۲-۳ و جایگزینی برای $\boldsymbol{a_i}$ در ۱۳-۲-۳، رگرسیون روی $\boldsymbol{S_i}$ و $\boldsymbol{a_{li}}$ عبارت است از:

$$Y_i = (\alpha - \gamma\frac{\pi_0}{\pi_2}) + (\rho - \gamma\frac{\pi_1}{\pi_2})S_i + \frac{\gamma}{\pi_2}a_{li} + \varepsilon_i. \tag{3.2.15}$$

در این سناریو، $\boldsymbol{\gamma}$ ، $\boldsymbol{\pi_1}$ و $\boldsymbol{\pi_2}$ مثبت هستند، بنابراین $\boldsymbol{\rho - \gamma\frac{\pi_1}{\pi_2}}$ بسیار کوچک است، مگر اینکه $\boldsymbol{\pi_1}$ در حد صفر باشد. به عبارت دیگر، استفاده از کنترل پروکسی که با متغیر مطلوب افزایش می‌یابد ضریبی پایین‌تر از تأثیر مطلوب ایجاد می‌کند. مهم‌تر از این، $\boldsymbol{\pi_1}$ می‌تواند تا حدی مورد تحقیق قرار گیرد: اگر رگرسیون $\boldsymbol{a_{li}}$ روی $\boldsymbol{S_i}$ صفر باشد، ممکن است درباره فرض اینکه $\boldsymbol{\pi_1}$ صفر باشد احساس بهتری داشته باشید (۱۴-۲-۳).

در اینجا ابهام جالبی در قضیه کنترل پروکسی وجود دارد که در قضیه کنترل بد اولیه وجود ندارد. کنترل متغیرهای نتیجه به طور ساده گمراه‌کننده است: اگر رگرسیون باید تفسیری علّی داشته باشد و شما نمی‌خواهید کنترلی بر اشتغال در رگرسیون تحصیلات داشته باشید. در قضیه کنترل پروکسی، زمانی که کنترل پروکسی ایجاد ضریب رگرسیون مطلوب نمی‌کند، ممکن است بهبودی در عدم کنترل کل آن باشد. به یاد آورید که انگیزه کنترل پروکسی معادله ۱۳-۲-۳ است. بر حسب پارامترهای این معادله، فرمول **OVB** به ما می‌گوید که رگرسیون روی $\boldsymbol{S_i}$ بدون کنترل ایجاد ضریب $\boldsymbol{\rho + \gamma\delta_{as}}$ می‌کند، که $\boldsymbol{\delta_{as}}$ ضریب شیب حاصل از رگرسیون $\boldsymbol{a_i}$ روی $\boldsymbol{S_i}$ است. ضریب تحصیلات در ۱۵-۲-۳ ممکن است نزدیک‌تر به $\boldsymbol{\rho}$ نسبت به ضریبی باشد که شما بدون کنترل تخمین زده باشید. علاوه بر این، با فرض مثبت بودن $\boldsymbol{\delta_{as}}$، شما می‌توانید با اطمینان بگویید که تأثیر علّی مطلوب بین این دو قرار دارد.

یکی از خصوصیات کنترل بد و کنترل پروکسی این است که وقتی درباره کنترل فکر می‌کنیم، زمان اهمیت دارد. متغیرهای اندازه‌گیری شده قبل از اینکه متغیرهای مطلوب تعیین شده باشند معمولاً کنترل خوبی هستند. خصوصاً، به این علت که متغیرها قبل از متغیرهای مطلوب تعیین شده‌اند و نمی‌توانند به خودی خود نتیجه‌ای در رابطه علّی باشند. در بسیاری موارد، در هر حال، زمان نامطمئن و ناشناخته است. در چنین مواردی، استدلال روشن درباره کانال‌های علّی مستلزم فرض صریح درباره آن چه در ابتدا اتفاق افتاده یا ادعایی که هیچ یک از متغیرهای کنترل به خودی خود توسط رگرسیون مطلوب موجب نشده‌اند، است.[2]

---

1. Deterministic Link

۲. گریلیچز و میسون (۱۹۷۲) منبع اصلي بررسي استفاده از کنترل‌هاي مقدم و مؤخر در روابط مرتبط با تحصیل است. براي مطالعات بیشتر مرتبط به چمبرلین (۱۹۷۷ و ۱۹۷۸) مراجعه کنید. روزنبام (۱۹۸۴) بحث دیگري در مورد ایده کنترل پروکسي با استفاده از نشانه‌گذاري بسیار متفاوت، خارج از چارچوب رگرسیون، مطرح مي‌کند.

۳-۳. ناهمگنی[1] و خطی نبودن

همان طوری که در فصل قبلی مشاهده کردیم، یک مدل علّی خطی در ترکیب با **CIA** به **CEF** خطی با تفسیر علّی منجر می‌شود. با فرض اینکه **CEF** خطی است، رگرسیون جامعه محسوب می‌شود. در عمل، در هر حال، فرض **CEF** خطی در حقیقت برای تفسیر علّی رگرسیون لازم نیست. به خاطر موضوعی که در بخش ۲-۱-۳ بحث شد، می‌توانیم تصور کنیم که رگرسیون $\boldsymbol{Y}_{\mathbf{i}}$ روی $\boldsymbol{X}_{\boldsymbol{i}}$ و $\boldsymbol{S}_{\boldsymbol{i}}$ بهترین تقریب خطی را برای **CEF** بدون توجه به شکل آن فراهم می‌کند. بنابراین، اگر **CEF** علّی باشد، این حقیقت که رگرسیون آن را تقریبی می‌کند به ضریب رگرسیون حالتی علّی می‌دهد. این ادعا اندکی مبهم است. در هر حال طبیعت ارتباط بین رگرسیون و **CEF** ارزش بررسی بیشتری دارد. این بررسی به درک رگرسیون به عنوان تخمین‌زننده جالب محاسباتی و مقایسه‌ای منجر می‌شود.

۱-۳-۳. ملاقات رگرسیون و جورشدگی[2] (جورسازی)

یک یا دو دهه قبل شاهد افزایش تمایلات به عنوان ابزاری تجربی بوده است. انگیزه جورسازی، عنوان راهبرد کنترل برای متغیرهای مشترک توسط **CIA**، همچنین برای رگرسیون علّی در فصل قبلی بحث شد. به عنوان مثال، آنگریست (۱۹۹۸) از جورسازی برای تخمین تأثیرات داوطلبانه جهت خدمات نظامی در درآمدهای بعدی سربازان استفاده کرده است. این جورسازی تخمینی دارای تفسیر علّی است که، مشروط به خصوصیات فردی است که ارتش برای انتخاب سربازان به کار می‌برد (سن، تحصیلات، امتیاز آزمون) و وضعیت سربازان وابسته به درآمدهای بالقوه است.

تصویر جالب از راهبردهای جورسازی این است که آنها همراه با بیان صریح فرض وابستگی مشروط هستند که برای ارائه تخمین جورسازی به صورت تفسیر علّی لازم است. به طور هم‌زمان، ما مشاهده کردیم که تفسیر علّی ضریب رگرسیون بر اساس فرض کاملاً مشابهی صورت می‌گیرد. به عبارت دیگر، جورسازی و رگرسیون هر دو راهبرهای کنترلی هستند. از آنجا که فرض اصلی استنتاج علّی برای هر دو راهبرد یکسان است، ارزش دارد که بررسی شود که آیا یا در چه حدی جورسازی با رگرسیون تفاوت دارد. دیدگاه ما این است که رگرسیون می‌تواند به عنوان ابزار محاسباتی برای نوع خاصی از تخمین‌کننده وزنی جورسازی به کار گرفته شود و بنابراین تفاوت بین رگرسیون و جورسازی احتمالاً اهمیت تجربی اساسی ندارد.

برای گسترش این عقیده، مشاهده عمیق‌تر در ساختار ریاضی تخمین جورسازی و رگرسیون، مثلاً کمیت جامعه که این روش‌ها در تلاش برای تخمین هستند، مفید است. در مورد رگرسیون، البته، تخمین، یک بردار ضریب‌های رگرسیون جامعه است. تخمین جورسازی نوعاً یک میانگین وزنی خاص از تمرکز یا مقایسه در میان سلول‌های تعریف شده توسط متغیرهای مشترک است. این موضوع در مورد

---

1. Heterogeneity
2. Matching

متغیرهای مشترک گسسته به راحت‌ترین شکل ممکن دیده می‌شود، که نمونه آن خدمت سربازی است که در آن برای رگرسور گسسته وضعیت سربازان، در اینجا با نماد موهومی D_i نشان داده شده است. از آنجا که تیمار کار فقط در دو کمیت است، می‌توانیم از نمادهای $Y_i = f_i(1)$ و $Y_{0i} = f_i(0)$ برای نشان دادن نتایج بالقوه استفاده کنیم. یک پارامتر مورد نظر در این زمینه تأثیر متوسط تیمار در تیمار شده است، $E[Y_{1i} - Y_{0i}|D_i = 1]$. این قضیه به ما می‌گوید که تفاوت بین درآمد متوسط سرباز، $\mathrm{E}[Y_{1i}|D_i = 1]$، که یک کمیت مشهود است و درآمد متوسط واقع نشده که می‌توانستند در صورت عدم خدمت کسب کنند، $\mathrm{E}[Y_{0i}|D_i]$، چه مقدار است. با مقایسه ساده درآمدهای مشهود متفاوت توسط سربازان در موقعیت‌های مختلف یک ارزیابی اریب از تأثیر تیمار بر تیمارشونده است، مگر اینکه D_i مستقل از Y_{0i} باشد. در حالتی خاص داریم:

$$\begin{aligned} E\left[\mathrm{Y}_i|\mathrm{D}_i=1\right]-E\left[\mathrm{Y}_i|\mathrm{D}_i=0\right] &= E\left[\mathrm{Y}_{1i}-\mathrm{Y}_{0i}|\mathrm{D}_i=1\right] \\ &\quad +\left\{E\left[\mathrm{Y}_{0i}|\mathrm{D}_i=1\right]-E\left[\mathrm{Y}_{0i}|\mathrm{D}_i=0\right]\right\}. \end{aligned}$$

به عبارت دیگر، تفاوت مشهود در درآمد توسط موقعیت سربازان معادل تأثیر متوسط تیمار بر تیمار شده به علاوه اریب گزینش است. این موضوع به موازات بحث انتخاب اریب‌آمیز در فصل ۲ است.

با توجه به **CIA**، اریب گزینش پس از مشروط شدن X_i محو می‌شود، بنابراین تأثیر تیمار بر تیمار شده می‌تواند توسط پیش‌بینی‌های تکراری روی X_i ایجاد شود:

$$\begin{aligned} \delta_{TOT} &\equiv E[\mathrm{Y}_{1i}-\mathrm{Y}_{0i}|\mathrm{D}_i=1] \\ &= E\{E[\mathrm{Y}_{1i}|\mathrm{X}_i,\mathrm{D}_i=1]-E[\mathrm{Y}_{0i}|\mathrm{X}_i,\mathrm{D}_i=1]|\mathrm{D}_i=1\}. \end{aligned}$$

البته، $\mathrm{E}[Y_{0i}, D_i = 1]$ انجام نشدنی است. با توجه به **CIA**، در هر حال داریم:

$$E[\mathrm{Y}_{0i}|\mathrm{X}_i,\mathrm{D}_i=0]=E[\mathrm{Y}_{0i}|\mathrm{X}_i,\mathrm{D}_i=1].$$

بنابراین:

$$\begin{aligned} \delta_{TOT} &= E\left\{E\left[\mathrm{Y}_{1i}|\mathrm{X}_i,\mathrm{D}_i=1\right]-E\left[\mathrm{Y}_{0i}|\mathrm{X}_i,\mathrm{D}_i=0\right]|\mathrm{D}_i=1\right\} \\ &= E[\delta_X|\mathrm{D}_i=1], \end{aligned} \tag{3.3.1}$$

در جایی که:

$$\delta_X \equiv E[\mathrm{Y}_i|\mathrm{X}_i,\mathrm{D}_i=1]-E[\mathrm{Y}_i|\mathrm{X}_i,\mathrm{D}_i=0],$$

تفاوت تصادفی مختص $\mathbf{X}$ در متوسط درآمد توسط سربازان در هر مقدار X_i است.

تخمین‌زننده جورسازی در آنگریست (۱۹۹۸) از این حقیقت استفاده می‌کند که X_i در ایجاد نمونه مشابه در سمت راست ۱-۳-۳ متمایز است. در مورد تمایز، تخمین جورسازی می‌تواند به شکل زیر نوشته شود:

$$E[\mathrm{Y}_{1i}-\mathrm{Y}_{0i}|\mathrm{D}_i=1]=\sum \delta_x P(\mathrm{X}_i=x|\mathrm{D}_i=1), \tag{3.3.2}$$

که در آن $P(X_i = x|D_i = 1)$ تابع جرم احتمال[1] برای X_i با توجه به $D_i = 1$ است.[2] در این مورد، X_i، مقادیر تعیین شده توسط کلیه ترکیبات محتمل از سال تولد، گروه امتیاز آزمون، سال تقاضا دادن به ارتش و تحصیلات در زمان تقاضاست. امتیاز آزمون در این مورد از **AFQT** گرفته می‌شود، که توسط ارتش برای طبقه‌بندی قابلیت‌های ذهنی متقاضیان به کار می‌رود (ما این موضوع را به عنوان کنترل در رگرسیون تحصیلی مورد بحث در فصل ۳-۲-۲ شامل کردیم). تخمین‌زننده جورسازی آنگریست (۱۹۹۸) به سهولت δ_X را با تفاوت درآمدهای سربازان و غیرسربازان برای هر ترکیب از متغیرهای مشترک جایگزین کرده و سپس آنها را در متوسط وزنی با استفاده از توزیع تجربی متغیرهای مشترک در میان سربازان ترکیب کرده است.[3] متذکر می‌شود که می‌توانیم تأثیر تیمار متوسط غیرمشروط را به سادگی ایجاد کنیم:

$$\begin{aligned} \delta_{ATE} &= E\{E[Y_{1i}|X_i, D_i = 1] - E[Y_{0i}|X_i, D_i = 0]\} \\ &= \sum_x \delta_x P(X_i = x) \\ &= E[Y_{1i} - Y_{0i}], \end{aligned} \tag{3.3.3}$$

که عبارت است از پیش‌بینی δ_x با استفاده از توزیع حاشیه‌ای X_i به جای توزیع در میان تیمار شده‌ها. δ_{TOT} به ما می‌گوید یک نمونه از سرباز در اثر خدمت سربازی چه مقدار به دست آورده یا از دست داده، در حالی که δ_{ATE} به ما می‌گوید چه مقدار متقاضی خدمت در ارتش به دست آورده یا از دست داده است (از آنگریست ۱۹۹۸، جامعه شامل متقاضیان است).

ارتش آمریکا خصوصاً پس از کوچک کردن ارتش پس از پایان جنگ سرد تمایل دارد درباره سربازانش منصف باشد. در اغلب بخش‌ها، ارتش اکنون فقط فارغ‌التحصیلان دبیرستان را با امتیاز آزمون و در نیمه بالای توزیع امتیازات جذب می‌کند. این غربالگری مثبت با اریب گزینش در نتیجه مقایسه درآمدهای سربازان و غیرسربازان است. این امر می‌تواند در جدول ۱-۳-۳ دیده شود، که گزارشگر تفاضل متوسط،[4] جورسازی و تخمین رگرسیون حاصل از تأثیر خدمت داوطلبانه سربازی در سال‌های ۱۹۸۸ تا ۱۹۹۱ و درآمدهای تأمین اجتماعی با مالیات مردانی است که برای پیوستن به ارتش بین سال‌های ۱۹۷۹ تا ۱۹۸۲ تقاضا کرده‌اند. تخمین جورسازی از مشابه نمونه (۲-۳-۳) گرفته شده است. اگرچه سربازان سفیدپوست ۱۲۳۳ دلار بیشتر از غیرسربازان درآمد داشتند، این تفاوت وقتی تفاوت در متغیرهای مشترک مطابقت داشته باشند منفی است. به طور مشابه، وقتی سربازان غیر سفیدپوست ۲۴۴۹ دلار بیشتر از غیرسربازان درآمد دارند، کنترل متغیرهای مشترک آن را به ۸۴۰ دلار کاهش می‌دهد.

---

1 Probability Mass Function

۲. این تخمین‌زن جورسازی در روبین (۱۹۷۷) بحث شده و توسط کارد و سولیوان (۱۹۸۸) برای تخمین تأثیر تحصیل یارانه‌ای در اشتغال استفاده شده است.

۳. جورسازی دقیق با متغیرهای کمکی پیوسته، ممکن نیست و گونه‌ای از تقریب، حقیقتی که به اریب منجر می‌شود، مورد نیاز است. برای این منظور به آبادی و ایمبن (۲۰۰۶) که نتایج جورسازی تقریبی برای محدود کردن توزیع تخمین‌زن‌های جورسازی را استخراج می‌کنند، نگاه کنید.

4. Differences-in-means

جدول ۱-۳-۳. تخمین‌های غیرکنترلی، جورسازی و رگرسیونی تأثیرات خدمات داوطلبانه سربازی بر درآمدها

Race	Average earnings in 1988-1991	Differences in means by veteran status	Matching estimates	Regression estimates	Regression minus matching
	(1)	(2)	(3)	(4)	(5)
Whites	14537	1233.4	-197.2	-88.8	108.4
		(60.3)	(70.5)	(62.5)	(28.5)
Non-whites	11664	2449.1	839.7	1074.4	234.7
		(47.4)	(62.7)	(50.7)	(32.5)

نکته: برگرفته از آنگریست (۱۹۹۸، جداول ۲ و ۵). خطاهای استاندارد در پرانتزها گزارش شده است. این جدول تخمین‌های تأثیر خدمت داوطلبانه سربازی را بر درآمدهای مالیات‌پذیر مردان در سال‌های ۱۹۸۸ تا ۱۹۹۱ که برای ورود به نیروهای ارتش بین ۱۹۷۹ تا ۱۹۸۲ اعمال شده، نشان می‌دهد. تخمین‌های جورسازی و رگرسیونی کنترل برای متقاضیان شامل سال تولد، تحصیلات در زمان تقاضا و امتیاز AFQT است.

جدول ۱-۳-۳ نیز تخمین‌های رگرسیون تأثیر خدمت داوطلبانه سربازی، برای کنترل مجموعه مشابه متغیرهای مشترک که برای ایجاد تخمین جورسازی به کار رفته را نشان می‌دهد. اینها تخمین‌های $\boldsymbol{\delta_R}$ در معادله زیر هستند.

$$Y_i = \sum_x d_{ix}\beta_x + \delta_R D_i + \varepsilon_i, \qquad (3.3.4)$$

که در آن $\boldsymbol{d_{ix}}$ یک متغیر موهومی است که $\boldsymbol{X_i = x}$ را نشان می‌دهد و $\boldsymbol{\beta_x}$ تأثیر رگرسیون برای $\mathbf{X_i = x}$ و $\boldsymbol{\delta_R}$ تخمین رگرسیون است. توجه شود که این مدل رگرسیون پارامتر جداگانه‌ای را برای هر مقدار برگرفته توسط متغیر مشترک را مجاز می‌کند. این مدل می‌تواند به عنوان اشباع شده در $\boldsymbol{X_i}$ خوانده شود، زیرا شامل پارامتری برای هر مقدار از $\boldsymbol{X_i}$ است (این کاملاً اشباع شده نیست، در هر حال، به علت اینکه تأثیر افزایش یک واحد برای $\boldsymbol{D_i}$ بدون تعامل $\boldsymbol{D_i . X_i}$ وجود دارد).

علی‌رغم این حقیقت که تخمین‌های رگرسیون و جورسازی برای کنترل متغیرهای یکسان است، تخمین‌های رگرسیون در جدول ۱-۳-۳ تا حدی بیشتر از تخمین‌های جورسازی برای سفیدپوستان و غیر سفیدپوستان است. در حقیقت، تفاوت بین نتایج جورسازی و رگرسیونی از حیث آماری مهم است. به طور هم‌زمان، دو راهبرد تخمینی تصویر یکسانی از تأثیرات خدمت سربازی نشان می‌دهند. دلیل اینکه تخمین‌های رگرسیونی و جورشدگی مشابهند این است که رگرسیون نیز می‌تواند به عنوان نوعی از تخمین جورسازی در نظر گرفته شود: تخمین رگرسیونی با تخمین جورشدگی فقط در وزن به کار رفته برای مجموع تأثیرات خاص متغیر مشترک، $\boldsymbol{\delta_x}$ در تأثیری واحد تفاوت می‌کند. خصوصاً، تخمین جورشدگی از توزیع متغیرهای مشترک در میان تیمار شده‌ها برای توزین تخمین‌های خاص متغیرهای مشترک در تخمین تأثیر تیمار در تیمار شده‌ها استفاده می‌کند، در صورتی که رگرسیون میانگین وزنی متغیر از این تأثیرات را ایجاد می‌کند.

برای مشاهده این موضوع، با استفاده از فرمول آناتومی رگرسیون برای نوشتن ضریب $\boldsymbol{D_i}$، در رگرسیون $\boldsymbol{Y_i}$ روی $\boldsymbol{X_i}$ و $\boldsymbol{D_i}$ شروع می‌کنیم:

$$\delta_R = \frac{Cov(Y_i, \tilde{D}_i)}{V(\tilde{D}_i)} \qquad (3.3.5)$$

$$= \frac{E[(D_i - E[D_i|X_i])Y_i]}{E[(D_i - E[D_i|X_i])^2]}$$

$$= \frac{E\{(D_i - E[D_i|X_i])E[Y_i|D_i, X_i]\}}{E[(D_i - E[D_i|X_i])^2]}. \qquad (3.3.6)$$

معادله دوم در این مجموعه عبارات از این حقیقت استفاده می‌کند که اشباع مدل $\boldsymbol{X_i}$ به معنای خطی بودن $\mathbf{E}[\boldsymbol{D_i}|\boldsymbol{X_i}]$ است. بنابراین $\tilde{D}_i$، که به عنوان پسماند رگرسیون $\boldsymbol{D_i}$ روی $\boldsymbol{X_i}$ بیان می‌شود، تفاوت بین $\boldsymbol{D_i}$ و $\mathbf{E}[\boldsymbol{D_i}|\boldsymbol{X_i}]$ است. سومین تساوی از این حقیقت استفاده می‌کند که رگرسیون $\boldsymbol{Y_i}$ روی $\boldsymbol{D_i}$ و $\boldsymbol{X_i}$ همانند رگرسیون $\boldsymbol{Y_i}$ روی $\mathbf{E}[\boldsymbol{Y_i}|\boldsymbol{D_i}, \boldsymbol{X_i}]$ است.

برای ساده‌سازی بیشتر، ما **CEF**، $\mathbf{E}[\boldsymbol{Y_i}|\boldsymbol{D_i}, \boldsymbol{X_i}]$ را گسترده می‌کنیم تا به دست آوریم:

$$E[Y_i|D_i, X_i] = E[Y_i|D_i = 0, X_i] + \delta_X D_i.$$

به عبارت دیگر، اگر متغیرهای مشترک لازم نباشند، **CIA** به طور غیرمشروط بر قرار است، مثل اینکه آزمایش تصادفی انجام شده باشد ـ و **CEF** خواهد شد:

$$E[Y_i|D_i, X_i] = E[Y_i|D_i = 0] + E[Y_{1i} - Y_{0i}]D_i,$$

که در آن نتیجه می‌شود که رگرسیون $\boldsymbol{Y_i}$ روی $\boldsymbol{D_i}$ تأثیر تیمار متوسط جامعه در این مورد را تخمین می‌زند (مثل تجربه بحث شده در فصل ۳-۲)، اما در اینجا ما تمایل به سناریوی کلی‌تری داریم که شرط $\boldsymbol{X_i}$ برای حذف انتخاب ماریبانه ضروری است.

برای ارزیابی تخمین رگرسیون کلی‌تر، ۵-۳-۳، با جایگزینی به جای $\mathbf{E}[\boldsymbol{Y_i}|\boldsymbol{D_i}, \boldsymbol{X_i}]$ در صورت کسر شروع می‌کنیم. در نتیجه خواهیم داشت:

$$E\{(D_i - E[D_i|X_i])E[Y_i|D_i, X_i]\} = E\{(D_i - E[D_i|X_i])E[Y_i|D_i = 0, X_i]\} + E\{(D_i - E[D_i|X_i])D_i\delta_X\}.$$

اولین جمله در سمت راست صفر است، زیرا $\mathbf{E}[\boldsymbol{Y_i}|\boldsymbol{D_i} = \mathbf{0}, \boldsymbol{X_i}]$ تابعی از $\boldsymbol{X_i}$ است و بنابراین تعاملی با $(\boldsymbol{D_i} - \boldsymbol{E}[\boldsymbol{D_i}|\boldsymbol{X_i}])$ ندارد. به دلیلی مشابه، جمله دوم به شکل زیر ساده می‌شود:

$$E\{(D_i - E[D_i|X_i])D_i\delta_X\} = E\{(D_i - E[D_i|X_i])^2\delta_X\}.$$

در اینجا، نشان دادیم که:

$$\delta_R = \frac{E[(D_i - E[D_i|X_i])^2\delta_X]}{E[(D_i - E[D_i|X_i])^2]} = \frac{E\{E[(D_i - E[D_i|X_i])^2|X_i]\delta_X\}}{E\{E[(D_i - E[D_i|X_i])^2|X_i]\}} = \frac{E[\sigma_D^2(X_i)\delta_X]}{E[\sigma_D^2(X_i)]}, \qquad (3.3.7)$$

در حالی که:

$$\sigma_D^2(X_i) = E[(D_i - E[D_i|X_i])^2|X_i]$$

متغیر مشروط $\boldsymbol{D_i}$ با داشتن $\boldsymbol{X_i}$ است. این موضوع بیان می‌کند که مدل ۴-۳-۳، یک متوسط وزنی تیمار متغیر $\boldsymbol{\delta_X}$ را ایجاد می‌کند.

به دلیل این که رگرسیون مطلوب، $\boldsymbol{D_i}$، یک متغیر موهومی است، آخرین مرحله می‌تواند انجام شود. در این مورد، $\boldsymbol{\sigma_D^2(X_i) = P(D_i = 1|X_i)(1 - P(D_i = 1|X_i))}$

بنابراین:

$$\delta_R = \frac{\sum_x \delta_x \left[P(\mathrm{D}_i = 1|\mathrm{X}_i = x)(1 - P(\mathrm{D}_i = 1|\mathrm{X}_i = x))\right] P\left(\mathrm{X}_i = x\right)}{\sum_x \left[P(\mathrm{D}_i = 1|\mathrm{X}_i = x)(1 - P(\mathrm{D}_i = 1|\mathrm{X}_i = x))\right] P\left(\mathrm{X}_i = x\right)}$$

این امر نشان می‌دهد که وزن‌های تخمینی رگرسیون هر تیمار اختصاصی متغیر مشترک تحت تأثیر $\boldsymbol{[\mathrm{P}(X_i = x|D_i = 1)\big(1 - P(X_i = x|D_i = 1)\big)P(X_i = x)}$ قرار دارد. بر عکس، تخمین جورشدگی برای تأثیر تیمار روی تیمار شده می‌تواند به شکل زیر نوشته شود:

$$E[\mathrm{Y}_{1i} - \mathrm{Y}_{0i}|\mathrm{D}_i = 1] = \sum_x \delta_x P(\mathrm{X}_i = x|\mathrm{D}_i = 1) = \frac{\sum_x \delta_x P(\mathrm{D}_i = 1|\mathrm{X}_i = x) P(\mathrm{X}_i = x)}{\sum_x P(\mathrm{D}_i = 1|\mathrm{X}_i = x) P(\mathrm{X}_i = x)}$$

زیرا:

$$P(\mathrm{X}_i = x|\mathrm{D}_i = 1) = \frac{P(\mathrm{D}_i = 1|\mathrm{X}_i = x) \cdot P(\mathrm{X}_i = x)}{P(\mathrm{D}_i = 1)}.$$

بنابراین وزن به کار رفته در ایجاد $\boldsymbol{\mathrm{E}[Y_{1i} - Y_{0i}|D_i = 1]}$ متناسب با احتمال تیمار هر مقدار از متغیرهای مشترک است.

نکته قابل توجه در این استنتاج این است که تخمین تیمار روی تیمار شده[1] بیشترین وزن را روی سلول‌های متغیرهای مشترکی می‌گذارد که در آنها متغیرهای مشروط تیمار بالاترین موقعیت را دارند. قانوناً، این متغیر وقتی $\boldsymbol{\mathrm{P}(D_i = 1|X_i = x) = 1/2}$ باشد، حداکثر است، به عبارت دیگر، برای سلول‌هایی که تعداد یکسانی از تیمار شده و کنترل مشهود وجود دارد، این وضعیت صادق است. البته، اگر $\boldsymbol{\delta_x}$ در سلول‌ها تغییری نکند، تفاوت در شکل وزنی دارای اهمیت کمتری است (اگرچه وزن دادن هنوز بر کارایی تخمین اثر دارد). در این مثال، در هر حال، کسانی که به احتمال زیاد در ارتش خدمت می‌کردند حداقل از خدمات خود بهره‌مند می‌شدند. این احتمال وجود دارد که دلیل آن باشد که کسانی که احتمال بالاتری دارند، کیفیت بهتری دارند، اما همچنین دارای درآمد شخصی بالاتری هستند و بنابراین از خدمت سربازی کمتر بهره‌مند می‌شوند. این حقیقت منجر به تخمین جورشدگی از تأثیر خدمت سربازی مشابه تخمین رگرسیونی بر اساس برداری یکسان از متغیرهای کنترل می‌شود.[2]

1. Treatment-on-the-treated

۲. جای تعجب نیست که رگرسیون بیشترین وزن را به سلول‌هایی می‌دهد که در آنها $P(\mathrm{D}_i = 1|\mathrm{X}_i = x) = 1/2$،

قابل ذکر است که نه تخمین رگرسیون و نه تخمین متغیر مشترک جورشدگی هیچ وزنی به سلول‌های متغیرهای مشترک نمی‌دهند که شامل مشاهدات تیمار شده و کنترل نباشند. یک مقدار $\boldsymbol{X_i}$ مثل $\boldsymbol{x}^*$ را در نظر بگیرید، که هیچ یک یا همه تیمار شده‌اند. سپس، $\boldsymbol{\delta_{x^*}}$ تعریف نشده است، در حالی که وزن‌های رگرسیون، $[\mathbf{P}(\boldsymbol{D_i}=\mathbf{1}|\boldsymbol{X_i}=\boldsymbol{x}^*(\mathbf{1}-\boldsymbol{P}(\boldsymbol{D_i}=\mathbf{1}|\boldsymbol{X_i}=\boldsymbol{x}^*))]$ صفر هستند. در زبان ادبیات اقتصادسنجی در جورسازی، هم تخمین رگرسیونی و هم تخمین جورشدگی *حمایت مشترکی*[1] اعمال می‌کنند، یعنی اینکه، آنها محدود به مقادیر متغیرهای مشترک هستند که در آنها مشاهدات تیمار شده و کنترل یافت می‌شوند.[2]

مرحله مورد برآورد[3] تا برآوردگر[4] اندکی پیچیده است. در عمل، هم برآوردگر رگرسیونی و هم جورشدگی با استفاده از فرضیات مدل‌سازی که به طور ضمنی شامل مقداری برون‌یابی در سلول‌هاست، اجرا می‌شوند. به عنوان مثال، برآوردگرهای جورشدگی اغلب سلول‌های متغیر مشترک را با مشاهدات کمتر ترکیب می‌کنند. این امر، اگر سلول‌های ترکیبی هیچ‌یک دارای مشاهدات تحت تیمار و تیمار نشده نباشند، ناقض حمایت مشترک است. مدل‌های رگرسیونی که در $\boldsymbol{X_i}$ اشباع نشده‌اند ممکن است ناقض حمایت مشترک باشند، زیرا سلول‌های متغیر مشترک بدون مشاهدات تیمار شده و کنترلی می‌توانند کمک به برآورد را توسط برون‌یابی پایان دهند. در اینجا، در هر حال، تقارنی بین راهبردهای جورشدگی و رگرسیونی مشاهده می‌کنیم: آنها در یک طبقه هستند و در اصل، مستلزم نوع مشابهی از سازش در عمل‌اند.[5]

موارد بیشتر درباره رگرسیون و جورشدگی: تیمارهای سفارشی و پیوسته

آیا تفسیر شبه‌جورشدگی[6] از رگرسیون که در بالا برای تیمار دوگانه ذکر شد برای مدل‌های تیمار مرتب[7] و پیوسته به کار می‌روند؟ پاسخ طولانی بسیار فنی است و ممکن است بیشتر از چیزی باشد که شما می‌خواهید بدانید. پاسخ کوتاه، تا حدی، بله است.

همان طوری که قبلاً بحث شد، یک تفسیر از رگرسیون این است که بردار شیب جامعه **OLS**

چراکه رگرسیون براي الگوهاي خطي با اثر ثابت همگن كاراست. بايد انتظار داشته باشيم كه يك تخمين‌زن كارا بيشترين وزن را به سلول‌هايي بدهد كه اثر تيمار عمومي با بيشترين دقت برآورد شده باشد. با وجود پسماندهاي همگن، دقيق‌ترين تأثيرات تيمار مربوط به سلول‌هايي است كه احتمال تيمار يك‌دوم است.

1. Common Support

۲. حمايت (Support) يك متغير تصادفي به مجموعه‌اي از تعميم‌ها گفته مي‌شود كه با احتمال مثبت اتفاق مي‌افتند. براي اين منظور و براي مطالعه بحثي از حمايت عمومي در جورسازي به هكمن، ايچيمورا، اسميت و تود (۱۹۹۸) و اسميت و تود (۲۰۰۱) مراجعه كنيد.

3. Estimand

4. Estimator

۵. مسائل جورسازي شامل متغيرهاي ايكس به طور متناهي توزيع شده، غالباً توسط تجميع مقادير با هدف ايجاد گروه‌هاي بزرگ‌تر و يا جفت كردن مشاهداتي كه مشابهت دارند، و نه الزاماً مشاهدات يكسان، صورت مي‌گيرد. براي بحثي در مورد اين رويكرد به كوچران (۱۹۶۵)، روبين (۱۹۷۳) يا روزنبام (فصل ۳، ۱۹۹۵) مراجعه كنيد. با وجود متغيرهاي كمكي داراي توزيع پيوسته، تخمين‌زن‌هاي جورسازي شده، اريب هستند، چراكه جورشده‌ها كامل نيستند. آبادي و ايمبنز (۲۰۰۸) اخيراً نشان داده‌اند كه يك تصحيح اريب مبتني بر رگرسيون مي‌تواند اريب مجانبي را از جورشده‌هاي ناكامل، حذف كند.

6. Pseudo-matching Interpretation

7. Ordered and Continuous Treatments

تقریب خطی **MMSE** را برای **CEF** تأمین می‌کند. البته این برای رگرسیون‌های مرتب و پیوسته همانند دوگانه کار می‌کند. یکی از خواص مربوطه این حقیقت است که ضرایب رگرسیون دارای تفسیر «مشتق متوسط»[1] هستند. در مدل‌های رگرسیون چندمتغیری، این تفسیر متأسفانه توسط این حقیقت به صورت پیچیده درمی‌آید که بردار شیب **OLS** متوسط وزنی ماتریس از گرادیان **CEF** است. متوسط‌های ماتریس وزنی[2] به استثنای موارد خاص (چامبرلین و لیمر، ۱۹۷۶) دارای مشکل تفسیر هستند. یک مورد خاص و مهم وقتی خواص مشتق متوسط نسبتاً ساده هستند در مدل‌های رگرسیونی برای تیمار مرتب و پیوسته با مدل اشباع شده برای متغیرهای مشترک یافت می‌شود. برای اجتناب از مشتق‌گیری طولانی، ما به سادگی این فرمول را شرح می‌دهیم. هر مشتق‌گیری در ضمیمه این فصل نشان داده شده است. برای جزئیات بیشتر، ضمیمه آنگریست و کروگر را مشاهده کنید (۱۹۹۹).

برای این بحث، شدت تیمار،[3] $\boldsymbol{S_i}$، فرض می‌شود که یک متغیر تصادفی توزیع شده پیوسته باشد، و لزوماً غیرمنفی نباشد. فرض کنید که **CEF** مطلوب می‌تواند به شکل $\mathbf{h(t)=E}[\boldsymbol{Y_i}|\boldsymbol{S_i}=\boldsymbol{t}]$ با مشتق **h’(t)** باشد. سپس:

$$\frac{E[Y_i(S_i-E[S_i])]}{E[S_i(S_i-E[S_i])]}=\frac{\int h'(t)\,\mu_t dt}{\int \mu_t dt} \qquad (3.3.8)$$

در جایی که

$$\mu_t \equiv \{E[S_i|S_i\geq t]-E[S_i|S_i<t]\}\{P(S_i\geq t)[1-P(S_i\geq t)\}, \qquad (3.3.9)$$

و انتگرال در ۸-۳-۳ روی مقادیر محتمل $\boldsymbol{S_i}$ قرار دارد. این فرمول هر مقدار محتمل $\boldsymbol{S_i}$ را متناسب با تفاوت متوسط مشروط $\boldsymbol{S_i}$ در بالا و پایین آن مقدار وزن‌دهی می‌کند. وزن بیشتر نیز به نقاط نزدیک به متوسط $\boldsymbol{S_i}$ داده می‌شود، زیرا $\mathbf{P}(\boldsymbol{S_i}\geq \boldsymbol{t}).[\mathbf{1}-\boldsymbol{P}(\boldsymbol{S_i}\geq \boldsymbol{t})]$ وقتی بالاترین مقدار است که $\mathbf{P}(\boldsymbol{S_i}\geq \boldsymbol{t})=\mathbf{1/2}$ باشد.

با متغیر مشترک $\boldsymbol{X_i}$، وزن‌ها در ۸-۳-۳ تبدیل به وزن‌های مختص **X**[4] می‌شوند. یک نسخه متوسط متغیر مشترک از فرمول مشابه بر ضریب رگرسیون چند متغیری $\boldsymbol{Y_i}$ روی $\boldsymbol{S_i}$، پس از تقسیم $\boldsymbol{X_i}$ اعمال می‌شود. خصوصاً:

$$\frac{E[Y_i(S_i-E[S_i|X_i])]}{E[S_i(S_i-E[S_i|X])]}=\frac{E\left[\int h'_X(t)\mu_{tX}\,dt\right]}{E\left[\int \mu_{tX}\,dt\right]}, \qquad (3.3.10)$$

که در آن $h'_X(t)\equiv \frac{\partial E[Y_i|X_i,S_i=t]}{\partial t}$ و

$$\mu_{tX}\equiv\{E[S_i|X_i,S_i\geq t]-E[S_i|X_i,S_i<t]\}\{P(S_i\geq t|X_i)[1-P(S_i\geq t|X_i)]$$

1. Average Derivative
2. Matrix-weighted Averages
3. Treatment Intensity
4. X-specific

لازم است تأکید شود که معادله ۱۰-۳-۳ دو نوع از میانگین را نشان می‌دهد: یک انتگرال که در طول CEF خطی در مقادیر متغیر مشترک ثابت متوسط‌گیری می‌کند و یک پیش‌بینی که در طول سلول‌های متغیر مشترک میانگین می‌گیرد. نکته مهم در این زمینه این است که ضرایب رگرسیون جامعه شامل هیچ اطلاعاتی درباره تأثیر S_i بر CEF برای مقادیر X_i با شرط $\mathrm{P}(S_i \geq t|X_i)$ مساوی ۰ و ۱ نمی‌باشد. این شامل مقادیر X_i است وقتی S_i ثابت باشد. در مفهومی مشابه، هیچ ارزشی ندارد که اگر S_i یک متغیر موهومی باشد، ما بتوانیم معادله ۷-۳-۳ را از فرمول کلی‌تر ۱۰-۳-۳ استخراج کنیم.

آنگریست و کروگر (۱۹۹۹) تابع متوسط وزنی را برای رگرسیون تحصیلی با متغیرهای مشترک وضعیت تولد و سال تولد ایجاد کردند. اگرچه معادلات ۸-۳-۳ و ۱۰-۳-۳ ممکن است پنهان یا حداقل ناروشن باشند، در این مثال وزن‌های متوسط، $E[\mu_{tX}]$ ، توابع متقارن یکنواخت از $\mathbf{t}$ هستند که در وجه S_i متمرکز شده‌اند.

پیامدهای ۸-۳-۳ یا ۱۰-۳-۳ می‌توانند با توجه به مدلی برای توزیع رگرسیون‌ها بیشتر بررسی شوند. فرض کنید، به عنوان مثال، S_i توزیع نرمال شده باشد. فرض کنید $z_i = \frac{S_i - E(S_i)}{\sigma_s}$ ، در جایی که σ_s انحراف استاندارد از S_i است، به طوری که z_i استاندارد نرمال است. سپس:

$$E[\mathrm{s}_i|\mathrm{s}_i \geq t] = E(\mathrm{s}_i) + \sigma_s E\left[z_i|z_i \geq \frac{\mathrm{t} - E(\mathrm{s}_i)}{\sigma_s}\right] = E(\mathrm{s}_i) + \sigma_s E\left[z_i|z_i \geq t^*\right].$$

از فرمول‌های نرمال کوتاه شده (جانسون و کوتز، ۱۹۷۰)، می‌دانیم که:

$$E[z_i|z_i > t^*] = \frac{\phi(t^*)}{[1-\Phi(t^*)]} \text{and } E[z_i|z_i < t^*] = \frac{-\phi(t^*)}{\Phi(t^*)}.$$

در جایی که $\emptyset(\mathbf{0})$ و $\mathbf{\Phi(0)}$ تابع چگالی نرمال استاندارد و تابع توزیع هستند. با جایگزینی در فرمول برای μ_t، ۹-۳-۳ داریم:

$$\mu_t = \sigma_s \left\{\frac{\phi(t^*)}{[1-\Phi(t^*)]} - \frac{-\phi(t^*)}{\Phi(t^*)}\right\}[1-\Phi(t^*)]\Phi(t^*) = \sigma_s \phi(t^*).$$

بنابراین ما نشان دادیم که:

$$\frac{Cov(\mathrm{Y}_i, \mathrm{s}_i)}{V(\mathrm{s}_i)} = E[h'(\mathrm{s}_i)].$$

به عبارت دیگر، رگرسیون Y_i روی S_i عبارت است از: مشتق متوسط جامعه (بدون وزن)، $\mathrm{E}[h'(S_i)]$، وقتی S_i توزیع نرمال داشته باشد. البته، این نتیجه یک مورد خاص از موارد خاص است. با این حال، به نظر می‌رسد عاقلانه باشد تصور شود که نرمال بودن در اینجا ممکن است زیاد مهم نباشد و در تجربه ما، مشتق متوسط (تأثیر حاشیه‌ای) ساخته شده از مدل‌های غیرخطی پارامتری برای متغیرهای محدود وابسته (پروبیت یا توبیت) معمولاً بدون توجه به توزیع رگرسیون از ضرایب رگرسیون

مربوطه غیرقابل تشخیص‌اند. ما درباره این نکته در بخش ۲-۴-۳ بحث بیشتری می‌کنیم.

۲-۳-۳. کنترل متغیرهای مشترک با استفاده از نمره تمایل[1]

مهم‌ترین نتیجه در نظریه رگرسیون عبارت است از: فرمول اریب متغیرهای حذف شده: ضرایب متغیرهای شامل شده توسط حذف متغیرها وقتی متغیرهای حذف شده با متغیرهای شامل شده تعاملی ندارند، تحت تأثیر قرار نمی‌گیرند. قضیه نمره تمایل، به دلیل نظر روزنبائوم و روبین (۱۹۸۳) این ایده را، در جایی که متغیر علّی مطلوب، یک متغیر تیمار موهومی هستند، به راهبردهای برآوردی که بر جورسازی به عوض رگرسیون متکی هستند گسترش می‌دهد.[2]

قضیه نمره تمایل بیان می‌کند که اگر نتایج بالقوه مستقل از وضعیت تیمار، مشروط به بردار متغیر مشترک چندمتغیره، $\boldsymbol{X_i}$، باشد، بنابراین نتایج بالقوه مستقل از وضعیت تیمار، مشروط به تابع اسکالر متغیرهای مشترک، یعنی نمره تمایل، طبق تعریف $\boldsymbol{p(X_i) \equiv E[D_i|X_i]}$ است. داریم:

قضیه ۱-۳-۳. قضیه نمره تمایل

فرض کنید **CIA** برای $\boldsymbol{Y_{j,i}}$، با شرط **j=0,1** بر قرار باشد. سپس داریم $Y_{ji} \amalg D_i | p(X_i)$.

اثبات: این ادعا درست است اگر $P[D_i = 1|Y_{ji}, p(X_i)]$ بستگی به $\boldsymbol{Y_{ji}}$ نداشته باشد.

$$
\begin{aligned}
P[D_i = 1|Y_{ji}, p(X_i)] &= E[D_i|Y_{ji}, p(X_i)] \\
&= E\{E[D_i|Y_{ji}, p(X_i), X_i]|Y_{ji}, p(X_i)\} \\
&= E\{E[D_i|Y_{ji}, X_i]|Y_{ji}, p(X_i)\} \\
&= E\{E[D_i|X_i]|Y_{ji}, p(X_i)\}, \text{by the CIA.}
\end{aligned}
$$

اما $E\{E[D_i|X_i]|Y_{ji}, p(X_i)\} = E\{p(X_i)|Y_{ji}, p(X_i)\}$، که به وضوح معادل $\mathbf{p}(\boldsymbol{X_i})$ است.

همانند فرمول **OVB** برای رگرسیون، قضیه نمره تمایل بیان می‌کند که شما فقط لازم است متغیرهای مشترکی را کنترل کنید که احتمال تیمار را تحت تأثیر قرار می‌دهد، اما همچنین نکته بیشتری را بیان می‌کند: تنها متغیر مشترکی که شما لازم است کنترل کنید احتمال تیمار به خودی خود است. در عمل، قضیه نمره تمایل معمولاً برای تخمین در دو مرحله استفاده می‌شود: اول، $\mathbf{p}(\boldsymbol{X_i})$ با استفاده از بعضی انواع مدل‌های پارامتری مثل لاجیت یا پروبیت تخمین زده می‌شود. سپس تخمین‌های تأثیر تیمار با جورشدگی در مقادیر مناسب از مرحله اول، یا توسط روش وزنی بیان شده در (ایمبنز، ۲۰۰۴) محاسبه می‌شود.

در عمل روش‌های بسیاری برای استفاده از قضیه نمره تمایل برای تخمین وجود دارد. جورشدگی

1. Propensity Score

۲. موارد خاص‌تر در این زمینه در رود (۱۹۸۶) در مورد برآوردهاي آزاد از توزیع (Distribution-free Estimation) الگوهاي با متغیر وابسته محدود با رگرسورهاي توزیع شده نرمال، قابل مطالعه است.

نمره تمایل مستقیم مثل جورشدگی متغیر مشترک عمل می‌کند، به استثنای اینکه ما جورشدگی در امتیاز را به عوض متغیر مشترک مستقیم انجام دهیم. توسط قضیه نمره تمایل و **CIA** داریم:

$$E[Y_{1i} - Y_{0i}|D_i = 1] = E\{E[Y_i|p(X_i), D_i = 1] - E[Y_i|p(X_i), D_i = 0]|D_i = 1\}.$$

تخمین‌های تأثیر تیمار بر تیمار شده می‌تواند توسط طبقه‌بندی تخمین‌های $\mathbf{p}(\boldsymbol{X_i})$ و جایگزینی متوسط نمونه مشروط به جای پیش‌بینی‌ها یا توسط جورشدگی هریک از مشاهدات تیمار شده برای کنترل مقادیر مشابه نمره تمایل به دست آید (هر دو این رویکردها توسط دهیجا و واهبا ۱۹۹۹ استفاده شده است). از سوی دیگر، تخمین مبتنی بر مدل یا تخمین غیرپارامتری از $\mathbf{E}[\boldsymbol{Y_i}|\boldsymbol{p}(\boldsymbol{X_i}), \boldsymbol{D_i}]$ می‌توانند جایگزین این توابع شرطی متوسط شوند و پیش‌بینی‌ها با مجموع جایگزین گردند (مثل مورد هکمن، ایشیمورا و تاد، ۱۹۹۸).

رویکرد بهتر وزنی برای تخمین نمره تمایل، مرحله پُرزحمت جورسازی را با به‌کارگیری این حقیقت حذف می‌کند که **CIA** شامل $\mathbf{E}[\frac{Y_iD_i}{p(X_i)}] = E[Y_i]$ و $\mathbf{E}\left[\frac{Y_i(1-D_i)}{1-p(X_i)}\right] = E[Y_{0i}]$ است. بنابراین، با توجه به روش تخمین $\mathbf{p}(\boldsymbol{X_i})$، می‌توانیم تخمین‌های تأثیر متوسط تیمار را به این شکل به دست بیاوریم:

$$\begin{aligned} E[Y_{1i} - Y_{0i}] &= E\left[\frac{Y_iD_i}{p(X_i)} - \frac{Y_i(1-D_i)}{1-p(X_i)}\right] \\ &= E\left[\frac{(D_i - p(X_i))Y_i}{p(X_i)(1-p(X_i))}\right]. \end{aligned} \tag{3.3.11}$$

این عبارت نهایی یک تخمین از شکلی است که توسط نیو وی (۱۹۹۰) و رابینز، مارک و نیو وی (۱۹۹۲) پیشنهاد شده است. تأثیر تیمار را بر تیمار شده از نمونه مشابه می‌توانیم به دست آوریم:

$$E[Y_{1i} - Y_{0i}|D_i = 1] = E\left[\frac{(D_i - p(X_i))Y_i}{(1-p(X_i))P(D_i)}\right]. \tag{3.3.12}$$

این ایده که می‌توان نمونه‌گیری غیرتصادفی را توسط وزن‌دهی با معکوس احتمال انتخاب[1] اصلاح کنید به هورویتز و تامپسون برمی‌گردد (۱۹۵۲). البته، برای ایجاد رویکردی امکان‌پذیر، و برای اینکه تخمین نتیجه شده پایدار باشد، ما نیاز به تخمین‌زننده برای $\mathbf{p}(\boldsymbol{X_i})$ داریم.

نسخه هورویتز و تامپسون برای رویکرد نمره تمایل جذاب است، زیرا تخمین‌زننده به طور خاص خودکار است و هیچ جورشدگی پُرزحمتی لازم ندارد. رویکرد هورویتز تامپسون همچنین ارتباط نزدیک بین جورشدگی و رگرسیون نمره تمایل را نشان می‌دهد، ارتباطی که برای جورشدگی متغیر مشترک در بخش ۱-۳-۳ مورد بحث قرار گرفته است. مجدداً تخمین رگرسیونی $\boldsymbol{\delta_R}$ را برای رگرسیون جامعه $\boldsymbol{Y_i}$ روی $\boldsymbol{D_i}$ با کنترل مدل اشباع برای متغیرهای مشترک در نظر بگیرید. این تخمین می‌تواند به شکل زیر

1. Weighting by the Reciprocal of the Probability of Selection

نوشته شود:

$$\delta_R = \frac{E[(D_i - p(X_i))Y_i]}{E[p(X_i)(1-p(X_i))]}. \qquad (3.3.13)$$

دو تخمین جورشدگی و رگرسیونی هورویتز و تامپسون اعضای طبقه تخمین متوسط وزنی مورد نظر توسط هیرانو، ایمبنز و ریدر (۲۰۰۳) هستند:

$$E\left\{g(X_i)\left[\frac{Y_iD_i}{p(X_i)} - \frac{Y_i(1-D_i)}{(1-p(X_i))}\right]\right\}, \qquad (3.3.14)$$

که در آن $\mathbf{g}(X_i)$ تابع مشخص وزنی است (برای حرکت از تخمین به تخمین‌زننده، جایگزین $\mathbf{p}(X_i)$ با یک تخمین‌زننده متناسب و پیش‌بینی توسط مجموع‌ها). برای تأثیر تیمار متوسط، $\mathbf{g}(X_i)$ را برابر صفر قرار دهید: برای تأثیر بر تیمار شده، $\mathbf{g}(X_i) = \frac{p(X_i)}{P(D_i)}$ قرار دهید و برای رگرسیون عبارت زیر قرار می‌گیرد:

$$g(X_i) = \frac{p(X_i)(1-p(X_i))}{E[p(X_i)(1-p(X_i))]}.$$

این مشابهت دوباره حقیقتی را نشان می‌دهد که رگرسیون و جورشدگی ـ شامل جورشدگی نمره تمایل ـ در واقع دو چیز متفاوت نیستند، حداقل نه تا اینکه ما مدلی را برای نمره تمایل مشخص کنیم. سؤال بزرگ در اینجا این است که چگونه به بهترین وجهی مدل‌سازی و تخمین $\mathbf{p}(X_i)$ انجام شود، یا چه میزان هموارسازی یا طبقه‌بندی در زمان تخمین $\mathbf{E}[Y_i|p(X_i), D_i]$ استفاده می‌شود، مخصوصاً اگر متغیرهای مشترک پیوسته هستند.

مشابه رگرسیون این سؤال این است که چگونه متغیرهای کنترل پارامترگذاری می‌شوند (مثلاً چند جمله‌ای‌ها یا تأثیرات اصلی و جملات تعاملی ـ اگر متغیرهای مشترک به صورت گسسته کدبندی شده باشند). پاسخ به این سؤال ذاتاً دارای خاصیت کاربردی است. ادبیات تجربی بالنده نشان می‌دهد که مدل لاجیت برای نمره تمایل با چند جمله‌ای متغیرهای مشترک پیوسته به خوبی عمل می‌کند ـ اگرچه این موضوع نمی‌تواند یک قضیه باشد (دهیجیا و وهبا، ۱۹۹۹).

ادبیات نظری گسترده، تعدادی قضایای تفکربرانگیز را برای استفاده مؤثر از نمره تمایل ایجاد کرده است. اولاً، از دیدگاه کارایی مجانبی، معمولاً هزینه‌ای برای جورشدگی با نمره تمایل به عوض جورشدگی متغیر مشترک کامل وجود دارد. ما می‌توانیم خطاهای استاندارد تقریبی کمتری توسط جورشدگی متغیرهای مشترک که نتایج را بیان می‌کنند به دست آوریم، چه اینکه به نمره تمایل تبدیل شود یا نشود. ما از تحقیقات هان (۱۹۹۸) با حداکثر دقت می‌دانیم که امکان دارد تخمین‌های تأثیرات تیمار تحت **CIA** را با آگاهی به نمره تمایل به دست آورد یا نه. در آنگریست (۱۹۹۸)، کارایی حاصل از جورسازی در سال تولد، حتی اگر احتمال خدمت در ارتش ربطی به سال تولد نداشته باشد، وجود دارد.

مشابه رگرسیون در اینجا نتیجه آن است که حتی در یک سناریوی بدون اریب از متغیرهای حذف شده، رگرسیون طولانی، در زمانی که این متغیرها دارای قدرت پیش‌بینی نتایج هستند، تخمین‌های دقیق‌تری از ضرایب متغیرهای شامل شده در رگرسیون کوتاه ایجاد می‌کند، چراکه این متغیرهای مشترک منجر به متغیرهای پسماند کوچک‌تری می‌شوند (فصل ۳-۱-۳).

نتایج هان (۱۹۹۸) این سؤال را مطرح می‌کند که چرا ما باید خودمان را با تخمین‌هایی اذیت کنیم که از نمره تمایل استفاده می‌کنند. یک استدلال فیلسوفانه این است که نمره تمایل توجه محققان را بر مدل‌های تخصیص تیمار متمرکز می‌کند، یعنی آنچه ممکن است ما درباره آن به شکل منطقی اطلاعات خوبی داشته باشیم، به عوض اینکه خودمان را درگیر فرایندهای نوعاً پیچیده‌تر و پُررمز و رازتر منتهی به خروجی بکنیم. این دیدگاه به خصوص وقتی متقاعدکننده به نظر می‌رسد که در حالی که نتایج تعیین‌کننده روندی ناشناخته‌تر هستند (مثلاً بازار)، تخصیص تیمار نتیجه نهادهای انسانی یا مقررات دولتی است. به عنوان مثال، در ارزیابی سری زمانی تأثیرات علّی سیاست پولی، آنگریست و کوراستاینر (۲۰۰۴) استدلال می‌کنند که ما بیشتر از اینکه در مورد روندهای تعیین‌کننده **GDP** اطلاع داشته باشیم، در مورد اینکه چگونه فدرال رزرو نرخ بهره را تعیین می‌کند، مطلع هستیم. به عبارت دیگر، ممکن است اعتباربخشی یک مدل برای اختصاص تیمار نسبت به اعتباربخشی به مدلی برای نتایج آسان‌تر باشد (برای نسخه‌ای از این استدلال مراجعه کنید به روزنبایو و رابین، ۱۹۸۵).

تفکری دقیق‌تر و استدلال آماری محض برای استفاده از نمره تمایل توسط آنگریست و هان (۲۰۰۴) پایه‌گذاری شده است. این مقاله نشان می‌دهد که حتی اگرچه هیچ دستاورد مؤثری تقریبی از استفاده تخمین‌زنندگان بر اساس نمره تمایل وجود نداشته باشد، اغلب یک دستاورد با دقت در نمونه‌های محدود وجود خواهد داشت. از آنجا که همه مجموعه‌های داده‌های حقیقی محدود هستند، این نتیجه دارای ربط تجربی است. به طور مستقیم، اگر متغیرهای مشترک حذف شده از نمره تمایل اندکی از تغییرات در دستاوردها را بیان می‌کنند (به مفهوم مطلق آماری)، ممکن است پس از آن بهتر باشد که آنها را نادیده بگیریم تا اینکه دردسر آماریِ نیاز به تخمین تأثیرات آنها را تحمل کرد. وقتی صدها متغیر مشترک وجود دارد که ممکن است دستاوردها را پیش‌بینی کند، بسیار آسان است که بررسی‌هایی را مشاهده کرد که از مجموعه داده‌هایی مثل **NLSY** استفاده شود. در عمل، ما روی زیرمجموعه‌های متغیرهای مشترک محتمل تمرکز می‌کنیم. این زیرمجموعه‌ها با در نظر گرفتن پیش‌بینی‌های تیمار و نتایج انتخاب می‌شوند.

در نهایت، هیرانو، ایمبنز و ریدر (۲۰۰۳) یک راه‌حل جایگزین مجانبی از قضیه پارادوکس نمره تمایل ایجاد شده توسط هان (۱۹۹۸) ارائه کردند. آنان نشان می‌دهند که حتی اگرچه تخمین‌های تأثیرات تیمار بر اساس نمره تمایل مشخص دارای کارایی نباشد، در مورد مدل‌هایی با متغیرهای مشترک پیوسته،

تخمین‌زننده وزنی از نوع هورویتس ـ تامپسون،[1] وقتی وزن‌دهی از تخمین غیرپارامتری نمره[2] استفاده می‌کند، دارای کارایی است. این حقیقت که نمره تمایل، تخمین زده شده و همچنین این حقیقت که این نمره، به صورت غیرپارامتری تخمین زده شده است، نقشی کلیدی برای هیرانو، ایمبنز و ریدر داشته است.

آیا دستاوردهای هیرانو، ایمبنز و ریدر (۲۰۰۳) پارادوکس نمره تمایل را حل می‌کند؟ تا این لحظه، ما راه‌حل نمونه محدود ارائه شده توسط آنگریست و هان (۲۰۰۴) را ترجیح داده‌ایم. نتایج آنان حقیقتی را نشان می‌دهد که این تمایل محققان بوده است که بعضی محدودیت‌هایی روی نمره قائل شده‌اند که این محدودیت‌ها قدرت آماری و مفهومی استنباط مبتنی بر نمره تمایل[3] را به دست می‌دهد. در مورد آنگریست (۱۹۹۸)، به عنوان مثال، تقاضایی با ابعاد بالا اگر چه متغیرهای مشترک متمایز داشته باشد، تخمین‌زننده غیرپارامتری نامحدود نمره احتمال تجربی تیمار در هر سلول متغیر مشترک را داراست. با این تخمین‌زننده غیرپارامتری برای $\mathbf{p}(X_i)$، بسیار ساده می‌توان نشان داد که مشابه‌های نمونه (۳-۳-۱۱) و (۱۲-۳-۳) از حیث جبری معادل تخمین‌زننده‌های جورشدگی متغیر مشترک کامل مربوطه هستند. بنابراین، جای تعجب نیست که تخمین مبتنی بر نمره کاراست، چراکه جورسازی متغیر مشترک کامل دارای جدول زمانی کارامدی مجانبی است. یکی از عناصر ذاتی روش‌های نمره تمایل استفاده از دانش زمینه برای کاهش ابعاد است. این دستاورد آماری، بهبودی در رفتار نمونه‌های محدود[4] است. اگر شما آمادگی هموار کردن، محدود کردن، یا در غیر این صورت کاهش ابعاد مسئله جورشدگی را به روشی که پیامدهای حقیقی تجربی دارد، ندارید، در نتیجه ممکن است دنبال جورسازی متغیر مشترک کامل یا کنترل رگرسیون اشباع بروید.

۳-۳-۳. روش‌های نمره تمایل در مقابل رگرسیون

روش‌های نمره تمایل توجه را از تخمین $E[Y_i|X_i, D_i]$ به سمت تخمین نمره تمایل $\mathbf{p}(X_i) \equiv E[D_i|X_i]$ برده است. این موضوع در کاربردهایی که مورد اخیر برای مدل‌سازی یا انگیزش آنها آسان‌تر است، جذاب است. به عنوان مثال، آشنفلتر (۱۹۷۸) نشان داد که شرکت‌کنندگان در برنامه‌های آموزشی دولتی اغلب دچار مشکل کاهش درآمد هستند و این پدیده‌ای است که در بررسی‌های بعدی مکرراً مشاهده شده است. اگر این کاهش درآمد تنها چیزی باشد که آموزش‌دیدگان را خاص می‌کند، پس می‌توانیم تأثیر علّی آموزش بر درآمد را توسط کنترل پویای درآمدهای گذشته برآورد کنیم. به هر حال، در عمل، جورسازی با پویایی درآمدها سخت است، زیرا سابقه درآمدها پیوسته و چندبُعدی است. دهیجیا و واهبا (۱۹۹۹) در این زمینه استدلال می‌کنند که تأثیرات علّی برنامه‌های آموزشی توسط شرایط نمره تمایل نسبت به شرایط سابقه درآمدها بهتر برآورد می‌شود.

1. Horvitz-thompson-type Weighting Estimator of the Score
2. Non-parametric Estimate
3. Propensity-score-based Inference
4. Finite-sample Behavior

برآوردهای نمره تمایل گزارش شده توسط دهیجیا و واهبا به شکل قابل توجهی نزدیک به برآوردهای آزمون تصادفی است که معیار سنجش آنها را تشکیل می‌دهند. با این حال، ما باور داریم که رگرسیون باید نقطه شروع اکثر پروژه‌های تجربی باشد. البته این امر یک قاعده نیست و بدون شک مواردی وجود دارد که جورسازی نمره تمایل برآوردهای متوسط تأثیرات علّی را تأمین می‌کند. اولین دلیلی که ما خود را در دامنه عمل نمره تمایل پیدا نمی‌کنیم، این است که جزئیات زیادی وجود دارد که باید در زمان اجرای جورسازی نمره تمایل وارد شود ـ مثل چگونگی مدل‌سازی نمره تمایل و اینکه این جزئیات تاکنون استاندارد نشده است. محققان مختلف، حتی وقتی از داده‌ها و متغیرهای مشابهی استفاده می‌کنند، ممکن است به نتایج بسیار متفاوتی برسند. علاوه بر این، همان طوری که درباره برآوردهای هورویتس ـ تامپسون دیدیم، تمایز زیادی بین رگرسیون و وزن‌دهی نمره تمایل وجود ندارد. اگر مدل رگرسیون برای متغیرهای مشترک به خوبی قابل انعطاف باشد، یا نزدیک به اشباع باشد، رگرسیون می‌تواند به عنوان نوعی از وزن نمره تمایل در نظر گرفته شود. بنابراین، تفاوت مذکور بیشتر در زمینه اجراست. در عمل شما ممکن است از اشباع فاصله داشته باشید، اما با وجود متغیرهای مشترک صحیح، این موضوع اهمیتی ندارد.

تفاوت بین رگرسیون و جورسازی نمره تمایل در نمونه برگرفته از طرح کار حمایت شده ملی (**NSW**[1]) توسط دهیجیا و واهبا (۱۹۹۹) نشان داده شده است.[2] **NSW** برنامه‌ای در سال‌های میانی دهه ۱۹۷۰ بود که برای نیروی کار کم‌تجربه، تجربه کاری فراهم می‌کرد. برنامه، **NSW** توسط آزمون تصادفیده ارزیابی شده است، چیزی که برای آن زمان تا حدودی غیرمتعارف به نظر می‌رسید. تحلیل لالوند (۱۹۸۶)، نتایج حاصل از آزمون تصادفیده **NSW** را با نتایج حاصل از اقتصادسنجی با استفاده از گروه‌های کنترل غیرتجربی از **PSID** و **CPS** مورد مقایسه قرار داده است. او پس از این تحلیل دید بدبینانه‌ای پیدا کرد، چراکه روش‌های غیرتجربی دامنه وسیعی از نتایج را به بار آورد، که بسیاری از آنها دور از برآوردهای تجربی بودند. علاوه بر این، لالوند استدلال کرد که هر محقق کاربردی، که نتایج آزمون تصادفی را نمی‌داند، احتمالاً بهترین مشخصات اقتصادسنجی و مشاهدات گروه‌های کنترل را در نظر نخواهد گرفت.

در اقدام ثانوی روی یافته‌های لالوند (۱۹۸۶)، دهیجیا و واهبا (۱۹۹۹) دریافتند که می‌توانند توسط جورسازی گروه تیمار **NSW** با گروه‌های کنترل مشهود انتخاب شده با استفاده از نمره تمایل، به نتایج تجربی **NSW** نزدیک شوند. آنان این موضوع را با استفاده از مقایسه گروه‌های مختلف نشان داده‌اند. پس از دهیجیا و واهبا (۱۹۹۹)، ما دوباره دو گروه مقایسه **CPS** را در نظر می‌گیریم، اول یک نمونه بزرگ

1 National Supported Work

۲. یک رویارویی مشابه اما گسترده‌تر در مورد نمره تمایل در مباحثه میان اسمیت و تاد (۲۰۰۵) و دهجا (۲۰۰۵) دیده می‌شود.

انتخاب نشده (**CPS-1**) و سپس یک گروه مقایسه کوچک انتخاب شده از بیکاران اخیر **(CPS-3)**.

جدول ۲-۳-۳ (تکرار شده جدول ۱ از کار دهیجیا و واهبا، ۱۹۹۹) آمارهای توصیفی از گروه تیمار **NSW** را گزارش می‌دهد، که به طور تصادفی از گروه کنترل **NSW** و از دو گروه کنترل مشاهده انتخاب شده است. گروه تیمار **NSW** و گروه کنترل **NSW** که به طور تصادفی انتخاب شده‌اند، جوان‌تر، با تحصیلات کمتر و به احتمال بیشتر سفیدپوست و دارای درآمدهای پایین‌تر نسبت به جامعه عمومی نشان داده شده توسط نمونه **CPS-1** هستند. نمونه **CPS-3** دارای جورسازی نزدیک‌تری با گروه تیمار **NSW** است، اما با وجود این تفاوت‌هایی، خصوصاً بر حسب نژاد و درآمدهای قبل از برنامه دارند.

جدول ۳-۳-۳ برآوردهای تأثیر تیمار **NSW** را گزارش می‌دهد. متغیر وابسته درآمدهای سالیانه در سال ۱۹۷۸ است که یک سال یا دو سال پس از تیمار است. ردیف‌های جدول نتایج مجموعه‌های کنترل جایگزین را نشان می‌دهد: هیچ، کلیه متغیرهای جامعه در جدول ۲-۳-۳، درآمدهای تأخیری (۱۹۷۵)، درآمدهای جامعه به علاوه تأخیری، درآمدهای جامعه و دو تأخیر. کلیه برآوردها حاصل از درآمد رگرسیون در سال ۱۹۷۸ در درمان موهومی به علاوه کنترل‌هاست (تفاوت کنترل تیمار خام در ردیف اول آمده است).

برآوردهایی که از گروه کنترل تجربی استفاده کردند، در ستون یک گزارش شده است و در حد ۱۶۰۰ تا ۱۸۰۰ دلار است. جای تعجب نیست که این تخمین‌ها با مشخصات اندکی تغییر می‌کنند. بر عکس، فاصله درآمدی خام بین مشارکت‌کنندگان **NSW** و نمونه **CPS-1**، گزارش شده در ستون دو، تقریباً ۸۵۰۰ دلار است که نشان می‌دهد این مقایسه با اریب شدید انتخاب شده است. اضافه کردن کنترل جامعه و درآمدهای تأخیری به طور قابل توجهی این فاصله را کم می‌کند، تأثیر تیمار تخمین زده شده (مثبت) ۸۰۰ دلار و در ردیف آخر قرار دارد. نتایج حتی در ستون سه بهتر هستند، که از گروه مقایسه **CPS-3** استفاده می‌کنند. خصوصیات این گروه خیلی نزدیک‌تر به مشارکت‌کنندگان **NSW** است و با آن مطابقت دارد. تفاوت درآمدهای خام فقط ۶۳۵ دلار است. برآورد کنترل شده کامل، که در ردیف آخر گزارش شده، نزدیک به ۱۴۰۰ دلار است که با تأثیر تیمار تجربی فاصله دارد.

یک ضعف در فرایندی که ما را از **CPS-1** به **CPS-3** می‌رساند، در ماهیت قواعد به کار رفته در ایجاد این گروه مقایسه **CPS-3** است که با دقت انتخاب شده و کوچک‌تر است. شاخص انتخاب **CPS-3** می‌تواند توسط قواعد برنامه **NSW** برانگیخته شود، که به نفع افراد با درآمد کم و تجربه کاری اندک است، اما در عمل، روش‌های متعددی برای اجرای آن موجود است. بنابراین، ما علاقه داریم رویکرد سیستماتیک بیشتری برای غربالگری داشته باشیم. در مقالات اخیر، کرامپ، هوتز، ایمبنز و میتنیک (۲۰۰۶) نشان دادند که نمره تمایل به عنوان پیشگام در برآورد رگرسیونی می‌تواند مورد استفاده سیستماتیک در انتخاب نمونه قرار گیرد. این مسئله با بحث قبلی ما در باره نمره تمایل به عنوان مبنای یک تخمین‌زن، در تناقض است.

ما پیشنهاد کرامپ و همکاران (۲۰۰۶) را توسط تخمین نمره تمایل در تیمار **NSW** و نمونه

مقایسه قابل مشاهده به اجرا درآوردیم و سپس فقط مشاهداتی را در فاصله $\mathbf{0.1<P(X_i) < 0.9}$ در نظر گرفتیم. به عبارت دیگر، نمونه تخمینی محدود به مشاهداتی با پیش‌بینی احتمال تیمار معادل حداقل ۱۰ درصد است، اما بیش از ۹۰ درصد نیست. این امر تضمین می‌کند که رگرسیون با نمونه‌ای شامل فقط سلول‌های متغیر مشترک که حداقل مشاهدات کنترل و تعدادی تیمار دارند، تخمین زده شده است. برآوردهای با استفاده از نمونه‌های غربال شده مستلزم برون‌یابی به سلول‌های بدون حمایت مشترک، به سلول‌هایی که هیچ هم‌پوشانی در توزیع متغیر مشترک بین تیمار و کنترل ندارند، است. آمارهای توصیفی برای نمونه‌های غربال شده بر اساس نمره (برآورد شده با استفاده از مجموعه کامل متغیرهای مشترک لیست شده در جدول) در دو ستون آخر جدول ۳-۳-۲ آمده است. متوسط متغیرهای مشترک در **CPS-1** و **CPS-3** غربال شده، به متوسط **NSW** در ستون یک نسبت به متوسط متغیرهای مشترک نمونه‌های غربال نشده، خیلی نزدیک هستند.

ما با همان متغیرهای مشترک که برای غربالگری و تخمین تأثیرات تیمار در هر تکرار انجام دادیم، غربال‌کننده حمایتی مشترک را با استفاده از مجموعه‌های متغیرهای مشترک جایگزین مشخص کردیم. برآوردهای نتیجه شده در دو ستون آخر جدول ۳-۳-۳ نشان داده شده است. با کنترل متغیرهای جمعیت شناختی، یا، یک تأخیر در درآمد، این نتایج دارای اندکی تفاوت با ستون‌های ۲-۳ بوده است. با کنترل متغیرهای جمعیت‌شناختی، و، یک تأخیر در درآمد، برآوردهای **CPS-1** غربال شده کاملاً نزدیک به برآوردهای تجربی در مقایسه با نتایج غربال نشده است. برآوردهای **CPS-1** غربال شده با دو تأخیر در درآمد، در نزدیکی شاخص معیار تجربی هستند. از طرف دیگر، با یک تأخیر در درآمدها، غربالگر حمایتی مشترک نتایج **CPS-3** را فقط اندکی بهبود بخشیده است و به نظر می‌رسد با در نظر گرفتن دو تأخیر، قدمی هم به عقب باشد.

این تحقیق ایمان ما را به رگرسیون (که قبلاً قوی بود) بیشتر کرده است. کنترل رگرسیون برای متغیرهای مشترک به خوبی حذف اریب گزینش را در نمونه **CPS-1** علی‌رغم فاصله زیاد در خط مبنا عملی می‌کند. محدود کردن نمونه با استفاده از دانش ما نسبت به معیارهای پذیرش در برنامه مذکور، حتی بهتر از برآوردهای رگرسیون در **CPS-3** نتیجه داده است و در حد نتیجه جورشدگی نمره تمایل دهیجیا و واهبا (۱۹۹۹) با دو تأخیر در درآمد بوده است. به نظر می‌رسد غربالگری قبلی سیستماتیک برای حمایت مشترک مثل یک الحاق مفید در برآورد رگرسیون در **CPS-1**، و یک نمونه اولیه انتخاب شده بزرگ باشد. برآوردها در **CPS-1** غربال شده به خوبی **CPS-3** غربال نشده هستند. در هر حال، قابل ذکر است که خطاهای استاندارد در برآوردها با استفاده از نمونه‌های غربال شده نمره تمایل، برای نشان دادن متغیرهای نمونه‌گیری در برآوردهای نمره، تنظیم نشده است. یکی از مزیت‌های پیش‌غربالگری با استفاده از اطلاعات قبلی، مثل مرحله **CPS-1** تا **CPS-3** این است که چنین تنظیمی لازم نیست.

۴-۳. جزئیات رگرسیون

۱-۴-۳. وزن‌دهی به رگرسیون

چیزهای اندکی مثل نقش وزن‌های نمونه‌ها برای محققان کاربردی گیج‌کننده است. حتی اکنون، ۲۰ سال پس از دوره .Ph.D، ما راهنمای نرم‌افزار Stata در وزن‌دهی را تا حدی با ناامیدی مطالعه می‌کنیم. وزن‌ها می‌توانند به روش‌های متعددی به کار روند و چگونگی به کار بردن آنها برای نتیجه‌گیری اهمیت دارد. متأسفانه، موارد له یا علیه وزن دادن اغلب روشن نیست، همان طوری که خصوصیات برنامه‌ریزی برای وزن‌ها مشخص نیست. بحث در جزئیات مورد اشاره طرفداران و مخالفان وزن‌دهی خارج از محدوده این کتاب است. برای این دو دیدگاه به فایفرمن (۱۹۹۳) و دیتون (۱۹۹۷) مراجعه کنید. در این زیرفصل کوتاه، ما چند راهنمایی و منطق کار رویکردمان به وزن دادن را ارائه کرده‌ایم.

یک قانون ساده سرانگشتی برای وزن‌دهی رگرسیون این است که موقعی از وزن‌دهی استفاده کنید که احتمال اینکه رگرسیون تخمینی شما نزدیک به جامعه هدفی است که شما درصدد تخمین هستید، زیاد باشد. اگر به عنوان مثال، هدف (یا متغیر تخمین زده شده) تابع رگرسیون جامعه باشد و نمونه به کار رفته نسبت به وزن نمونه غیرتصادفی، $\boldsymbol{w_i}$، باشد، و نیز مساوی عکس احتمال مشاهده نمونه $\mathbf{i}$ باشد، پس معنادار است که از حداقل مربعات موزون با $\boldsymbol{w_i}$ استفاده کنیم (برای این کار شما می‌توانید از دستور *pweights* در استتا یا WEIGHT در **SAS** استفاده کنید). حتی اگر نمونه‌ای که شما به کار می‌گیرید یک نمونه ساده تصادفی نباشد، وزن‌دهی با احتمال نمونه‌گیری معکوس تخمین‌هایی ایجاد می‌کند که برای تابع رگرسیون جامعه جورسازی دارد.

یکی از سناریوهای وزن‌دهی، در داده‌های گروهی است. فرض کنید که شما، با این فرض قبلی که می‌خواهید درباره بردار رگرسیون جامعه $\boldsymbol{\beta = E[X_iX_i']^{-1}E[X_iY_i]}$ اطلاع داشته باشید، می‌خواهید رگرسیون $\boldsymbol{Y_i}$ را روی $\boldsymbol{X_i}$ در نمونه تصادفی داشته باشید. به عوض نمونه تصادفی، در هر حال، شما دارای داده‌هایی گروهی در سطح $\boldsymbol{X_i}$ هستید. یعنی اینکه، شما تخمین $\mathrm{E}[\boldsymbol{Y_i}|\boldsymbol{X_i} = \boldsymbol{x}]$ را برای هر $\mathbf{x}$ با استفاده از داده‌های نمونه تصادفی دارید. فرض کنید این متوسط توسط $\overline{\boldsymbol{y_x}}$ نشان داده شود و شما $\boldsymbol{n_x}$ را می‌دانید و $\frac{\boldsymbol{n_x}}{\boldsymbol{N}}$ نیز فرکانس نسبی $\mathbf{x}$ در نمونه تصادفی باشد. همان طور که در فصل ۳-۱-۲ دیدیم، رگرسیون $\overline{\boldsymbol{y_x}}$ روی $\mathbf{x}$ ، وزن شده توسط $\boldsymbol{n_x}$ در رگرسیون نمونه تصادفی یکسان است. بنابراین، اگر هدف شما برگشت به رگرسیون داده‌های خُرد باشد، معنادار است که این رگرسیون با اندازه گروه وزن‌دهی شود. قابل ذکر است که در هر حال، اقتصادسنجی کارها، عادت دارند با متوسط‌های منتشره کار کنند و داده‌های خُرد را در نظر نگیرند، یا ممکن است با آن موافقت نکنند، یا شاید نکته اصلی را در نظر بگیرند، اما علاقه‌ای ندارند که آنها را در دستور کار قرار دهند، که به نفع تحلیل بدون وزن‌دهی سرجمع‌هاست.

اگر از سوی دیگر، منطق وزن‌دهی با ناهمگنی ارتباط داشته باشد، مثل اغلب مباحث کتاب‌های درسی در مورد وزن‌دهی، ما حتی حساسیت کمتری به وزن‌دهی نسبت به اقتصاددانان کلان داریم.

جدول ۲-۳-۳. متوسط‌های متغیرهای مشترک در NSW و نمونه‌های کنترل مشاهده شده

Variable	NSW		Full Samples		P-score Screened Samples	
	Treated	Control	CPS-1	CPS-3	CPS-1	CPS-3
	(1)	(2)	(3)	(4)	(5)	(6)
Age	25.82	25.05	33.23	28.03	25.63	25.97
Years of schooling	10.35	10.09	12.03	10.24	10.49	10.42
Black	0.84	0.83	0.07	0.20	0.96	0.52
Hispanic	0.06	0.11	0.07	0.14	0.03	0.20
Dropout	0.71	0.83	0.30	0.60	0.60	0.63
Married	0.19	0.15	0.71	0.51	0.26	0.29
1974 earninigs	2,096	2,107	14,017	5,619	2,821	2,969
1975 earnings	1,532	1,267	13,651	2,466	1,950	1,859
Number of Obs.	185	260	15,992	429	352	157

نکته: برگرفته از دهیجیا و واهبا (۱۹۹۹)، جدول ۱. نمونه‌ها در اولین چهار ستون طبق شرح ستون نمره تمایل بین ۰/۱ و ۰/۹ (.) نمونه‌ها در دو ستون آخر محدود به مشاهدات نمره تمایل بین ۰/۱ و ۰/۹ تخمین‌های نمره تمایل از کل متغیرهای مشترک لیست جدول استفاده کرده است.

جدول ۳-۳-۳. تخمین‌های رگرسیونی تأثیرات آموزش NSW با استفاده از کنترل گزینه‌های جایگزین

Specification	Full Samples			P-Score Screened Samples	
	NSW	CPS-1	CPS-3	CPS-1	CPS-3
	(1)	(2)	(3)	(4)	(5)
Raw Difference	1,794 (633)	-8,498 (712)	-635 (657)		
Demographic controls	1,670 (639)	-3,437 (710)	771 (837)	-3,361 (811) [139/497]	890 (884) [154/154]
1975 Earnings	1,750 (632)	-78 (537)	-91 (641)	no obs. [0/0]	166 (644) [183/427]
Demographics, 1975 Earnings	1,636 (638)	623 (558)	1,010 (822)	1,201 (722) [149/357]	1,050 (861) [157/162]
Demographics, 1974 and 1975 Earnings	1,676 (639)	794 (548)	1,369 (809)	1,362 (708) [151/352]	649 (853) [147/157]

نکته: این جدول، تخمین‌های رگرسیون تأثیرات آموزش با استفاده از داده‌های دهیجیا و واهبا (۱۹۹۹) با مجموعه گزینه کنترل را گزارش می‌کند. کنترل‌های جامعه عبارتند از: سن، سال‌های تحصیلی و نمونه‌های انسانی سیاه‌پوست، اسپانیایی، ترک تحصیلکرده‌های دبیرستانی و ازدواج‌کرده‌ها.

خطاهای استاندارد در پرانتزها گزارش شده‌اند، شمارش‌های مشاهده شده در کروشه‌ها (تیمار شده/کنترل) گزارش شده‌اند.

استدلال وزن‌کشی تحت چندین آشفتگی به طور تقریب به این شکل ادامه می‌یابد: فرض کنید شما به یک **CEF** خطی تمایل دارید، $E[X_i|Y_i] = X_i'\beta$. جمله خطا، تعریف شده به شکل $e_i \equiv \mathbf{Y_i - X_i'\beta}$ ممکن است آشفتگی چندگانه داشته باشد. به عبارت دیگر، تابع شرطی متغیر، $E[e_i^2|X_i]$ لازم نیست ثابت باشد. در این مورد، در حالی که تابع رگرسیون جامعه هنوز مساوی با $E[X_iX_i']^{-1}E[X_iY_i]$ است، مشابه نمونه کارایی ندارد. یک تخمین‌زننده دقیق‌تر از **CEF** خطی با

حداقل مربعات وزن‌کشی شده است، یعنی حداقل شدن مجموع خطاهای مربعات وزن شده توسط یک تخمین $E\left[e_i^2|X_i\right]^{-1}$ است.

همان طور که در فصل ۳-۱-۳ ذکر شد، یک سناریوی تصادفی ذاتی عبارت است از: **LPM**، وقتی که Y_i یک متغیر موهومی است. با فرض اینکه **CEF** در حقیقت خطی است، همان طوری که مدل اشباع است، سپس $E\left[e_i^2|X_i\right] = X_i'\beta(1 - X_i'\beta)$ و بنابراین $1[X_i] = X_i'\beta P[Y_i = 1|X_i]$=، که بدیهی است تابعی از X_i است. این مثالی از آشفتگی مبتنی بر مدل است که اصولاً، تابع شرطی متغیر به سهولت از تخمین تابع رگرسیون تشکیل می‌گردد. تخمین‌زننده حداقل مربعات وزنی ـ یک مورد خاص از حداقل مربعات کلی است (**GLS**) ـ وزن آن توسط $[X_i'\beta(1 - X_i'\beta)]^{-1}$ گرفته می‌شود. در عمل، چون **CEF** فرض شده که خطی است، این وزن‌ها می‌توانند با اولین عبور از **OLS** تخمین زده شوند.

دو دلیل وجود دارد که چرا ما ترجیح می‌دهیم در این مورد وزن‌کشی نکنیم (اگرچه ما یک ماتریس متغیر مشترک پایدار آشفتگی را به کار می‌بریم). اولاً، در عمل، تخمین $E\left[e_i^2|X_i\right]$ ممکن است خیلی خوب نباشد. اگر مدل متغیر مشترک شرطی دارای تقریب ضعیفی باشد/ یا تخمین آن خیلی پُردردسر است (در **LPM**، به معنای این است که **CEF** در واقع خطی نیست)، تخمین حداقل مربعات وزن شده ممکن است دارای خواص نمونه محدود غلط نسبت به تخمین غیروزنی باشد. استنتاج شما بر مبنای نظریه تقریب ممکن است گمراه‌کننده باشد و امید به کسب کارایی تحقق پیدا نمی‌کند. (۲۶) ثانیاً، اگر **CEF** خطی نباشد، تخمین‌زننده حداقل مربعات وزنی دیگر احتمال ندارد که تخمین **CEF** را به عوض تخمین‌زننده غیروزنی بزند. علاوه بر این، تخمین‌زننده غیروزنی هنوز چیزی را که آسان است تخمین می‌زند: این تخمین شامل **MMSE** تقریب خطی برای جامعه **CEF** است.

البته، تخمین‌زننده **GLS** نیز نوعی از تقریب را فراهم می‌کند، اما ذات این تقریب بستگی به وزن‌ها دارد. در حداقل، مقایسه نتایج شما برای تخمین توسط سایر محققان سخت است و راه‌های اضافی را برای جستجوی مشخصات وقتی نتایج بستگی به وزن‌کشی دارد، باز می‌کند. در نهایت، یک تیمار قدیمی به ذهن می‌رسد: اگر این قابل شکستن نیست، آن را ثابت نکنید. تفسیر بردار رگرسیون جامعه توسط آشفتگی تأثیر نمی‌پذیرد، بنابراین چرا باید درباره آن نگران بود؟ هرگونه کسب کارایی از وزن‌کشی مثل اعتدال است و تخمین‌های نادرست یا ضعیف از وزن‌ها می‌تواند خیلی مضرتر از خوب باشد.

۲-۴-۳. آثار حاشیه‌ای[1] و متغیرهای وابسته محدود

در اغلب مطالعات تجربی متغیرهایی وجود دارد که تعداد مقادیر قابل اختصاص به آنها محدود است. به عنوان مثال، در بخش ۲-۴-۳ از این فصل و در فصل متغیرهای ابزاری به تلاش آنگریست و ایوانز (۱۹۹۸) در جهت بررسی اثر توانایی فرزندآوری بر عرضه نیروی کار زنان کارگر اشاره می‌شود. این پژوهش به بررسی

1. Marginal Effects

آثار علّی توانایی فرزندآوری بر کار و درآمد خانواده‌ها می‌پردازد. با توجه به همبستگی احتمالی میان توانایی فرزندآوری و درآمدهای بالقوه خانواده، در این پژوهش متغیرهای ابزاری ارائه شده است که بر اساس ترکیب جنسیتی فرزندان، تعداد تولدها و همچنین برآوردهای **OLS** تخمین زده می‌شوند. تقریباً کلیه نتایج این پژوهش به یکی از دو شکل دودویی (مانند وضعیت شغلی) یا نامنفی[1] (مانند ساعات کاری، هفته‌های کاری و درآمد) هستند. بنابراین، آیا می‌توان چنین نتیجه گرفت که یک متغیر وابسته تأثیر روش تجربی را محدود می‌کند؟ با توجه به مباحث مطرح در بسیاری از کتب مرجع اقتصادسنجی، اگرچه رویکرد **OLS** برای متغیرهای وابسته پیوسته مناسب است، اما استفاده از مدل‌های رگرسیون خطی در حالتی که مطلوب یک متغیر وابسته محدود (**LDV**[2]) باشد، مناسب نیست و استفاده از مدل‌های غیرخطی مانند پروبیت و توبیت ترجیح داده می‌شود. در مقابل، دیدگاه ما نسبت به رگرسیون همانند وارثی است که مشروعیت خود را از **CEF** وام گرفته است، روشی که **LDV** بودن را از مرکز توجه کنار زد.

مثل همیشه، آزمون تصادفی یک معیار مفید است (زیرا در این حالت، عمل رگرسیون به یک اختلاف ساده میان مقادیر کنترلی ـ تیماری تبدیل می‌شود). نتیجه رگرسیون متغیرهای نتیجه مختلف وابسته به یک رگرسور با تخصیص تصادفی را در نظر بگیرید. این رگرسور معرف یکی از گروه‌های درمانی در آزمایش بیمه درمانی رند است (**HIE**، منینگ و همکاران، ۱۹۸۷). در این آزمایش بلندپروازانه (که احتمالاً پُرهزینه‌ترین مطالعه علمی در حوزه علوم اجتماعی آمریکاست)، شرکت رند یک شرکت بیمه درمانی کوچک تأسیس کرد که بیمه‌شدگان آن هیچ حق بیمه‌ای نمی‌پرداختند. حدود ۶ هزار نفر در این پژوهش مشارکت داشتند که به طور تصادفی در هریک از برنامه‌های مختلف بیمه درمانی گروه‌بندی شدند.

یکی از مهم‌ترین ویژگی‌های هر برنامه بیمه‌ای سهم افراد بیمه شده در پرداخت هزینه‌های تیمار سلامت بود. HIE به طور تصادفی افراد را در برنامه‌های بیمه‌ای مختلف گروه‌بندی کرد. در یکی از این برنامه‌ها خدمات درمانی کاملاً رایگان انجام می‌شد. در سایر برنامه‌ها ترکیبی از برنامه‌های مختلف مشارکت در پرداخت هزینه، پوشش هزینه‌ها و برنامه‌های تخفیفی ارائه می‌شد؛ به نحوی که بخشی از هزینه‌های تیمار سلامت بیماران از طریقی جز پرداخت هزینه توسط بیمار تأمین می‌شد. هدف اصلی از انجام این آزمایش آگاهی از حساسیت استفاده از خدمات تیمار پزشکی به هزینه و نحوه تأثیر آن بر سلامت بود. بر اساس نتایج حاصل از HIE میزان استفاده اعضای دو گروه تیمار پزشکی رایگان و ارزان نسبت به اعضای سایر گروه‌ها بالاتر بود، اما در اکثر موارد این موضوع به سلامت بیشتر این افراد منجر نشده بود. این یافته‌ها به هموار شدن مسیر تیمار سلامت مدیریت شده و طراحی برنامه‌های بیمه‌ای درمانی حساس به هزینه کمک کرد.

1. Non-negative
2. Limited Dependent Variable

اکثر نتایج حاصل از HIE از نوع LDV بود. این یافته‌ها از تحمیل هزینه‌های پزشکی یا بستری در یک سال مشخص به افراد شرکت‌کننده و نتایج نامنفی مانند تعداد دفعات ویزیت حضوری و هزینه‌های هنگفت سالیانه پزشکی (پرداختی توسط بیمار یا شرکت بیمه) حکایت دارند. متغیر هزینه برای حدود ۲۰ درصد از جامعه نمونه صفر است. نتایج مربوط به دو گروه درمانی HIE در جدول ۱-۴-۳ از تخمین موارد مندرج در جدول ۲ از پژوهش منینگ و همکاران (۱۹۸۷) به دست آمده است. متوسط نتایج مربوط به گروه‌های تیماری رایگان و تخفیفی فردی در جدول ۱-۴-۳ ارائه شده است. در گروه آخر، افراد از تخفیف سالیانه ۱۵۰ دلاری برای هر فرد یا ۴۵۰ دلاری برای هر خانواده در حوزه تیمارهای سرپایی برخوردارند و در سایر موارد کلیه هزینه‌ها توسط شرکت بیمه پوشش داده می‌شود (لذا، پرداختی بیمار در موارد بستری صفر است). اندازه جامعه نمونه در این دو گروه اندکی بیش از ۳۰۰۰ نفر در نظر گرفته شده بود.

جدول ۱-۴-۳. میانگین نتایج حاصل در دو گروه درمانی HIE

Plan	Face-to-face visits	Outpatient Expenses (1984$)	Admis-sions	Prob. Any Medical (%)	Prob. Any Inpatient (%)	Total Expenses (1984$)
Free	4.55	340	.128	86.8	10.3	749
	(.168)	(10.9)	(.0070)	(.817)	(.45)	(39)
Individual	3.02	235	.115	72.3	9.6	608
Deductible	(.171)	(11.9)	(.0076)	(1.54)	(.55)	(46)
Deductible	-1.53	-105	-0.013	-14.5	-0.7	-141
minus free	(.240)	(16.1)	(.0103)	(1.74)	(.71)	(60)

نکته: برگرفته از مقاله منینگ (۱۹۸۷)، جدول ۲. کلیه خطاهای استاندارد (نوشته شده داخل پرانتز) با توجه به همبستگی بین زمانی و درون‌خانوادگی اصلاح شده‌اند. مقادیر به نرخ دلار ژوئیه ۱۹۸۴ بیان شده است. ویزیت‌ها به صورت حضوری با **MD**، **DO** یا سایر ارائه‌دهندگان خدمات بهداشتی انجام شده است. ویزیت‌های انحصاری فقط مخصوص خدمات رادیولوژی، بیهوشی یا آسیب‌شناختی است. تیمارهای دندانپزشکی و روان‌درمانی سرپایی جزء موارد ویزیت و هزینه‌ها محسوب نشده است.

به منظور ساده‌سازی بررسی LDV چنین فرض می‌شود که تنها مقایسه مدنظر محققان این پژوهش مقایسه میان نتایج دو گروه برنامه تیماری رایگان و تخفیفی است و این کار با استفاده از تخصیص تصادفی ساده انجام می‌شود.[1] $D_i = 1$ به معنای تخصیص به گروه تخفیفی است. به واسطه تخصیص تصادفی اختلاف میان اعضای دو گروه $D_i = 1$ و $D_i = 0$ نشان‌دهنده اثر روش اتخاذی در درمان است. با توجه به نکاتی که پیش از این در مورد آزمایش‌ها مطرح شد (فصل ۲):

۱. HIE واقعی بسیار پیچیده‌تر از آن چیزی است که در این سطور تشریح شده است. در این برنامه بیمه درمانی ۱٤ روش مختلف از جمله خدمت پیش‌پرداخت شبه HMO وجود دارد. طرح آزمایشی از تخصیص تصادفی ساده استفاده نمی‌کند، بلکه از یک رویکرد تصادفی پیچیده‌تر بهره می‌گیرد که به معنای تضمین توازن متغیرهای تصادفی کمکی در میان گروه‌هاست.

$$E[\mathrm{Y}_i|\mathrm{D}_i=1]-E[\mathrm{Y}_i|\mathrm{D}_i=0] \quad (3.4.1)$$
$$= E[\mathrm{Y}_{1i}|\mathrm{D}_i=1]-E[\mathrm{Y}_{0i}|\mathrm{D}_i=1]$$
$$= E[\mathrm{Y}_{1i}-\mathrm{Y}_{0i}]$$

زیرا D_i از نتایج بالقوه مستقل است. همچنین، مشابه قبل، $E[\mathrm{Y}_i|\mathrm{D}_i=1]-E[\mathrm{Y}_i|\mathrm{D}_i=0]$ ضریب شیب رگرسیون Y_i روی D_i است.

بنابر رابطه (۱-۴-۳) برآورد اثرات علّی در آزمایش‌ها فارغ از نحوه توزیع Y_i اعم از دودویی، نامنفی یا پیوسته هیچ چالش ویژه‌ای ایجاد نمی‌کند. تفسیر سمت راست به ازای انواع مختلف متغیرهای وابسته متفاوت خواهد بود، اما تعیین متوسط اثر علّی به انجام اقدامات پیچیده نیازی ندارد. به عنوان مثال، یکی از نتایج HIE متغیری است که به هزینه‌های پزشکی اشاره دارد. با توجه به اینکه توزیع این نتیجه از نوع برنولی است، داریم:

$$E[\mathrm{Y}_{1i}-\mathrm{Y}_{0i}]=E[\mathrm{Y}_{1i}]-E[\mathrm{Y}_{0i}]=P[\mathrm{Y}_{1i}=1]-P[\mathrm{Y}_{0i}=1]. \quad (3.4.2)$$

هرچند ممکن است این ارتباط بر زبان مورد استفاده ما در تشریح نتایج اثر بگذارد، اما به هیچ‌وجه بر محاسبات اساسی اثرگذار نیست. به عنوان مثال، همان طور که سمت چپ رابطه (۱-۴-۳) نشان می‌دهد، نتیجه مقایسه گروه‌های آزمایشی در HIE حاکی از آن است که ۸۷ درصد از اعضای گروه تیماری رایگان سالیانه حداقل چند بار از امکانات تیماری استفاده می‌کنند؛ این در حالی است که آمار استفاده سالیانه اعضای گروه تخفیفی از امکانات تیماری حدود ۷۲ درصد است. بنابراین، اثر تخفیف معمولی ۱۵۰ دلاری بر میزان استفاده از امکانات تیماری قابل ملاحظه است. اختلاف میان این دو نرخ ۰/۱۵- (**s.e.=0.017**) معادل برآورد $E[\mathrm{Y}_{1i}-\mathrm{Y}_{0i}]$ است که در آن، Y_i یک متغیر ظاهری است. متوسط اثر علّی نیز معادل اثر علّی بر احتمال یا نرخ استفاده است.

بنابراین، در این مورد به جای استفاده از پروبیت برای جورشدگی CEF متغیر نتیجه به صورت یک احتمال در نظر گرفته می‌شود. این موضوع هیچ اشکالی در روند انجام کار ایجاد نمی‌کند. مدل پروبیت معمولاً با فرض تعیین مشارکت توسط یک متغیر پنهان Y_i^* که در رابطه زیر صدق می‌کند، به دست می‌آید.

$$\mathrm{Y}_i^*=\beta_0^*+\beta_1^*\mathrm{D}_i-\nu_i, \quad (3.4.3)$$

که در آن، ν_iمعرف توزیع $N(0,\sigma^2)$ است. شایان ذکر است که این متغیر در ارائه هزینه واقعی درمانی ناتوان است؛ زیرا هزینه یک متغیر نامنفی است و در نتیجه توزیع نرمال ندارد. در حالی که توزیع متغیرهای با توزیع نرمال روی محور حقیقی پیوسته است و در نتیجه امکان منفی شدن آن نیز وجود دارد. با توجه به مدل شاخص پنهان،

$$\mathrm{Y}_i = 1[\mathrm{Y}_i^* > 0],$$

می‌توان CEF را به صورت زیر نوشت:

$$E[\mathrm{Y}_i|\mathrm{D}_i] = \Phi[\frac{\beta_0^* + \beta_1^*\mathrm{D}_i}{\sigma}],$$

که در آن، $\Phi[\cdot]$ معرف **CDF** نرمال است. بنابراین،

$$E[\mathrm{Y}_i|\mathrm{D}_i] = \Phi[\frac{\beta_0^*}{\sigma}] + \{\Phi[\frac{\beta_0^* + \beta_1^*}{\sigma}] - \Phi[\frac{\beta_0^*}{\sigma}]\}\mathrm{D}_i$$

رابطه فوق یک تابع خطی از رگرسور D_i است، بنابراین ضریب شیب رگرسیون Y_i روی D_i دقیقاً با اختلاف مقادیر منطبق شده روی مدل پروبیت $\Phi[\frac{\beta_0^*+\beta_1^*}{\sigma}] - \Phi[\frac{\beta_0^*}{\sigma}]$ برابر است. با این حال، توجه داشته باشید که ضرایب پروبیت، یعنی $\frac{\beta_0^*}{\sigma}$ و $\frac{\beta_1^*}{\sigma}$، (با وجود علامت درست) بدون ورود به **CDF** نرمال هیچ اطلاعاتی از میزان اثر D_i روی مشارکت ارائه نمی‌دهند.

یکی از مهم‌ترین دستاوردهای HIE هزینه‌های هنگفت پزشکی و به بیان دیگر هزینه‌های تیمارهای سلامت است. آیا از نظر هزینه، میزان بهره‌گیری اعضای گروه‌های تخفیفی از امکانات تیماری نسبت به سایرین کمتر بوده است؟ متوسط اختلاف میان هزینه دو گروه تیماری رایگان و تخفیفی در HIE ۱۴۱- دلار (s.e.=60) است که حدوداً معادل ۱۹ درصد از مخارج درمانی گروه تیماری رایگان است. بنابراین، با پرداخت بخشی از هزینه‌های درمان از سوی بیماران مخارج طرح تنها اندکی کاهش می‌یابد، اما این تخمین خیلی دقیق نیست.

از آنجایی که متغیرهای خروجی هزینه متغیرهای تصادفی نامنفی و گاهی صفر هستند، امید این متغیرها برابر است با:

$$E[\mathrm{Y}_i|\mathrm{D}_i] = E[\mathrm{Y}_i|\mathrm{Y}_i > 0, \mathrm{D}_i]P[\mathrm{Y}_i > 0|\mathrm{D}_i]$$

اختلاف میان هزینه‌های انجام شده در این گروه‌های درمانی برابر است با:

$$\begin{aligned}
& E[\mathrm{Y}_i|\mathrm{D}_i = 1] - E[\mathrm{Y}_i|\mathrm{D}_i = 0] \qquad (3.4.4)\\
= \; & E[\mathrm{Y}_i|\mathrm{Y}_i > 0, \mathrm{D}_i = 1]\, P[\mathrm{Y}_i > 0|\mathrm{D}_i = 1] - E[\mathrm{Y}_i|\mathrm{Y}_i > 0, \mathrm{D}_i = 0]\, P[\mathrm{Y}_i > 0|\mathrm{D}_i = 0]\\
= \; & \underbrace{\{P[\mathrm{Y}_i > 0|\mathrm{D}_i = 1] - P[\mathrm{Y}_i > 0|\mathrm{D}_i = 0]\}}_{\text{اثر مشارکت}} E[\mathrm{Y}_i|\mathrm{Y}_i > 0, \mathrm{D}_i = 1]\\
& + \underbrace{\{E[\mathrm{Y}_i|\mathrm{Y}_i > 0, \mathrm{D}_i = 1] - E[\mathrm{Y}_i|\mathrm{Y}_i > 0, \mathrm{D}_i = 0]\}}_{\text{اثر هزینه}} P[\mathrm{Y}_i > 0|\mathrm{D}_i = 0].
\end{aligned}$$

بنابراین، اختلاف کلی میان متوسط هزینه‌های انجام شده به دو بخش قابل تفکیک است: یکی بخشی که به اختلاف میان احتمال مثبت بودن هزینه‌ها اختصاص دارد (و اغلب اثر مشارکت نامیده می‌شود) و دیگری بخشی که به اختلاف در مقادیر متوسط مشروط به مشارکت اختصاص دارد و به اثر

مشروط بر مثبت بودن (**COP**[1]) موسوم است. با این حال، در این مورد نیز این موضوع هیچ محدودیت خاصی در فرایند تخمین اثرات علّی تحمیل نمی‌کند و رابطه (۱-۴-۳) همچنان صادق است و اثر درمانی متوسط جامعه برای هزینه‌ها از رگرسیون Y_i روی D_i به دست می‌آید.

COP خوب، COP بد: اثرات مشروط بر مثبت بودن

با توجه به دوبخشی بودن اثر مذکور بر یک متغیر تصادفی نامنفی مانند هزینه، بسیاری از محققان به طور جداگانه به بررسی این بخش‌ها می‌پردازند. در حقیقت، اغلب از «مدل دوبخشی» استفاده می‌کنند که بخش نخست آن یک ارزیابی از اثر مشارکت و بخش دوم اثر COP است (به منظور شناخت این چنین مدل‌های مورد استفاده در HIE به مقاله دوان و همکاران، ۱۹۸۳ و ۱۹۸۴ مراجعه فرمایید). بخش اول رابطه (۴-۴-۳) هیچ مشکل خاصی ایجاد نمی‌کند، زیرا همان طور که پیش از این نیز مطرح شد، متغیر موهومی Y_i در حقیقت فقط به معنای اثرات متوسط درمانی است که میان احتمالات آنها نیز اختلافاتی وجود دارد. مشکل مدل دوبخشی عدم وجود تفسیر علّی برای اثرات COP (حتی در آزمایش‌های تصادفی) است. این موضوع دقیقاً مشابه مشکل انتخاب به وجود آمده در بخش ۳-۲-۳ در کنترل بد است. به منظور تحلیل بیشتر اثرات COP رابطه را به شکل زیر می‌نویسیم.

$$
\begin{aligned}
& E\left[\mathrm{Y}_i \mid \mathrm{Y}_i>0, \mathrm{D}_i=1\right]-E\left[\mathrm{Y}_i \mid \mathrm{Y}_i>0, \mathrm{D}_i=0\right] \qquad (3.4.5) \\
= \; & E\left[\mathrm{Y}_{1i} \mid \mathrm{Y}_{1i}>0\right]-E\left[\mathrm{Y}_{0i} \mid \mathrm{Y}_{0i}>0\right] \\
= \; & \underbrace{E\left[\mathrm{Y}_{1i}-\mathrm{Y}_{0i} \mid \mathrm{Y}_{1i}>0\right]}_{\text{اثر علی}}+\underbrace{\left\{E\left[\mathrm{Y}_{0i} \mid \mathrm{Y}_{1i}>0\right]-E\left[\mathrm{Y}_{0i} \mid \mathrm{Y}_{0i}>0\right]\right\}}_{\text{اثر انتخاب}} .
\end{aligned}
$$

بنابراین، اثر COP از دو جمله تشکیل شده است: یک اثر علّی برای آن دسته از افراد جامعه نمونه که فقط در صورت رایگان بودن از امکانات تیمار پزشکی استفاده می‌کنند و اختلاف Y_{0i} بین آن دسته از افراد جامعه نمونه که فقط در صورت رایگان بودن از امکانات تیمارپزشکی استفاده می‌کنند و آن دسته از افراد جامعه نمونه که تنها در صورت الزام به پرداخت وجه از امکانات تیمارپزشکی استفاده می‌کنند. این جمله دوم نوعی اریب انتخاب است، اگرچه این اریب در مقایسه با اریب انتخاب مطرح در فصل ۲ دقیق‌تر است.

در این بخش، اریب انتخاب از تغییر ترکیب گروه‌ها با هزینه مثبت توسط آزمایش نشئت می‌گیرد. جامعه $\mathrm{Y}_{0i} > 0$ احتمالاً شامل مصرف‌کنندگان کم‌هزینه‌ای است که در صورت الزام به پرداخت کسری گزینه خودداری از درمان را انتخاب می‌کنند. به عبارت دیگر، اندازه این گروه نسبت به گروه با $\mathrm{Y}_{1i} > 0$ بزرگ‌تر و متوسط هزینه آن نسبت به گروه با $\mathrm{Y}_{1i} > 0$ کمتر است. بنابراین جمله اریب ناشی از انتخاب مثبت است و در نتیجه، اثرات COP نسبت به اثر علّی منفی به صفر نزدیک‌تر

1. Conditional-on-positive Effect

است $E[Y_{1i}-Y_{0i}|Y_{1i} > 0]$. این مورد یک نسخه از مشکل کنترل بد بخش ۳-۲-۳ است: در تنظیم اثرات علّی، $Y_i > 0$ یک متغیر نتیجه است و لذا برای شایسته‌سازی مناسب نیست؛ مگر اینکه روش انتخابی بر احتمال مثبت بودن Y_i هیچ تأثیری نداشته باشد. یکی از نتایج غیرعلّی بودن اثرات COP به مدل‌های رگرسیون سانسور شده مثل توبیت بستگی دارد. بر اساس این مدل‌ها باید یک مؤلفه هزینه پنهانی برای افراد غیر از افراد مشارکت‌کننده در نظر گرفت (به عنوان مثال هی و اولسن، ۱۹۸۴). بر اساس فرمول‌بندی سنتی مسائل هزینه در مدل توبیت Y_i مشاهده شده از رابطه زیر به دست می‌آید:

$$Y_i = 1[Y_i^* > 0]Y_i^*$$

که در آن، Y_i^* توزیع نرمال متغیر پنهان هزینه است که امکان پذیرفتن مقادیر منفی را نیز دارد. از آنجایی که Y_i^* یک LDV نیست، طرفداران مدل توبیت به راحتی و با استفاده از یک مدل خطی کلاسیک میان آن و D_i پیوند می‌زنند (ر.ک. رابطه ۳-۴-۳). در این حالت، β_1^* اثر علّی D_i روی متغیر هزینه پنهان Y_i^* است. این رابطه فارغ از علامت Y_i به ازای تمام موارد تعریف شده است. اگر به مطالعه اثرات روی Y_i^* بسنده کنیم، هیچ مشکل انتخاب سبک COP وجود نخواهد داشت.

اما اگر مطالعه اثرات روی Y_i^* بسنده نکنیم، اولین مشکل مسئله معماگونه «هزینه پنهان تیمار سلامت» است.[1] هزینه تیمار سلامت برای برخی مردم واقعاً صفر است، این حقیقت یک ساخته آماری یا ناشی از حدی از سانسور داده‌ها نیست. بنابراین، درک مفهوم پنهان و بالقوه Y_i^* منفی دشوار است. هیچ داده‌ای از Y_i^* وجود ندارد و البته هرگز وجود نخواهد داشت. مشکل دوم تبدیل مفروضات توزیع شده مربوط به متغیر پنهان به واسطه ارتباط میان پارامتر β_1^* در مدل پنهان و اثرات علّی در نتیجه مشاهده شده است. برای ایجاد این پیوند به ارزیابی امید Y_i به ازای D_i معلوم می‌پردازیم.

$$E[Y_i|D_i] = \Phi\left[\frac{\beta_0^* + \beta_1^* D_i}{\sigma}\right][\beta_0^* + \beta_1^* D_i] + \sigma\phi\left[\frac{\beta_0^* + \beta_1^* D_i}{\sigma}\right] \qquad (3.4.6)$$

که در آن، σ معرف انحراف استاندارد ν_i است (ر.ک. ای. جی. ام سی دونالد و مافیت، ۱۹۸۰). این بیان با فرض نرمال بودن و همواریانسی ν_i و همچنین با فرض امکان ارائه Y_i به شکل $1[Y_i^* > 0]Y_i^*$ و نیز ضرایب پنهان همراه است.

CEF توبیت بیانی از اثر روش درمانی بر هزینه‌های مشاهده شده ارائه می‌دهد. به طور مشخص، این بیان یک بیان نسبتاً دشوار است، اما از آنجایی که تنها متغیر شرطی یک متغیر موهومی D_i است، مشخص بودن هیچ‌یک از این موارد برای تخمین $E[Y_i|D_i = 1] - E[Y_i|D_i = 0]$ ضروری

---

۱. مدل انتخاب نمونه یک بسط از مدل توبیت است که در آن، متغیر پنهان تعیین‌کننده مشارکت متغیری جز متغیر هزینه پنهان است. ر.ک. ای. جی. مادالا (۱۹۸۳). در مدل انتخاب نمونه نیز مانند مدل توبیت مشکلات مفهومی مربوط به تفسیر اثرات متغیرهای پنهان به وجود می‌آید.

نیست. فارغ از اتخاذ یا عدم اتخاذ مدل توبیت برای تشریح ساختار اساسی، ضریب شیب حاصل از رگرسیون **OLS** Y_i روی D_i اختلاف **CEF** سمت چپ رابطه ۷-۴-۳ را بهبود می‌بخشد.

$$E[Y_i|D_i=1]-E[Y_i|D_i=0] \tag{3.4.7}$$
$$= \left\{\Phi\left[\frac{\beta_0^*+\beta_1^*}{\sigma}\right][\beta_0^*+\beta_1^*]+\sigma\phi\left[\frac{\beta_0^*+\beta_1^*}{\sigma}\right]\right\}-\left\{\Phi\left[\frac{\beta_0}{\sigma}\right][\beta_0^*]+\sigma\phi\left[\frac{\beta_0^*}{\sigma}\right]\right\}$$

گاهی اثرات **COP** به واسطه احساس پژوهشگر برانگیخته می‌شوند. اگر توزیع نتیجه یک نقطه متراکم داشته باشد (یعنی روی یک مقدار خاص مثل صفر انباشته شده باشد) یا تغییرات توزیع نتیجه شدید باشد، آنگاه در صورت تحلیل اثرات با استفاده از مقدار متوسط اثر برخی اطلاعات موجود نادیده گرفته خواهد شد. در واقع، با تحلیل اثرات با استفاده از مقادیر متوسط برخی موارد مانند تغییر احتمال مقادیر ویژه یا دور شدن چارک‌ها از میانه از بین می‌روند، اما چرا به طور مستقیم به بررسی این اثرات توزیع نپردازیم؟ یک گزینه حساس به اثرات **COP** بررسی مستقیم اثرات روی توزیع یا چارک‌هاست. نتایج توزیع شامل احتمال افزایش مخارج سالیانه پزشکی از صفر، ۱۰۰ دلار، ۲۰۰ دلار و ... است. این به معنای قرار دادن $1[Y_i > c]$ به ازای مقادیر مختلف **c** در سمت چپ رگرسیون مطلوب است. از نظر اقتصادی، کلیه این نتایج در گروه معادله ۲-۴-۳ دسته‌بندی می‌شوند. ایده بررسی مستقیم اثرات توزیع با مدل‌های احتمال خطی توسط آنگریست (۲۰۰۱) در تحلیل اثرات توان فرزندآوری بر ساعات کاری افراد مطرح شد. در عوض، اگر چارک‌ها یک نقطه کانونی به وجود آورند، می‌توان از رگرسیون چارکی برای مدلسازی آنها استفاده کرد. این ایده در فصل هفتم با جزئیات بیشتر مطرح خواهد شد.

آیا مدل‌های متغیر پنهان از نوع توبیت نیز حساسیت ایجاد می‌کنند؟ بله، اگر داده‌های مورد استفاده شما به درستی سانسور شده باشند. سانسور درست به معنای وجود یک همتای تجربی برای متغیر پنهان است که حاصل سهم اولیه است. داده‌های درآمد پیمایش جاری جمعیت **CPS**[1] یک مثال پیشرو از اقتصاد کار است که درآمد بسیار بالای رمزنگاران ارشد آن (سانسورکننده‌ها) به جهت حفاظت از موارد محرمانه مربوطه است. معمولاً، یکی از موضوعات مورد علاقه ما بررسی اثر علّی تحصیل بر درآمد افراد است، زیرا این عامل اثر خود را در بازپرداخت‌های مالیاتی افراد و نه در دریافتی‌های رمزنگار ارشد **CPS** نشان می‌دهد. بر اساس یافته‌های چمبرلین (۱۹۹۴)، در طول سالیان اخیر، رمزنگاری سطح بالای **CPS** بازدهی‌های اندازه‌گیری شده تحصیل را به شدت کاهش داده است. او همچنین اصلاحیه‌ای برای سانسور اطلاعات بر اساس مدل توبیت رگرسیون چارکی ارائه نمود. موضوع استفاده از رگرسیون چارکی برای مدلسازی داده‌های سانسور شده نیز در فصل ۷ ارائه خواهد شد.[2]

---

1. Current Population Survey

۲. توجه داشته باشید که ممکن است نمونه رگرسیون مورد نظر ما ـ رگرسیون کسی درآمد روی سطح تحصیل ـ یک مشکل COP داشته باشد؛ زیرا معمولاً افراد با درآمد صفر از نمونه کسب درآمد حذف می‌شوند. اگر تحصیلات بر احتمال کسب شغل مؤثر باشد، این موضوع به مشکل بایاس انتخاب سبک COP منجر خواهد شد. بنابراین، در عمل، بر نمونه‌هایی از مردان در

متغیرهای کمکی غیرخطی‌ساز[1]

سانسور درست مانند آنچه توسط رمزنگار ارشد **CPS** انجام می‌شود به ندرت اتفاق می‌افتد. همین حقیقت باعث باقی ماندن حوزه‌های کاربردی سازنده محدود برای مدل‌های از نوع توبیت در کارهاست. با این حال، باید در این نقطه اندکی تعلل کنیم. بخشی از سادگی بحث آزمایش‌ها از این حقیقت نشئت می‌گیرد که $E[Y_i|D_i]$ الزاماً یک تابع خطی از D_i است، به نحوی که رگرسیون و **CEF** یکسان و مشابه هم هستند. در حقیقت، این **CEF** برای هر تابع Y_i شامل شاخص توزیع $1[Y_i > c]$ خطی است. البته در عمل، متغیر توصیفی مورد نظر همیشه یک متغیر موهومی نیست و معمولاً متغیرهای کمکی اضافی در **CEF** وجود دارد که در این حال $E[Y_i|X_i,D_i]$ برای **LDV**ها تقریباً غیرخطی است. در واقع، هرچه مقادیر پیش‌بینی شده به مرزهای متغیر وابسته نزدیک‌تر شود (به خاطر نزدیک بودن برخی سلول‌های متغیرهای کمکی به مرزها)، مشتقات **CEF** بر حسب **LDV**ها کوچک‌تر خواهد شد (به عنوان مثال، در مورد نحوه هموار شدن **CDF** نرمال در مقادیر اکسترمم فکر کنید).

بنابراین، در مدل‌های **LDV** شامل متغیرهای کمکی، رگرسیون باید دقیقاً بر **CEF** منطبق باشد. با این حال، موضوع قائل شدن یک تفسیر علّی برای **CEF** اساسی در شرایط وجود **CIA** همچنان صادق خواهد بود و اگر **CEF** یک تفسیر علّی داشته باشد، تلقی وجود یک تفسیر علّی برای رگرسیون نیز منطقی است؛ زیرا هنوز تقریب **MMSE** را برای **CEF** تأمین می‌کند. افزون بر این، اگر این مدل برای متغیرهای کمکی اشباع شده باشد، رگرسیون نیز یک اثر متوسط وزن‌دار مشابه (۱-۳-۳) و (۳-۳-۳) برآورد خواهد کرد. به همین ترتیب، اگر رگرسور مورد نظر چندمقداری یا پیوسته باشد، حاصل یک مشتق متوسط وزنی مشابه فرمول ارائه شده در بخش ۱-۳-۳ خواهد بود.

در عین حال، اغلب اطلاعات ما از مشخصات رگرسیون متغیرهای کمکی اشباع شده به حد کافی نیست و این موضوع از جذابیت این حوزه می‌کاهد. بنابراین، رگرسیون برخی ویژگی‌های **CEF** را از بین می‌برد. ممکن است در یک مورد مقادیر جورشدگی تولید شده خارج از مرزهای **LDV** قرار گیرند. این موضوع بسیاری از محققان را به دردسر می‌اندازد و مطمئناً مشکلات زیادی برای مدل احتمال خطی ایجاد می‌کند. یک ویژگی جذاب مدل‌های غیرخطی مانند پروبیت و توبیت تولید **CSF**هایی است که درون مرزهای **LDV** قرار می‌گیرند. به طور خاص، مقادیر انطباق یافته پروبیت همواره بین صفر و یک قرار دارند در حالی که مقادیر انطباق یافته توبیت مثبت هستند (این موضوع از رابطه ۶-۴-۳ مشخص نیست). بنابراین، در موارد ساده از انطباق منحنی استفاده از مدل‌های غیرخطی ترجیح داده می‌شود.

نکته تصدیق شد. با این حال تأکید بر لزوم تبدیل خروجی مدل‌های غیرخطی به اثرات حاشیه‌ای

سنین اولیه کار متمرکز خواهیم بود که نرخ مشارکت آنها در گروه‌های تحصیلی بالاست و از پایداری منطقی برخوردار است.
1. Covariates Lead to Nonlinearity

ضروری است. اثرات حاشیه‌ای تغییرات (متوسط) **CEF** هستند که به وسیله یک مدل غیرخطی به دست می‌آیند. بدون وجود اثرات حاشیه‌ای بحث در مورد اثر متغیرهای وابسته مشاهده شده دشوار خواهد بود. با فرض اینکه همچنان D_i به عنوان رگرسور مورد نظر عمل می‌کند، می‌توان متوسط اثرات حاشیه‌ای جامعه را با اختلاف زیر نیز به دست آورد:

$$E\{E[Y_i|X_i, D_i = 1] - E[Y_i|X_i, D_i = 0]\},$$

یا با مشتق‌گیری:

$$E\left\{\frac{\partial E[Y_i|X_i, D_i]}{\partial D_i}\right\}$$

بسیاری از افراد هنگام کار با رگرسورهای چندمقداری یا پیوسته نیز از مشتق‌ها استفاده می‌کنند.

برآوردهای رگرسیون OLS چقدر به اثرات حاشیه‌ای ناشی از مدل‌های غیرخطی مانند توبیت و پروبیت نزدیک است؟ در پاسخ به این سؤال نخست اثرات حاشیه‌ای را به دست می‌آوریم و سپس یک مثال تجربی ارائه خواهیم کرد. CSF پروبیت برای یک مدل با متغیرهای کمکی عبارت است از:

$$E[Y_i|X_i, D_i] = \Phi\left[\frac{X_i'\beta_0^* + \beta_1^* D_i}{\sigma}\right].$$

بنابراین، میانگین اختلاف متناهی برابر است با:

$$E\left\{\Phi\left[\frac{X_i'\beta_0^* + \beta_1^*}{\sigma}\right] - \Phi\left[\frac{X_i'\beta_0^*}{\sigma}\right]\right\}. \qquad (3.4.8)$$

که در عمل می‌توان آن را به وسیله مشتق متوسط تقریب زد.

$$E\left\{\Phi\left[\frac{X_i'\beta_0^* + \beta_1^* D_i}{\sigma}\right]\right\} \cdot (\beta_1^*/\sigma)$$

(Stata اثرات حاشیه‌ای هر دو مسیر را محاسبه می‌کند، اما در انجام (۸-۴-۳) برای رگرسورهای موهومی ناتوان است).

به همین ترتیب، با تعمیم رابطه ۶-۴-۳ به مدلی با متغیرهای کمکی رابطه زیر برای یک LDV نامنفی برقرار است.

$$E[Y_i|X_i, D_i] = \Phi\left[\frac{X_i'\beta_0^* + \beta_1^* D_i}{\sigma}\right][X_i'\beta_0^* + \beta_1^* D_i] + \sigma\phi\left[\frac{X_i'\beta_0^* + \beta_1^* D_i}{\sigma}\right]$$

اثرات حاشیه‌ای توبیت تقریباً همواره در قالب مشتق متوسط (که یک بیان به شدت ساده است) مطرح می‌شوند (ر.ک. وولدریج (۲۰۰۶)).

$$E\left\{\phi\left[\frac{X_i'\beta_0^* + \beta_1^* D_i}{\sigma}\right]\right\} \cdot \beta_1^*. \qquad (3.4.9)$$

یک برداشت سریع از رابطه ۹-۴-۳ آن است که ضریب مدل توبیت β_1^* همواره نسبت به اثر D_i روی Y_i بسیار بزرگ است (در یک مدل غیرخطی معلوم برای Y_i پنهان). این موضوع در تغییرات دائم خروجی پنهان در هنگام یک و صفر شدن D_i ریشه دارد، اما Y_i حقیقی باید تغییر کند: برای بسیاری از افراد این متغیر در هر حالتی صفر است.

در جدول ۲-۴-۳، به منظور بررسی اثر اشتغال زنان و ساعات کاری آنان (هر دو LDV) بر شاخص‌های باروری، اثرات حاشیه‌ای غیرخطی و رگرسیون مقایسه می‌شوند. برآوردها با استفاده از یکی از نمونه‌های سرشماری ۱۹۸۰ مورد استفاده آنگریست و ایوانز (۱۹۹۸) انجام شده است. این نمونه شامل زنان متأهل ۳۵-۲۱ساله با حداقل دو فرزند است. متغیرهای فرزندآوری متشکل از یک متغیر پنهان است که این متغیر معرف قدرت فرزندآوری بیش از دو فرزند یا تعداد کل تولدهاست. متغیرهای کمکی شامل جملات خطی بر حسب سن مادر، سن مادر در نخستین زایمان، نژاد مادر (سیاه‌پوست یا اسپانیولی) و سطح تحصیلات مادر (دبیرستان، دانشگاهی و تحصیلات عالی) است. نه تنها مدل متغیرهای کمکی اشباع نشده است، بلکه اثرات خطی دیده می‌شود و تعاملی وجود ندارد؛ بنابراین، مطمئناً CEF اساسی این مثال غیرخطی است. اثرات حاشیه‌ای مدل پروبیت یک متغیر پنهان حاکی از آن است که وجود بیش از دو فرزند با استفاده از برآوردهای OLS همین رابطه قابل تشخیص است. این موضوع در ستون‌های ۲، ۳ و ۴ جدول ۲-۴-۳ دیده می‌شود. اولین ردیف از جدول برآوردهای حاصل از روش‌های مختلف را برای کل جامعه نمونه ۱۹۸۰ مقایسه می‌کند. برآورد OLS از اثر فرزند سوم ۰/۱۶۲- و اثرات حاشیه‌ای متناظر مدل پروبیت ۰/۱۶۳- و ۰/۱۶۲- است. این برآوردها با استفاده از رابطه ۸-۴-۳ در دو حالت اول و دوم انجام شده است (و لذا، یک اثر حاشیه‌ای برای روش درمانی است).

$$E\left\{\Phi\left[\frac{X_i'\beta_0^*+\beta_1^*}{\sigma}\right]-\Phi\left[\frac{X_i'\beta_0^*}{\sigma}\right]\middle| D_i=1\right\}$$

اثرات حاشیه‌ای میان باروری و ساعات کاری نیز با وجود غیرقابل تشخیص بودن به برآوردهای OLS متناظر بسیار نزدیکند. این موضوع را می‌توان از مراجعه به ستون‌های ۵ و ۶ دریافت. به عنوان مثال، برآوردهای مدل توبیت یعنی ۶/۵۶- و ۵/۸۷- را با برآورد OLS در ستون ۲ یعنی ۵/۹۲- مقایسه کنید. اگرچه قدر مطلق یکی از برآوردهای مدل توبیت ۱۰ درصد بزرگ‌تر است، با این حال اهمیت داشتن این موضوع بعید به نظر می‌رسد. در سایر ستون‌های جدول با جایگزینی متغیر موهومی با متغیر وصفی فرزندآوری به مقایسه OLS با اثرات حاشیه‌ای پرداخته شده است. در کلیه این محاسبات برای محاسبه اثرات حاشیه‌ای (با برچسب MFX) از مشتق‌ها استفاده شده است. همچنین، برآورد OLS و اثرات حاشیه‌ای غیرخطی برای هر دو مدل پروبیت و توبیت یکسان است.

گاهی چنین گفته می‌شود که چنانچه مقادیر انطباق یافته به ۰/۵ نزدیک باشند، به علت تقریباً خطی بودن CEF غیرخطی در ناحیه مرکزی می‌توان انتظار داشت اثرات حاشیه‌ای تولیدی مدل‌های پروبیت به

OLS نزدیک باشند. بنابراین، مقایسه OLS و اثرات حاشیه‌ای را در یک نمونه فرعی با نرخ متوسط اشتغال نسبتاً بالا متشکل از زنان رنگین‌پوست با سن بالای سی سال و دارای تحصیلات عالی که نخستین زایمان آنان در سن کمتر از ۲۰ سالگی انجام شده است، تکرار می‌کنیم. با وجود بالا بودن متوسط نرخ اشتغال در این گروه (۸۳٪)، همچنان برآوردهای حاصل از OLS و اثرات حاشیه‌ای مشابه است.

جدول ۲-۴-۳. مقایسه برآوردهای مختلف از اثر قدرت فرزندآوری بر LDVها

		Right-hand side variable								
		More than two children					Number of children			
	Mean	OLS	Probit		Tobit		OLS	Probit MFX	Tobit MFX	
			Avg effect, full sample	Avg effect on treated	Avg effect, full sample	Avg effect on treated		Avg effect, full sample	Avg effect, full sample	Avg effect on treated
Dependent variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: Full Sample										
Employment	.528	-.162	-.163	-.162	-	-	-.113	-.114	-	-
	(.499)	(.002)	(.002)	(.002)			(.001)	(.001)		
Hours worked	16.7	-5.92	-	-	-6.56	-5.87	-4.07	-	-4.66	-4.23
	(18.3)	(.074)			(.081)	(.073)	(.047)		(.054)	(.049)
Panel B: Non-white College Attendees over 30, first birth before age 20										
Employment	.832	-.061	-.064	-.070	-	-	-.054	-.048	-	-
	(.374)	(.028)	(.028)	(.031)			(.016)	(.013)		
Hours worked	30.8	-4.69	-	-	-4.97	-4.90	-2.83	-	-3.20	-3.15
	(16.0)	(1.18)			(1.33)	(1.31)	(.645)		(.670)	(.659)

نکته: موارد ارائه شده در جدول فوق عبارتند از: تخمین OLS، متوسط اثرات تیمار و اثرات حاشیه‌ای (MFX) مربوط به اثر قدرت فرزندآوری بر عرضه نیروی کار مادران کارگر. اندازه جامعه نمونه در پنل الف ۲۵۴۶۵۴ نفر و معادل اندازه نمونه زنان متأهل در سرشماری سال ۱۹۸۰ مورد استفاده آنگریست و ایوانز است (۱۹۹۸). متغیرهای کمکی شامل سن، سن در زمان اولین تولد و متغیرهای موهومی برای پسران در اولین و دومین تولد است. اندازه جامعه نمونه در پنل ب ۷۴۶ نفر رنگین‌پوست با حداقل سن بالای ۳۰ سال است که اولین زایمان آنان در کمتر از سن ۲۰ سالگی رخ داده است. انحرافات استاندارد در ستون یک داخل پرانتز گزارش شده‌اند. خطاهای استاندارد در سایر ستون‌ها داخل پرانتز نوشته شده‌اند. از همین نمونه برای تخمین متوسط اثرات درمان در جامعه زنان با بیش از دو فرزند نیز استفاده شده است.

نتیجه این بحث آن است که اگر میزان انطباق CEF با LDVها از طریق یک مدل غیرخطی از یک مدل خطی بیشتر باشد، احتمالاً اهمیت موضوع در مورد اثرات حاشیه‌ای کمتر خواهد بود. این نتیجه خوش‌بینانه یک نظریه نیست، اما با توجه به مثال تجربی مطرح شده در قالب این پژوهش به نظر می‌رسد که یک نتیجه درست است.

پس چرا باید زحمت کار با مدل‌های غیرخطی و اثرات حاشیه‌ای را متقبل شویم؟ در پاسخ می‌توان به سادگی اثرات حاشیه‌ای برای انجام محاسبات خودکار توسط بسته‌های نرم‌افزاری مانند Stata اشاره کرد، اما در این میان در مورد مواردی (مثل روش وزن‌دهی، مشتقات بر حسب اختلافات متناهی) نیز باید تصمیم گرفت. ضمن اینکه OLS نیز استاندارد شده است. افزون بر این، در صورت توجه به نکات

مطرح در بخش ۴ و داده‌های پنل، به پیچیده‌تر شدن اوضاع به واسطه دوام ویژگی غیرخطی بودن پی می‌بریم. در نهایت، مرحله استدلال نیز پیچیده‌تر می‌شود، زیرا اثرات حاشیه‌ای به خطاهای استاندارد نیاز دارند. بنابر اصل تیغ اوکمز، «نهادها نباید بدون ضرورت تعدد یابند». در این زمینه به نقل قولی از آنگوس دیتون معلم سابق خود (۱۹۹۷) در باب تابع رگرسیون غیرخطی تولید شده توسط مدل‌های نوع توبیت اشاره می‌کنیم: با عدم آگاهی از F (توزیع خطاها) این تابع رگرسیون حتی قادر به تشخیص β (ضرایب مدل توبیت) نیست (ر.ک. پاول، ۱۹۸۹)، اما باید بپرسیم که چنین مفهوم زمخت، دشوار و غیرمحکمی چگونه به وجود آمده است؟

۳-۴-۳. وجه تسمیه رگرسیون چیست و رگرسیون به متوسط[1] به چه معناست؟

اصطلاح رگرسیون از مطالعات اس. فرانسیس گالتون (۱۸۸۶) در حوزه قد سرچشمه گرفته است. گالتون، کسی که با نمونه‌هایی از داده‌های با توزیع تقریباً نرمال خانواده‌ها و فرزندان کار می‌کرد، خاطرنشان می‌سازد که بر اساس پارامترهای به دست آمده از تقاطع و شیب رگرسیون دومتغیره CEF قد کودک با توجه به قد والدین او خطی است. با توجه به ثابت بودن قد (عدم تغییرات قابل توجه آن در طول زمان) شیب رگرسیون دومتغیره نیز ضریب همبستگی است، یعنی مقدار آن بین دو مقدار صفر و یک قرار دارد.

تنها رگرسور مورد استفاده گالتون x_i میانگین قد والدین و قد فرزند بزرگسال متغیر وابسته Y_i است. مانند همیشه، ضریب شیب رگرسیون برابر $\beta_1 = \frac{Cov(\mathrm{Y}_i, x_i)}{V(x_i)}$ و نقطه نقاطع $\alpha = E[\mathrm{Y}_i] - \beta_1 E[\mathrm{X}_i]$ است، اما از آنجایی که معمولاً قد در میان نسل‌ها تغییر نمی‌کند، متوسط و واریانس Y_i و x_i یکسان است. بنابراین:

$$\begin{aligned}\beta_1 &= \frac{Cov(\mathrm{Y}_i, x_i)}{V(x_i)} = \frac{Cov(\mathrm{Y}_i, x_i)}{\sqrt{V(x_i)}\sqrt{V(\mathrm{Y}_i)}} = \rho_{xy} \\ \alpha &= E[\mathrm{Y}_i] - \beta_1 E[\mathrm{X}_i] = \mu(1-\beta_1) = \mu(1-\rho_{xy})\end{aligned}$$

که در آن، ρ_{xy} معرف ضریب همبستگی میان‌نسلی قد و $\mu = E[\mathrm{Y}_i] = E[\mathrm{X}_i]$ معرف میانگین قدی افراد جامعه است. بنابراین، CEF خطی عبارت است از:

$$E[\mathrm{Y}_i|x_i] = \mu(1-\rho_{xy}) + \rho_{xy}x_i,$$

بنابراین، با توجه به ارتباط قد فرزند با قد والدین خود، قد فرزندان با میانگین وزنی قد والدین میانگین قدی جامعه برابر است. لذا، به طور متوسط قد فرزندان والدین بلندقد با قد پدر و مادر خود برابر نیست. این موضوع در مورد قد فرزندان والدین کوتاه‌قد نیز صادق است. به عنوان مثال، پیشک با قد ۶/۳ می‌تواند انتظار یک فرزند بلندقد را داشته باشد، هرچند بلندی قد فرزند او به بلندی قد خود او نخواهد

1. Regression-to-the-mean

بود. با این حال، آنگریست با قد ۵/۶ می‌تواند انتظار داشته باشد که قد فرزندش از قد او بلندتر شود. گالتون به این ویژگی «رگرسیون نسبت به میانه در قد وراثتی» می‌گوید. امروز این ویژگی به «رگرسیون به سوی مقدار متوسط» موسوم است.

نام گالتون، پسرعموی چارلز داروین، به خاطر تأسیس انجمن Eugenics در خاطرها باقی مانده است. این انجمن در حوزه تلاش برای پرورش بهتر مردم فعالیت داشت. در واقع، علاقه او به رگرسیون عمدتاً حاصل این تلاش بود. بنابراین می‌توان نتیجه گرفت که ارزش ایده‌های علمی توسط سیاست‌های نویسندگان آن قابل قضاوت نیست.

به نظر نمی‌رسد گالتون به رگرسیون چندگانه (نگرانی اصلی ما در این فصل) علاقه‌ای نشان داده باشد. در واقع، منظور از رگرسیون در کار گالتون همان ویژگی‌های مکانیکی توزیع متغیرهای تصادفی ایستاست که قطعاً علّی نیست. گالتون در اعتراض به ایده لامارک (که بعدها در روسیه استالین تبلیغ شد) چنین می‌گوید که ویژگی‌های به دست آمده قابل توارث است.

ایده امکان استفاده از رگرسیون در کنترل آماری از یک پژوهش که توسط جورج آدنی یول (۱۹۹۸) در زمینه تعیین نرخ فقر انجام شد، سرچشمه می‌گیرد. یول، آماردان و دانشجوی کارل پیرسون (شاگرد گالتون) به این نکته پی برد که ضریب رگرسیون گالتون با حل معادلات حداقل مربعات نرمال (که پیش‌تر توسط گاوس و لاگرانژ به دست آمده بود) به متغیرهای متعدد قابل تعمیم است. مقاله یول (۱۸۹۹) اولین مقاله منتشر شده در زمینه برآوردهای رگرسیون چند متغیره است. مدل او در عین کنترل رشد جامعه و توزیع سنی آن، میان نرخ فقر در یک ناحیه و تغییر در مدیریت قوانین ضعیف انگلیسی ارتباط برقرار می‌کند. او به طور خاص به بررسی این موضوع علاقه داشت که آیا کمک‌های غیرقانونی، ارائه پرداختی‌های حمایتی بدون الزام برای تغییر وضعیت موجود خود در افزایش نرخ فقر نقشی ندارد. این یک سؤال علّی مشهور در این زمینه است که تا امروز نیز ذهن ما را به خود مشغول نگه داشته است.[1]

در نهایت، یادآوری می‌شود که تاریخچه رگرسیون با جزئیات زیبا در کتاب استیون استایلر (۱۹۸۶) مطرح شده است. استایلر یک آمارگر مشهور در دانشگاه شیکاگو است، اما شهرت او به پدرش جورج استایلر (اقتصاددان و برنده جایزه نوبل) نمی‌رسد.

۵-۳. پیوست: فرمول استخراج مشتق متوسط

فرایند مربوطه با محاسبه رگرسیون Y_i روی S_i آغاز می‌شود:

$$\frac{Cov(Y_i, S_i)}{V(S_i)} = \frac{E[h(S_i)(S_i - E[S_i])]}{E[S_i(S_i - E[S_i])]}.$$

۱. اولین مقاله یول در حوزه قانون فقر در سال ۱۸۹۵ در مجله اکونومیک به چاپ رسید (که پیشک افتخار همکاری در آن را داشت). نظریه رگرسیون چندگانه در مقاله یول (۱۸۹۷) مطرح شد.

$\kappa_{-\infty} = \lim_{t\to -\infty} h(t)$ در نظر گرفته می‌شود. بر اساس نظریه اساسی محاسبه داریم:

$$h(s_i) = \kappa_{-\infty} + \int_{-\infty}^{s_i} h'(t)\,dt.$$

با جایگذاری $h(s_i)$ صورت کسر به شکل زیر تبدیل می‌شود.

$$E[h(s_i)(s_i - E[s_i])] = \int_{-\infty}^{+\infty}\int_{-\infty}^{s} h'(t)\,(s - E[s_i])g(s)dtds$$

که در آن، $g(s)$ به چگالی s_i در **s** اشاره دارد. با معکوس کردن مرتبه انتگرال‌گیری، داریم:

$$E[h(s_i)(s_i - E[s_i])] = \int_{-\infty}^{+\infty} h'(t)\int_{t}^{+\infty}(s - E[s_i])g(s)dsdt.$$

حاصل انتگرال داخلی برابر $\mu_t \equiv \{E[s_i|s_i \geq t] - E[s_i|s_i < t]\}\{P(s_i \geq t)[1 - P(s_i \geq t)\}$ و به وضوح مقداری نامنفی است. اگر $s_i = Y_i$ قرار دهیم، به سادگی می‌توان مخرج را معادل انتگرال این مقادیر وزنی قرار داد. بنابراین با ارائه مشتق متوسط وزنی ضریب رگرسیون دومتغیره یعنی $\frac{Cov(Y_i,s_i)}{V(s_i)}$ مواجهیم (ر.ک. رابطه ۸-۳-۳). در پیوست مقاله منتشر شده توسط آنگریست و کروگر (۱۹۹۹) نیز فرمول مشابهی برای رگرسیون با کوواریانس X_i به دست آمده است.

۴. متغیرهای ابزاری در عمل: گاهی اوقات آنچه را که می‌خواهی به دست می‌آوری

هرچه که باید اتفاق افتد، اتفاق می‌افتد،

هرچه که اتفاق می‌افتد، سبب اتفاق دیگری می‌شود، و آن اتفاق رخداد دیگری را به دنبال دارد.

هر تصادفی که با رخدادش، دوباره همان اتفاق را به همراه داشته باشد، دوباره اتفاق می‌افتد.

البته این اتفاقات لزوماً به ترتیب نیست.

داگلاس آدامز،[1] تقریباً بی‌ضرر[2] (۱۹۹۵)

دو موضوع رشته اقتصادسنجی را از خواهر قدیمی‌اش، آمار، جدا می‌کند. نخست خجالت نکشیدن از علّیت است. استنتاج علّی همیشه یک اسم بازی در اقتصادسنجی کاربردی بوده است. پل هلند[3] آمارشناس (۱۹۸۶) هشدار می‌دهد که «هیچ علّیتی بدون دستکاری نیست»[4]، اصلی که به نظر می‌رسد، استنتاج علّی را از داده‌های غیرتجربی حذف می‌کند. شاهدانی که دقت کمتری دارند بر این گفته بدیهی

1. Douglas Adams
2. Mostly Harmless
3. Paul Holland
4. "No Causation Without Manipulation"

تکیه می‌کنند: «همبستگی، علّیت نیست».[1] همانند کسانی که شغلشان کار با داده‌هاست، ما باور داریم که همبستگی گاهی اوقات می‌تواند سند خوبی از رابطه علّی باشد، حتی وقتی که پژوهشگران و آزمایشگران در متغیرهای مورد ملاحظه دست برده‌اند.[2]

دومین مسئله که ما را از سایر آمارشناسان و به طور حتم سایر دانشمندان علوم اجتماعی ـ متمایز می‌کند مجموعه‌ای از ابزارهای آماری ـ مهمّات آماری ـ است که از پژوهش‌های نخستین اقتصادسنجی سرچشمه گرفتند. قدرتمندترین ابزار در این مجموعه روش متغیر ابزاری،[3] یعنی موضوع این بخش است. همان طور که مشخص شد، **IV** فراتر از فقط ارزیابی مداوم پارامترهای دستگاه معادلات است، گرچه این امکان را نیز فراهم می‌کند.

در دهه ۱۹۲۰، فیلیپ رایت و پسرش سیوال با مطالعه بازارهای کشاورزی به چالش استنتاج علّی توجه نشان دادند: هنگامی که داده‌ها درباره قیمت مشاهده شدند و کمیت‌ها توسط تقاطع این دو منحنی تعیین شدند، چگونه شیب عرضه و منحنی تقاضا را ارزیابی کنند. به عبارت دیگر، تعادل کمیت و قیمت ـ تنها دو چیزی که امکان مشاهده آنها وجود دارد ـ این دو معادله تصادفی را هم‌زمان حل می‌کند. بنابراین، نمودار پراکندگی مشاهده شده کمیت و قیمت بر کدام منحنی می‌نشیند؟ فلیپ رایت این حقیقت را درک کرد که ضریب رگرسیون جامعه، شیب هیچ‌کدام از معادلات را در دستگاه معادلات هم‌زمان نشان نمی‌دهد. روش **IV** که اولین بار توسط رایت (۱۹۲۸) مطرح شد مشکل دستگاه معادلات هم‌زمان را با متغیرهایی حل می‌کند که در یک معادله وجود دارند تا آن معادله را تغییر دهند و معادله دیگر را ردیابی کنند. به متغیرها که این تغییر را انجام می‌دهند، متغیرهای ابزاری[4] می‌گویند (ریرسول،[5] ۱۹۴۱). در تحقیقات جداگانه، روش‌های **IV** برای اولین بار برای حل مشکل خطای اریبی در مدل‌های رگرسیون استفاده شد.[6] یکی از مهم‌ترین نتایج در نظریه آماری مدل خطی در زمانی که رگرسور مورد نظر با خطاهای تصادفی اندازه‌گیری می‌شود این است که اریب ضریب رگرسیون به سمت صفر است (برای اینکه دلیل آن را بدانید تصور کنید که رگرسور شامل خطای تصادفی است و در نتیجه رگرسور با متغیر وابسته ناهمبسته می‌شود). روش متغیرهای ابزاری می‌تواند برای حذف چنین اریبی استفاده شود.

در پیشینه تفکر اقتصادسنجی، الگوهای معادلات هم‌زمان (SEM)[7] مطلبی مهم است. هم‌زمان، تعداد کمی از مقالات کاربردی تأثیرگذار بر چارچوب SEM تکیه داشتند، اگرچه زبان تکنیکی استفاده

---

1. "correlation is not causality "

۲. در سال‌های اخیر علاقه کارشناسان آمار به بحث درباره مدل‌های آماری برای داده‌های مشاهداتی در چارچوب علّی روشن افزایش یافته است، برای مثال بررسی فریمن در (۲۰۰۵)

3. Instrumental Variable (IV)

4. instrumental variables

5. Reiersol

۶. مرجع مهم در اینجا (Wald,1940) و دوربین (Durbin,1954) هستند که در ادامه درباره آنها بحث شده است

7. Simultaneous equations models (SEMs)

شده برای بحث درباره IV هنوز از این چارچوب نشئت می‌گیرند. امروزه، IV بیشتر از ارزیابی پارامترهای SEM برای حل مشکلات خطای اندازه‌گیری استفاده می‌شود. اما، بدون شک، مهم‌ترین استفاده کنونی IV حل مشکل اریب متغیرهای حذف شده است. IV مشکلات متغیرهای کنترل ناشناخته یا گم شده را حل می‌کند، همان طور که آزمایش تصادفی نیاز کنترل گسترده در رگرسیون را برطرف می‌کند.[۱]

۴-۱. متغیرهای ابزاری (IV) و علّیت

می‌خواهیم درباره **IV** به دو شکل صحبت کنیم. نخست، در مدلی مقیّد با اثر ثابت، سپس در چارچوبی با خروجی‌های بالقوه ناهمگن نامقیّد که در آن صورت اثر علّی باید ناهمگن باشد. بدون تغییر سازوکار اساس روش‌های **آماری** که ما از آنها زیاد در کار استفاده می‌کنیم، معرفی اثرات ناهمگن، تفسیرِ برآوردهای **IV** را غنی می‌سازد (به خصوص، در مورد حداقل مربعات دومرحله‌ای).[۲] تمرکز اولیه بر اثرات ثابت به ما امکان می‌دهد که مکانیک **IV** را با حداقل مشکلات حل می‌کنیم.

برای برانگیختن راه‌اندازی اثر ثابت به عنوان چارچوبی برای پیوند علّی بین تحصیل و دستمزدها، دوباره تصور کنید که خروجی‌های بالقوه می‌توانند این گونه نوشته شوند:

$$\mathrm{Y}_{si} \equiv f_i(s),$$

و همانند معرفی رگرسیون در بخش ۳:

$$f_i(s) = \pi_0 + \pi_1 s + \eta_i, \qquad (4.1.1)$$

همچنین، **همانند** بحث اولیه ما، تصور کنید که بردار متغیرهای کنترلی، یعنی A_i، که «قابلیت (توانایی)[۳]» نامیده می‌شود، گزارشی گزینشی بر اساس موارد قابل مشاهده[۴] می‌دهد:

$$\eta_i = A_i'\gamma + v_i,$$

که در آن γ دوباره بردار ضریب رگرسیون جامعه می‌شود، بنابراین v_i و A_i از لحاظ ساختاری ناهمبسته هستند. هم‌اکنون، تصور می‌شود که متغیرهای A_i تنها دلیل برای این است که η_iو s_i همبسته هستند، بنابراین:

$$E[\mathrm{s}_i v_i] = 0.$$

به عبارت دیگر، اگر A_i مشاهده شود، ما با کمال میل آن را وارد رگرسیون دستمزد بر تحصیل می کنیم و بنابراین رگرسیونی دوربرد تولید می‌کنیم که می‌توان به صورت زیر نوشت:

$$\mathrm{Y}_i = \alpha + \rho \mathrm{s}_i + A_i'\gamma + v_i. \qquad (4.1.2)$$

---

۱. آنگریست و کروگر (۲۰۰۱) برای شرح خلاصه‌ای از تاریخ و استفاده IV، (Stock and Trebbi, 2003) برای شرح کامل پیدایش IV و (Morgan,1990) برای تاریخ گسترده نظریات اقتصادسنجی ببینید، که شامل مدل معادلات همزمان می‌شود.

2. Two-stage least Squares (2SLS)

3. Ability

4. Observable

معادله ۲-۱-۴ نسخه‌ای از مدل علّی خطی است (۹-۲-۳). جمله خطا در این معادله، بخش تصادفی خروجی‌های بالقوه است، یعنی v_i، که بعد از کنترل A_i باقی ماندند. فرض بر این است که این جمله خطا با تحصیل ناهمبستگی دارد. اگر این فرض درست باشد، جامعه رگرسیون Y_i روی S_i و A_iضریب معادله ۲-۱-۴ را تولید می‌کنند.

چگونگی ارزیابی ضریب رگرسیون دوربرد، یعنی**ρ**، زمانی که A_i مشاهده نشده است مشکلی است که ما ابتدا قصد حل آن را داریم. هنگامی که محقق به متغیری دسترسی دارد (متغیر ابزاری که ما Z_i می‌نامیم) که با متغیر علّی[1] مورد نظر، یعنی S_i، همبسته است، اما با دیگر تعیین‌کننده‌های[2] متغیرهای وابسته ناهمبسته است، روش متغیرهای ابزاری می‌تواند برای حل این مسئله استفاده شود. در اینجا عبارت «ناهمبسته» با دیگر تعیین‌کننده‌های متغیرهای وابسته مثل این است که بگوییم $Cov(\eta_i,z_i) = 0$ یا اینکه Z_i با A_i و V_i ناهبسته است. از آنجایی که Z_i می‌تواند از مدل علّی مورد نظر حذف شود، این اصل «قید عدم شمول»[3] نامیده می‌شود. قید عدم شمول نسخه‌ای از فرض استقلال مشروط[4] در بخش قبل است، به استثنای اینکه در حال حاضر ابزاری است که از خروجی بالقوه مستقل است، به جای تحصیل (هنگامی که مدل IV را با متغیرهای کمکی در نظر بگیریم، «مشروط» در استقلال مشروط وارد بحث می‌شود).

با توجه به قید عدم شمول، از معادله ۱-۲-۴ برداشت می‌شود که:

$$\rho = \frac{Cov(Y_i, Z_i)}{Cov(S_i, Z_i)} = \frac{Cov(Y_i, Z_i)/V(Z_i)}{Cov(S_i, Z_i)/V(Z_i)}. \qquad (4.1.3)$$

تساوی دوم در ۳-۱-۴ سودمند است، زیرا معمولاً فکر کردن درباره ضریب رگرسیون آسان‌تر از کوواریانس‌هاست. ضریب مورد نظر، یعنی **ρ**، نسبت رگرسیون جامعه $\mathbf{Y_i}$ روی $\mathbf{Z_i}$ بر رگرسیون جامعه $\mathbf{S_i}$ روی $\mathbf{Z_i}$ می‌شود (اولین مرحله). برآوردگر **IV** مشابه نمونه‌ای از معادله ۳-۱-۴ است. به یاد داشته باشید برآوردهای **IV** مبتنی بر این نظر است که مرحله اول صفر نیست. به عنوان قانون، اگر اولین مرحله تفاوت معنادار کمی داشته از صفر باشد، غیرمحتمل است که برآوردهای **IV** به دست آمده حاوی اطلاعات مفید باشند، نکته‌ای که ما بعد به آن برمی‌گردیم. تکرار فرض مورد نیاز برای نسبت کوواریانس‌ها در معادله ۳-۱-۴ برای تساوی با اثر علّی، **ρ**، امری بجاست. نخست، این متغیر ابزاری باید اثر روشنی بر $\mathbf{S_i}$ داشته باشد. این اولین مرحله است. دوم، تنها دلیل برای ارتباط بین $\mathbf{Y_i}$ و $\mathbf{Z_i}$ همان مرحله اول است. فعلاً ما این فرض دوم را قید عدم شمول می‌نامیم. اگرچه، همان طور که در بحث مدل‌ها با اثر ناهمگن

1. Causal Variable
2. Determinant
3. Exclusion Restriction
4. Conditional Independence Assumption

می‌بینیم، این فرض در واقع دو قسمت دارد: ابتدا گزاره‌ای است که این متغیر ابزاری به همان خوبی متغیر تخصیص یافته به طور تصافی[1] است (به عبارت دیگر مستقل از خروجی بالقوه که مشروط بر متغیرهای کمکی است)، در حالی که قسمت دوم این است که این متغیر ابزاری اثری بر خروجی‌ها ندارد، مگر اینکه از طریق گذر از مرحله اول باشد.

بنابراین کجا می‌توانید متغیر ابزاری پیدا کنید؟ متغیر ابزاری خوب از دانش نهادی و نظر شما درباره فرایندهای تعیین متغیر مورد نظر نشئت می‌گیرد. برای مثال، مدل اقتصادی تحصیل نشان می‌دهد که تحصیل کردن توسط مقایسه هزینه و امتیازهای انتخاب‌های دیگر مشخص می‌شود. بنابراین، یک منبع احتمالی متغیر ابزاری برای تحصیل تفاوت در هزینه‌هاست، به طور مثال، رویه وام دادن یا دیگر کمک‌های مالی که مستقل از قابلیت یا پتانسیل درآمدی تنوع دارند. منبع دوم تغییرات در تحصیل قیدهای نهادی است. دسته‌ای از قیدهای نهادیِ مرتبط با تحصیل، قوانین تحصیل اجباری است. آنگریست و کروگر[2](۱۹۹۱) در مقاله‌ای که از کاربرد «آزمایش‌های طبیعی» نمونه‌ای نشان داده بود از تغییرات به وجود آمده توسط تحصیل اجباری بهره‌برداری کرد تا اریب متغیرهای حذف شده را به طور کامل حذف کند.

نکته اول برای راهبرد فصل تولد[3] یا آنگریست و کروگر (۱۹۹۱) مشاهده این بود که بیشتر ایالت‌ها نیاز دارند تا دانش‌آموزان وقتی در تقویم سالیانه به سن ۶ سالگی رسیدند وارد مدرسه شوند. بنابراین، سن آغاز مدرسه تابع تاریخ تولد است. به خصوص، آنهایی که در آخرین ماه‌های سال به دنیا آمدند زود وارد مدرسه می‌شوند. در ایالت‌هایی که محدوده تاریخ تولد تا ۳۱ام دسامبر است، فرزندان متولد شده در فصل چهارم، زودتر از ۶ سالگی وارد مدرسه می‌شوند، در حالی که آنهایی که در فصل اول به دنیا آمدند، در سن حدود ۶ سال و نیمی وارد مدرسه می‌شوند. به علاوه، به این دلیل که اقتضای قوانین تحصیل اجباری این است که دانش‌آموزان تا ۱۶ سالگی در مدرسه تحصیل کنند، هنگامی که به سن خروج از مدرسه رسیدند، این گروه دانش‌آموزان در پایه تحصیلی متفاوتی خواهند بود یا در حال ورود به مقطع تحصیلی متفاوتی هستند. در حقیقت، ترکیب تدابیر سن شروع مدرسه و قوانین تحصیل اجباری آزمایشی طبیعی را تولید می‌کند که در آن دانش‌آموزان طبق تاریخ تولدشان، در طول متفاوت زمانی مجبور به مدرسه رفتن هستند. آنگریست و کروگر با استفاده از داده‌های سرشماری آمریکا ارتباط بین دستاورد تحصیلی و فصل تولد را بررسی کردند. پنل **A** در شکل ۱-۱-۴ (اتخاد شده از آنگریست و کروگر، ۲۰۰۱) مدل تحصیل ـ فصل تولد برای مردانی که در دهه ۱۹۳۰ در سرشماری ۱۹۸۰ نشان می‌دهد. این شکل نشان می‌دهد مردانی که در ماه‌های اولیه سال متولد شدند میانگین سطح تحصیلی کمتری داشتند. پنل **A** نمایش تصویری از مرحله اول است. مرحله اول در چارچوب کلی **IV** رگرسیون

1. Randomly Assigned Variable
2. Angrist and Krueger
۳. راهبرد Quarter-of-birth Strategy به فصلی که تولد در آن صورت می‌گیرد مرتبط است.

متغیرهای علّی مورد نظر روی متغیرهای کمکی و متغیر(ها) ابزاری است. این تصویر به طور خلاصه این رگرسیون را نشان می‌دهد، زیرا میانگین سال‌های تحصیل و فصل تولد چیزی است که شما از مقدار برازانده شده از رگرسیون تحصیل روی مجموعه کامل از متغیرهای موهومی[1] سال و فصل تولد دریافت می‌کنید.

پنل **B** در شکل ۱-۱-۴ میانگین درآمد را، طبق فصل تولد برای همان نمونه که در ایجاد پنل **A** استفاده شد، نشان می‌دهد. این پنل آنچه را کارشناسان اقتصادسنجی «فرم تقلیل یافته»[2] می‌نامند یا به عبارتی ارتباط بین متغیرهای ابزاری و متغیرهای وابسته را نشان می‌دهد. فرم تقلیل یافته رگرسیون متغیرهای وابسته روی متغیرهای کمکی در این مدل و متغیر(ها)ی ابزاری است. پنل **B** نشان می‌دهد که هم‌گروهی‌های بزرگ‌تر با احتمال بیشتری درآمد بالاتری دارند، زیرا درآمد با سابقه کاری بیشتر افزایش می‌یابد. این شکل همچنین نشان می‌دهد مردانی که در ماه‌های نخستین متولد شدند نسبت به آنهایی که در ماه‌های آخر متولد شدند حتی بعد از جورشدگی سال تولد که نقش متغیرهای کمکی برون‌زا[3] را در طرح آنگریست و کروگر بازی می‌کنند، میانگین درآمد کمتری دارند. به خصوص، این رابطه تقلیل یافته شبیه به مدل فصل تولد در تحصیل است که نشان می‌دهد هر دو مدل ارتباط نزدیکی با هم دارند. به این دلیل که تاریخ تولد یک فرد ممکن است با توانایی‌های درونی، انگیزه، یا ارتباط فامیلی مرتبط نباشد، می‌توان گفت تنها دلیل برای مدل غیرثابت فصل تولد در درآمد، مدل غیرثابت فصل تولد در تحصیل است. این فرض بسیار مهمی است که مدل **IV** فصل تولد را به میان می‌آورد.[4]

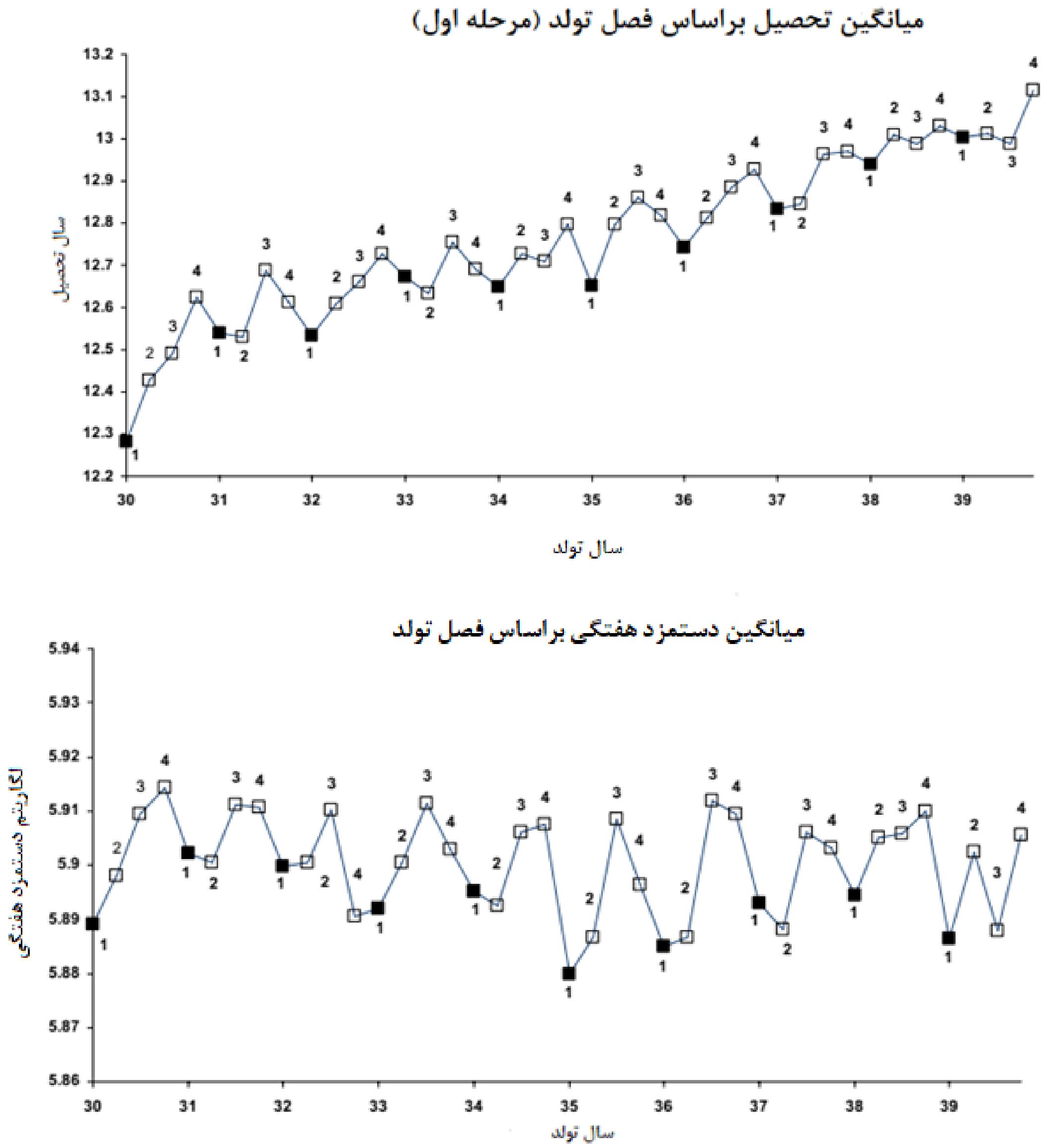
نمایش ریاضی اتفاقی که شکل ۱-۱-۴ بیان کرد از معادلات رگرسیون تقلیل یافته ناشی می‌شود که در پایین ذکر شده است.

$$s_i = X_i'\pi_{10} + \pi_{11}z_i + \xi_{1i} \quad (4.1.4a)$$

$$Y_i = X_i'\pi_{20} + \pi_{21}z_i + \xi_{2i} \quad (4.1.4b)$$

1. Dummies or Dummy Variables
2. Reduced Form
3. Exogenous Covariate
٤. دیگر توضیحات نیز ممکن هستند، در حالی که اثر پیشینه خانوادگی مرتبط با فصل تولد بیشترین احتمال را دارد (برای مثال Bound, Jaeger, and Baker, 1995 را ببینید). در مقایسه با احتمال اثر حذف شده پیشینه خانوادگی، در سطوح تحصیل که تحت تأثیر قانون تحصیل اجباری است الگوی فصل تولد در میانگین تحصیل برجسته‌تر است. دیگر توضیح ممکن اثر سن خالص در هنگام ورود است که از طریق مجراهایی بجز بالاترین سال تحصیلی که کامل شده است (برای مثال دستاورد). اثر علّی سن در هنگام ورود بر یادگیری برای جدا شدن از اثر سن خالص، اگر غیرممکن نباشد، کاری مشکل است، (همان طور که در فصل ۱ صحبت شد). پژوهش اخیر (Elder and Lubotsky, 2008) نشان می‌دهد که نسبت به امتیاز یادگیری در مدرسه برای دانش‌آموزان بزرگ‌تر، تکامل اثر مفروض سن در هنگام ورود در طول زمان به خودی خود با اثرات وابسته به تفاضل سن سازگارتر است.

شکل ۱-۱-۴. نمایش تصویری مرحله اول و فرم تقلیل‌یافته برای برآوردهای IV بازده اقتصادی تحصیل با استفاده از فصل تولد (از آنگریست و کروگر، ۱۹۹۱)



پارامتر π_{11} در معادله **4.1.4a** اثر مرحله اول $\mathbf{Z_i}$ روی $\mathbf{S_i}$ را نشان می‌دهد که متغیر کمکی $\mathbf{X_i}$ را جورشدگی می‌دهد. پارامتر π_{21} در معادله **4.1.4b** اثر فرم تقلیل‌یافته $\mathbf{Z_i}$ روی $\mathbf{Y_i}$را بیان می‌کند، که همین متغیرهای کمکی را جورشدگی می‌دهد. به زبان **SEM**، متغیرهای وابسته در این دو معادله را «متغیرهای درونزا»[1] می‌نامیم، یعنی این متغیرها به طور مشترک در داخل سیستم تعیین می‌شوند، در حالی که متغیرهای سمت راست «متغیرهای برونزا» نامیده می‌شوند (بیرون از سیستم تعیین می‌شوند). متغیرهای ابزاری، یا $\mathbf{Z_i}$، زیرمجموعه‌ای از متغیرهای برونزا هستند. متغیرهای برون‌زایی که

1. Endogenous Variables

متغیر ابزاری نیستند «متغیر کمکی برون‌زا» نامیده می‌شود. اگرچه ما در این مورد سیستم عرضه و تقاضای سنتی را بررسی نمی‌کنیم، این برچسب‌های متغیر **SEM** هنوز به طور گسترده در شیوه تجربی استفاده می‌شود.

برآوردگر **IV** تعدیل شده[1] ـ مشابه نمونه‌ای از نسبت $\frac{\pi_{21}}{\pi_{11}}$ است. به یاد داشته باشید که مخرج کسر آثار مرحله اول و فرم تقلیل‌یافته یکسان هستند. بنابراین نسبت آنها می‌شود:

$$\rho = \frac{\pi_{21}}{\pi_{11}} = \frac{Cov(Y_i, \tilde{z}_i)}{Cov(S_i, \tilde{z}_i)}, \qquad (4.1.5)$$

که در آن $\tilde{Z}_i$ مانده رگرسیون $\mathbf{Z_i}$ روی متغیرهای کمکی $\mathbf{X_i}$ هستند، بنابراین، طرف راست معادله ۵-۱-۴ متغیر $\tilde{Z}_i$ را با $\mathbf{Z_i}$ در فرمول کلی **IV** عوض می‌کند. کارشناسان اقتصادسنجی مشابه نمونه طرف چپ معادله ۵-۱-۴ را برآوردگر حداقل مربعات غیرمستقیم[2] غیرمستقیم $\boldsymbol{\rho}$ در مدل علّی با متغیرهای کمکی می‌نامند:

$$Y_i = \alpha' X_i + \rho S_i + \eta_i, \qquad (4.1.6)$$

که در آن $\boldsymbol{\eta_i}$ جمله خطای مرکب است یعنی $A_i'\gamma + v_i$[3] از آنجایی که $\tilde{Z}_i$ بر اساس ساختار با $\mathbf{X_i}$ و طبق فرض با $\boldsymbol{\eta_i}$ ناهمبسته است، استفاده از معادله ۶-۱-۴ برای تأیید مستقیم $cov(Y_i, \tilde{Z}_i) = \rho Cov(S_i, \tilde{Z}_i)$ بسیار آسان است. در پژوهش آنگریست و کروگر (۱۹۹۱)، متغیر ابزاری $\mathbf{Z_i}$ فصل تولد است (یا متغیرهای موهومی هستند که به فصل تولد اشاره دارند) و متغیرهای موهومی برای سال تولد، وضعیت تولد و نژاد است.

۱-۱-۴. حداقل مربعات دومرحله‌ای (2SLS)

معادله فرم تقلیل‌یافته، یعنی **4.1.4b**، می‌تواند از جایگزین کردن معادله مرحله اول یعنی **4.1.4a** در رابطه علّی مورد نظر یعنی ۶-۱-۴ که در دستگاه معادلات، «معادله ساختاری»[4] نامیده می‌شود به دست آید. بنابراین داریم:

$$\begin{aligned} Y_i &= \alpha' X_i + \rho[X_i'\pi_{10} + \pi_{11} z_i + \xi_{1i}] + \eta_i \qquad (4.1.7) \\ &= X_i'[\alpha + \rho\pi_{10}] + \rho\pi_{11} z_i + [\rho\xi_{1i} + \eta_i] \\ &= X_i'\pi_{20} + \pi_{21} z_i + \xi_{2i}, \end{aligned}$$

که در آن و در معادله **4b**-۱-۴، $\pi_{20} \equiv \alpha + \rho\pi_{10},\ \pi_{21} \equiv \rho\pi_{11}$ و $\xi_{2i} \equiv \rho\xi_{1i} + \eta_i$.

1. Covariate-adjusted IV
2. Indirect Least Squares (ILS)
۳. اثبات مستقیم برای اینکه ٥-١-٤ برابر با ρ در ٦-١-٤ است جایگزین کردن ٦-١-٤ به جای Yi در $\frac{COV(Y_i, \tilde{Z}_i)}{Cov(S_i, \tilde{Z}_i)}$ است.
4. Structural Equation

معادله ۷-۱-۴ بار دیگر نشان می‌دهد که چرا $\rho = \frac{\pi_{21}}{\pi_{11}}$. همچنین توجه کنید که یک بازآرایی اندک ۷-۱-۴ این نتیجه را به ما می‌دهد:

$$Y_i = \alpha' X_i + \rho[X_i'\pi_{10} + \pi_{11} Z_i] + \xi_{2i}, \qquad (4.1.8)$$

که در آن $[X_i'\pi_{10} + \pi_{11}Z_i]$ مقدار برازش شده از رگرسیون مرحله اول S_i روی X_i و Z_i است. به این خاطر که X_i و Z_i با خطای فرم تقلیل‌یافته ξ_{2i} ناهمبسته هستند، ضریب روی $[X_i'\pi_{10} + \pi_{11}Z_i]$ در رگرسیون جامعه Y_i روی X_i و $[X_i'\pi_{10} + \pi_{11}Z_i]$ برابر است با ρ.

به حتم در عمل، ما همیشه با داده‌های گرفته شده از نمونه کار می‌کنیم. با توجه به نمونه تصادفی، مقادیر برازانده شده مرحله اول در جامعه به طور پیوسته به این شکل برآورد می‌شود:

$$\hat{s}_i = X_i'\hat{\pi}_{10} + \hat{\pi}_{11} Z_i,$$

در معادله بالا $\hat{\pi}_{10}$ و $\hat{\pi}_{11}$ برآوردهای **OLS** برای معادله **4.1.4a** است. ضریب روی $\hat{S}_i$ در رگرسیون Y_i روی X_i و $\hat{S}_i$ برآوردگر کمترین حداقل مربعات دومرحله‌ای[1] ρ نامیده می‌شود. به عبارت دیگر، برآوردگرهای **2SLS** می‌توانند با برآورد **OSL** «معادله مرحله دوم» ساخته شوند.

$$Y_i = \alpha' X_i + \rho\hat{s}_i + [\eta_i + \rho(s_i - \hat{s}_i)], \qquad (4.1.9)$$

این معادله **2SLS** نامیده می‌شود، زیرا می‌تواند در دو مرحله انجام شود، اولین مرحله برآورد کردن $\hat{S}_i$ با استفاده از معادله **4.1.4a** و دومین مرحله برآورد معادله ۹-۱-۴ است. برآورد به دست آمده برای ρ سازگار است زیرا:

برآوردهای اولین مرحله سازگار هستند.

متغیرهای کمکی، یعنی Xi، و متغیرهای ابزاری، یعنی Zi با ηi و $(S_i - \hat{S}_i)$ همبسته هستند.

علی‌رغم نام 2SLS، ما همیشه برآوردهای 2SLS را در دو مرحله نمی‌سازیم. دلیل نخست این است که خطاهای استاندارد به دست آمده اشتباه هستند، که درباره آن صحبت خواهیم کرد. معمولاً، ما اجازه می‌دهیم که نرم‌افزارهای متخصصی (مثل SAS یا Stata) این محاسبات را برای ما انجام دهند. این کار خطاهای استاندارد درست به ما می‌دهد و کمک می‌کند که تمام مشکلات را حل کنیم (بخش ۱-۶-۴ را ببینید)، اما این حقیقت که برآوردگر 2SLS می‌تواند با توالی رگرسیون‌های OLS محاسبه شود، تنها راه برای یادآوری آن است که چرا این برآوردگر درست عمل می‌کند. به طور طبیعی، اگر 2SLS مشروط بر متغیرهای کمکی باشد، تنها تغییرات در Si را که توسط تغییرات شبه‌آزمایشی به وجود می‌آیند نگاه می‌دارد، به عبارت دیگر تغییراتی را نگاه می‌دارد که توسط متغیر ابزاری Zi به وجود می‌آید.

2SLS رویکرد بسیار بزرگی است. نخست، برآوردگر متغیری ابزاری است: برآورد 2SLS برای ρ در معادله ۹-۱-۴ مشابه نمونه‌ای $\frac{cov(Y_i,\hat{S}_i^*)}{cov(S_i,\hat{S}_i^*)}$ است، که در آن $\hat{S}_i^*$ مانده رگرسیون $\hat{S}_i$ روی X_i است.

---

1. Two-stage Least Squares (2SLS)

این قضیه نتیجه فرمول تشریحی رگرسیون چندمتغیره و این حقیقت است که $Cov\left(S_i\,,\widehat{S}_i^*\right) = V(,\widehat{S}_i^*)$. همچنین نشان دادن آنکه در مدلی با متغیر درون‌زای واحد و یک متغیر ابزاری، برآوردگر 2SLS همانند برآوردگر متناظر ILs است.[۱]

پیوند بین 2SLS و IV اطلاعات بیشتری را درباره چندابزاری می‌دهد. اگر هر متغیر ابزاری یک اثر علّی را نشان می‌دهد، شاید ما این برآوردهای IV جایگزین را در یک برآورد دقیق‌تر ترکیب کنیم. در مدل‌های با چند متغیر ابزاری، 2SLS ترکیبی خطی را با ترکیب متغیرهای ابزاری مختلف در یک متغیر ابزاری درست می‌کند.

برای مثال اگر سه متغیر ابزاری Z_{1i}، Z_{2i} و Z_{3i} را داریم، در پژوهش آنگریست و کروگر، اینها متغیر موهومی برای فصل اول و دوم و سوم است. بنابراین معادله مرحله اول می‌شود:

$$s_i = X_i'\pi_{10} + \pi_{11}z_{1i} + \pi_{12}z_{2i} + \pi_{13}z_{3i} + \xi_{1i}, \qquad (4.1.10a)$$

در حالی که مرحله دوم همانند ۹-۱-۴ است، به استثنای اینکه مقادیر برازانده شده به جای 4.1.4a از 4.1.10a هستند. تفسیر IV از این برآوردگر 2SLS همانند قبل است. متغیر ابزاری، مانده رگرسیون مرحله اول مقدار برازانده شده روی متغیر کمکی است. قید عدم شمول در این مورد این نظریه است که همه متغیرهای موهومی فصل تولد در 4.110a با ηi در معادله ۶-۱-۴ ناهمبسته است.

نتیجه برآورد 2SLS برای معادله تحصیل با استفاده از متغیرهای موهومی فصل تولد و نیز دیگر اآثار متقابل در جدول ۱-۱-۴ بیان شده‌اند که برآوردهای OLS و 2SLS از مدل‌ها را شبیه به مدل‌هایی که توسط آنگریست و کروگر (۱-۱-۴) برآورد شدند نشان داده است. هر ستون در این جدول شامل برآوردهای OLS و 2SLS از ρ در معادله‌ای شبیه به ۶-۱-۴ است که با ترکیبات مختلفی از متغیرهای ابزاری و متغیرهای کنترلی برآورد می‌شوند. برآورد OLS در ستون اول از رگرسیون لگاریتم دستمزد با هیچ گونه متغیر کنترلی است، در حالی که برآوردهای OLS در ستون دوم از مدلی است که به عنوان متغیر کنترلی، متغیرهای موهومی برای سال تولد و وضعیت تولد اضافه می‌کنند. در هر دو مورد بازگشت برآورد شده به تحصیل حدود ۰/۷۵ است.

---

۱. به یاد داشته باشید که $\widehat{S}_i^* = \tilde{Z}_i\hat{\pi}_{11}$ که در آن $\tilde{Z}_i$ مانده رگرسیون Zi بر Xi است، بنابراین برآوردگر 2SLS قیاس نمونه $\frac{\frac{Cov(Y_i\,,\tilde{Z}_i)}{V(,\tilde{Z}_i)}}{\hat{\pi}_{11}}$ است، اما قیاس نمونه صورت کسر، یعنی $\frac{Cov(Y_i\,,\tilde{Z}_i)}{V(,\tilde{Z}_i)}$، برآورد OLS π_{21} در فرم تقلیل‌یافته 4.1.4b است، در حالی که $\hat{\pi}_{11}$ برآورد OLS اثر مرحله اول، یعنی π_{11}، در 4.1.4a است. بنابراین 2SLS با یک متغیر ابزاری ILS است، به عبارت دیگر نسبت اثر فرم تقلیل‌یافته متغیر ابزاری بر اثر مرحله اول متناظر که در آن مرحله اول و فرم تقلیل‌یافته هر دو شامل متغیر کمکی هستند.

جدول ۱-۱-۴. برآورد 2SLS از بازده اقتصادی تحصیل

	OLS		2SLS					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Years of education	0.075 (0.0004)	0.072 (0.0004)	0.103 (0.024)	0.112 (0.021)	0.106 (0.026)	0.108 (0.019)	0.089 (0.016)	0.061 (0.031)
Covariates:								
Age (in quarters)								✓
Age (in quarters) squared								✓
9 year of birth dummies		✓			✓	✓	✓	✓
50 state of birth dummies		✓			✓	✓	✓	✓
Instruments:			dummy for QOB=1	dummy for QOB=1 or QOB=2	dummy for QOB=1	full set of QOB dummies	full set of QOB dummies int. with year of birth dummies	full set of QOB dummies int. with year of birth dummies

نکات: این جدول برآوردهای **OLS** و **2SLS**های بازده تحصیل با استفاده از نمونه سرشماری ۱۹۸۰ آنگریست و کروگر (۱۹۹۱) را نشان می‌دهد. این نمونه شامل مردان بومی متولد شده در ۱۹۳۰ تا ۱۹۳۹ با درآمد مثبت و مقادیر تخصیص‌نیافته برای متغیرهای کلیدی است. این اندازه نمونه ۳۲۹،۵۰۹ است و خطاهای استاندارد استوار در پرانتزها گزارش شدند.

اولین جفت برآوردهای IV در ستون سوم و چهارم از مدل‌های بدون متغیرهای کنترلی هستند. متغیرهای ابزاری استفاده شده برای ایجاد برآورد در ستون اول، متغیر موهومی برای تولد در سه‌ماهه اول است در حالی که متغیرهای ابزاری استفاده شده برای ایجاد برآوردها در ستون دوم دو متغیر موهومی هستند که بر تولدهایی است که به فصل اول و دوم تولد اشاره دارد. طیف برآورد خطای استاندارد از ۰/۱۰ تا ۰/۱۱ است. از آنجایی که فصل تولد ارتباط نزدیکی با هیچ کدام این متغیر کنترلی ندارد، تعجبی ندارد که نتایج مدل‌های شامل متغیرهای موهومی سال تولد و وضعیت تولد به عنوان متغیر کنترلی شبیه هستند. به طور کل برآوردهای 2SLS اکثراً بزرگ‌تر از برآوردهای متناظر OLS هستند.

ستون هفتم جدول ۱-۱-۴ نتایج اضافه کردن اثر متقابل را به لیست متغیرهای ابزاری نشان می‌دهد. به خصوص، برای مجموع ۳۰ متغیر ابزاری حذف شده، هر تشخیص[1] اثر متقابلی را با ۹ متغیر موهومی برای سال تولد اضافه می‌کند (نمونه شامل هم‌گروهی‌های متولد شده در ۱۹۳۹-۱۹۳۰ است). لذا معادله مرحله اول می‌شود:

$$
\begin{aligned}
s_i &= X_i'\pi_{10} + \pi_{11}z_{1i} + \pi_{12}z_{2i} + \pi_{13}z_{3i} \\
&\quad + \sum_j (B_{ij}z_{1i})\kappa_{1j} + \sum_j (B_{ij}z_{2i})\kappa_{2j} + \sum_j (B_{ij}z_{3i})\kappa_{3j} + \xi_{1i}
\end{aligned}
\qquad (4.1.10b)
$$

که در آن اگر فرد i در سال j برابر با ۱۹۳۹-۱۹۳۱ متولد شده باشد، Bij متغیر موهومی است که برابر می‌شود با یک. ضریب‌های K1j، K3j، K3j اثرهای متقابل سال تولد هستند. این جمله‌های اثر متقابل تفاوت در ارتباط بین فصل تولد و تحصیل را بین هم‌گروهی‌ها نشان می‌دهد. دلیل اضافه

1. Specification

کردن جمله‌های اثر متقابل به فرمول، افزایش دقتی است که از افزایش R2 مرحله اول سرچشمه می‌گیرد و افزایش R2 به این دلیل است که الگوی فصل در تحصیل در هم‌گروهی‌ها تفاوت دارد. در این مثال، اضافه کردن جمله اثر متقابل به فهرست متغیر ابزاری به افزایش نسبتاً کمی در دقت می‌انجامد و خطای استاندارد از ۰/۱۹۴ به ۰/۱۶۱ کاهش می‌یابد.[۱]

آخرین مدل 2SLS که در جدول ۱-۱-۴ آمده است شامل متغیرهای کنترلی برای جمله‌های خطی و درجه دوم در سن ـ در ـ فصل در فهرست متغیرهای کمکی Xi است. به عبارت دیگر، کسی که در فصل اول ۱۹۳۰ متولد شده باشد در روز سرشماری، یعنی اول آپریل سال ۱۹۸۰، به عنوان فردی ۵۰ساله ثبت می‌شود، در صورتی که فردی که در فصل چهارم متولد می‌شود، به عنوان فردی ۴۹ ـ ۲۵ ساله ثبت می‌شود. به این خاطر که تفاضل سنی کم ممکن است متغیری حذف شده باشد که راهبرد شناسایی فصل تولد را مغشوش می‌کند، این متغیر سن به درستی کدگذاری شده، که در مدلی با جمله خطی و درجه دو وارد شده است، کنترلی نسبی را فراهم می‌کند. تا آنجایی که اثر متغیر سن به طور مشابه هموار است، مدل سن ـ در ـ فصل درجه دوم آن را مشاهده می‌کند.

این تغییرات در راه‌اندازی 2SLS اثر متقابل بین شناسایی و برآورد را نشان می‌دهد. برای اینکه رویه 2SLS کار کند، باید تغییراتی در مقادیر برازانده شده مرحله اول وجود داشته باشد که مشروط بر هرگونه متغیرهای کنترلی (متغیرهای کمکی) است که در مدل وارد شده است. اگر مقدار برازانده مرحله اول ترکیب خطی مقادیر کمکی مشمول در مدل باشد، به وضوح برآورد 2SLS وجود ندارد. در معادله ۹-۱-۴ این مسئله توسط همبستگی بین متغیرهای مستقل[۲] نشان داده می‌شود. برآوردهای 2SLS با متغیر درجه دوم سن وجود دارد، اما وقتی که متغیرهای کمکی متغیرهایی مثل سن در فصل را که بسیار نزدیک به متغیرهای ابزاری (متغیرهای موهومی فصل تولد) هستند دربرمی‌گیرند، تغییرپذیری که در مقدار برازانده مرحله اول «پس‌ماند»[۳] است کاهش می‌یابد. به این دلیل که این تغییرپذیری تعیین‌کننده اصلی خطای استاندارد 2SLS است، برآورد ستون هشتم نسبت به برآورد ستون هفتم دقت بسیار کمتری دارد، اگرچه این برآورد هنوز نزدیک به برآورد OLS متناظر نزدیک است.

یادآوری IV و 2SLS

همان طور که مشاهده کردیم، متغیرهای درون‌زا شامل دو نوع متغیر است: الف) متغیر وابسته و ب) متغیر(های) مستقلی هستند که متغیر ابزاری شوند. در مدل معادلات هم‌زمان، متغیرهای درون‌زا با حل سیستم معادلات خطی تصادفی تعیین می‌شود. برای تیمار متغیرهای مستقل به عنوان متغیر برون‌زا باید آنها

۱. از آنجا که استفاده از متغیرهای ابزاری اضافی بسیار احتمال افزایش اریب را به همراه دارد، این دستاورد ممکن است بدون هزینه نباشد. درباره این موضوع در بخش ۸ بحث شده است.

2. Multicollinearity

3. Left Over

را متغیر ابزاری کرد. به عبارت دیگر، در مرحله دوم رویه **2SLS**، آنها را با مقدار برازانده شده جایگزین کنیم. متغیرهای درون‌زای مستقل در پژوهش آنگریست و کروگر (۱۹۹۱) تحصیل است. متغیرهای برون‌زا شامل خود متغیرهای ابزاری و همچنین متغیرهای کمکی برون‌زا هستند که متغیر ابزاری نشده‌اند. در مدل معادلات هم‌زمان، متغیرهای برون‌زا بیرون سیستم مشخص می‌شوند. متغیرهای کمکی برون‌زا در پژوهش آنگریست و کروگر (۱۹۹۱) متغیرهای موهومی برای سال تولد و وضعیت تولد هستند. فرض بر این است که متغیرهای کنترلی برون‌زا متغیر کنترلی هستند. طرفداران **2SLS** در دنیایی زندگی می‌کنند که برچسب‌های منحصربه‌فرد دوطرفه وجود دارد: در هر پژوهش تجربی که شامل متغیرهای ابزاری هستند، متغیرهای تصادفی که باید مطالعه شوند یا متغیرهای وابسته، مستقل درون‌زا، و متغیرهای ابزاری هستند یا متغیرهای کمکی برون‌زا هستند. گاهی اوقات ما این عبارت را به این شکل کوتاه می‌کنیم: متغیرهای وابسته و درون‌زا، متغیرهای ابزاری و کمکی (با پنهان کردن این حقیقت که متغیرهای وابسته در **SEM** سنتی درون‌زا هستند).

۲-۱-۴. برآوردگر والد[1]

آسان‌ترین برآوردگر **IV** از متغیر ابزاری دودویی (باینری) (۱-۰) برای برآورد مدلی با یک متغیر درون‌زا و هیچ متغیر کمکی است. بدون متغیر کمکی، مدل رگرسیون علّی این گونه است:

$$Y_i = \alpha + \rho S_i + \eta_i, \qquad (4.1.11)$$

که در آن η_i و S_i ممکن است همبسته باشند. با توجه به ساده‌سازی بعدی که Z_i متغیر موهومی است که احتمال p را برابر با یک می‌کند، ما می‌توانیم نشان دهیم:

$$Cov(Y_i, Z_i) = \{E[Y_i|Z_i = 1] - E[Y_i|Z_i = 0]\}p(1-p),$$

با فرمول متناظر برای $cov(S_i, Z_i)$. بنابراین:

$$\rho = \frac{E[Y_i|Z_i = 1] - E[Y_i|Z_i = 0]}{E[S_i|Z_i = 1] - E[S_i|Z_i = 0]}. \qquad (4.1.12)$$

مسیر مستقیم به این نتیجه از معادله ۱۱-۱-۴ و این حقیقت که $E[\eta_i|Z_i] = 0$ استفاده می‌کند، بنابراین داریم:

$$E[Y_i|Z_i] = \alpha + \rho E[S_i|Z_i]. \qquad (4.1.13)$$

که حل این معادله برای رویه ρ معادله (۱۲-۱-۴) را تولید می‌کند.

معادله ۱۲-۱-۴ مشابه جامعه برآورد والد (۱۹۴۰) برای رگرسیون دومتغیره با رگرسورهای بد اندازه‌گیری شده است.[2] برآورد والد مشابه نمونه این عبارت است. در شرایط ما، فرمول والد اجرای بسیار

1. Wald Estimator

۲. همان طور که در مقدمه این بخش ذکر شد، خطای اندازه‌گیری شده در رگرسور تمایل به کاهش ضریب رگرسیون تا صفر دارد. برای حذف این اریب، والد (۱۹٤۰) پیشنهاد می‌دهد که داده به روش مستقل از خطای اندازه‌گیری تقسیم شود و ضریب مورد نظر که به عنوان نسبت تفاضل‌های میانگین همانند ۱۲-۱-۴ برآورد شود. (Durbin,1954) نشان داد که روش والد در برازاندن خطوط مستقیم (Fitting Straight Lines) برآورد IV است که در آن متغیر موهومی است که تقسیم داده والد را نشانه‌گذاری می‌کند. (Hausman, 2001) راهبردهای اقتصادسنجی را در برخورد با خطای اندازه‌گیری بررسی می‌کند.

شفافی از راهبرد IV برای حذف اریب متغیر حذف شده است. ادعای مهمی که برآورد IV اثر علّی را برمی‌انگیزد، این است که تنها دلیل برای رابطه متغیرهای وابسته و متغیر ابزاری اثر متغیر ابزاری بر متغیر علّی مورد نظر است. بنابراین در شرایط متغیر ابزاری باینری، تقسیم تفاضل فرم تقلیل‌یافتهِ میانگین‌ها به تفاضل مرحله اول متناظر میانگین‌ها امری طبیعی است. پژوهش آنگریست و کروگر (۱۹۹۱) با استفاده از فصل تولد برای برآورد بازده اقتصادی تحصیل، برآورد والد را هنگام اجرا نشان می‌دهد. جدول ۱۲-۴ پارامترهای برآورد والد را که با استفاده از سرشماری ۱۹۸۰ به دست آمدند نشان می‌دهد. تفاضل درآمد بین مردان متولد شده در نیمه اول و دوم سال ۰/۱۳۴۹- (S.e.=.00337) است. در حالی که تفاضل متناظر در تحصیل ۰/۱۵۱۴- است. برآورد والد از ارزش اقتصادی تحصیل به ازای هر سال، نسبت این دو تفاضل میانگین است که در اینجا می‌شود ۰/۸۹۱ (S.e.=021). تعجبی ندارد که این برآورد از برآورد 2SLS در جدول ۱-۱-۴ بیشتر باشد. اینکه انتظار داریم برآورد 2SLS و والد شبیه باشد به این خاطر است که هر دو از یک نوع اطلاعات به وجود آمدند: تفاضل درآمدها بر اساس فصل تولد.

پژوهش آنگریست (۱۹۹۰) درباره اثر خدمت سربازی در ویتنام بر سربازان، اجرای برآوردگر والد را نشان می‌دهد.

جدول ۲-۱-۴. برآورد والد از بازده تحصیل با استفاده از متغیر ابزاری فصل تولد

	(1) Born in the 1st or 2nd quarter of year	(2) Born in the 3rd or 4th quarter of year	(3) Difference (std. error) (1)-(2)
ln (weekly wage)	5.8916	5.9051	-0.01349 (0.00337)
Years of education	12.6881	12.8394	-0.1514 (0.0162)
Wald estimate of return to education			0.0891 (0.0210)
OLS estimate of return to education			0.0703 (0.0005)

نکته: این جدول از تحلیل دوباره آنگریست و کروگر (۱۹۹۱) توسط آنگریست و ایمبنز[1] (۱۹۹۵) اتخاذ شده است. این نمونه شامل مردان بومی متولد ۱۹۳۹-۱۹۳۰ با درآمد مثبت در سند ۵ درصدی سرشماری ۱۹۸۰ است. اندازه نمونه ۳۲۹,۵۰۹ است.

1. Imbens

در دهه ۱۹۶۰ و ۱۹۷۰، ریسک اعزام به خدمت سربازی برای مردان جوان وجود داشت. نگرانی‌ها درباره سیاست خدمت اجباری آمریکا به قرعه‌کشی خدمت سربازی در ۱۹۷۰ انجامید که برای تعیین اولویت‌های خدمت اجباری استفاده شد. متغیر ابزاری خوب برای وضعیت سربازان موجود در ویتنام واجد شرایط بودن یا مشمولیت است، زیرا این مشمولیت توسط قرعه‌کشی بر اساس تاریخ تولد است. به خصوص، از سال ۱۹۷۰ تا ۱۹۷۲، هر سال شماره‌های ترتیب تصادفی (**RSNs**) به طور تصادفی به تاریخ تولد افراد ۱۹ساله تخصیص داده شد. مردان با شماره قرعه‌کشی پایین‌تر از سقف مشمولیت برای این قرعه‌کشی واجد شرایط بودند. در عمل، خیلی از مردان واجد شرایط قرعه‌کشی به دلیل عدم سلامت و دیگر دلایل معاف شدند. در حالی که بسیاری از مردانی که معاف بودند هنوز برای خدمت داوطلب شدند. بنابراین وضعیت سربازان کاملاً توسط مشمولیت سربازی تصادفیده تعیین نشد، اما مشمولیت متغیر ابزاری باینری ایجاد کرد که با وضعیت سربازان در ویتنام بسیار همبسته بود. این مسئله در جدول ۳-۱-۴ که در آن ستون دو حاکی از اثر وضعیت مشمولیت تصادفیده بر میانگین درآمد تأمین اجتماعی مشمول مالیات است نشان داده شده است. ستون اول این جدول بیانگر میانگین درآمد سالیانه است که هدف آن مقایسه می‌باشد. برای مردانی که ۱۹۵۰ متولد شدند وقتی که این مردان اکثراً داشتند خدمت سربازی خود را آغاز می‌کردند، اثر بسیار منفی شرایط مشمولیت بر درآمد در سال ۱۹۷۱ و شاید با کمال تعجب در ۱۹۸۱ یعنی ده سال دیرتر وجود دارد. در مقابل، مدرکی از رابطه بین شرایط مشمولیت و درآمد در سال ۱۹۶۹، یعنی سالی که قرعه‌کشی برای مردان متولد ۱۹۵۰ پیش از آنکه مردان متولد ۱۹۵۰ واقعاً اعزام شوند برگزار شد، وجود ندارد. به این خاطر که شرایط مشمولیت به صورت تصادفی تخصیص یافتند، به نظر می‌آید این ادعا که برآوردها در ستون دوم اثر مشمولیت بر درآمد را نشان می‌دهد، قابل قبول است. اطلاعات مورد نیاز برای تغییر اثر مشمولیت به اثر وضعیت سرباز مخرج برآورد والد است که اثر مشمولیت بر احتمال اعزام به خدمت سربازی است. این اطلاعات در ستون سوم جدول ۴-۱-۳ گزارش شده است که نشان می‌دهد مردان واجد شرایط ۱۶ درصد بیشتر احتمال داشت که به ویتنام اعزام شده باشند. برآورد والد از اثر خدمت سربازی بر درآمد ۱۹۸۱، که در ستون چهارم بیان شده است، برابر با حدود ۱۵ درصد میانگین بود. وقتی سربازان تحت تأثیر هنوز در ارتش بودند، این آثار حتی در سال ۱۹۷۱ بزرگ‌تر بودند (از لحاظ درصدی).

جدول ۳-۱-۴. برآورد والد از اثر خدمت سربازی بر درآمد مردان سفیدپوست متولد ۱۹۵۰

Earnings year	Earnings: Mean (1)	Earnings: Eligibility Effect (2)	Veteran Status: Mean (3)	Veteran Status: Eligibility Effect (4)	Wald Estimate of Veteran Effect (5)
1981	16,461	-435.8 (210.5)	0.267	0.159 (0.040)	-2,741 (1,324)
1971	3,338	-325.9 (46.6)			-2050 (293)
1969	2,299	-2.0 (34.5)			

نکته: این جدول از جدول ۲و ۳ آنگریست (۱۹۹۰) گرفته شده است. خطای استاندار در پرانتز نشان داده شده است. داده درآمد از اطلاعات ثبت شده اداره تأمین اجتماعی گرفته شده است. ارقام به دلار هستند. داده‌های وضعیت سرباز از منبع اطلاعات آماری[1] اداره آمار آمریکا گرفته شده است. در این نمونه ۱۳۵۰۰ فرد حضور دارند.

مهم‌ترین ویژگی برآوردگر والد **IV** این است که ارزیابی و تفسیر فرضیات شاخص آسان است. تصور کنید که $\mathbf{D_i}$ وضعیت سربازان موجود در ویتنام را نشان می‌دهد و $\mathbf{Z_i}$ مشمولیت را نشان می‌دهد. نظریه پایه که تفسیر ما را از برآوردگر والد به عنوان نشان‌دهنده اثر علّی $\mathbf{D_i}$ توجیه می‌کند این است که تنها دلیل اینکه چرا وقتی $\mathbf{Z_i}$ تغییر می‌کند، $E[Y_i|Z_i]$ تغییر می‌کند تغییرات در $E[D_i|Z_i]$ است. بررسی ساده در این مورد یافتن رابطه بین Z_i و ویژگی‌های شخصی است که نباید تحت تأثیر D_i باشند، به طور مثال سن، نژاد، جنسیت، یا هر ویژگی دیگری که قبل از مشخص شدن D_i مشخص شد. بررسی مفید دیگر، جستجوی ارتباط بین متغیر ابزاری و درآمدها در نمونه‌هایی است که در آن هیچ ارتباطی بین $\mathbf{D_i}$ و Z_i وجود ندارد. اگر تنها دلیل اثر مشمولیت بر درآمد وضعیت سرباز باشد، در مثال‌هایی که اثر مشمولیت با وضعیت سرباز ناهمبسته است، اثر مشمولیت بر درآمد باید صفر باشد.

این نظر در پژوهش آنگریست (۱۹۹۰) درباره قرعه‌کشی سربازی توسط بررسی درآمدهای ۱۹۶۹، برآوردی که در ستون آخر جدول ۳-۱-۴ نشان شد، بیان شده است. از آنجایی که درآمد کسب شده در سال ۱۹۶۹ پیش از قرعه‌کشی خدمت سربازی بوده است، اینکه تیمار اثر شمولیت بر درآمد در سال ۱۹۶۹ صفر است مسئله‌ای اطمینان‌بخش است. دومین تغییرات در این نظریه مردان متولد ۱۹۵۳ را بررسی می‌کند. اگرچه در قرعه‌کشی به متولدین فوریه ۱۹۵۳، **RSNs** تخصیص یافت، در واقع هیچ کدام از متولدین ۱۹۵۳ اعزام نشدند (اعزام در سال ۱۹۷۳ به طور رسمی پایان یافت). بنابراین ارتباط مرحله اول بین مشمولیت و وضعیت مردان ۱۹۵۳ تفاضل کمی را در احتمال خدمت کردن بر اساس

1. Survey of Program Participation

وضعیت مشمولیت نشان می‌دهد. همچنین ارتباط معناداری بین درآمد و وضعیت مشمولیت برای مردان متولد ۱۹۵۳ وجود ندارد که این نظریه را که تنها دلیل اثرات مشمولیت، خدمت سربازی است تأیید می‌کند. ما بحث برآوردگر والد را با مجموعه‌ای از برآوردهای **IV** اثر بزرگی خانواده بر شغل مادر است. همانند تحصیل و خدمت سربازی، این برآوردها برای نمایش دادن در جایی از کتاب استفاده شده است. ارتباط بین باروری و عرضه کار[1] همیشه مورد توجه کارشناسان اقتصاد کار بوده است، در حالی که مورد اریب متغیرهای حذف شده در این زمینه روشن است: مادری که نیروی کاری ضعیفی در محل کار محسوب می‌شود یا پتانسیل درآمد پایین دارد، نسبت به مادرانی که نیروی کاری خوبی محسوب می‌شوند و پتانسیل درآمد بالایی دارند، احتمال دارد که فرزندان بیشتری داشته باشند. از آنجا که مادران با فرزندان زیاد ممکن است کمتر کار کرده باشند، این مسئله تفسیر ارتباط مشاهده شده بین بزرگی خانواده و شغل را مشکل می‌کند. در اینجا آنگریست و ایوانز[2] (۱۹۹۸) مشکل متغیر حذف شده را با استفاده از دو متغیر ابزاری حل می‌کنند، که هر دو برای راهبردهای برآورد نوع والد مناسب هستند.

اولین برآوردگر والد از چندقلوزایی استفاده می‌کند که راهبرد شناسایی برای اثر بزرگی خانواده است و مبتکر آن روزنتسوایگ و ولپین[3] (۱۹۸۰) بود. دوقلو به عنوان متغیر ابزاری در پژوهش آنگریست و ایوانز (۱۹۹۸) در نمونه‌ای که مادر حداقل دو فرزند دارد، متغیر موهومی برای سومین تولد است. اولین مرحله متغیر ابزاری دوقلو ۰/۶۲۵ است که در ستون سوم جدول ۴-۱-۴ نمایش داده است. این مسئله نشان می‌دهد که ۳۷/۵ درصد مادران با دو فرزند یا بیشتر، به هر حال فرزند سوم را نیز به دنیا آوردند و سومین تولد این نسبت را به یک افزایش می‌دهد. متغیر ابزاری دوقلو بر پایه این نظریه است که رخداد چندقلوزایی لزوماً تصادفی است، که به خروجی بالقوه یا ویژگی‌های جامعه‌شناختی مرتبط نیست.

برآوردگر دوم والد در جدول ۴-۱-۴ از ترکیب جنسیت خواهر و برادرها به متغیری ابزاری استفاده می‌کند. برانگیزاننده این متغیر ابزاری این حقیقت است که والدین آمریکایی با دو فرزند همجنس (یعنی دو دختر یا دو پسر)، نسبت به والدینی که یک دختر و پسر دارند، احتمال بسیار زیاد دارد که صاحب فرزند سوم شوند. این مطلب در ستون پنجم جدول ۱-۱-۴ نشان داده شده است که نشان می‌دهد والدین با فرزندان همجنس با احتمال ۶/۷ درصد بیشتر صاحب فرزند سوم می‌شوند (احتمال فرزند سوم میان والدین که هم فرزند دختر دارند و هم فرزند پسر ۰/۳۸ است). متغیر ابزاری دوجنسی بر اساس این نظریه است که داشتن فرزند از هر دو جنس دختر و پسر لزوماً تصادفی است و عرضه کار خانواده را با افزایش باروری تحت تأثیر قرار می‌دهد.

هر دو متغیر ابزاری دوقلویی و دوجنسی بیانگر آن هستند که تولد سومین فرزند بر نرخ اشتغال،

1. Labor Supply
2. Evans
3. Rosenzweig and Wolpin

هفته‌ها و ساعات کاری اثر بسیار زیادی دارند. برآوردگر والد با استفاده از متغیر ابزاری دوقلو برآورد دقیقی از کاهش اشتغال حدود ۰/۸ را نشان می‌دهد، در حالی که هفته کاری کاهش ۳/۸ و ساعات کاری در هفته کاهش ۳/۴ دارد. قدر مطلق این نتایج، که در ستون چهارم ۴-۱-۴ نشان داده شدند، کوچک‌تر از برآورد **OLS** متناظر گزارش شده در ستون دوم هستند. این قضیه بیانگر آن است که برآورد **OLS** توسط اریبی گزینش اغراق‌آمیز شده است. جالب توجه است برآوردگرهای ساخته شده با استفاده از متغیر موهومی همجنسی که در ستون ششم جدول ۴-۱-۴ آمده است بسیار بیشتر از برآوردهای متغیر ابزاری دوقلو است. کنار هم گذاشتن متغیرهای ابزاری دوقلو و دوجنسی در جدول ۴-۱-۴ نشان می‌دهد که ابزارهای متفاوت نیاز به تولید برآوردهای مشابه از اثر علّی ندارند، حتی اگر هردو معتبر باشند. ما این مطلب بسیار مهم را در بخش ۴-۴ شرح خواهیم داد، اما در حال حاضر تمرکز ما چارچوب اثر پایدار است.

جدول ۴-۱-۴. برآورد والد از اثر عرضه کار

			IV Estimates using:			
			Twins		Sex-composition	
Dependent variable	Mean (1)	OLS (2)	First stage (3)	Wald estimates (4)	First stage (5)	Wald estimates (6)
Employment	0.528	-0.167	0.625	-0.083	0.067	-0.135
		(0.002)	(0.011)	(0.017)	(0.002)	(0.029)
Weeks worked	19.0	-8.05	"	-3.83	"	-6.23
		(0.09)	"	(0.758)	"	(1.29)
Hours/week	16.7	-6.02	"	-3.39	"	-5.54
		(0.08)	"	(0.637)	"	(1.08)

نکته: این جدول برآورد والد و OLS از اثر تولد سوم بر عرضه کار را با استفاده از دوقلو و دوجنسی متغیرهای ابزاری نشان می‌دهد. داده‌ها از قسمت آنگریست و ایوانز (۱۹۹۸) گرفته شدند که شامل زنان متأهل در سنین ۳۵-۲۱ با حداقل دو فرزند در سرشماری ۱۹۸۰ است. مدل‌های OLS شامل متغیر کنترلی برای سن مادران، سن در تولد فرزند اول، متغیرهای موهومی برای جنسیت فرزند اول و دوم، متغیر موهومی برای نژاد است.

۳-۱-۴. داده‌های گروهی و 2SLS

برآوردگر والد،[1] مادر تمام برآوردگرهای متغیرهای ابزاری است، زیرا برآوردگرهای پیچیده‌تر **2SLS** معمولاً از مجموعه اصلی برآوردگرهای والد تشکیل می‌شوند. رابطه بین برآوردگر والد و برآوردگر **2SLS** به داده‌های گروهی مربوط می‌شود: استفاده برآوردگر **2SLS** از متغیرهای موهومی، مشابه **GLS** در مجموعه میانگین‌های گروهی است. **GLS** را می‌توان ترکیب خطی تمام برآوردگرهای والد دانست که از جفت‌های میانگین تشکیل می‌شود. به نظر می‌رسد کلیت این ارتباط به این احتمال محدود می‌شود که متغیرهای موجود، موهومی هستند. تمام متغیرهای ابزاری موهومی، یا حتی گسسته نیستند، ولی در واقع این موضوع اهمیتی ندارد. بسیاری از ابزارهای معتبر را می‌توان به صورت طبقه‌بندی شده از

1. Wald

جمله بر حسب فصل تولد تعریف کرد. به علاوه، متغیرهای ابزاری که پیوسته‌تر باشند (مانند ارقام قرعه‌کشی انتخاب سرباز که در بازه ۱-۳۶۵ هستند) معمولاً می‌توانند بدون اتلاف مقادیر زیادی از اطلاعات، گروهی شوند (برای مثال، یک متغیر موهومی برای وضعیت مشمولیت فراخوان سربازی یا متغیرهای موهومی برای گروه‌هایی با ۲۵ رقم قرعه‌کشی) [۹].

برای توضیح کامل‌تر رابطه گروهی یا والد و یا 2SLS، روی قرعه‌کشی فراخوان سربازی تحقیق می‌کنیم. پیش از این اشاره کردیم که مشمولیت در فراخوان سربازی، ابزار مناسبی برای بررسی وضعیت سربازان کهنه کار ویتنام است که برای مردان متولد ۱۹۵۰، حداکثر RSN[1] 195 برای مردان متولد ۱۹۵۱، RSN 125 و برای مردان متولد ۱۹۵۲، RSN 95 بود. با این همه، در عمل، بین ارقام قرعه‌کشی فراخوان سربازی (که آن را به طور خلاصه برای RSN، Ri می‌نامیم) و وضعیت کهنه سربازی (D_i) ارتباط بهتری بود تا تنها شامل شدن در فراخوان. با وجودی که مردان با ارقام بالای سقف شمولیت سربازی، انتخاب نمی‌شدند، این سقف از قبل معلوم نبود. بنابراین، برخی از مردان به امید خدمت کردن تحت شرایط بهتر و داشتن کنترل بر زمان‌بندی خدمتشان، داوطلب می‌شدند. فشار برای انتخاب شدن، برای مردان با شانس پایین انتخاب شدن در فراخوان، بالا بود، ولی برای مردان با ارقام شانس بالای انتخاب شدن، پایین بود. در نتیجه، تغییر در $P[D_i = 1|R_i]$ حتی برای مقادیر بالا یا پایین، راه میانبری برای شامل شدن در فراخوان بود. برای مثال، مردان متولد ۱۹۶۰ با شانس انتخاب ۲۰۰-۲۲۵۰ به احتمال بیشتری خدمت می‌کردند تا آنهایی که شانس انتخابشان بین ۲۲۶ و ۲۵۰ بود، هرچند در نهایت هیچ کدام از اعضای دو گروه انتخاب نمی‌شدند.

برآوردگر والد از فراخوان سربازی به عنوان ابزاری برای مقایسه مردان متولد ۱۹۵۰ استفاده می‌کند تا درآمد مردان با $R_i < 195$ با درآمد مردان $R_i > 195$. مقایسه شود. ولی بحث قبلی احتمال مقایسه‌های بیشتری را نشان می‌دهد، برای مثال مردان با $R_i \leq 25$ با مردان با $R_i \in [26 - 50]$ مقایسه می‌شوند؛ مردان با $R_i \in [51 - 75]$ با مردان با $R_i \in [76 - 100]$ مقایسه می‌شوند و به همین ترتیب این مقایسه‌ها ادامه پیدا می‌کند، تا این بازه‌های ۲۵ رقمی به پایان برسد. همچنین می‌توانیم بازه‌ها را کوچک‌تر بگیریم، تا به جای بازه‌های ۲۵ رقمی، مردان در بازه‌های پنج رقمی یا تک رقمی مقایسه می‌شوند. نتیجه این بسط در مجموعه مقایسه‌ها، مجموعه برآوردگرهای والد است. این مجموعه‌ها کامل هستند، به طوری که با بخش‌بندی بازه‌ها، ابزار اصلی تأیید می‌شود، در حالی که برآوردگرهای تکی به صورت خطی مستقل هستند و صورت کسرها به صورت خطی مستقل هستند. در آخر هریک از این برآوردگرهای والد به طور ثابت اثر علّی مشابهی را

1. RSN: Random Sequence Number

برآورد می‌کنند که در اینجا فرض می‌شود تا وقتی R_i از خروجی‌های بالقوه مستقل و با وضعیت کارآزموده همبستگی داشته باشد، ثابت است (یعنی مخرج کسرهای والد صفر نیست).

احتمال ساخت برآوردگرهای والد چندگانه با تأثیر علّی مشابه طبیعتاً این پرسش را به همراه دارد که با همه آنها چه باید کرد. ما مایلیم برآوردگری را ارائه دهیم که به نوعی اطلاعات را به شکل کارآمد در برآوردهای تکی والد، ترکیب می‌کند. همان طور که مشخص است، برای به دست آمدن این برآوردها، کارآمدترین ترکیب خطی مجموعه کامل برآوردهای مستقل خطی والد با انطباق خطی از طریق میانگین‌های گروهی به دست می‌آید.

برآوردگر داده‌های گروهی مستقیماً می‌توانند به صورت زیر به دست بیایند. همان طور که در رابطه ۱۱-۱-۴ دیدیم، ما با مدل دو متغیری[1] اثرات ثابت کار می‌کنیم که در این مورد به صورت زیر نوشته می‌شود:

$$Y_i = \alpha + \rho D_i + \eta_i, \qquad (4.1.14)$$

که در آن $\rho = Y_{1i} - Y_{0i}$ اثر علّی مورد نظر و $Y_{0i} = \alpha + \eta_i$ است. از آنجایی که R_i به صورت تصادفی تعیین می‌شود و فرض می‌شود ارقام قرعه‌کشی هیچ تأثیری روی دریافتی‌ها، غیر از مورد کهنه‌سربازها، ندارند $E[\eta_i|R_i] = 0$ است. بنابراین در ادامه گفته می‌شود که:

$$E[Y_i|R_i] = \alpha + \rho P[D_i = 1|R_i], \qquad (4.1.15)$$

زیرا $P[D_i = 1|R_i] = E[D_i|R_i]$. به بیان دیگر، شیب خطی که دریافتی‌های میانگین بر اساس ارقام قرعه‌کشی را با احتمال میانگین خدمات بر اساس ارقام قرعه‌کشی به هم وصل می‌کند مساوی است با تأثیر خدمات نظامی یا همان ρ. این موضوع علی‌رغم این حقیقت است که رگرسیون Y_i در D_i- در این حالت اختلاف میانگین‌ها بر حسب وضعیت کار ـ قطعاً با ρ متفاوت است، زیرا Y_{0i} و D_i احتمالاً همبستگی دارند.

معادله (۱۵-۱-۴) راهبرد برآورد بر مبنای برازش خطی با مشابه نمونه $E[Y_i|R_i]$ و $P[D_i = 1|R_i]$ را پیشنهاد می‌دهد. فرض کنید که R_i بر حسب مقادیر $j = 1, \ldots, J$ است. در اصل، j را می‌توان از ۱ تا ۳۶۵ گرفت، ولی در آنگریست (۱۹۹۰)، اطلاعات رقم قرعه‌کشی در ۶۹ بازه پنج رقمی، به اضافه یک بازه هفتادی برای ارقام ۳۴۶-۳۶۵ جمع شد. بنابراین، می‌توانیم R_i را از ۱ تا ۷۰ بگیریم. $\bar{y}_j$ و $\hat{p}_j$ به ترتیب نشان‌دهنده برآوردهای $E[Y_i|R_i = j]$ و $P[D_i = 1|R_i = j]$ هستند، در حالی که $\bar{\eta}_j$ نشان‌دهنده خطای میانگین در (۱۴-۱-۴) است. از آنجایی که گشتاورهای نمونه

1. Bivariate

به گشتاورهای جامعه همگرا می‌شوند، برآوردهای **OLS** از ρ در معادله گروهی:

$$\bar{y}_j = \alpha + \rho\hat{p}_j + \bar{\eta}_j \qquad (4.1.16)$$

ثابت هستند. با این همه، در عمل **GLS** ترجیح دارد، زیرا معادله گروهی، با ساختار واریانس معلوم، دارای واریانس ناهمسانی است. برآوردگر **GLS** کارا برای داده‌های گروهی در مدل خطی با اثرات ثابت، حداقل مربعات موزون است که این وزن‌دهی، با واریانس $\bar{\eta}_j$ انجام می‌شود (برای مثال به پرایس و آیچیسون، ۱۹۵۴ یا وولدریج، ۲۰۰۶ رجوع کنید). فرض می‌شود واریانس پسماند داده‌های خُرد با واریانس σ^2_η، همسان است. این واریانس $\frac{\sigma^2_\eta}{n_j}$ است که در آن n_j اندازه گروه است.

برآوردگر GLS (یا حداقل مربعات موزون) ρ در معادله (۱۶-۱-۴) به دو دلیل در این متن اهمیت دارد. نخست، برآورد شیب GLS از J مشاهده گروهی به طور مجانبی ترکیب خطی کارآمد مجموعه کامل J-1 برآوردگرهای والد مستقل خطی است (آنگریست، ۱۹۹۱). این مسئله را می‌توان بدون هرگونه محاسبه ریاضی مشاهده کرد: GLS و هر ترکیب خطی از برآوردگرهای جفتی والد هر دو ترکیب‌های خطی متغیر وابسته گروهی هستند. به علاوه، GLS برآوردگر خطی مؤثر مجانب برای داده‌های گروهی است. بنابراین، می‌توان نتیجه گرفت که هیچ ترکیب خطی بهتر (یعنی از نظر مجانبی مؤثرتر) برآوردگرهای والد از GLS وجود ندارد (مجدداً در اینجا فرض می‌شود ρ ثابت است). آنگریست فرمول ساخت برآوردگر GLS از مجموعه کامل برآوردگرهای والد مستقل خطی را ارائه داد (۱۹۸۸).

دوم، وقتی هر برآوردگر والد، برآوردگر IV نیز باشد، برآوردگر GLS (حداقل مربعات موزون) معادله ۱۶-۱-۴ نیز 2SLS است. در این مورد ابزارها مجموعه کامل متغیرهای موهومی برای تعیین هر سلول تعداد قرعه‌کشی هستند. برای مشاهده علت آن، مجموعه ابزارهای موهومی $Z_i \equiv \{r_{ji} = 1[\mathrm{R}_i = j]; j = 1, \ldots J-1\}$ توصیف می‌شوند. اکنون، رگرسیون مرحله اول D_i روی Z_i به اضافه یک ضریب در نظر گرفته می‌شوند. وقتی مرحله اول اشباع شد، مقادیر برازش شده ابزار نمونه مرسوم $\hat{p}_j$ می‌شوند که n_j بار برای هر j تکرار می‌شوند. برآورد شیب مرحله دوم دقیقاً مشابه برآورد حداقل مربعات موزون معادله گروهی، ۱۶-۱-۴ موزون بر حسب اندازه سلول است.

ارتباط بین داده‌های گروهی و 2SLS از اهمیت عملی و مفهومی برخوردار است. از نظر مفهومی، هر برآوردگر 2SLS با استفاده از مجموعه ابزارهای موهومی را می‌توان به عنوان ترکیب تمام برآوردگرهای والد در نظر گرفت که همزمان با این ابزارها به دست می‌آیند. برآوردگر والد چارچوب ساده‌ای فراهم می‌کند که بعداً در این فصل برای تفسیر برآوردهای IV در دنیای واقعی‌تر خروجی‌های بالقوه ناهمگون[1] به کار برده می‌شوند.

هرچند تمام ابزارها ذاتاً گسسته نیستند و بنابراین بلافاصله تابع تفسیر داده‌های گروهی یا والد

1. Heterogeneous

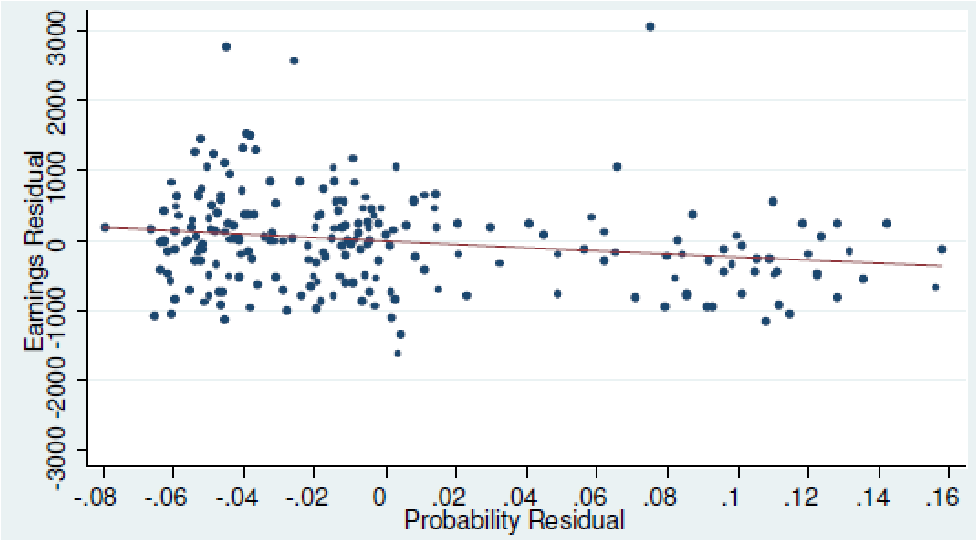
هستند، ولی بسیاری از ابزارها نیز چنین هستند. مثال‌ها شامل تعداد قرعه‌کشی‌های فراخوان سربازی، فصل تولد، دوقلوها و ابزارهای ترکیب جنسیت برادر و خواهری می‌شوند که پیش از این درباره آنها بحث کردیم. همچنین می‌توانید مطالعات اخیر بندسن و همکاران، ۲۰۰۷ و مایکل و آنانات ۲۰۰۸ را مشاهده کنید که هر دو از متغیرهای موهومی استفاده می‌کنند. به علاوه برآوردگرهایی که ویژگی پیوسته دارند، اغلب می‌توانند به شکل نتیجه‌بخشی به متغیرهای گسسته تبدیل شوند. برای مثال آنگریست ابزارهای آب و هوایی پیوسته را با سه متغیر موهومی، طوفانی، مخلوط (کمی تا قسمتی ابری) و صاف گروهی کرد و سپس برای برآورد تقاضای ماهی به کار برد. این پارامترسازی متغیر موهومی برای به دست آوردن ویژگی‌های اصلی رابطه بین شرایط آب و هوایی و قیمت ماهی است.

از نظر عملی، هم‌ارزی داده‌های گروهی 2SLS به ما ابزار ساده‌ای می‌دهد که می‌توانند برای توضیح و برآورد هر راهبرد IV به کار برده شوند. برای مثال در قرعه‌کشی فراخوان سربازی، مدل گروهی این فرض را به دست می‌دهد که تنها دلیل اینکه درآمدهای میانگین با ارقام قرعه‌کشی فرق دارند، تغییر در احتمال خدمات در گروه‌های ارقام قرعه‌کشی است. اگر رابطه موجه اصلی با اثرات ثابت خطی باشد، معادله ۱۶-۱-۴ باید به خوبی با میانگین گروهی همخوانی داشته باشد که آن را هم می‌توانیم در بخش بعدی با بحث و بررسی از طریق استنباط آماری برآورد کنیم.

گاهی اقتصاددانان حوزه اقتصاد کار، از نمودارهای داده‌های گروهی متغیرهای ابزاری گسسته به عنوان متغیرهای ابزاری تصویری (VIV[1]) یاد می‌کنند. آنگریست در سال ۱۹۹۰ مثالی ارائه کرده که در شکل ۲-۱-۴ ارائه شده است. این شکل رابطه بین درآمدهای میانگین در سلول‌های RSN پنج رقمی و احتمال‌پذیری خدمات را در این سلول‌ها برای دستمزدهای سال‌های ۱۹۸۴-۱۹۸۱ مردان سفید متولد ۱۹۵۳-۱۹۵۰ نشان می‌دهد. شیب خط گذری از این نقاط برآورد IV از اتلاف درآمد ناشی از خدمات نظامی است که در این مورد حدود ۲۴۰۰ دلار بود و با برآوردهای والد که قبلاً درباره آنها بحث شد تفاوتی ندارد، ولی خطای استاندارد پایین‌تری دارد (در این مورد حدود ۸۰۰ دلار است).

1. Visual Instrumental Variables (VIV)

شکل ۲-۱-۴. رابطه بین میانگین درآمدها و احتمال خدمت سربازی



این نمودار VIV میانگین درآمد سال‌های ۱۹۸۱ تا ۱۹۸۴ است که به وسیله گروه‌های ارقام متوالی انتخابی بر حسب احتمالات شرطی وضعیت کهنه‌سربازان در سلول‌های مشابه به دست می‌آید. نمونه شامل مردان سفیدپوست متولد ۱۹۵۰ تا ۱۹۵۳ است. نقاط رسم شده شامل میانگین پسماندهای رگرسیون و اثرات هم‌گروهی[1] است. شیب خط رگرسیون حداقل مربعات رسم شده از نقاط ۲۳۸۴- با خطای استاندارد ۷۷۸ است.

۲-۴. استنباط 2SLS مجانبی

۱-۲-۴. توزیع محدودکننده بردار ضرایب 2SLS[2]

ما می‌توانیم توزیع محدودکننده بردار ضرایب **2SLS** را با استفاده از بحث مشابه بخش ۳-۱-۳ برای **OLS** به دست آوریم. در این مورد، $V_i \equiv \left[\begin{array}{cc} X_i' & \hat{s}_i \end{array}\right]'$ بردار رگرسورها در مرحله دوم **2SLS** است. برآوردگر **2SLS** را می‌توان به صورت زیر نوشت:

$$\hat{\Gamma}_{2SLS} \equiv \left[\sum_i V_i V_i'\right]^{-1} \sum_i V_i Y_i,$$

که $\Gamma \equiv \left[\begin{array}{cc} \alpha' & \rho \end{array}\right]'$ بردار ضرایب مربوطه است. یادآور می‌شویم که:

1. Cohort Effects
2. The Limiting Distribution of the 2SLS Coefficient Vector

$$\begin{aligned}\hat{\Gamma}_{2SLS} &= \Gamma + \left[\sum_i V_i V_i'\right]^{-1} \sum_i V_i[\eta_i + \rho(s_i - \hat{s}_i)] \\ &= \Gamma + \left[\sum_i V_i V_i'\right]^{-1} \sum_i V_i \eta_i \qquad (4.2.1)\end{aligned}$$

معادله دوم از این حقیقت به دست می‌آید که پسماندهای مرحله اول، $(s_i - \hat{s}_i)$، در این نمونه متعامد بر V_i هستند. بنابراین، توزیع محدودکننده بردار ضرایب **2SLS** توزیع محدودکننده $[\sum_i V_i V_i']^{-1} \sum_i V_i \eta_i$ است. کار با این مقدار نسبت به مقدار **OLS** کمی سخت‌تر است، زیرا رگرسورها در این مورد شامل مقادیر برازش شده برآورد شده می‌شوند. با این همه، استدلال نوع اسلاتسکی[1] نشان می‌دهد که ما توزیع محدودکننده مشابهی داریم که مقادیر برازش شده برآورد شده را با مقادیر برازش شده جامعه مربوطه جایگزین می‌کند (یعنی جایگزینی $\hat{s}_i$ با $[X_i'\pi_{10} + \pi_{11}z_i]$). بنابراین $\hat{\Gamma}_{2SLS}$ توزیع نرمال مجانبی دارد، با حد احتمال Γ و ماتریس کوواریانس که با $[\sum_i V_i V_i']^{-1} [\sum_i V_i V_i' \eta_i^2] [\sum_i V_i V_i']^{-1}$ برآورد می‌شود. این فرمول ساندویچی[2] مانند خطاهای استاندارد **OLS** است. همانند **OLS**، اگر η_i مشروط به هم واریانس بودن با متغیرها و ابزارهای داده شده باشد، برآوردگر ماتریس کوواریانس به $[\sum_i V_i V_i']^{-1} \sigma_\eta^2$ ساده می‌شود.

در اینجا اطلاعات جدید کم، ولی یک نقطه سخت وجود دارد. برآوردهای **2SLS** به صورت دستی به وسیله برآورد مرحله اول **a**۴-۱-۴ و سپس وارد کردن مقادیر برازش شده در معادله ۹-۱-۴ و برآورد آن به وسیله **OLS**، به نظر طبیعی می‌رسد. تا وقتی برآوردهای ضرایب ادامه داشته باشد این درست است، ولی خطاهای استاندارد برایند نادرست خواهند شد. نرم‌افزارهای رگرسیون سنتی نمی‌دانند که شما سعی دارید برآورد **2SLS** را انجام دهید، بنابراین برآوردگر واریانس پسماند که در فرمول استاندارد قرار می‌گیرد، نادرست خواهد بود. وقتی خطاهای استاندارد تشکیل می‌شوند، نرم‌افزار واریانس پسماند معادله را با **OLS** در مرحله دوم برآورد خواهد کرد:

$$Y_i - [\alpha' X_i + \rho \hat{s}_i] = [\eta_i + \rho(s_i - \hat{s}_i)],$$

ضرایب با برآوردهای مربوطه جایگزین می‌شوند. با این همه، برآوردگر واریانس پسماند درست از رگرسور درونی اصلی برای ساخت پسماندها استفاده می‌کند و نه مقادیر برازش شده مرحله اول. به بیان دیگر، پسماندهایی که شما می‌خواهید $Y_i - [\alpha' X_i + \rho s_i] = \eta_i$ است، به همین ترتیب σ_η^2 برآورد می‌شود و نه $\eta_i + \rho(s_i - \hat{s}_i)$. هرچند این مسئله به راحتی قابل حل شدن است (شما می‌توانید

1. Slutsky-type Argument
2. Sandwich Formula

برآوردگر واریانس پسماند مناسب را در محاسبه‌ای جداگانه به دست بیاورید)، نرم‌افزار طراحی شده برای **2SLS** به صورت خودکار این کار را انجام می‌دهد و می‌تواند کمک کند تا از اشتباهات متداول دیگر **2SLS** جلوگیری کنید.

۲-۲-۴. بیش‌شناسایی[1] و حداقل شوند (مینیماند)[2] *2SLS

مدل‌های اثر ثابت[3] با ابزارهای بیشتر نسبت به رگرسورهای درون‌زا، بیش از حد لازم شناسایی می‌شوند. از آنجایی که ابزارهایی بیش از حد لازم برای شناسایی پارامترهای مورد نظر وجود دارند، این مدل‌ها مجموعه محدودیت‌هایی را اعمال می‌کنند که می‌توانند به عنوان بخشی از فرایند آزمایش مشخص‌سازی،[4] برآورد شوند. با استفاده از این فرایند مشخص می‌شود که آیا خط رسم شده در تصویر نوع **VIV** با میانگین مربوطه با دقت برآورد همخوانی دارد یا خیر. جزئیات پشت این ایده مفید به ساده‌ترین شکل با استفاده از نشانگذاری ماتریس و مدل خطی سنتی بیان می‌شوند.

$Z_i \equiv \begin{bmatrix} X_i' & z_{1i} & \dots & z_{Qi} \end{bmatrix}'$ بردار تشکیل شده به وسیله به هم پیوستن متغیرهای کمکی برون‌زا و **Q** متغیرهای ابزاری را نشان می‌دهد و $W_i \equiv \begin{bmatrix} X_i' & s_i \end{bmatrix}'$ بردار تشکیل شده با به هم پیوستن متغیرهای کمکی و متغیر درون‌زا مورد نظر را نشان می‌دهد. برای مثال در مقاله فصل تولد، متغیرهای کمکی سال تولد و وضعیت تولد هستند، ابزارها فصل تولد هستند و متغیر درون‌زا تحصیل است. بردار ضرایب همانند بخش قبلی همچنان $\Gamma \equiv [\alpha', \rho]'$ است. پسماندها برای مدل علی می‌توانند به صورت تابعی از Γ به صورت زیر توصیف شوند:

$$\eta_i(\Gamma) \equiv Y_i - \Gamma' W_i = Y_i - [\alpha' X_i + \rho s_i].$$

این پسماند با بردار ابزاری Z_i ناهمبسته است. به بیان دیگر، η_i شرط تعامد را برآورده می‌کند.

$$E[Z_i \eta_i(\Gamma)] = 0.$$

با این همه، در هر نمونه، این معادله دقیقاً حفظ نخواهد شد، زیرا شرایط تکانه‌ای[5] بیشتری نسبت به عناصر Γ وجود دارند. مشابه نمونه ۲-۲-۴ جمع روی **i** است.

$$\frac{1}{N}\sum Z_i \eta_i(\Gamma) \equiv m_N(\Gamma). \qquad (4.2.3)$$

1. Over-identification
2. Minimand
3. Constant-effects Models
4.Specification Testing
5. Moment Conditions

2SLS را می‌توان به عنوان برآوردگر روش عمومی شده تکانه‌ها (**GMM**[1]) در نظر گرفت که مقداری را به وسیله مشابه نمونه ۲-۲-۴ نزدیک به صفر، انتخاب می‌کند.

طبق نظریه حد مرکزی، بردار تکانه نمونه $\sqrt{N}m_N(\Gamma)$ ماتریس کوواریانس مجانبی معادل $E[Z_i Z_i' \eta_i(\Gamma)^2]$ دارد که ما آن را Λ می‌نامیم. هرچند در ابتدا عجیب به نظر می‌رسد، این فقط یک ماتریس تکانه چهارم است، همانند فرمول‌های ساندویچی که برای تشکیل خطاهای استاندارد قطعی به کار برده می‌شوند. همان طور که هانسن نشان داد، برآوردگر **GMM** بهینه بر مبنای ۲-۲-۴ شکل درجه دوم را بر حسب بردار تکانه نمونه $m_N(\hat{g})$ به حداقل می‌رساند که در آن $\hat{g}$ برآوردگر کاندیدای Γ است. ماتریس موزون بهینه در میانه شکل درجه دوم **GMM** Λ^{-1} است. البته در عمل، Λ نامعلوم است و باید برآورد شود. نسخه محتمل روش **GMM** از برآوردگر سازگار Λ در ماتریس موزون استفاده می‌کند. از آنجایی که برآوردگر از Λ معلوم و برآورد شده استفاده می‌کند که دارای توزیع محدودکننده مشابهی است، ما این تمایز را نادیده خواهیم گرفت. بنابراین، شکل درجه دوم کمینه شده را دوباره می‌نویسیم:

$$J_N(\hat{g}) \equiv N m_N(\hat{g})' \Lambda^{-1} m_N(\hat{g}), \qquad (4.2.4)$$

که در آن عبارت **N** از نرمالیزاسیون $\sqrt{N}$ تکانه نمونه به دست می‌آید. همان طور که در زیر می‌توان دید، وقتی پسماندها همواریانس هستند، کمینه‌ساز $J_N(\hat{g})$ برآوردگر **2SLS** است. بدون همواریانسی، برآوردگر **GMM** که ۴-۲-۴ را کمینه می‌کند، **IV** دو مرحله‌ای است، به طوری که منطقی به نظر می‌رسد که $J_N(\hat{g})$ مینیماند **2SLS** نامیده شود.

در اینجا جزئیات مربوط به تفسیر **GMM** از **2SLS** ارائه می‌شود. همواریانسی معمول به این معناست که:

$$E[Z_i Z_i' \eta_i(\Gamma)^2] = E[Z_i Z_i']\sigma_\eta^2.$$

با جایگذاری برای Λ^{-1} و با استفاده از Z،Y و W برای نشان دادن بردارها و ماتریس‌ها، شکل درجه دوم کمینه به صورت زیر درمی‌آید:

$$J_N(\hat{g}) = (N\sigma_\eta^2)^{-1} \times (\text{Y} - W\hat{g})' Z E[Z_i Z_i']^{-1} Z'(\text{Y} - W\hat{g}). \qquad (4.2.5)$$

در آخر، با جایگذاری ماتریس مقطع تولید نمونه $\left[\frac{Z'Z}{N}\right]$ برای $E[Z_i Z_i']$، داریم:

$$\hat{J}_N(\hat{g}) = (1/\sigma_\eta^2) \times (\text{Y} - W\hat{g})' P_Z (\text{Y} - W\hat{g}),$$

که $P_Z = Z(Z'Z)^{-1}Z$. از اینجا راه‌حل را به دست می‌آوریم:

1. Generalized Method of Moments

$$\hat{g} = \hat{\Gamma}_{2SLS} = [W'P_Z W]^{-1} W'P_Z Y.$$

از آنجایی که اپراتور P_Z مقادیر برازش شده را تولید می‌کند و P_Z ماتریس خودتوان[1] است، می‌توان آن را به عنوان برآوردگر **OLS** معادله مرحله دوم در نظر گرفت که در ماتریس نوشته می‌شود. در کل، بدون همواریانسی می‌توانیم برآوردگر نوع **2SLS** کارآمد محتمل را با کمینه‌سازی ۴-۲-۴ و با استفاده از برآوردگر $E[Z_i Z_i' \eta_i(\hat{g})^2]$ برای تشکیل $\hat{J}_N(\hat{g})$ به دست آورد. برای نمونه، می‌توان از تکانه‌های چهارم تجربی، $\sum Z_i Z_i' \hat{\eta}_i^2$ استفاده کرد که در آن $\hat{\eta}_i$ پسماند محاسبه شده بدون نگرانی درباره همواریانسی است.

آماره آزمون بیش‌شناسایی با حداقل شوند **2SLS** کمینه شده مفروض می‌شود. این آماره به ما می‌گوید که با این فرض که $E[Z_i \eta_i] = 0$ قابل قبول است، آیا بردار تکانه نمونه $m_N(\hat{g})$ به اندازه کافی به صفر نزدیک است یا نه. به ویژه، با این فرض صفر که پسماندها و ابزارها متعامد هستند، کمینه $J_N(\hat{g})$ توزیع $\chi^2(Q-1)$ دارد. بنابراین، می‌توانیم مقدار تجربی **2SLS** کمینه شده را با جداول کای اسکور، به روش آزمون قبلی برای $H_0 : E[Z_i \eta_i] = 0$ مقایسه کنیم.

به دلایلی که به زودی روشن خواهند شد، ما به بیش‌شناسایی مستقیم علاقه‌ای نداریم. وقتی ابزارها مجموعه کاملی از متغیرهای موهومی که به طور متقابل منحصربه‌فرد هستند، درست مانند برآوردگرهای والد و راهبردهای برآورد داده‌های گروهی که در بالا درباره آنها بحث شد، هدف اصلی ما حداقل شوند (مینیماند) **2SLS** است. در این مورد مهم خاص، در حالی که مینیماند **2SLS** مجموع موزون مناسب مربعات کمینه است، **2SLS** حداقل مربعات موزون معادله گروهی مانند ۶-۱-۴ است. برای مشاهده آن، یادآور می‌شویم که پیش‌بینی بر روی مجموعه کاملی از متغیرهای متقابلاً منحصربه‌فرد، روی مقادیر **J**، یک بردار **N*1** از مقادیر برازش شده مساوی با مقادیر مقید **J** در هر مقدار فراهم می‌سازد (از جمله متغیرهای کمکی که به عنوان ابزار به حساب می‌آیند)، هرکدام n_j بار، که n_j اندازه گروه و $\sum n_j = N$ است. ماتریس محصول مقطع $[Z'Z]$ در این مورد یک ماتریس قطری **J*J** با n_j عنصر است. با ساده‌سازی داریم:

$$\hat{J}_N(\hat{g}) = (1/\sigma_\eta^2) \times \sum n_j (\bar{y}_j - \hat{g}'\bar{W}_j)^2, \qquad (4.2.6)$$

که در آن $\bar{W}_j$ میانگین نمونه ردیف های ماتریس W در گروه j اســـت. بنابراین، $\hat{J}_N(\hat{g})$ مینیماند حداقل مربعات موزون **GLS** برای برآورد رگرسیون گروهی است: $\bar{y}_j$ روی $\bar{W}_j$. با کمی کار بیشتر، ما به سادگی نشان می‌دهیم که روش **IV** دومرحله‌ای مناسب بدون همواریانسی به حداقل می‌رسد.

1. Idempotent Matrix

$$\hat{J}_N(\hat{g}) = \sum_j \left(\frac{n_j}{\sigma_j^2}\right)(\bar{y}_j - \hat{g}'\bar{W}_j)^2, \qquad (4.2.7)$$

که σ_j^2 واریانس η_i در گروه j است. برآورد با استفاده از ۷-۲-۴ آسان است، زیرا می‌توانیم σ_j^2 را در مرحله اول، با استفاده از **2SLS** ناکارآمد، ولی هنوز سازگار که ناهمواریانسی را ندیده می‌گیرد، برآورد کنیم. برآوردگرهای **IV** دو مرحله‌ای کارا توسط آنگریست ساخته شدند.

ساختار **GLS** مینیماند **2SLS** به ما اجازه می‌دهد آماره آزمون بیشناسایی را برای ابزارهای موهومی به عنوان مقیاس ساده خوبی برازش خط متصل‌کننده $\bar{y}_j$ و $\bar{W}_j$ ببینیم. به بیان دیگر، این آماره، آماره خوبی برازش کای اسکوور برای خط موجود در نمودار **VIV** همانند شکل ۲-۱-۴ است. پارامتر درجات آزای کای اسکوور به وسیله اختلاف بین تعداد مقادیر گرفته شده توسط ابزار و تعداد پارامترهای ارزیابی شده داده می‌شود.

همانند مسیرهای مختلفی که به برآوردگر **2SLS** منتهی می‌شوند، مسیرهای بسیاری برای آماره آزمون ۷-۲-۴ وجود دارند. در اینجا دو مسیر دیگر وجود دارند که خوب است با آنها آشنا شویم. نخست آزمون آماری بر مبنای مینیماند **GMM** برای **IV**، خواه ابزارها متغیرهای موهومی باشند یا نه، مشابه آزمون آماری بیش‌شناسایی در بسیاری از منابع اقتصادی در معادلات شبیه‌سازی است. برای مثال، این آماره در فصلی از هاوسمن (۱۹۸۳) درباره معادلات شبیه‌سازی مشخص شد که روش محاسباتی ساده‌ای را نیز پیشنهاد می‌کند: برای مدل‌های هم‌واریانسی، مینیماند کمینه شده **2SLS** اندازه نمونه **R2** از رگرسیون پسماندهای **2SLS** در ابزارهاست. فرمول آن $N\left[\frac{\hat{\eta}' P_Z \hat{\eta}}{\hat{\eta}'\hat{\eta}}\right]$ است که $\hat{\eta} = Y - W\hat{\Gamma}_{2SLS}$ بردار پسماندهای **2SLS** است.

دوم، باید تأکید کرد که بیش‌شناسایی را می‌توان ماهیتاً به عنوان ابزاری دارای بیش از یک کاربرد در اقتصادسنجی مورد توجه قرار داد. به بیان دیگر، بیش از یک ابزار برای رابطه علّی مشابه وجود دارد و ما می‌توانیم برآوردگرهای ساده **IV** ساخته شده را در یک زمان در نظر بگیریم و آنها را مقایسه کنیم. این مقایسه مستقیم بیش‌شناسایی را بررسی می‌کند: اگر هر برآوردگر شناسایی شده سازگار باشد، فاصله بین آنها باید نسبت به واریانس نمونه کوچک باشد و باید همراه با اندازه نمونه کوچک شود و بنابراین دقت این برآوردها افزایش پیدا می‌کند. در حقیقت، ما می‌توانیم این گونه در نظر بگیریم که آیا تمام برآوردگرهای تازه شناسایی شده مشابه هستند یا خیر. این گونه بیان می‌شود که آماره آزمون نتیجه شده، یک آزمون والد بر این فرض صفر تولید می‌کند، در حالی که آماره آزمون بر مبنای مینیماند **2SLS** آزمون مضرب لاگرانژ[1] گفته می‌شود، چراکه می‌توان آن را به بردار امتیاز در نسخه احتمال

1. Lagrange Multiplier (LM) Test

بیشینه مجموعه **IV** نسبت داد.

در نسخه داده‌های گروهی **IV** ، مقادیر آزمون والد برای آزمون تساوی مجموعه تمام برآوردگرهای والد مستقل خطی احتمالی استفاده می‌شود. برای مثال، اگر ارقام قرعه‌کشی به چهار گروه تقسیم شوند، در این صورت سه برآوردگر والد مستقل خطی را می‌توان تشکیل داد. به این ترتیب، برآوردگر داده‌های گروهی مؤثر می‌توانند به وسیله **GLS** در این چهار میانگین مرسوم تشکیل شوند. چهار گروه به معنای این است که سه برآوردگر والد احتمالی و دو محدودیت تساوی غیراضافی روی این سه وجود دارند؛ بنابراین، آمار والد مناسب دو درجه آزای دارد. از سوی دیگر، چهار گروه به معنای سه ابزار و یک متغیر ثابت برای برآورد مدلی با دو پارامتر است. بنابراین مینیماند **2SLS** آزمون آمار بیش‌شناسایی را با دو درجه آزای می‌دهد و در حقیقت، با استفاده از روش برآورد مشابه برای ماتریس موزون به شکل مربع مربوطه، این دو آزمون آماری نه فقط مورد مشابهی را آزمون می‌کنند، بلکه از نظر عددی معادل هستند. این به نظر منطقی می‌رسد، زیرا قبلاً دیدیم که **2SLS** ترکیب خطی مناسب برآوردگرهای والد است.

در آخر و در عمل، یک پیش‌بینی احتیاطی درباره آزمون‌های بیش‌شناسایی وجود دارد: در تجربه ما، آماره بیش‌شناسایی معمولاً در کارهای عملی مقدار کمی دارد. از آنجایی که $J_N(\hat{g})$ سازگاری واریانس نرمال شده را اندازه‌گیری می‌کند و وقتی برآوردهای اساسی دقیق نباشند، آماره آمار بیش‌شناسایی، پایین است. از آنجایی که برآوردهای **IV** اغلب بسیار غیردقیق هستند، ما نمی‌توانیم از این حقیقت راضی باشیم که یک برآورد داخل واریانس نمونه دیگری باشد، حتی اگر برآوردهای تکی به اندازه کافی دقیق به نظر برسند. از سوی دیگر، در موردی که در آن برآوردهای **IV** کاملاً دقیق هستند، حقیقت این است که رد آماره بیش‌شناسایی به خطای شناسایی اشاره نمی‌کند. بیشتر، این نشانه عدم تجانس اثر مداخله باشد، احتمالی که در ادامه درباره آن بحث خواهیم کرد. با این همه، از نظر عملی، شناخت ساختار تشریحی مینیماند **2SLS** از آنجا که که یک بار دیگر اهمیت رابطه بین داده‌های گروهی و **IV** را نشان می‌دهد، بی‌نهایت ارزشمند است. این ارتباط از برآورد و آزمون با متغیرهای ابزاری رمزگشایی می‌کند و ما را وادار می‌کند تا با تکانه‌های دست اولی مواجه شویم که مبنای استنباط علّی هستند.

۳-۴. IV دو نمونه‌ای[1] و IV نمونه ــ شکاف[2]*

تفسیر **GMM** از **2SLS** این حقیقت را نشان می‌دهد که برآوردگر **IV** می‌تواند تنها از تکانه‌های نمونه، بدون داده خُرد، تشکیل شود. با برگشت به شرط تکانه نمونه‌ای، ۳-۲-۴ و با آرایش مجدد آن، معادله رگرسیون مانند شامل تکانه‌های دوم ایجاد می‌شود:

1. Two-sample IV
2. Split-sample IV

$$\frac{Z'\Upsilon}{N} = \frac{Z'W}{N}\Gamma + \frac{Z'\eta}{N} \qquad (4.3.1)$$

GLS برآوردی از Γ در ۱-۳-۴ ثابت است، زیرا: $E\left[\frac{Z'\Upsilon}{N}\right] = E\left[\frac{Z'W}{N}\right]\Gamma$.

مینیماند **2SLS** می‌تواند یک **GLS** به کار رفته برای معادله ۱-۳-۴ پس از ضرب در $\sqrt{N}$ باشد تا پسماند در حال ناپدید شدن با بزرگ شدن اندازه نمونه، حفظ شود. به بیان دیگر، **2SLS** شکل درجه دوم را در پسماندها از ۱-۳-۴ با ماتریس موزون کمینه می‌کند. نگرش مهمی که از نوشتن مسئله **2SLS** به دست می‌آید این است که ما نیازی به مشاهدات تکی خود در نمونه برای برآورد ۱-۳-۴ نداریم. با بردار ضرایب **OLS**، که از تابع میانگین نمونه تشکیل می‌شود، برآوردگرهای **IV** نیز می‌توانند از تکانه‌های نمونه تشکیل شوند. تکانه‌های مورد نیاز برای **IV**، $\frac{Z'\Upsilon}{N}$ و $\frac{Z'W}{N}$ هستند. متغیر وابسته $\frac{Z'\Upsilon}{N}$ بردار بعد $[K+Q] \times 1$ است. ماتریس رگرسور $\frac{Z'W}{N}$، از بعد $[K+Q] \times [K+1]$ است. معادله تکانه دوم نمی‌تواند دقیق حل شود، مگر آنکه **Q=1** باشد، بنابراین منطقی است که سازگاری به اندازه کمینه‌سازی شکل درجه دوم در پسماندها امکانپذیر باشد. مناسب‌ترین ماتریس برای این کار ماتریس کوواریانس مجانبی $\frac{Z'\eta}{\sqrt{N}}$ است. این ماتریس مجدد مینیماند **2SLS** را می‌سازد.

یک نگرش مناسب این واقعیت است که ماتریس‌های تکانه در سمت چپ و راست معادلات در معادله ۱-۳-۴ از مجموعه داده‌های مشابه به دست آمده در این مجموعه‌ها از جامعه مشابه به دست نمی‌آیند. این مشاهده به برآوردگر متغیرهای ابزاری دو نمونه‌ای (**TSIV**[1]) به کار رفته توسط آنگریست منتهی می‌شود. به طور خلاصه Z_1 و Υ_1 ماتریس کوواریانس و یا وسیله و بردار متغیر مستقل در مجموعه داده‌های یک با اندازه N_1 هستند و Z_2 و W_2 ماتریس کوواریانس و یا ابزار و ماتریس کوواریانس و یا متغیر بیرونی را در مجموعه داده‌های دو با اندازه N_2 نشان می‌دهد. فرض می‌شود $plim\left(\frac{Z_2'W_2}{N_2}\right) = plim\left(\frac{Z_1'W_1}{N_1}\right)$، **GLS** برآوردی از معادله تکانه دو نمونه‌ای است.

$$\frac{Z_1'\Upsilon_1}{N_1} = \frac{Z_2'W_2}{N_2}\Gamma + \left\{\left[\frac{Z_1'W_1}{N_1} - \frac{Z_2'W_2}{N_2}\right]\Gamma + \frac{Z_1'\eta_1}{N_1}\right\}$$

که همچنین برای Γ ثابت هستند. توزیع محدودکننده این برآوردگر با نرمال کردن آن به وسیله $\sqrt{N_1}$ و با فرض ثابت بودن $plim\left(\frac{N_2}{N_1}\right)$ به دست می‌آید.

فایده **TSIV** از این حقیقت به دست می‌آید که قالب برآورد **IV** را برای شرایطی توسعه می‌دهد که در آن مشاهدات روی متغیرهای وابسته، ابزارها و متغیر برونی مورد نظر به سختی در نمونه تکی

1. Two-sample Instrumental Variables Estimator

یافت می‌شوند. یافتن یک مجموعه داده که اطلاعاتی درباره خروجی‌ها و ابزارها داشته باشد، کار ساده‌ای به نظر می‌رسد، که با آن فرم کاهش (تقلیل یافته) یافته را می‌توان برآورد کرد و مجموعه داده‌های دیگری به دست می‌آیند که اطلاعاتی درباره متغیرها و ابزارها داشته باشند که با آن مرحله اول برآورد امکانپذیر شود. برای مثال، طبق نظر آنگریست، گزارش‌های به دست آمده از سازمان تأمین اجتماعی،[1] اطلاعاتی درباره متغیر وابسته (درآمدهای سالیانه) و ابزارها (ارقام قرعه‌کشی فراخوان سربازی کدگذاری شده از تاریخ‌های تولد و کوواریانس‌هایی برای نژاد و سال تولد) فراهم می‌کنند. با این همه، سازمان تأمین اجتماعی وضعیت کهنه‌سربازان شرکت‌کننده را دنبال نمی‌کند. این اطلاعات از گزارش‌های نظامی به دست می‌آیند که حاوی تاریخ تولد نیز هستند که می‌توانند برای کدگذاری ارقام قرعه‌کشی به کار برده شوند. آنگریست از این گزارش‌های نظامی برای تشکیل $\frac{Z_2'W_2}{N_2}$، رابطه مرحله اول بین ارقام قرعه‌کشی و وضعیت کهنه‌سربازان از نظر نژاد و سال تولد استفاده کرد، در حالی که داده‌های **SSA** برای تشکیل $\frac{Z_1'Y_1}{N_1}$ به کار برده می‌شوند.

دو ساده‌سازی دیگر باعث می‌شوند **TSIV** به راحتی قابل استفاده باشد. نخست، همان طور که پیش از این اشاره شد، وقتی ابزارها از مجموعه کامل متغیرهای موهومی متقابل تشکیل شوند، معادله تکانه دوم، ۱-۳-۴، به مدلی برای ابزار سنتی ساده می‌شود. به ویژه، مینیماند **2SLS** برای مسئله دو نمونه می‌شود:

$$\hat{J}_N(\hat{g}) = \sum_j \omega_j \left(\bar{y}_{1j} - \hat{g}'\bar{W}_{2j}\right)^2, \qquad (4.3.2)$$

که $\bar{y}_{1j}$ میانگین متغیر وابسته در مقدار کوواریانس و ابزار **j** در یک نمونه و $\bar{W}_{2j}$ میانگین متغیرهای درونی و کوواریانس در مقدار ابزاری **j** در نمونه دوم و ω_j وزن مناسب است. این مقادیر برای برآورد حداقل مربعات موزون معادله **VIV** هستند، بجز اینکه متغیرهای وابسته و مستقل از نمونه مشابهی به دست نمی‌آیند. مجدداً آنگریست تصاویری ارائه کرد. ارزش‌های بهینه برای **TSIV** مؤثر مجانبی با واریانس $\bar{y}_{1j} - \hat{g}'\bar{W}_{2j}$ به دست می‌آیند. این واریانس تحت تأثیر این واقعیت است که تکانه‌ها از نمونه‌های مختلف خطاهای استاندارد **TSIV** هستند که به راحتی در مورد ابزار موهومی محاسبه می‌شوند، زیرا برآوردگر معادل حداقل مربعات موزون هستند.

دوم، آنگریست و کروگر (۱۹۹۵) یک برآوردگر نوع **TSIV** جذاب از نظر محاسباتی ارائه کردند که به ضرب ماتریسی نیاز نداشت و می‌توانست با نرم‌افزار رگرسیون معمولی انجام شود. این برآوردگر **IV** نمونه ـ شکاف[2] (**SSIV**) نامیده می‌شود که به صورت زیر توضیح داده می‌شود. برآوردهای مرحله اول

1. Social Security Administration
2. Split-sample IV (SSIV)

در مجموعه داده‌های دو، با $(Z_2'Z_2)^{-1}Z_2'W_2$ به دست می‌آیند. این مقادیر برازش شده به مجموعه داده‌های یک با تشکیل مقادیر برازش شده بین‌نمونه‌ای[1] نسبت داده می‌شوند $\hat{W}_{12} \equiv Z_1(Z_2'Z_2)^{-1}Z_2'W_2$. مرحله دوم **SSIV** رگرسیون Y_1 روی $\hat{W}_{12}$ است. توزیع محدودکننده درست برای این برآوردگر توسط اینوآ و سولون به دست آمد که نشان دادند توزیع محدودکننده آنگریست و کروگر به این فرض نیاز دارند که $Z_1'Z_1 = Z_2'Z_2$ (که در صورتی درست است که توزیع حاشیه‌ای نهادها و کوواریانس‌ها در نمونه‌های تکرار شده ثابت باشند). با این همه باید گفت که توزیعات محدود **SSIV** و **2SLS** وقتی مشابه هستند که ضریب در متغیر درونی صفر باشد. خطاهای استاندارد برای این نمونه خاص به سادگی تشکیل می‌شوند و احتمالاً تقریب خوب قابل قبولی برای یک مورد کلی به ما می‌دهند.

۴-۴. IV یا نتایج بالقوه نامتجانس[2]

تا اینجا می‌توان گفت که بحث در مورد متغیرهای ابزاری، اثر قابل قبول ثابت مرتبط با علیت دارد. در مورد متغیر موهومی مانند وضعیت کهنه‌سربازان این به معنا $Y_{1i} - Y_{0i} = \rho$ برای تمام **i** هاست، در حالی که یک تیمار دارای چند مقدار، مانند تحصیلات، این میانگین برای تمام **s** ها و تمام **i**ها معادل $Y_{si} - Y_{s-1,i} = \rho$ می‌شود. هر دو این موارد، نمای آشکارشده‌ای[3] از جهان هستند، به ویژه در مورد مقادیر چندگانه که خطی‌سازی مشابه تجانس دارد. برای تمرکز روی این موضوع در مدل آثار عدم تجانس، ما با متغیر علّی صفر و یکی شروع می‌کنیم. در این زمینه، ما به بیان دیگر عدم تجانس اثر تیمار،[4] یعنی توزیع اثرات علّی در افراد را در نظر می‌گیریم.

چرا عدم تجانس اثر تیمار مهم است؟ پاسخ در تمایز بین دو نوع اعتباری وجود دارد که طرح تحقیقاتی را مشخص می‌کنند. اعتبار داخلی[5] به این مسئله مربوط می‌شود که آیا طرح داده شده با موفقیت اثرات علّی را برای جامعه بررسی شده روشن می‌کند یا خیر. اعتبار خارجی[6] مقدار پیش‌بینی یافته‌های تحقیق، برای یک زمینه متفاوت است. برای مثال، اگر جامعه مطالعه در آزمون تصادفی از تیمار بهره ببرد، برآوردهای به دست آمده اعتبار خارجی کمتری دارند. به همین ترتیب، برآوردهای قرعه‌کشی انتخاب مشمولان در مورد اثرات خدمت اجباری در دوره جنگ ویتنام، سنجه خوبی از

---

1. Cross-sample Fitted Value
2. Heterogeneous Potential Outcomes
3. Stylized Views
4. Treatment-effect Heterogeneity
5. Internal Validity
6. External Validity

پیامدهای خدمت نظامی داوطلبانه نیست. یک چارچوب اقتصادسنجی با آثار تیمار نامتجانس به ما کمک می‌کند تا هم اعتبار خارجی و هم اعتبار داخلی برآوردهای متغیرهای ابزاری را بسنجیم.

۱-۴-۴. آثار موضعی میانگین تیمار[1]

در چارچوب **IV**، موتوری که باعث استنباط علّی می‌شود، ابزار **Zi** است، ولی هنوز متغیر مورد نظر **Di** است. این ویژگی **IV** ما را به سمت پذیرش مفهوم خروجی‌های بالقوه جمع‌بندی شده و شاخص‌بندی شده بر حسب ابزارها و وضعیت تیمار هدایت می‌کند. $Y_i(d,z)$ نتیجه بالقوه یک **i** را نشان می‌دهد که این فرد را به وضعیت تیمار $D_i = d$ می‌رساند و مقدار ابزار $Z_i = z$ می‌شود. برای مثال، این به ما می‌گوید که درآمدهای **i** ترکیبات جایگزین وضعیت کهنه‌سربازان و وضعیت شایستگی انتخاب را به ما نشان می‌دهند. اثر علّی وضعیت کهنه‌سربازان با وضعیت مفروض محقق شده **i** در فراخوان، $Y_i(1,Z_i)-Y_i(0,Z_i)$ است، در حالی که اثر علّی وضعیت شایستگی انتخاب با وضعیت کهنه‌سربازان **i** $Y_i(D_i,1)-Y_i(D_i,0)$ است.

ما می‌توانیم متغیرهای ابزاری را برای آغاز زنجیره علّی در نظر بگیریم که در آن ابزار Z_iروی متغیر مورد نظر D_i تأثیر می‌گذارد که در برگشت روی خروجی‌های Y_iتأثیر می‌گذارد. برای اینکه دقیق‌تر بگوییم، باید به این ایده توجه داشته باشیم که ابزار اثر علّی روی D_iدارد. D_iوضعیت تیماری است، وقتی $Z_i = 1$، در حالی که D_i وضعیت تیمار وقتی است که $Z_i = 0$باشد. بنابراین، وضعیت تیمار مشاهده شده می‌شود:

$$D_i = D_{0i} + (D_{1i} - D_{0i})Z_i = \pi_0 + \pi_{1i}Z_i + \xi_i. \tag{4.4.1}$$

در ضرایب تصادفی، $\pi_0 \equiv E[D_{0i}]$ و $\pi_{1i} \equiv (D_{1i}-D_{0i})$ هستند، بنابراین π_{1i} اثر علّی نامتجانس ابزار روی D_i است. با نتایج بالقوه، تنها یکی از مقادیر تعیین شده D_i و D_{0i} برای هر شخص مشاهده می‌شوند. در مثال قرعه‌کشی انتخاب سرباز D_{0i} به ما می‌گوید که اگر رقم درآمده در قرعه‌کشی بالا باشد، آیا **i** در ارتش به خدمت گرفته می‌شود یا خیر، در حالی که D_{0i} به ما می‌گوید که اگر رقم درآمده در قرعه‌کشی پایین باشد، آیا **i** به خدمت گرفته می‌شود یا خیر. ما یکی از این ارقام تعیین شده یا دیگری را بنا به Z_i می‌بینیم. میانگین تأثیر علّی Z_iروی D_i، $E[\pi_{1i}]$ است.

فرض اول در چارچوب نامتجانس این است که ابزار به اندازه مقدار تعیین شده تصادفی خوب است: این مقدار از بردار نتایج بالقوه و مقادیر بالقوه اختصاص داده شده تیمار مستقل است. این مقدار را می‌توان به صورت زیر نوشت:

$$[\{Y_i(d,z); \forall\, d,z\}, D_{1i}, D_{0i}] \amalg Z_i, \tag{4.4.2}$$

---

1. Local Average Treatment Effects

مستقل بودن برای تفسیر علّی شکل کاهش یافته[1] کافی است، یعنی رگرسیون Y_iدر Z_i.

به ویژه،

$$E[Y_i|Z_i=1]-E[Y_i|Z_i=0] = E[Y_i(D_{1i},1)|Z_i=1]-E[Y_i(D_{0i},0)|Z_i=0]$$
$$= E[Y_i(D_{1i},1)-Y_i(D_{0i},0)],$$

اثر علّی ابزار روی Y_i. همچنین عدم وابستگی بدین معناست که:

$$E[D_i|Z_i=1]-E[D_i|Z_i=0] = E[D_{1i}|Z_i=1]-E[D_{0i}|Z_i=0]$$
$$= E[D_{1i}-D_{0i}],$$

به بیان دیگر، اولین مرحله از بحث قبلی ما درباره **2SLS** اثر علّی Z_iروی D_i را نشان می‌دهد.

فرض کلیدی دوم در چارچوب خروجی‌های نامتجانس این استنباط است که $Y_i(d,z)$ تنها تابعی از **d** است. وقتی صلاحیت انتخاب شدن روی وضعیت کهنه‌سربازان تأثیر می‌گذارد، درآمدهای بالقوه فردی به عنوان یک کهنه‌سرباز بر حسب وضعیت صلاحیت انتخاب شدن بدون تغییر می‌مانند؛ در حالی که درآمدهای بالقوه به عنوان غیر کهنه‌سرباز بدون تأثیر می‌مانند. در کل، این ادعا که ابزار از طریق کانال علّی معلومی عمل می‌کند، محدودیت نامیده می‌شود. در مدل خطی با اثرات ثابت، محدودیت به وسیله حذف ابزار از معادله علّی مورد نظر بیان می‌شود، یا معادل آن در معادله ۴-۱-۴ $E[Z_i\eta_i]=0$ است. باید یادآور شویم که عبارت خطای به کار رفته برای مدل‌های معادلات لحظه‌ای خود را در اختیار تمایز بین استقلال و عدم شمول قرار نمی‌دهند. ما نیاز داریم Z_i و η_i در این معادله ناهمبسته بمانند، ولی استدلالی که پشت این فرض قرار دارد روشن نیست، تا هم استقلال و هم محدودیت‌ها را در نظر بگیریم.

اگر مردان با ارقام قرعه‌کشی انتخاب سرباز به شیوه‌ای غیر از احتمال افزایش یافته خدمات تحت تأثیر قرار بگیرند، محدودیت عدم شمول[2] برای ابزارهای قرعه‌کشی انتخاب سرباز ناموفق می‌شود. برای مثال، آنگریست و کروگر (۱۹۹۲) به دنبال رابطه بین ارقام قرعه‌کشی انتخاب سرباز و تحصیل بودند. ایده آنها این بود که تأخیرات انتخابی تحصیلی منجر به این می‌شوند که مردان با امتیازات قرعه‌کشی پایین طولانی‌تر از میزانی که مایل بودند، در کالج بمانند. اگر این طور باشد، حداقل به دو دلیل ارقام قرعه‌کشی انتخاب سرباز با درآمدها همبستگی دارند: احتمال افزایش یافته خدمات نظامی و احتمال افزایش یافته خدمات در کالج. این حقیقت که رقم قرعه‌کشی به صورت تصادفی تعیین می‌شود (و بنابراین فرض استقلال را برآورده می‌کند) این احتمال را پایین نمی‌آورد. محدودیت عدم شمول از این ادعا جداست که ابزار به صورت تصادفی تعیین می‌شود. این ادعا بیشتر درباره کانال منحصربه‌فرد اثرات علّی ابزار است.

با استفاده از محدودیت عدم شمول، می‌توانیم نتایج بالقوه تعیین شده بر حسب وضعیت تیماری را

---

1. Reduced Form
2. Exclusion Restriction

با استفاده از یک شاخص (Y_{1i}, Y_{0i}) توصیف کنیم. به ویژه:

$$\begin{aligned} Y_{1i} &\equiv Y_i(1,1) = Y_i(1,0); \\ Y_{0i} &\equiv Y_i(0,1) = Y_i(0,0). \end{aligned} \tag{4.4.3}$$

بنابراین نتیجه مشاهده شده Y_i را می‌توان بر حسب نتایج بالقوه به صورت زیر نوشت:

$$\begin{aligned} Y_i &= Y_i(0,Z_i) + [Y_i(1,Z_i) - Y_i(0,Z_i)]D_i \\ &= Y_{0i} + (Y_{1i} - Y_{0i})D_i. \end{aligned} \tag{4.4.4}$$

نمادسازی ضرایب تصادفی می‌شود:

$$Y_i = \alpha_0 + \rho_i D_i + \eta_i,$$

نسخه خلاصه ۴-۴-۴ با $\alpha_0 \equiv E[Y_{0i}]$ و $\rho_i \equiv Y_{1i} - Y_{0i}$ است.

فرض نهایی برای مدل‌های **IV** نامتجانس این است که یا $\pi_{1i} \geq 0$ و یا $\pi_{1i} \leq 0$ برای تمام **i**ها. این فرض یکنوایی[1] ارائه شده توسط آمبنز و آنگریست به معنای این است که وقتی ابزار هیچ اثری روی برخی از افراد نداشته باشد، تمام آنهایی که تحت تأثیر قرار می‌گیرند، به شیوه مشابهی تحت تأثیر هستند. به بیان دیگر، برای تمام **i**ها یا $D_{1i} \geq D_{0i}$ یا $D_{1i} \leq D_{0i}$. در ادامه، یکنوایی را با $D_{1i} \geq D_{0i}$ در نظر می‌گیریم. در مثال قرعه‌کشی انتخاب سرباز، این بدین معناست که هرچند صلاحیت انتخاب شدن هیچ اثری روی احتمال خدمات نظامی برای برخی از مردان ندارد، هیچ کسی در واقع با صلاحیت انتخاب شدن از ارتش بیرون نمی‌ماند. بدون یکنوایی، تضمین نمی‌شود برآوردگرهای متغیرهای ابزاری، میانگین اثرات علّی تکی اصلی موزون را برآورد کنند $Y_{1i} - Y_{0i}$.

با توجه به محدودیت عدم شمول، عدم وابستگی ابزارها و نتایج بالقوه، وجود مرحله اول و یکنوایی، برآورد والد می‌تواند به صورت اثر وضعیت کهنه‌سربازان روی شرایط تیمار تغییر یافته توسط ابزار، توصیف شود. این پارامتر اثر موضعی میانگین تیمار (**LATE**[2]) نامیده می‌شود؛ (ایمبنز و آنگریست). در اینجا شرح قبلی آورده می‌شود.

نظریه ۱-۴-۴ LATE

فرض کنید:

(**A1**، عدم وابستگی) $\{Y_i(D_{1i},1), Y_{0i}(D_{0i},0), D_{1i}, D_{0i}\} \amalg Z_i$

(**A2**، عدم شمول) $Y_i(d,0) = Y_i(d,1) \equiv Y_{di}$ $for\ d = 0,1;$

(**A3**، مرحله اول) $E[D_{1i} - D_{0i}] \neq 0$

1. Monotonicity
2. Local Average Treatment Effect

(**A4**، یکنوایی) $D_{1i} - D_{0i} \geq 0 \forall i$ یا برعکس.

سپس:

$$\frac{E[Y_i|Z_i = 1] - E[Y_i|Z_i = 0]}{E[D_i|Z_i = 1] - E[D_i|Z_i = 0]} = E[Y_{1i} - Y_{0i}|D_{1i} > D_{0i}] = E[\rho_i|\pi_{1i} > 0]$$

اثبات. از محدودیت عدم شمول برای نوشتن $E[Y_i|Z_i = 1] = E[Y_{0i} + (Y_{1i} - Y_{0i})D_i|Z_i = 1]$ استفاده کنید که مساوی با $E[Y_{0i} + (Y_{1i} - Y_{0i})D_{1i}]$ با عدم وابستگی است. در این صورت، $E[Y_i|Z_i = 0] = E[Y_{0i} + (Y_{1i} - Y_{0i})D_{0i}]$، بنابراین صورت برآوردگر والد $E[(Y_{1i} - Y_{0i})(D_{1i} - D_{0i})]$ می‌شود. یکنوایی به معنای این است که $D_{1i} - D_{0i}$ مساوی با یک یا صفر است، بنابراین:

$$E[(Y_{1i} - Y_{0i})(D_{1i} - D_{0i})] = E[Y_{1i} - Y_{0i}|D_{1i} > D_{0i}]P[D_{1i} > D_{0i}].$$

بحث مشابهی نشان می‌دهد.

$$E[D_i|Z_i = 1] - E[D_i|Z_i = 0] = E[D_{1i} - D_{0i}] = P[D_{1i} > D_{0i}].$$

این نظریه می‌گوید که ابزاری که به اندازه تعیین شده به صورت تصادفی خوب است و از طریق کانال معلوم روی خروجی تأثیر می‌گذارد، یک مرحله اول دارد و روی کانال علّی مورد نظر تنها در یک جهت تأثیر می‌گذارد و می‌تواند برای برآورد میانگین اثر علّی روی گروه تأثیر گرفته به کار برده شود. بنابراین، برآوردهای **IV** از اثرات خدمات نظامی با استفاده از قرعه‌کشی انتخاب سرباز، تأثیر خدمات نظامی را روی مردانی که به خدمت گرفته می‌شوند، برآورد می‌کند، زیرا آنها صلاحیت انتخاب شدن را دارند وگرنه به خدمت گرفته نمی‌شدند. روشن است که داوطلبان و مردانی استخراج می‌شوند که به دلایلی از خدمات نظامی معاف شده باشند، ولی این شامل مردانی می‌شود که خط سیاست انتخاب شدن آنها اجباری باشد.

LATE چقدر کاربرد دارد؟ هیچ پاسخی برای این سؤال وجود ندارد، ولی همواره بحث درباره آن ارزشمند است. بخشی از توجه به اثرات خدمت سربازی در ویتنام حول این سؤال است که آیا کهنه‌سربازان (به ویژه خدمت‌کنندگان اجباری) به شکل مناسبی به خاطر خدماتشان مورد قدردانی قرار می‌گیرند یا نه. قرعه‌کشی انتخاب معتبر پاسخ به این سؤال را برآورد می‌کند. برآوردهای قرعه‌کشی انتخاب سرباز اثرات خدمت اجباری در ویتنام همچنین با بحث‌هایی درباره آینده سیاست خدمت اجباری مرتبط است. از سوی دیگر، ابزارهای قرعه‌کشی انتخاب سرباز برآوردهای معتبری از اثرات علّی خدمت اجباری در ویتنام، اعتبار خارجی را ارائه می‌دهند ـ یعنی ارزش پیش‌بینی این برآوردها برای خدمت نظامی در مکان‌ها و زمان‌های دیگر ـ مستقیم با این چارچوب **IV** بیان نمی‌شود. هیچ چیزی در فرمول **IV** وجود ندارد که بتواند توضیح دهد چرا خدمات ویتنام روی درآمدها تأثیر می‌گذارند.

ممکن است تعجب کنید که چرا ما برای نظریه **LATE** به یکنوایی نیاز داریم، فرضی که هیچ نقشی در چارچوب معادلات همزمان سنتی با اثرات ثابت ندارد. خطای یکنوایی به معنای این است که این ابزار اشخاص را به سمت تیمار هل می‌دهد در حالی که بقیه را به بیرون هل می‌دهد. آنگریست، ایمبنز و روبین گروه آخری را *مخالفت‌کننده*[۱] نامیدند. مخالفت‌کنندگان رابطه بین **LATE** و شکل کاهش یافته را پیچیده می‌کنند. برای مشاهده علت آن، به مرحله اثبات نظریه **LATE** برگردید که شکل کاهش یافته را به صورت زیر نشان می‌دهد:

$$E[Y_i|Z_i=1]-E[Y_i|Z_i=0]=E[(Y_{1i}-Y_{0i})(D_{1i}-D_{0i})].$$

بدون یکنوایی، این مساوی است با:

$$E[Y_{1i}-Y_{0i}|D_{1i}>D_{0i}]P[D_{1i}>D_{0i}]-E[Y_{1i}-Y_{0i}|D_{1i}<D_{0i}]P[D_{1i}<D_{0i}].$$

بنابراین ممکن است صحنه‌ای داشته باشیم که در آن اثرات تیمار برای هرکسی مثبت هستند، هرچند شکل کاهش‌یافته، صفر می‌شود، زیرا اثرات روی موافقت‌کننده‌ها با اثرات روی مخالفت‌کننده‌ها رد می‌شوند. این موضوع به شکل مدل اثرات ثابت[۲] درنمی‌آید، زیرا شکل کاهش‌یافته همواره اثر ثابت، ضربدر مرحله اول است، علی‌رغم اینکه آیا مرحله اول شامل رفتار مخالفتی می‌شود.

شناخت عمیق‌تر **LATE** می‌تواند با ارتباط دادن آن به برخی زحمات انجام شده در اقتصادسنجی معاصر، مدل شاخص پنهانی برای متغیرهای شایستگی موهومی مانند تعیین تیمار به دست آورد. این مدل‌ها انتخاب‌های فردی را به صورت تعیین شده از طریق مقایسه فواید و هزینه‌های تا حدودی مشاهده شده و تاحدودی نامعلوم (پنهان)[۳] توصیف می‌کنند. معمولاً، این *غیرقابل مشاهده‌ها* با خروجی‌هایی رابطه دارند، که در آن متغیر تیمار درون‌زاست (هرچند در مفهوم معادلات همزمان، واقعاً درون‌زا نیستند). برای مثال (متغیرهای کمکی نادیده گرفته شده)، می‌توانیم وضعیت کهنه‌سربازان را به صورت زیر مدل‌سازی کنیم.

$$D_i=\begin{cases}1 & \text{if } \gamma_0+\gamma_1 Z_i>v_i\\ 0 & \text{در غیر اینصورت}\end{cases},$$

که v_i عامل تصادفی از جمله هزینه‌های مشاهده نشده و مزایای خدمت نظامی مستقل از Z_i است.

این مدل شاخص پنهانی[۴] سهم‌بندی‌های تیماری بالقوه را به صورت زیر نشان می‌دهد:

$$D_{0i}=1[\gamma_0>v_i] \text{ and } D_{1i}=1[\gamma_0+\gamma_1>v_i].$$

یادآور می‌شویم که در این مدل، یکنوایی به صورت خودکار برآورده می‌شود، زیرا γ_1 ثابت است.

با فرض $\gamma_1>0$ می‌توان، **LATE** را به صورت زیر نوشت:

1. Defiers
2. Constant-effects Model
3. Latent
4. Latent-index Model

$$E[\mathrm{Y}_{1i}-\mathrm{Y}_{0i}|\mathrm{D}_{1i}>\mathrm{D}_{0i}]=E[\mathrm{Y}_{1i}-\mathrm{Y}_{0i}|\gamma_0+\gamma_1>v_i>\gamma_0],$$

که تابع پارامترهای مرحله اولیه پنهانی، γ_0 و γ_1 است، همانند توزیعات پیوندی $\mathrm{Y}_{1i}-\mathrm{Y}_{0i}$ و v_i. در کل، این مشابه اثر تیمار میانگین جامعه $E[\mathrm{Y}_{1i}-\mathrm{Y}_{0i}]$ یا اثر روی تیمار $E[\mathrm{Y}_{1i}-\mathrm{Y}_{0i}|\mathrm{D}_i=1]$ شده نیست. ما توزیع بین اثرات علّی میانگین مختلف را در بخش ۲-۴-۴ استخراج می‌کنیم.

۲-۴-۴. جامعه فرعی موافق[1]

چارچوب **LATE** هر جامعه را با ابزاری در مجموعه زیرگروه‌های وابسته به ابزار تقسیم‌بندی می‌کند که به صورتی توصیف می‌شود که در آن اعضای جامعه به ابزار واکنش نشان می‌دهند:

تعریف ۱-۴-۴. موافقت‌کننده‌ها،[2] جامعه فرعی با $\mathrm{D}_{1i}=1$ و $\mathrm{D}_{0i}=0$.

همواره گیرنده.[3] جامعه فرعی با $\mathrm{D}_{1i}=\mathrm{D}_{0i}=1$.

هرگز گیرنده.[4] جامعه فرعی با $\mathrm{D}_{1i}=\mathrm{D}_{0i}=0$.

LATE اثر تیمار روی جامعه موافقت‌کنندگان (پذیرندگان) است. عبارت *موافقت‌کننده (پذیرنده)*[5] از قیاس با آزمون‌های تصادفی به دست می‌آید که در آن برخی اشخاص تجربی با پروتکل مداخله تعیین شده به صورت تصادفی همخوانی دارند و برخی ندارند، در حالی که برخی از اشخاص کنترل به درمان تجربی دسترسی پیدا می‌کنند، هرچند در معرض آن نبوده‌اند. آنهایی که دارو نگرفتند، وقتی به صورت تصادفی برای انجام این کار تعیین شدند، هرگز گیرنده بودند، در حالی که آنهایی که دارو گرفتند، حتی در گروه کنترل، همواره گیرنده بودند. بدون افزودن فرض دیگری، **LATE** درباره اثرات هرگز گیرنده و همواره گیرنده، اطلاعاتی اضافه نمی‌کند، زیرا طبق تعریف، وضعیت درمان برای این دو گروه، با ابزار، بدون تغییر باقی می‌ماند. قیاس بین **IV** و آزمون تصادفی با سازگاری جزئی بیشتر تمثیلی است ـ **IV** مسئله استنباط علّی را در آزمون تصادفی با سازگاری جزئی حل می‌کند. این نکته مهم یک بخش فرعی گسسته در زیر به وجود می‌آورد.

پیش از آنکه به این مورد خاص و مهم برگردیم، چند نکته کلی را ارائه می‌دهیم. نخست، میانگین اثر علّی روی موافقت‌کننده‌ها همواره مشابه میانگین اثر درمان روی تیمارشده‌ها نیست. از این حقیقت ساده که $\mathrm{D}_i=\mathrm{D}_{0i}+(\mathrm{D}_{1i}-\mathrm{D}_{0i})\mathrm{Z}_i$، یاد می‌گیریم که جامعه تیمار شده از دو گروه بدون هم‌پوشانی تشکیل می‌شوند. با یکنوایی، نمی‌توانیم هم $\mathrm{D}_{0i}=1$ و هم $\mathrm{D}_{1i}-\mathrm{D}_{0i}=1$ را داشته باشیم، زیرا $\mathrm{D}_{0i}=1$ نشان‌دهنده $\mathrm{D}_{1i}=1$ است. بنابراین تیمار $\mathrm{D}_{0i}=1$ یا $\mathrm{D}_{1i}-\mathrm{D}_{0i}=1$ و $\mathrm{Z}_i=1$ را دارد و بنابراین، D_i را می‌توان به صورت مجموع دو متغیر موهومی انحصاری D_{i0} و $(\mathrm{D}_{1i}-\mathrm{D}_{0i})\mathrm{Z}_i$ در نظر گرفت.

1. The Compliant Subpopulation
2. Compliers
3. Always-takers
4. Never-takers
5. Compliers

تیمارشده‌ها از همواره گیرنده یا موافقت‌کننده‌ها با ابزار متصل تشکیل می‌شوند. از آنجایی که ابزار به صورت تصادفی تعیین می‌شود، موافقت‌کننده‌ها با ابزار متصل نماینده تمام موافقت‌کننده‌ها هستند. از اینجا، داریم:

$$\underbrace{E\left[Y_{1i}-Y_{0i}|D_i=1\right]}_{\text{تاثیر روی تیمار شده}} \tag{4.4.5}$$
$$= E[Y_{1i}-Y_{0i}|D_{0i}=1]P[D_{0i}=1|D_i=1]$$
$$+E\left[Y_{1i}-Y_{0i}|D_{1i}>D_{0i},Z_i=1\right]P\left[D_{1i}>D_{0i},Z_i=1|D_i=1\right]$$
$$= \underbrace{E[Y_{1i}-Y_{0i}|D_{0i}=1]}_{\text{تاثیر روی همواره گیرندگان}}P[D_{0i}=1|D_i=1]$$
$$+\underbrace{E\left[Y_{1i}-Y_{0i}|D_{1i}>D_{0i}\right]}_{\text{تاثیر روی موافقان}}P\left[D_{1i}>D_{0i},Z_i=1|D_i=1\right]$$

زیرا $P[D_{0i}=1|D_i=1]$ و $P[D_{1i}>D_{0i},Z_i=1|D_i=1]$ به یک اضافه می‌شوند و این بدین معناست که اثر درمان روی تیمار شده‌ها میانگین اثرات را روی موافقت‌کننده‌ها و همیشه گیرنده‌ها دارد.

به این ترتیب، **LATE** میانگین اثر علّی تیمار (در اینجا درمان با دارو) روی تیمارنشدگان (درمان‌نشدگان) نیست، $E[Y_{1i}-Y_{0i}|D_i=0]$. در مثال قرعه‌کشی انتخاب سرباز، میانگین اثر روی درمان‌نشدگان میانگین اثر علّی خدمت نظامی روی جامعه غیر کهنه‌سرباز از همراهان در ویتنام است. میانگین اثر درمان روی درمان‌نشدگان میانگین موزون اثرات روی پذیرندگان و هرگز درمان‌نشدگان است. به ویژه:

$$\underbrace{E\left[Y_{1i}-Y_{0i}|D_i=0\right]}_{\text{تاثیر روی تیمارنشدگان}} \tag{4.4.6}$$
$$= \underbrace{E\left[Y_{1i}-Y_{0i}|D_{1i}=0\right]}_{\text{تاثیر روی هرگز گیرندگان}}P\left[D_{1i}=0|D_i=0\right]$$
$$+\underbrace{E\left[Y_{1i}-Y_{0i}|D_{1i}>D_{0i}\right]}_{\text{تاثیر روی موافقان}}P\left[D_{1i}>D_{0i},Z_i=0|D_i=0\right],$$

که ما از این حقیقت استفاده می‌کنیم که یکنوایی با $D_{1i}=0$ باید هرگز گیرندگان باشد.

در آخر، میانگین‌گیری (۵-۴-۴) و (۶-۴-۴) با استفاده از،

$$E[Y_{1i}-Y_{0i}]=E[Y_{1i}-Y_{0i}|D_i=1]P[D_i=1]+E[Y_{1i}-Y_{0i}|D_i=0]P[D_i=0]$$

اثر درمان میانگین جامعه کلی را با میانگین موزون روی پذیرندگان، همواره گیرندگان و هرگز گیرندگان نشان می‌دهد. البته، این نتیجه‌ای است که مستقیم با یکنوایی و توصیف در ابتدای این بخش به آن می‌رسیم.

از آنجایی که یک متغیر ابزاری مستقیم اطلاعاتی درباره اثرات روی همواره گیرنده‌ها و هرگز گیرنده‌ها ندارد، ابزارها همواره میانگین اثر علّی را به دست نمی‌آورند. همواره استثناهای مهمی برای این

قانون وجود دارند: متغیرهای ابزاری که به هیچ همواره گیرنده یا هیچ هرگز گیرنده‌ای اجازه نمی‌دهند. هرچند این سناریو معمول نیست، یک مورد خاص مهم است. یک مثال، ابزار باروری دوقلو[1] است که توسط رزنزویگ و ولپین (۱۹۸۰)، برونارس و گروگر (۱۹۹۴)، آنگریست و ایوانس (۱۹۹۸) و انگریست، لوی و شلوسر (۲۰۰۶) به کار برده شد. مورد دیگر اورئوپولوس (۲۰۰۶) است که اخیراً با استفاده از تغییر قوانین خدمت سربازی به عنوان ابزاری برای تحصیل در بریتانیا به کار برده می‌شود.

برای اینکه ببینیم این مورد خاص چگونه عمل می‌کند T_i را متغیر موهومی می‌گیریم که تاریخ‌های تولد را نشان می‌دهد. آنگریست و ایوانس (۱۹۹۸) از این ابزار برای برآورد اثر علّی داشتن سه فرزند روی درآمد جامعه زنان با حداقل دو کودک استفاده کرد. کودک سوم به این دلیل مورد توجه است که باروری کاهش یافته زنان آمریکایی در دهه ۱۹۶۰ و ۱۹۷۰ به معنای تبدیل سه کودک به دو کودک است. وجود چندین تولد دوم، تغییر شبه‌تجربی در این حاشیه ایجاد می‌کند. فرض کنید که Y_{0i} نشان‌دهنده درآمد بالقوه یک زن موقعی باشد که فقط دو فرزند داشته باشد، در حالی که Y_{1i} نشان‌دهنده درآمد او برای موقعی که سه فرزند داشته باشد، رویدادی که توسط D_i نشان داده می‌شود. فرض می‌شود T_i به صورت تصادفی تعیین می‌شود، یعنی باروری با حداکثر یک کودک نسبت به چندتولدی افزایش پیدا می‌کند و این چندتولدی تنها با افزایش باروری روی خروجی تأثیر می‌گذارد.

LATE از ابزار دوقلوها استفاده کرد $E[Y_{1i}-Y_{0i}|D_i = 0]$ که به معنای اثر میانگین علّی روی زنانی است که تنها دو فرزند دارند. این بدین دلیل است که زنانی که چندقلو دارند، سه فرزند دارند، یعنی هیچ هرگز گیرنده‌ای نسبت به ابزار دوقلوها وجود ندارد.

اورئوپولوس (۲۰۰۶) نیز از **IV** برای برآورد میانگین اثر علّی تیمار روی تیمارنشدگان استفاده کرد. تحقیقات او بازده اقتصادی تحصیل را با استفاده از افزایش سن خدمت در بریتانیا از ۱۴ به ۱۵ سال استفاده کرد. طبق قانون خدمات جدید بریتانیا، بسیاری از نوجوانان در سن ۱۴ سالگی از مدرسه بیرون می‌آیند. اثر علّی مورد نظر در این مورد پاداش درآمدی برای یک سال دیگر مدرسه رفتن است. به پایان بردن این سال می‌تواند به عنوان یک تیمار در نظر گرفته شود. از آنجایی که همه در نمونه اورئوپولوس زمانی سال اضافی را به پایان می‌رسانند که قوانین مدرسه سخت‌تر شد، او میانگین اثر علّی به دست آوردن یک سال بیشتر در دبیرستان را روی تمام کسانی که در ۱۴ سالگی مدرسه را ترک می‌کنند، مدنظر قرار داد. این بدین حقیقت برمی‌گردد که نوجوانان بریتانیا به قانون متعهد هستند ـ راهبرد **IV** اثر تیمار روی تیمارنشده‌ها را مدارسی مدنظر قرار نمی‌دهد که در آن نوجوانان به هنگام حضور در مدرسه اجباری آزادی عملی بیشتری دارند. اقتصاددانان از تغییرات حضور اجباری به عنوان ابزاری برای کار با **LATE** استفاده می‌کنند.

---

1. Twins Instrument

۳-۴-۴. IV در آزمون‌های تصادفی

زبان چارچوب **LATE** بر مبنای مقایسه میان **IV** و آزمون‌های تصادفی است. ولی برخی از ابزارها واقعاً از آزمون‌های تصادفی به دست می‌آیند. اگر ابزار یک پیشنهاد تیمار تصادفی باشد، **LATE** اثر تیمار روی کسانی است که با پیشنهاد سازگاری دارند، ولی تیمار نمی‌شوند. یک مورد مهم وقتی است که ابزار با آزمون تصادفی بدون سازگاری جانبی جمع‌بندی می‌شود. در بسیاری از آزمون‌های تصادفی شرکت‌کنندگان داوطلب دریافت تیمار هستند. از سوی دیگر، هیچ‌کس در گروه کنترل دسترسی به مداخلات تجربی ندارد. از آنجایی که گروهی تیمار تعیین شده را دریافت می‌کند که زیرمجموعه انتخاب شده تیمار پیشنهادی است، بین گروه کنترل و گروه واقعاً تیمارشده مقایسه‌ای انجام می‌شود که اشتباه است. اریب انتخاب در این مورد تقریباً همواره مثبت است: آنهایی که در آزمون تصادفی دارو دریافت می‌کنند گرایش به سالم بودن دارند و آنهایی که از مزیت مداخلات اقتصادی تصادفی مانند برنامه‌های آموزشی بهره می‌برند، درآمد بیشتری دارند.

IV از تیمار تعیین شده تصادفی به عنوان متغیر ابزاری برای تیمار دریافتی استفاده می‌کند که برخی از مشکلات پذیرش را برطرف می‌کند. به علاوه، **LATE** اثر تیمار روی تیمارشده‌ها در این مورد است. فرض کنید ابزار Z_i یک متغیر موهومی است که تخصیص تصادفی به گروه تیمار را تعیین می‌کند در حالی که D_i یک متغیر موهومی است که تعیین می‌کند آیا تیمار واقعاً دریافت شده است یا خیر. در عمل، به دلیل عدم پذیرش، D_i مساوی با Z_i نیست. مثالی از برآورد تصادفی برنامه آموزشی **JTPA**[1] است که در آن تنها ۶۰ درصد از اشخاص اختصاص داده شده به گروه دریافت‌کننده آموزش، آموزش را دریافت کردند، در حالی که ۲ درصد از اشخاص اختصاص داده شده به گروه کنترل، آموزش را دریافت کرده بودند. ناسازگاری مشاهده شده در **JTPA** به دلیل عدم علاقه کافی در شرکت‌کنندگان و عدم توانایی اپراتورهای برنامه برای تشویق مشارکت‌کنندگان به وجود آمده بود. از آنجایی که مسئله پذیرش در این مورد با گروه تیمار سازگاری دارد، **LATE** از تخصیص تصادفی استفاده می‌کند، Z_i ابزاری برای تیمار دریافتی و D_i اثر تیمار روی تیمار شده‌هاست.

این کاربرد **IV** برای حل مسائل پذیرش در جدول ۱-۴-۴ نشان داده شده است که در آن نتایج آزمون **JTPA** ارائه می‌شوند. متغیر خروجی اصلی در آزمون **JTPA** کل درآمد در سی ماه پس از تخصیص تصادفی است. ستون‌های ۱-۲ از جدول اختلاف در درآمد را بین کسانی که آموزش دیدند و کسانی که آموزش ندیدند نشان می‌دهد (برآوردهای ستون دو از مدل رگرسیون هستند که برای تعدادی از مشخصات فردی اندازه‌گیری شده در ابتدای آزمون تعیین می‌شوند. تباین گزارش شده در ستون ۱-۲ از درجه ۴۰۰۰ دلار برای مردان و ۲۲۰۰ دلار برای زنان است که هر دو تحت تأثیر شدید تیمار

---

۱. برنامه آموزشي مطابق قانون Joint Training Partnership Act

هستند که حدود ۲۰ درصد میانگین درآمد است. ولی این برآوردها گمراه‌کننده هستند، زیرا افراد را مطابق D_i مقایسه می‌کنند که واقعاً تیمار را دریافت کردند. از آنجایی که افراد اختصاص داده شده به گروه تیمار آزاد بودند که تیمار را رد کنند (و ۴۰ درصد همین کار را کردند)، این مقایسه تخصیص تصادفی را کنار می‌گذارد مگر اینکه تصمیم برای پذیرش تیمار خود مستقل از خروجی‌های بالقوه باشد که به نظر بعید می‌رسد.

جدول ۱-۴-۴. نتایج آزمون JTPA: برآوردهای OLS و IV از تأثیرات آموزشی

	Comparisons by Training Status		Comparisons by Assignment Status		Instrumental Variable Estimates	
	Without Covariates (1)	With Covariates (2)	Without Covariates (3)	With Covariates (4)	Without Covariates (5)	With Covariates (6)
A. Men	3,970	3,754	1,117	970	1,825	1,593
	(555)	(536)	(569)	(546)	(928)	(895)
B. Women	2,133	2,215	1,243	1,139	1,942	1,780
	(345)	(334)	(359)	(341)	(560)	(532)

نکته: جدول **OLS**، شکل کاهش‌یافته و برآوردهای **IV** از اثر آموزش یارانه‌ای را روی درآمدهای تجربه **JTPA** گزارش می‌دهد. ستون‌های (۱) و (۲) اختلافات در درآمدها را با وضعیت آموزش نشان می‌دهند. ستون‌های (۳) و (۴) اختلافات بر حسب وضعیت تخصیص تصادفی را نشان می‌دهند. ستون‌های (۵) و (۶) تحصیلات بالا یا **GED**، سیاه، اسپانیایی، متأهل، با کار کمتر از ۱۳ ساعت در سال، **AFDC** (برای زنان)، به اضافه شاخص‌هایی برای راهبرد خدمات توصیه شده، گروه سنی و بررسی پیگیری هستند. خطاهای استاندارد قطعی در پرانتز نشان داده می‌شوند.

ستون‌های ۳ و ۴ جدول ۱-۴-۴ افراد را بر حسب اینکه به آنها پیشنهاد به مداخله درمانی شده مقایسه می‌کند. به عبارت دیگر، این مقایسه بر مبنای عبارت z_i، که به شکل تصادفی اختصاص داده شده، می‌باشد. در زبان آزمایش‌های بالینی، تقابل موجود در ستون‌های ۴-۳ به عنوان اثر میل به درمان[1] (**ITT**) شناخته شده است. آثار میل به درمان در جدول از مرتبه ۱۲۰۰ دلار هستند (تا حدی کمتر با کمی متغیر). از آنجایی که z_i به شکل تصادفی انتخاب شده، اثر **ITT** دارای تفسیری بر مبنای علت است: آنها به ما درباره اثر علّی پیشنهاد به درمان، بر پایه این حقیقت که بسیاری از این پیشنهادها رد می‌شوند، می‌گویند. به همین دلیل، اثر **ITT** در مقایسه با میانگین اثر علّی درباره کسانی که تحت درمان قرار گرفته‌اند بسیار کوچک است. ستون‌های ۵ و ۶ تکه‌ها را کنار هم گذاشته و جالب‌ترین اثر را به ما نشان می‌دهند: میل به درمان تقسیم بر اختلاف در نرخ‌های انطباق بین گروه‌های کنترل و درمان که به شکل ابتدایی انتخاب شده است (در حدود ۶). این اعداد، تقریباً ۱۸۰۰ دلار، اثر درمان روی افراد تحت درمان را تخمین می‌زند.

چطور می‌دانیم **ITT** صورت تقسیم شده جورسازی اثر درمان روی فرد تحت درمان است؟ ما **ITT**

1. Intention-to-treat

را به صورت اثر تقلیل‌یافته پیشنهاد تصادفی برای درمان، که ابزار ما در این مورد است می‌شناسیم. نرخ انطباق اولین مرحله مرتبط با این ابزار است و برآورد شیب والد، مثل همیشه، فرم کاهش‌یافته تقسیم شده با مرحله اول است. به طور کلی این موضوع معادل با **LATE** است، اما چون (تقریباً) دریافت‌کنندگان همیشگی نداریم، جامعه مورد مداخله (تقریباً) شامل تمام پذیرندگان است. بنابراین تخمین‌های **IV** در ستون ۵ و ۶ جدول ۱-۴-۴ تخمین‌های ثابت اثر مداخله روی افراد تحت درمان است.

نتیجه تا حدی مهم است که استنتاجی جایگزین را تضمین می‌کند. تا جایی که اطلاع داریم، اولین شخصی که نشان داد می‌توان از فرمول **IV** جهت تخمین اثر درمان روی افراد تحت درمان در آزمایش تصادفی با عدم انطباق یک‌طرفه بهره جست، هوارد بلوم (۱۹۸۴) بود. در اینجا نتیجه بلوم با اثبات ساده مستقیم را می‌آوریم.

قضیه ۲-۴-۴. نتیجه بلوم.[1] تصور کنید فرضیات قضیه **LATE** برقرار باشد و $E[D_i|Z_i=0]=0$، بنابراین،

$$\frac{E[Y_i|Z_i=1]-E[Y_i|Z_i=0]}{E[D_i|Z_i=1]}=E[Y_{1i}-Y_{0i}|D_i=1].$$

اثبات. $E[Y_i|Z_i=1]=E[Y_{i0}+(Y_{1i}-Y_{0i})D_i|Z_i=1],$ در حالی که $E[Y_i|Z_i=0]=E[Y_{i0}|Z_i=0]$ به دلیل اینکه $E[D_i|Z_i=0]=0$، بنابراین به واسطه استقلال داریم:

$$E[Y_i|Z_i=1]-E[Y_i|Z_i=0]=E[(Y_{1i}-Y_{0i})D_i|Z_i=1]$$

اما،

$$E[(Y_{1i}-Y_{0i})D_i|Z_i=1]=E[Y_{1i}-Y_{0i}|D_i=1,Z_i=1]P[D_i=1|Z_i=1]$$

در حالی که $E[D_i|Z_i=0]=0$ بدان معناست که $D_i=1$ دلالت دارد $Z_i=1$. پس

$$E[Y_{1i}-Y_{0i}|D_i=1,Z_i=1]=E[Y_{1i}-Y_{0i}|D_i=1]$$

چارچوب **LATE** علاوه بر اینکه به ما می‌گوید چگونه آزمایش‌های تصادفی را با عدم انطباق تحلیل نماییم، مسیری به روی آزمایش‌های تصادفی هوشمندانه در تنظیماتی که غیرممکن یا غیراخلاقی است وادار به انطباق درمانی شوند ایجاد می‌نماید. مثالی مشهور در حوزه جرم‌شناسی آزمایش خشونت خانگی مینیاپلیس[2] (**MDVE**) است. آزمایش **MDVE** اقدامی پیشگامانه در جهت تعیین بهترین پاسخ پلیس به خشونت خانگی بود (شرمن و برک، ۱۹۸۴). به طور کلی، پلیس بعد از دریافت تماسی درباره خشونت خانگی استراتژی‌های مختلفی اتخاذ می‌کند. این استراتژی‌ها شامل ارجاع به مشاور، دستورات جدایی و بازداشت است. مناظره‌ای سخت حول این سؤال که آیا پاسخی قاطع، مثل بازداشت

1. The Bloom Result
2. Minneapolis Domestic Violence Experiment

یا حداقل حبس موقت، سودبخش است در جریان می‌باشد، خصوصاً با توجه به اینکه اتهامات مربوط به خشونت‌های خانگی اغلب پس گرفته می‌شوند.

در نتیجه این مناظره، شهر مینیاپلیس مجوز برگزاری آزمایشی تصادفی را داد که در آن واکنش پلیس به مزاحمت خانگی به شکل تصادفی انتخاب می‌شد. در طراحی تحقیق از برگه‌های اتهام کدگذاری شده با رنگ که به شکل تصادفی با هم مخلوط شده بودند استفاده شده بود که به بعضی از افسران دستور بازداشت مجرم داده شده بود در حالی که دیگر افسران باید اعضای درگیر را از هم جدا کرده یا آنها را نزد مشاور ارجاع می‌دادند. با این حال، در عمل پلیس این امکان را داشت که مأموریت محوله را لغو نماید. برای مثال، مجرم خطرناک یا مست در هر صورت بازداشت می‌شد. در نتیجه، واکنش واقعی اغلب از واکنش محوله تصادفی متفاوت بود، با این حال به میزان زیادی با هم متناسب هستند.

بیشتر تحلیل‌های انتشاریافته از اطلاعات **MDVE** مسئله انطباق را شناخته و روی اثرات **ITT** متمرکز است، یعنی تحلیلی که از مأموریت تصادفی ابتدایی و نه درمانی که در واقع اجرا شده استفاده کرده است، اما همچنان می‌توان از اطلاعات **MDVE** برای دریافت اثر علّی متوسط روی پذیرندگان استفاده کرد، در این حالت شامل افرادی که بازداشت شدند چون به شکل تصادفی تعیین شده بود، اما در غیر این صورت بازداشت نمی‌شدند. تحلیل **MDVE** به این صورت در آنگریست (۲۰۰۶) آورده شده است. چون تمامی افرادی که در **MDVE** تعیین شده بود باید بازداشت شوند در واقع بازداشت شدند، عدم دریافت‌کنندگان وجود نداشتند. این موضوع ویژگی جالب و روی دیگر سناریوی بلوم است: در اینجا برای همه $D_{I1} = 1$ است. در نتیجه، **LATE** اثر درمان روی افرادی است که مورد درمان قرار نگرفته‌اند، یعنی:

$$E[Y_{1i} - Y_{0i}|D_{1i} > D_{0i}] = E[Y_{1i} - Y_{0i}|D_i = 0],$$

که در آن D_i نشان‌دهنده بازداشت است. تخمین **IV** با استفاده از اطلاعات **MDVE** نشان می‌دهد بازداشت دفعات تکرار تهاجم را به شکل چشمگیری کاهش می‌دهد، در این مورد، میان زیرگروهی که دستگیر نشده بودند.[1]

۴-۴-۴. شمارش و توصیفِ پذیرندگان (اجابت‌کنندگان)[2]

ما مشاهده کردیم، بجز در موارد خاص، هر متغیر ابزاری پارامتر علّی یکتایی را شناسایی می‌کند، که

۱. کاربردی دیگر از IV در اطلاعات به دست آمده از آزمایشی تصادفی مربوط به کروگر (۱۹۹۹) است. این مطالعه از ابعاد کلاس انتخاب شده به شکل تصادفی به عنوان ابزاری جهت ابعاد واقعی کلاس با اطلاعات به دست آمده از آزمایش STAR تنسی استفاده می‌کند. برای دانش‌آموزان کلاس اول و بالاتر، ابعاد واقعی کلاس با ابعاد کلاس انتخاب شده به شکل تصادفی در آزمایش STAR متفاوت است، چون معلمان و والدین در طول سال‌های بعد از شروع آزمایش دانش‌آموزان را به مکان‌های مختلف منتقل می‌نمایند. همچنین کروگر ۱۹۹۹ نشان داد 2SLS برای مدلی با شدت درمان متغیر به کار رفته است، همان طور که در بخش ۳-۵-۴ بحث شد.

2. Compliers

برای زیرگروه پذیرندگان آن ابزار به خصوص است. بنابراین ابزارهای معتبر گوناگون حداقل در اساس، برای رابطه علّی مسائل مختلفی را تخمین می‌زنند (مورد استثنایی مهم ابزارهایی هستند که اجازه انطباق عالی در یک سمت یا سمتی دیگر را می‌دهند). با وجود اینکه تخمین‌های متفاوت **IV** توسط **2SLS** به خوبی «بررسی»[1] شده‌اند تا اثر علّی متوسط یکتایی ایجاد کنند، دسته‌ای از آزمایش‌های شناختی که در بخش ۲-۲-۴ درباره آن بحث شد، که در آن ابزارهای متعددی طبق اینکه آیا نتایج مشابهی را تخمین می‌زنند یا خیر اعتبارسنجی شده‌اند، در جهانی ناهمگون محو گشته‌اند.

تفاوت‌ها در زیرجامعه‌های سازگار ممکن است توصیف‌کننده گوناگونی در آثار تیمار از ابزاری به ابزار دیگر باشد. بنابراین ما خواهان آنیم تا جایی که امکان دارد درباره ابزارهای مختلف برای پذیرندگان بدانیم. علاوه بر این، اگر زیرجامعه سازگار مشابه سایر جامعه‌های مورد بررسی باشد، حالت مورد استفاده جهت برون‌زایی آثار علّی تخمینی برای این جامعه‌ها قوی‌تر است. در این صورت، آسمگلو و آنگریست (۲۰۰۰) استدلال می‌کنند ابزار یک‌چهارم تولد و قوانین اجباری ایالتی سن حضور در مدرسه (حداقل تحصیل مورد نیاز قبل از ترک مدرسه در ایالت شما زمانی که ۱۴ ساله هستید) روی گروه مشابهی از مردم و به دلایل یکسان اثرگذار است. بنابراین انتظار داریم تخمین‌های **IV** از تعداد ادامه‌دهندگان به تحصیل از این دو گروه ابزار یکسان باشد. همین طور انتظار داریم تخمین یک‌چهارم تولد اثر پیشنهادهای معاصر جهت تقویت قوانین اجباری سن حضور در مدرسه را پیش‌بینی نماید.

از سویی دیگر، اگر زیرجامعه‌های سازگار مرتبط با دو یا بیشتر از دو ابزار خیلی متفاوت باشند، همچنان تخمین‌های **IV** که ایجاد می‌کنند مشابه است و ممکن است آماده آن باشیم تا آثار همگن را به عنوان فرضیه‌ای کارآمد به کار ببریم. این مسئله احیاکننده ایده شناخت بیش از اندازه است، اما آن را در شرایطی قرار می‌دهد که نیاز به اعتبارسنجی اضافی دارد.[2] این استدلال به وسیله مطالعه آثار اندازه خانواده روی آموزش بچه‌ها توسط آنگریست، لوی و اسکلوسر (۲۰۰۶) شرح داده شده است. این ایده در مطالعه آنگریست، لوی و اسکولسر از آنجا به وجود آمد که مشاهده شد کودکانی با خانواده‌های بزرگ‌تر نسبت به خانواده‌های کوچک‌تر غالباً از آموزش کمتری برخوردارند. نگرانی بلندمدت در تحقیق درباره باروری این است که رابطه منفی مشاهده شده بین خانواده‌های بزرگ‌تر و نتیجه‌های بدتر علّی است. آن طور که معلوم است، تخمین **IV** اثر اندازه خانواده با استفاده از ابزارهای مختلف، هرکدام با زیرجامعه‌های سازگار کاملاً متفاوت، همه نتایجی مبنی بر عدم تأثیر اندازه خانواده نشان می‌دهند. استدلال آنگریست، لوی و اسکولسر (۲۰۰۶) این است که نتایج آنها به سمت درمان معمول صفر برای تقریباً همه در جامعه اسرائیلی تحت مطالعه نشانه رفته است.

تا به اینجا مشاهده کردیم تعداد گروه پذیرندگان به راحتی قابل اندازه‌گیری است. این تنها اولین

---

1. Weighted-up

۲. در واقع، با حفظ این فرضیه که تمام ابزارهای مدل شناخت بیش از اندازه معتبر هستند، آماره آزمون مرسوم شناسایی بیش از اندازه، به آزمایشی رسمی برای سنجش ناهمگونی اثر تیمار تبدیل می‌شود.

مرحله والد است، نظر به اینکه، به شکل یکنواختی داده شده است، خواهیم داشت:

$$\begin{aligned} P[D_{1i}>D_{0i}] &= E[D_{1i}-D_{0i}] \\ &= E[D_{1i}]-E[D_{0i}] \\ &= E[D_i|Z_i=1]-E[D_i|Z_i=0]. \end{aligned}$$

همچنین می‌توانیم بگوییم چه نسبتی از افراد تحت تیمار از پذیرندگان هستند چون، برای شرکت‌کنندگان، وضعیت تیمار تماماً توسط Z_i تعیین شده است. با شروع از تعریف احتمال مشروط:

$$\begin{aligned} P[D_{1i} > D_{0i}|D_i=1] &= \frac{P[D_i=1|D_{1i}>D_{0i}]P[D_{1i}>D_{0i}]}{P[D_i=1]} \\ &= \frac{P[Z_i=1](E[D_i|Z_i=1]-E[D_i|Z_i=0])}{P[D_i=1]}. \end{aligned} \tag{4.4.7}$$

تساوی دوم از این واقعیت بهره می‌برد که $P[D_i=1|D_{1i}>D_{0i}] = P[Z_i=1|D_{1i}>D_{0i}]$ و اینکه $P[Z_i=1|D_{1i}>D_{0i}] = P[Z_i=1]$ به وسیله استقلال. به عبارت دیگر، نسبت افرادی که مورد تیمار قرار گرفته و از پذیرندگان مرحله اول هستند، ضربدر احتمال اینکه ابزار به کار رفته باشد، تقسیم بر نسبت افرادی که تحت تیمار قرار گرفته‌اند، داده شده است.

در اینجا رابطه ۷-۴-۴ با محاسبه نسبتی از کهنه‌سربازان که جزء پذیرندگان در قرعه‌کشی فراخوانی به خدمت هستند نشان داده شده است. محتویات در جدول ۲-۴-۴ گزارش شده است. به عنوان مثال، برای مردان سفیدی که در سال ۱۹۵۰ به دنیا آمدند، اولین مرحله ۰/۱۵۹، احتمال واجد شرایط بودن برای فراخوانی به خدمت ۱۹۵/۳۶۶ و احتمال حاشیه درمان ۰/۲۶۷ است. از این آمار، می‌توان محاسبه کرد زیرجامعه سازگار ۰/۳۲ جامعه کهنه‌سربازان در این گروه است. نسبت کهنه‌سربازانی که جزء پذیرندگان قرعه‌کشی فراخوانی به خدمت هستند برای مردان غیرسفیدی که در سال ۱۹۵۰ متولد شدند به ۲۰ درصد کاهش یافت. چون اولین مرحله قرعه‌کشی فراخوانی به خدمت برای غیرسفیدها به مقدار قابل توجهی ضعیف‌تر بود چندان جای تعجب ندارد. آخرین ستون از جدول نسبت غیر کهنه‌سربازهایی را گزارش می‌کند که اگر واجد شرایط فراخوانی به خدمت بودند خدمت می‌کردند. این مقدار از ۳ درصد برای غیر سفیدپوستان تا ۱۰ برای سفیدپوستان متغیر است و منعکس‌کننده این حقیقت می‌باشد که بیشتر غیر کهنه‌سربازان برای خدمت سربازی دارای تأخیر، غیر مشمول یا فاقد صلاحیت بودند.

جدول ۲-۴-۴. احتمال انطباق در مطالعات متغیرهای ابزاری

Source (1)	Endogenous Variable (D) (2)	Instrument (Z) (3)	Sample (4)	$P[D=1]$ (5)	1st Stage, $P[D_1 > D_0]$ (6)	$P[Z=1]$ (7)	$P[D_1 > D_0 \mid D=1]$ (8)	$P[D_1 > D_0 \mid D=0]$ (9)
Angrist (1990)	Veteran Status	Draft eligibility	White men born in 1950	0.267	0.159	0.534	0.318	0.101
			Non-white men born in 1950	0.163	0.060	0.534	0.197	0.033
Angrist and Evans (1998)	More than 2 children	Twins at second birth	Married women aged 21-35 with two or more children in 1980	0.381	0.603	0.008	0.013	0.966
		First two children are of the same sex	Married women aged 21-35 with two or more children in 1980	0.381	0.060	0.506	0.080	0.048
Angrist and Krueger (1991)	High school graduate	Third or fourth quarter birth	Men born between 1930 and 1939	0.770	0.016	0.509	0.011	0.034
Acemoglu and Angrist (2000)	High school graduate	State requires 11 or more years of school attendance	White men aged 40-49	0.617	0.037	0.300	0.018	0.068

نکته در مورد جدول ۲-۴-۴: جدول تحلیلی از اندازه مطلق و نسبی جامعه پذیرندگان برای تعدادی از متغیرهای ابزاری را نشان می‌دهد. اولین مرحله، گزارش شده در ستون ۶، اندازه مطلق گروه پذیرنده را نشان می‌دهد. ستون‌های ۸ و ۹ اندازه جامعه پذیرندگان نسبت به افرادی که تحت تیمار قرار گرفته‌اند و نگرفته‌اند را نشان می‌دهد.

اثر خدمت سربازی/*اجباری*، پارامتری اولویت‌دار در مطالعه آنگریست (۱۹۹۰) است، بنابراین این حقیقت که پذیرندگان واجد شرایط فراخوانی به خدمت جزء اقلیت کهنه‌سربازان هستند در واقع محدودیتی در این مطالعه به حساب نمی‌آید. حتی در زمان جنگ ویتنام، بسیاری از سربازان داوطلب بودند، حقیقتی که کمتر در مورد کهنه‌سربازان جنگ ویتنام مورد تقدیر قرار گرفته است. تفسیر **LATE** تخمین **IV** با استفاده از قرعه‌کشی فراخوانی به خدمت این حقیقت را نمایان می‌سازد که سایر استراتژی‌های شناسایی نیاز به تخمین آثار خدمت سربازی روی داوطلبان دارد (بعضی از اینها در آنگریست ۱۹۹۸ به کار رفته است).

سطرهای باقی پسماند در جدول ۲-۴-۴ اندازه زیرجامعه سازگار برای دوقلوها و ابزارهای ترکیب جنسیتی ـ خویشاوندی مورد استفاده توسط آنگریست و اوانز (۱۹۹۸) جهت تخمین آثار وضع حمل و ابزار یک‌چهارم تولد و قوانین اجباری سن حضور در مدرسه به کارگیری شده توسط آنگریست و کروگر (۱۹۹۱) و آسمگلو و آنگریست (۲۰۰۰) به منظور تخمین بازگشت به تحصیل را نشان می‌دهد. در هرکدام از این مطالعات، زیرجامعه سازگار کسر کوچکی از گروه مورد رفتار است. برای مثال، کمتر از ۲ درصد افرادی که از دبیرستان فارغ‌التحصیل شدند به دلیل قوانین اجباری سن حضور در مدرسه یا مزیت تولد در یک‌چهارم پایانی بود.

این سؤال که آیا زیرجامعه سازگار کوچک جای نگرانی دارد توسط زمینه بررسی مشخص می‌شود. در بعضی حالت‌ها منصفانه است بگوییم «شما به چیزی که نیاز داشتید دست یافتید». برای مثال با تدابیر اقداماتی بسیار، این گروهی حاشیه‌ای است که در اولویت می‌باشد، نکته‌ای که در مطالعه برجسته **IV** مک کللان (۱۹۹۴) درباره بررسی اثر جراحی روی بیماران با حمله قلبی بر آن تأکید گشت. مک کللان از فاصله نسبی به تأسیسات تیمار قلبی بهره می‌جوید تا ابزاری جهت بررسی اینکه آیا بیمار مسن با حمله قلبی نیاز به اقدامات جراحی دارد یا خیر ایجاد نماید. در هر حال بسیاری از بیماران تحت درمان یکسانی قرار می‌گیرند، اما برای بعضی، مورد جراحی اساسی حاشیه‌ای است. در چنین مواردی، بیماران یا تأمین‌کنندگان استراتژی ملایم‌تری را در صورت دور بودن تأسیسات جراحی اتخاذ می‌کنند. مک کللان دریافت روش‌های جراحی برای این گروه حاشیه‌ای سود کمی دارد. به طور مشابه، افزایش در سن حضور اجباری به ۱۸ سال به شکل واضح برای بسیاری از دانش‌آموزان مقطع دبیرستان در آمریکا بی‌ربط است، اما روی تعداد کمی که ممکن است ترک تحصیل کنند اثرگذار است. تخمین **IV** پیشنهاد می‌دهد بازده اقتصادی آموزش برای این گروه حاشیه‌ای قابل توجه است.

آخرین ستون جدول ۲-۴-۴ ویژگی مخصوص ابزارهای دوقلو که در انتهای بخش قبل به آن اشاره شد را شرح می‌دهد. به مانند گذشته، برای زنانی با دو بچه در نمونه‌ای با زنان حداقل دارای دو بچه $D_i = 0$ قرار دهید، در حالی که $D_i = 1$ زنانی که بیشتر از دو بچه دارند را نشان می‌دهد. از آنجایی که عدم دریافتی در پاسخ به اتفاق وضع حمل وجود ندارد، یعنی همه مادرانی که در دومین وضع حمل خود صاحب دوقلو می‌شوند (حداقل) سه بچه دارند، احتمال انطباق میان افرادی با $D_i = 0$ به شکل مجازی برابر با یک است (جدول ورودی برابر ۰/۹۷ را نشان می‌دهد). بنابراین در این حالت **LATE**

اثر روی افرادی است که مورد درمان قرار نگرفته‌اند.

برخلاف اندازه گروه پذیرندگان، به نظر می‌رسد جمع‌آوری اطلاعات مربوط به ویژگی‌های پذیرندگان کار دشواری باشد چون آنها را نمی‌توان به شکل منحصربه‌فردی شناسایی نمود. از آنجایی که ما نمی‌توانیم D_{i1} و D_{i0} را برای هر فرد ببینیم، نمی‌توانیم فقط آنها را با $D_{i1} > D_{i0}$ مرتب کنیم و سپس توزیع ویژگی‌هایی برای این گروه را حساب نماییم. با این حال، توصیف توزیع ویژگی‌های پذیرندگان آسان است. برای ساده‌سازی، در اینجا ما روی ویژگی‌هایی مثل نژاد یا مدرک تحصیلی، که به کمک متغیرهای موهومی قابل توصیف هستند تمرکز می‌کنیم. در این حالت، هر چیزی که نیاز داریم بدانیم را می‌توان از تغییرات در اولین مرحله میان گروه‌های متغیر کمی آموخت.

بگذارید x_{1i} مشخصه توزیع برنولی باشد، مثلاً متغیر موهومی بیانگر فارغ‌التحصیلان کالج. آیا اینکه پذیرندگان ترکیب جنسیتی نسبت به سایر زنان با دو بچه، از فارغ‌التحصیلان کالج باشند از احتمال بیشتر یا کمتری برخوردار است؟ این سؤال با محاسبه‌ای که در ادامه می‌آید پاسخ داده می‌شود:

$$\frac{P[x_{1i}=1|D_{1i}>D_{0i}]}{P[x_{1i}=1]}=\frac{P[D_{1i}>D_{0i}|x_{1i}=1]}{P[D_{1i}>D_{0i}]}=\frac{E[D_i|Z_i=1,x_{1i}=1]-E[D_i|Z_i=0,x_{1i}=1]}{E[D_i|Z_i=1]-E[D_i|Z_i=0]}. \qquad (4.4.8)$$

به عبارت دیگر، احتمال اینکه پذیرنده فارغ‌التحصیل کالج باشد توسط نسبت اولین مرحله برای فارغ‌التحصیلان کالج نسبت به کل مرحله اول داده شده است.[۱]

این محاسبه در جدول ۳-۴-۴ نشان داده شده است، که نسبت‌های ویژگی‌های پذیرندگان بر اساس سن در اولین وضع حمل، نژاد غیر سفیدپوست و تکمیل تحصیلات با استفاده از ابزار دوقلو و همجنس را گزارش می‌کند. این جدول بر اساس سرشماری استخراج شده ۱۹۸۰ از کار آنگریست و اوانز (۱۹۹۸) ایجاد گشته است. احتمال اینکه پذیرندگان دوقلو بیشتر از ۳۰ سال باشند که بیشتر از متوسط مادران این نمونه است بالاتر بوده و منعکس‌کننده این حقیقت می‌باشد که زنان جوان‌تر که وضع حمل‌های متعددی داشتند در هر حال صاحب تعداد بچه‌های بیشتری می‌شدند. همچنین پذیرندگان دوقلو نسبت به متوسط مادران از تحصیلات بالاتری برخوردارند، در حالی که پذیرندگان ترکیب جنسیتی سطح سواد پایین‌تری دارند. این موضوع به توضیح درباره تخمین‌های کوچک‌تر **2SLS** که توسط ابزارهای دوقلو ایجاد شده‌اند کمک می‌کند (در اینجا در جدول ۴-۱-۴ گزارش شده است)، چون آنگریست و اوانز (۱۹۹۸) نشان دادند که عرضه نیروی کار در نتیجه وضع حمل با میزان تحصیل مادران کاهش می‌یابد.

۱. روشی کلی برای ایجاد مقدار متوسط یا سایر ویژگی‌های توزیع متغیرهای کمی برای پذیرندگان از طرح کاپای وزنی ابادی (۲۰۰۳) بهره می‌جوید. برای مثال،

$$E[X_i|D_{1i}>D_{0i}]=\frac{E[\kappa_i X_i]}{E[\kappa_i]},$$

که در آن،

$$\kappa_i=1-\frac{D_i(1-Z_i)}{1-P(Z_i=1|X_i)}-\frac{(1-D_i)Z_i}{P(Z_i=1X_i)}.$$

این رابطه به این دلیل کارآمد است که تابع وزنی K_1، «پذیرندگان را می‌یابد»، به نحوی که در بخش ۲-۵-۴ توضیح داده شده است.

جدول ۳-۴-۴. نسبت‌های ویژگی‌های پذیرندگان برای ابزارهای دوقلوها و ترکیب جنسیتی

Variable	$E[x]$ (1)	Twins at second birth $E[x\|D_1 > D_0]$ (2)	$P[x\|D_1 > D_0]/P[X]$ (3)	First two children are same sex $E[x\|D_1 > D_0]$ (4)	$P[x\|D_1 > D_0]/P[X]$ (5)
Age 30 or older at first birth	0.00291	0.00404	1.39 (0.0201)	0.00233	0.995 (0.374)
Black or hispanic	0.125	0.103	0.822 (0.00421)	0.102	0.814 (0.0775)
High school graduate	0.822	0.861	1.048 (0.000772)	0.815	0.998 (0.0140)
College graduate	0.132	0.151	1.14 (0.00376)	0.0904	0.704 (0.0692)

نکته در مورد جدول ۳-۴-۴: جدول گزارشی از تحلیل ویژگی‌های پذیرندگان برای ابزارهای دوقلوها و ترکیب جنسیتی ارائه می‌دهد. نسبت‌های ستون‌های ۳ و ۵ احتمال نسبی اینکه پذیرندگان خصوصیاتی که در هر سطر آورده شده را داشته باشند نشان می‌دهد. اطلاعات همانند آنگریست و اوانز (۱۹۹۸) از نمونه سرشماری ۵ درصد ۱۹۸۰، شامل مادران متأهل بین ۳۵-۲۱ سال با حداقل دو فرزند استخراج شده است. اندازه نمونه برای تمام ستون‌ها برابر ۲۵۴۶۵۴ است.

۵-۴. تعمیم LATE

تئوری **LATE** از مدل علّی تقلیل یافته به اصول که در آن از یک ابزار موهومی جهت تخمین اثر رفتار موهومی بدون متغیرهای کمی استفاده شده است بهره می‌جوید. ما می‌توانیم این مسئله را به سه روش تعمیم دهیم: چند ابزار (برای مثال مجموعه‌ای موهومی از یک‌چهارم تولد)، مدل‌هایی با متغیرهای کمی (برای مثال کنترل برای سال تولد) و مدل‌هایی با شدت رفتار متغیر و پیوسته (برای مثال سال‌های آموزش). در هر سه این موارد، برآورد **IV** میانگین وزنی آثار علّی برای پذیرندگان با ابزار مشخص است. ابزار اقتصادسنجی **2SLS** باقی پسماند و تفسیر به صورت بنیادی مشابه نتیجه ابتدایی **LATE** باقی می‌ماند، اما با جزئیات و نکات بیشتر. **2SLS** با ابزارهای متعدد اثری علّی ایجاد کرده که از برآوردهای **IV** با استفاده از هر ابزار در یک زمان میانگین می‌گیرد: **2SLS** با متغیرهای کمی میانگینی از **LATE**های به خصوص با متغیر کمی؛ و **2SLS** با شدت رفتار متغیر یا پیوسته مشتق میانگین وزنی در امتداد طول تابع پاسخ علّی غیرخطی بالقوه ایجاد می‌نماید.

۱-۵-۴. LATE با ابزارهای متعدد[1]

استفاده از بسط ابزارهای متعدد راحت است. این حالت از اساس مشابه نتیجه‌ای است که در زمینه اطلاعات گروهی درباره آن بحث کردیم. جفتی از ابزارهای موهومی z_{1i} و z_{2i} را در نظر بگیرید. بدون از دست دادن عمومیت، فرض کنید این ابزارهای موهومی متقابلاً منحصربه‌فرد هستند (اگر این گونه نبود، پس می‌توانیم با مجموعه منحصربه‌فرد متقابل سه ابزار موهومی $z_{1i}(1-z_{2i}), z_{2i}(1-z_{1i})$ و $z_{1i}z_{2i}$ کار کنیم). از دو ابزار موهومی می‌توان برای ساخت برآوردگر والد استفاده کرد. دوباره، بدون از دست دادن عمومیت فرض کنید برای هر مرحله اول مثبت یکنوایی ارضا شده است (اگر این گونه نبود می‌توان به نحوی ابزار موهومی را دوباره کدنویسی کرد که درست دربیاید). بنابراین هر دو نسخه‌ای از -- - را تخمین می‌زنند، با این حال جامعه با $D_{i1} > D_{i0}$ از z_{1i} و z_{2i} متفاوت است.

به جای برآوردگرهای والد، می‌توانیم از z_{1i} و z_{2i} با هم در روند **2SLS** استفاده کنیم. از آنجایی که این دو ابزار موهومی و یک ثابت اطلاعات موجود در مجموعه ابزار را تمام می‌کنند، این روند **2SLS** مشابه تخمین اطلاعات گروهی با استفاده از روش‌های شرطی تعریف شده با استفاده از z_{1i} و z_{2i} می‌باشد (ابزارها ممکن است همبسته باشند یا نباشند). همان طور که در آنگریست (۱۹۹۱) آورده شد، برآوردگر اطلاعات گروهی حاصل شده ترکیبی خطی از برآوردگرهای مبنای والد است. به عبارت دیگر، ترکیبی خطی از **LATE**های با ابزار خاص که از یک ابزار در هر بار استفاده می‌کنند است (در واقع، ترکیب خطی بهینه‌ای در مدل آثار ثابت خطی هم‌واریانس مرسوم می‌باشد).

این استدلال کاملاً کامل نیست چون هنوز نشان نداده‌ایم که ترکیب خطی **LATE**های ایجاد شده توسط **2SLS** میانگین وزنی است (یعنی وزن‌ها غیرمنفی هستند و مجموع به یک می‌شوند).

---

1. LATE with Multiple Instruments

فرمول‌های وزنی مربوط در ایمبنس و آنگریست (۱۹۹۴) و آنگریست و ایمبنس (۱۹۹۵) آورده شده‌اند. فرمول‌های قدری نامرتب هستند، بنابراین در اینجا نسخه‌ای ساده بر مبنای مثال دو ابزاری می‌آوریم. مثال نشان می‌دهد **2SLS** با استفاده از z_{1i} و z_{2i} با هم میانگین وزنی تخمین‌های **IV** با استفاده از z_{1i} و z_{2i} در هر دفعه است. اجازه دهید:

$$\rho_j = \frac{Cov(Y_i, Z_{ji})}{Cov(D_i, Z_{ji})}; j = 1, 2$$

دو برآورد **IV** با استفاده از z_{1i} و z_{2i} را تعیین نماید.

مقادیر متناسب مرحله اول برای **2SLS** برابر --- هستند. در نتیجه مفهوم **IV** برای **2SLS**، برآورد **2SLS** برابر:

$$\begin{aligned}
\rho_{2SLS} &= \frac{Cov(Y_i, \hat{D}_i)}{Cov(D_i, \hat{D}_i)} = \frac{\pi_{11}Cov(Y_i, Z_{1i})}{Cov(D_i, \hat{D}_i)} + \frac{\pi_{12}Cov(Y_i, Z_{2i})}{Cov(D_i, \hat{D}_i)} \\
&= \left[\frac{\pi_{11}Cov(D_i, Z_{1i})}{Cov(D_i, \hat{D}_i)}\right]\left[\frac{Cov(Y_i, Z_{1i})}{Cov(D_i, Z_{1i})}\right] + \left[\frac{\pi_{21}Cov(D_i, Z_{2i})}{Cov(D_i, \hat{D}_i)}\right]\left[\frac{Cov(Y_i, Z_{2i})}{Cov(D_i, Z_{2i})}\right] \\
&= \psi\rho_1 + (1-\psi)\rho_2,
\end{aligned}$$

است که در آن:

$$\psi = \frac{\pi_{11}Cov(D_i, Z_{1i})}{\pi_{11}Cov(D_i, Z_{1i}) + \pi_{21}Cov(D_i, Z_{2i})}$$

عددی بین صفر و یک است که به قدرت نسبی هر ابزار در مرحله اول بستگی دارد. بنابراین، نشان دادیم **2SLS** میانگین وزنی آثار علّی برای زیرجامعه‌های سازگار با ابزار خاص است. برای مثال فرض کنید، z_{1i} تولد دوقلوها را مشخص می‌کند و z_{2i} خویشاوندان همجنس درون خانواده با دو بچه یا بیشتر را نشان می‌دهد، هر دو ابزار برای اندازه خانواده کاربرد دارند همان طور که در آنگریست و اوانز (۱۹۹۸) آمده است. تولد متعدد دوم احتمال داشتن فرزند سوم را تا حدود ۰/۶ افزایش می‌دهد در حالی که داشتن جفت همجنس خواهر یا برادر احتمال تولد سوم را تا ۰/۷ افزایش می‌دهد. زمانی که از این دو ابزار با هم استفاده نماییم، تخمین‌های **2SLS** حاصله میانگین وزنی تخمین‌های والد تولید شده توسط ابزارها در هر بار هستند.[1]

۲-۵-۴. متغیرهای کمّی در مدل آثار ناهمگن[2]

شاید برایتان عجیب باشد که متغیرهای کمی کجا رفته‌اند. به هر حال، متغیرهای کمی نقش پُررنگی در بحث‌های اولیه‌مان در حوزه رگرسیون و جورشدگی ایفا نمودند. با این همه تئوری **LATE** شامل

---

۱. با استفاده از ابزار دوقلوها به تنهایی، تخمین IV اثر فرزند سوم روی مشارکت نیروی کار زن ۰/۰۸۴- است (s.e=.017). تخمین همجنس مشابه برابر ۰/۱۳۸- می‌باشد (s.e.=.029). استفاده از هر دو ابزار تخمین 2SLS برابر ۰/۰۹۸ - (۰/۰۱۵) تولید می‌کند. وزن 2SLS در این حالت برابر ۰/۷۴ برای دوقلوها، ۰/۲۶ برای همجنس‌هاست، که به دلیل دوقلوهای قوی‌تر در مرحله اول می‌باشد.

2. Covariates in the Heterogeneous-effects Model

متغیرهای کمی نمی‌باشد. این مسئله از این حقیقت ریشه می‌گیرد که وقتی ما متغیرهای ابزاری را به صورت مدلی از (طبیعی یا ساخته دسته بشر) آزمایش تصادفی می‌بینیم، متغیرهای کمی در جایگاه کم‌اهمیت‌تری قرار می‌گیرند. اگر ابزار به شکل تصادفی انتخاب شود، احتمال دارد از متغیرهای کمی مستقل باشد. با این حال، تمام ابزارها دارای چنین خاصیتی نیستند. همانند متغیرهای کمی در مدل‌های رگرسیون بخش قبل، دلیل اصلی اینکه چرا متغیرهای کمی در تحلیل‌های علّی با استفاده از متغیرهای ابزاری قرار دارند این است که استقلال شرطی و محدودیت‌های حذف نهفته در تخمین **IV** ممکن است احتمال بیشتری برای اعتبارسنجی بعد از شرط‌گذاری روی متغیرهای کمی داشته باشند. حتی ابزارهایی که به شکل تصادفی انتخاب شده‌اند، مثل وضعیت واجد شرایط بودن برای فراخوانی خدمت، ممکن است تنها بعد از شرط‌گذاری روی متغیرهای کمی اعتبارسنجی شوند. در مورد واجد شرایط بودن برای فراخوانی خدمت، احتمال اینکه افراد با سن بالاتر واجد شرایط شوند بیشتر بود چون محدودیت سنی بالاتر بود. چون هنگام جذب نیرو در سال تولد (یا سن) اختلاف وجود دارد، وضعیت واجد شرایط بودن برای فراخوانی تنها زمانی ابزاری معتبر است که شرط‌گذاری روی سال تولد صورت پذیرد.

به شکل رسمی‌تر، تخمین **IV** با متغیرهای کمی را می‌توان به کمک فرض استقلال شرطی توجیه کرد.

$$\{Y_{1i}, Y_{0i}, D_{1i}, D_{0i}\} \amalg Z_i | X_i \qquad (4.5.1)$$

به عبارت دیگر، ما متغیرهای ابزاری را به صورت «به خوبی به عنوان تصادفی انتخاب شده»، شرط‌گذاری روی متغیرهای کمی، X_i مدنظر قرار می‌دهیم (در اینجا ما به شکل ضمنی محدودیت حذف را حفظ می‌نماییم). دومین دلیل جهت درآمیختن متغیرهای کمی این است که شرط‌گذاری روی متغیرهای کمی از تغییرات در متغیر وابسته می‌کاهد.

ساده‌ترین مدل علّی با متغیرهای کمی در مدل آثار ثابت، با محدودیت‌های فرم عملکردی به شکل زیر،

$$E[Y_{0i}|X_i] \;=\; X_i'\alpha^* \text{ for a } K \times 1 \text{ vector of coefficients, } \alpha^*;$$

$$Y_{1i} - Y_{0i} \;=\; \rho.$$

است. در ترکیب با ۱-۵-۴، این رابطه موجب به دست آوردن تخمین **2SLS** معادله‌ای شبیه به ۶-۱-۴ همان طور که در بخش ۱-۴ بحث شد خواهد شد.

تعمیم سرراست مدل آثار ثابت منتج به رابطه زیر:

$$Y_{1i} - Y_{0i} = \rho(X_i),$$

می‌شود که در آن $\rho(X_i)$ تابعی قطعی از X_i می‌باشد. این مدل را می‌توان با اضافه نمودن اثرهای متقابل بین z_i و X_i به مرحله اول و (همان) اثرهای متقابل بین D_i و X_i به مرحله دوم تخمین زد. اکنون متغیرهای درون‌زای متعدد و در نتیجه چندین معادله مرتبه اول وجود دارد. اینها را می‌توان به شکل:

$$\begin{aligned} D_i &= X_i'\pi_{00} + \pi_{01}z_i + z_iX_i'\pi_{02} + \xi_{0i} \\ D_iX_i &= X_i'\pi_{10} + \pi_{11}z_i + z_iX_i'\pi_{12} + \xi_{1i} \end{aligned}$$

نوشت. در این حالت معادله مرتبه دوم به صورت:

$$Y_i = \alpha'X_i + \rho_0 D_i + D_iX_i'\rho_1 + \eta_i,$$

است. پس $\rho(X_i) = \rho_0 + \rho_1'X_i$. از سوی دیگر، نسخه غیرپارامتری از $\boldsymbol{\rho}(\boldsymbol{X_i})$ را می‌توان به وسیله **2SLS** در زیرنمونه‌های طبقه‌بندی شده روی $\boldsymbol{X_i}$ تخمین زد.

همچنین مدل آثار ناهمگن که بر اساس تئوری **LATE** است امکان شناسایی بر پایه استقلال شرطی همانند ۱-۵-۴ را فراهم می‌سازد، با این حال برآورد مقداری پیچیده‌تر است. برای هر مقدار $\boldsymbol{X_i}$، ما **LATE** با متغیر کمی خاص تعریف می‌نماییم،

$$\lambda(X_i) \equiv E[Y_{1i} - Y_{0i} | D_{1i} > D_{0i}, X_i].$$

روش «اشباع و وزن»[1] جهت تخمین با متغیرهای کمی در قضیه‌ای که در ادامه به آن می‌پردازیم آمده است.

قضیه ۱-۵-۴ اشباع و وزن. تصور کنید فرض‌های قضیه **LATE** به شکل مشروط روی $\boldsymbol{X_i}$ صادق باشند. به این معنا که،

(**CA1**، استقلال) $\{Y_i(D_{1i},1), Y_{0i}(D_{0i},0), D_{1i}, D_{0i}\} \amalg z_i | X_i$

(**CA2**، عدم شمول) $P[Y_i(d,0) = Y_i(d,1)|X_i] = 1 \ for \ d = 0,1$

(**CA3**، اولین مرحله) $E[D_{1i} - D_{0i}|X_i] \neq 0$

همچنین فرض می‌کنیم یکنوایی (**A4**) به مانند قبل برقرار باشد. برآورد **2SLS** را بر پایه معادله مرحله اول در نظر بگیرید

$$D_i = \pi_X + \pi_{1X}z_i + \xi_{1i} \qquad (4.5.3)$$

و معادله مرحله دوم

$$Y_i = \alpha_X + \rho_c D_i + \eta_i$$

که در آن $\boldsymbol{\pi_x}$ و $\boldsymbol{\alpha_x}$ مدل‌های اشباع برای متغیرهای کمی را مشخص می‌کنند (مجموعه کاملی از متغیرهای موهومی برای تمامی مقادیر $\boldsymbol{X_i}$) و $\boldsymbol{\acute{\pi}_{1x}}$ اثر مجزای مرحله اول $\boldsymbol{z_i}$ را برای هر مقدار $\boldsymbol{X_i}$ تعیین می‌نماید. بنابراین $\rho_c = E[\omega(X_i)\lambda(X_i)]$ که در آن

$$\begin{aligned} \omega(X_i) &= \frac{V\{E[D_i|X_i,z_i]|X_i\}}{E[V\{E[D_i|X_i,z_i]|X_i\}]} \qquad (4.5.4) \\ &= \frac{E\{P[D_i=1|X_i,z_i](1-P[D_i=1|X_i,z_i])|X_i\}}{E[E[D_i|X_i,z_i](1-P[D_i=1|X_i,z_i])]}. \end{aligned}$$

1. Saturate and Weight

این قضیه بیان می‌دارد **2SLS** با مرحله اول کاملاً اشباع و مدل اشباع برای متغیرهای کمی در مرحله دوم میانگین وزنی **LATE**های با متغیر کمی خاص تولید می‌کند. وزن‌ها متناسب با واریانس مشروط میانگین مقدار متناسب جامعه مرحله اول $[E\ D_i|X_i, z_i]$، در هر مقدار از X_i هستند.[۱] قضیه بیان شده برگرفته از این حقیقت است زمانی که معادله ۳-۵-۴ اشباع شود مرحله اول منطبق بر $[E\ D_i|X_i, z_i]$ خواهد شد.

در عمل، ممکن است ما نخواهیم با مدلی با پارامتر مرحله اول برای هر مقدار از متغیرهای کمی کار کنیم. اول، همان طور که در انتهای این فصل بحث خواهیم کرد احتمال تبعیض وجود دارد و دوم، توده بزرگی از تخمین‌های مرحله اول یکتای غیردقیق چندان زیبا به نظر نمی‌رسد. به نظر منطقی می‌رسد تصور کنیم مدل‌هایی با پارامترهای کمتر، مثلاً مدل مرحله اول محدود که ثابت π_{1x} را تحمیل می‌کند، در هر حال شکلی از میانگین متغیر کمی **LATE** را تخمین می‌زند. به نظر می‌رسد این مسئله درست باشد، اما شناسه آن به شکل جالبی غیرمستقیم است. دیدگاه **2SLS** به عنوان فراهم‌آورنده تخمین خطای **MMSE** به رابطه علّی اساسی توسط ابادی[۲] (۲۰۰۳) توسعه داده شد.

راه‌حل ابادی با تعریف تابع هدف به صورت $[E\ Y_i|D_i, X_i, D_{1i} > D_{0i}]$ شروع می‌شود، **CEF** برای Y_i وضعیت رفتار و متغیرهای تصادفی مربوط به پذیرندگان را می‌دهد. ویژگی مهم **CEF** این است که وقتی شرایط قضیه **LATE** مشروط به X_i باقی می‌مانند، دارای مفهوم علّی است. به عبارت دیگر، برای پذیرندگان، تقابل‌های کنترل رفتار مشروط به X_i مساوی با **LATE**های مشروط به X_i است:

$$E\left[Y_i|D_i=1, X_i, D_{1i}>D_{0i}\right]-E\left[Y_i|D_i=0, X_i, D_{1i}>D_{0i}\right]$$
$$= \quad E\left[Y_{1i}-Y_{0i}|X_i, D_{1i}>D_{0i}\right]$$

این مسئله پیرو این حقیقت می‌باشد، در بخش ۱-۵-۴ آورده شده، نتایج بالقوه مستقل از D_i، X_i داده شده و $D_{1i} > D_{0i}$ است.[۳] نتیجه این است که می‌توان اجرای رگرسیونی از Y_i روی D_i و X_i در جامعه پذیرندگان تصور کرد. با اینکه ممکن است این رگرسیون **CEF** مورد نظر را به ما ندهد (مگر

---

۱. توجه داشته باشید که تغییرپذیری در $[E\ D_i|X_i, z_i]$ مشروط به X_i از z_i می‌آید. پس فرمول وزنی وزن بیشتری به مقادیر متغیر کمی می‌بخشد که در آن ابزار در مقادیر متناسب تغییرپذیری بیشتری ایجاد می‌کند. اولین خط از فرمول وزن ۴-۵-۴، برای هر متغیر درونی در تنظیم 2SLS برقرار است. دومین خط نتیجه این حقیقت است که در اینجا متغیر درونی ساختگی است.

2. Abadie

۳. برای پذیرندگان،

$$P\left[D_i=1|\{Y_{1i}, Y_{0i}\}, X_i, D_{1i}>D_{0i}\right]$$
$$= \quad P\left[Z_i=1|\{Y_{1i}, Y_{0i}\}, X_i, D_{1i}>D_{0i}\right].$$

و با استقلال مشروط

$$P\left[Z_i=1|\{Y_{1i}, Y_{0i}\}, X_i, D_{1i}>D_{0i}\right]$$
$$= \quad P\left[Z_i=1|X_i, D_{1i}>D_{0i}\right].$$

اینکه خطی باشد یا مدل اشباع باشد)، مثل همیشه تخمین **MMSE** آن را فراهم می‌سازد. بنابراین رگرسیون $\boldsymbol{Y_i}$ روی $\boldsymbol{D_i}$ و $\boldsymbol{X_i}$ در جامعه پذیرندگان $[\boldsymbol{E\ Y_i|D_i, X_i, D_{1i} > D_{0i}}]$ را تخمین می‌زند دقیقاً همان طور که **OLS**، $[\boldsymbol{E\ Y_i|D_i, X_i}]$ را تخمین می‌زند. متأسفانه نمی‌دانیم پذیرندگان چه کسانی هستند، بنابراین نمی‌توانیم از آنها نمونه‌برداری کنیم.

قضیه ۲-۵-۴ کاپا ابادی.[1] تصور کنید فرضیات قضیه **LATE** روی متغیرهای کمی $\boldsymbol{X_i}$ مشروط باقی بماند. بگذارید $\mathbf{g}(\boldsymbol{Y_i, D_i, X_i})$ هر تابع قابل اندازه‌گیری از $(\boldsymbol{Y_i, D_i, X_i})$ با امید محدود باشد.

$$\kappa_i = 1 - \frac{D_i(1 - Z_i)}{1 - P(Z_i = 1|X_i)} - \frac{(1 - D_i)Z_i}{P(Z_i = 1|X_i)}.$$

بنابراین:

$$E[g(Y_i, D_i, X_i)|D_{1i} > D_{0i}] = \frac{E[\kappa_i g(Y_i, D_i, X_i)]}{E[\kappa_i]}.$$

این قضیه را می‌توان با محاسبه مستقیم و استفاده از این حقیقت که، با داشتن فرض‌های قضیه **LATE**، هر امید میانگین وزنی روش‌ها برای دریافت‌کنندگان همیشگی، ردکنندگان همیشگی و پذیرندگان می‌باشد، ثابت کرد. به کمک یکنوایی، افرادی با $\boldsymbol{D_i(1 - z_i) = 1}$ جزء دریافت‌کنندگان همیشگی هستند چون $\boldsymbol{D_{0i} = 1}$ دارند، در حالی که افرادی با $\mathbf{z_i(1 - D_i) = 1}$ جزء ردکنندگان همیشگی هستند چون $\boldsymbol{D_{1i} = 1}$ است. پس، پذیرندگان گروه باقی پسماند می‌باشند.

قضیه ابادی حاوی مفاهیم مهمی است؛ برای مثال دوباره در بحث‌های مربوط به آثار رفتار چندک پدیدار می‌گردد. در اینجا ما از آن برای تخمین $[\boldsymbol{E\ Y_i|D_i, X_i, D_{1i} > D_{0i}}]$ به وسیله رگرسیون خطی استفاده می‌نماییم. به طور مشخص، بگذارید $\boldsymbol{\alpha_\alpha}$ و $\boldsymbol{\beta_\alpha}$ رابطه زیر را حل کنند:

$$(\alpha_a, \beta_a) = \arg\min_{a,b} E\{(E[Y_i|D_i, X_i, D_{1i} > D_{0i}] - aD_i - X_i'b)^2|D_{1i} > D_{0i}\}.$$

به عبارت دیگر، $\boldsymbol{\alpha_\alpha D_i + \acute{X}_i\beta_\alpha}$ تخمین **MMSE** برای $[\boldsymbol{E\ Y_i|D_i, X_i, D_{1i} > D_{0i}}]$ است، یا در صورت خطی بودن به طور کامل با آن برابر است. نتیجه‌ای از قضیه ابادی این است که تابع تخمین‌زننده را می‌توان با حل

$$(\alpha_a, \beta_a) = \arg\min_{a,b} E\{\kappa_i(Y_i - aD_i - X_i'b)^2\}, \quad (4.5.5)$$

به دست آورد که حداقل مربعات وزنی کاپا است.[2]

ابادی استراتژی تخمینی (و قضیه توزیع را توسعه می‌دهد) برای روشی که شامل تخمین مرحله اول

1. ABADIE KAPPA

۲. لازم نیست کلاس توابع تخمین‌زننده خطی باشد. به جای $\alpha_\alpha D_i + \acute{X}_i\beta_\alpha$، منطقی‌تر است از تابعی غیرخطی مثل تابع نمایی (اگر متغیر وابسته منفی نباشد) یا پروبیت (اگر متغیر وابسته صفر ـ یک باشد) استفاده نماییم. در انتهای این فصل به این نقطه خواهیم رسید. همان طور که در بخش ۴-۴-۴ اشاره شد، از طرح وزنی کاپا می‌توان جهت توصیف کردن توزیع‌های متغیر کمی برای پذیرندگان و تخمین توزیع‌های خروجی استفاده کرد.

$\boldsymbol{K_i}$ با استفاده از مدل‌های پارامتریک یا نیمه‌پارامتریک برای تابع است پیشنهاد می‌دهد،= $\mathbf{P}(\boldsymbol{X_i})$ $\boldsymbol{P(z_i = 1|X_i)}$. سپس تخمین‌های حاصل از مرحله اول وارد مشابه نمونه ۵-۵-۴ در نمونه دوم می‌شوند. جای تعجب ندارد زمانی که تنها متغیر کمی ثابت باشد، روش ابادی به تخمین‌زننده والد ساده می‌شود. جالب‌تر آنکه، حداقل‌سازی ۵-۵-۴ تخمین‌زننده مرسوم **2SLS** را تا زمانی که مدل خطی برای $\mathbf{P}(\boldsymbol{X_i})$ در ایجاد $\boldsymbol{K_i}$ استفاده می‌شود تولید می‌کند. به عبارت دیگر اگر از $\boldsymbol{P(z_i = 1|X_i)}$=$\boldsymbol{\acute{X}_i\pi}$ در ایجاد تخمین $\boldsymbol{K_i}$ استفاده شود، برآورد ابادی برابر **2SLS** است. بنابراین، می‌توان نتیجه گرفت هر وقت $\mathbf{P}(\boldsymbol{X_i})$ را بتوان توسط مدلی خطی دقیقاً یا به طور تقریبی تخمین زد، منطقی است که **2SLS** را به عنوان تابع تقریب پاسخ علّی پذیرنده، $\boldsymbol{[E\ Y_i|D_i, X_i, D_{1i} > D_{0i}]}$ مدنظر قرار داد. از طرف دیگر، $\boldsymbol{\alpha_\alpha}$ به طور کلی برآورد **2SLS** نیست و $\boldsymbol{\beta_\alpha}$ هم به طور کلی بردار آثار متغیر کمی ایجاد شده توسط **2SLS** نمی‌باشد. همچنان، هم‌ارزی با **2SLS** برای حالت خطی $\boldsymbol{P(z_i = 1|X_i)}$ باعث می‌شود فکر کنیم محتمل است روش ابادی و **2SLS** در بسیاری از کاربردها تخمین‌های مشابهی ایجاد نمایند، با استنباط‌های بعدی بر این مبنا که **2SLS** را تخمین $\boldsymbol{[E\ Y_i|D_i, X_i, D_{1i} > D_{0i}]}$ در نظر می‌گیریم.

تحلیل دوباره آنگریست (۲۰۰۱) از آنگریست و اوانز (۱۹۹۸) مثالی است که در آن تخمین‌ها بر مبنای ۵-۵-۴ غیرقابل تمیز از تخمین‌های **2SLS** هستند. با به کارگیری ابزارهای دوقلوها جهت تخمین اثر فرزند سوم روی عرضه نیروی کار زن تخمین **2SLS** برابر ۰/۸۸- (**s.e.=0.017**) ایجاد می‌کند، در حالی که تخمین مشابه ابادی برابر ۰/۸۹- (**s.e.=0.017**) است. مشابها، تخمین‌های ابادی و **2SLS** از اثر روی ساعت‌های کاری مساوی با هم و برابر ۳/۵۵- (**s.e.=0.617**) است. این موضوع حمله‌ای به روش ابادی نیست، بلکه از این مفهوم حمایت می‌کند که، **2SLS** رابطه علّی مورد بررسی را تقریب می‌زند.[1]

۴-۵-۳. پاسخ متوسط علّی با شدت رفتار متغیر[2]*

تفاوت مهم میان آثار علّی متغیر موهومی و متغیری که مقادیر $\{0, 1, 2, \ldots\}$ را می‌پذیرد این است که در حالت اول، تنها یک اثر علّی برای هر شخص است، در حالی که برای حالت دوم تعداد زیادی وجود دارد: اثر تغییر از صفر به یک، اثر تغییر از یک به دو، و به همین ترتیب. علامتگذاری نتایج بالقوه که برای آموزش استفاده کرده‌ایم این امر را می‌شناسد. دوباره خواهیم داشت: بگذارید

$$Y_{si} \equiv f_i(s),$$

تحصیلات بالقوه (یا نهفته) که شخص **i** بعد از **s** سال آموختن کسب کرده است را مشخص کند. توجه داشه باشید تابع $\boldsymbol{f_i(s)}$ زیروند **i** دارد در حالی که **s** ندارد. تابع $\boldsymbol{f_i(s)}$ به ما می‌گوید شخص **i** برای هر مقدار آموزش **s**، و نه فقط برای مقدار شناخته شده $\boldsymbol{s_i}$، چه مقدار می‌آموزد. به عبارت دیگر،

---

۱. ابادی (۲۰۰۳) فرمول‌های مربوط به خطاهای استاندارد را ارائه می‌دهد و آلبرتو ابادی نرم‌افزاری برای محاسبه آنها فرستاده است. راه‌انداز خودکار جایگزین ساده‌ای فراهم می‌کند، که از آن برای ساخت خطاهای استاندارد برای تخمین‌های ابادی که در این پاراگراف آورده شده‌اند بهره می‌جوییم.

2. Average Causal Response with Variable Treatment Intensity

$f_i(s)$ به پرسش‌های علّی «چه می‌شود اگر» برای چندجمله‌ای s_i پاسخ می‌دهد.

فرض کنید s_i مقادیری در مجموعه $\{0, 1, ..., \bar{s}\}$ اتخاذ می‌نماید. در این صورت $\bar{s}$ آثار علّی واحد وجود دارد، $Y_{si} - Y_{s-1,i}$. در مدل علّی خطی فرض بر آن است این مقادیر برای تمامی **s** و تمامی **i** یکسان هستند، که فرض‌های غیرواقعی می‌باشند، اما لازم نیست عیناً این فرض‌ها را اعمال کنیم. **2SLS** ابزاری محاسباتی فراهم می‌سازد که میانگین وزنی از آثار علّی واحد ایجاد می‌نماید، به کمک تابعی وزنی می‌توان تخمین زد و تحقیق کرد، تا بیاموزیم هر عمل با ابزاری به خصوص از کجا می‌آید. این تابع وزنی به ما می‌گوید پذیرندگان چگونه در محدوده s_i توزیع شده‌اند. برای مثال به ما می‌گوید، تخمین بازگشت به تحصیل با استفاده از یک‌چهارم تولد یا قوانین اجباری سن حضور در مدرسه از جابه‌جایی‌هایی در توزیع نمره‌های دبیرستان می‌آید. سایر ابزارها، مثل ابزارهای فاصله استفاده شده توسط کارد (۱۹۹۵)، در ناحیه دیگری در توزیع مدرسه عمل می‌کنند و بنابراین دسته‌ای دیگر از نتایج را ثبت می‌کنند.

برای افزودن جزئیات بیشتر، فرض کنید از یک ابزار دودویی (باینری)، z_i، که متغیری موهومی برای نشان دادن اینکه در ایالتی با قوانین محدودکننده اجباری حضور در مدرسه به دنیا آمده‌اید، استفاده شده است تا میزان بازگشت به تحصیل را تخمین بزنند (همان طور که در آسمگلو و آنگریست آمده است، ۲۰۰۰). همچنین، بگذارید s_{1i}، آموزش **i** دریافت شده اگر $z_i = 1$ را مشخص کند و بگذارید s_{0i}، آموزش **i** دریافت شده اگر $z_i = 0$ را مشخص کند. قضیه‌ای که در ادامه می‌آید و برای آنگریست و ایمبنس (۱۹۹۵) است، مفهومی از برآورد والد با شدت رفتار متغیر در این حالت را ارائه می‌دهد. توجه کنید که ما محدودیت‌های استقلال و عدم شمول را به سادگی و با بیان اینکه نتایج بالقوه با ضریب **s** مستقل از ابزار هستند با هم ترکیب کرده‌ایم.

قضیه ۳-۵-۴ پاسخ علّی میانگین.[1] فرض کنید

(**ACR1**، استقلال و عدم شمول) $\{Y_{0i}, Y_{1i}, ..., Y_{\bar{s}i}; s_{0i}, s_{1i}\} \amalg z_i$

(**ACR2**، اولین مرحله) $E[s_{1i} - s_{0i}] \neq 0$

(**ACR3**، یکنوایی) $s_{1i} - s_{0i} \geq 0 \forall i$

بنابراین:

$$\frac{E[Y_i|z_i=1] - E[Y_i|z_i=0]}{E[s_i|z_i=1] - E[s_i|z_i=0]} = \sum_{s=1}^{\bar{s}} \omega_s E[Y_{si} - Y_{s-1,i}|s_{1i} \geq s > s_{0i}]$$

که در آن،

$$\omega_s = \frac{P[s_{1i} \geq s > s_{0i}]}{\sum_{j=1}^{\bar{s}} P[s_{1i} \geq j > s_{0i}]}$$

1. Average Causal Response (ACR)

وزن‌های $\boldsymbol{\omega_s}$ غیر منفی بوده و مجموعشان یک می‌شود.

قضیه پاسخ علّی میانگین (**ACR**) می‌گوید برآوردگر والد با شدت رفتار متغیر میانگین وزنی پاسخ علّی واحد در امتداد طول رابطه علّی غیرخطی بالقوه که توسط $\boldsymbol{f_i(s)}$ توصیف شده است می‌باشد. پاسخ علّی واحد، $E[Y_{si} - Y_{s-1,i}|s_{1i} \geq s > s_{0i}]$، تفاضل میانگین در نتایج بالقوه برای پذیرندگان در نقطه **s** است، یعنی افرادی که توسط ابزار از شدت رفتار کمتر از **s** به حداقل **s** برده شده‌اند. برای مثال، ابزار یک‌چهارم تولد استفاده شده توسط آنگریست و کروگر (۱۹۹۱) بعضی افراد را از پایه یازدهم به دوازدهم یا بالاتر و دیگران را از پایه دهم به پایه یازدهم یا بالاتر و به پایان رساندن آنها سوق داد. برآوردگر والد که از ابزار یک‌چهارم تولد استفاده می‌نماید تمامی این آثار را با یکدیگر ترکیب نموده و به یک پاسخ علّی میانگین تبدیل می‌کند.

اندازه میانگین گروه پذیرندگان در نقطه **s** برابر $P[s_{1i} \geq s > s_{0i}]$ است. در صورت یکنوایی، این مقدار می‌بایست غیر منفی بوده و توسط اختلاف در **CDF** مربوط به $\boldsymbol{s_i}$ در نقطه **s** داده شده است. برای مشاهده این موضوع، توجه کنید که:

$$\begin{aligned} P\left[s_{1i} \geq s > s_{0i}\right] &= P\left[s_{1i} \geq s\right] - P\left[s_{0i} \geq s\right] \\ &= P\left[s_{0i} < s\right] - P\left[s_{1i} < s\right], \end{aligned}$$

غیر منفی است، زیرا یکنوایی نیاز به این دارد که $\boldsymbol{s_{1i} \geq s_{0i}}$ باشد. به علاوه،

$$P[s_{0i} < s] - P[s_{1i} < s] = P[\mathrm{s}_i < s|\mathrm{z}_i = 0] - P[\mathrm{s}_i < s|\mathrm{z}_i = 1]$$

به صورت مستقل. در نهایت، توجه داشته باشید چون مقدار میانگین متغیر تصادفی غیر منفی یک عدد کمتر از **CDF** است، بنابراین خواهیم داشت:

$$\begin{aligned} & E\left[\mathrm{s}_i|\mathrm{z}_i = 1\right] - E\left[\mathrm{s}_i|\mathrm{z}_i = 0\right] \\ = \; & \sum_{j=1}^{\bar{s}} \left(P\left[\mathrm{s}_i < j|\mathrm{z}_i = 1\right] - P\left[\mathrm{s}_i < j|\mathrm{z}_i = 0\right]\right) = \sum_{j=1}^{\bar{s}} P\left[s_{1i} \geq j > s_{0i}\right] \end{aligned}$$

پس، تابع وزنی **ACR** را می‌توان به صورت مستمر با مقایسه **CDF**های متغیرهای درون‌زای (شدت رفتار) با به‌کارگیری و عدم به‌کارگیری ابزار تخمین زد. تابع وزنی در مرحله اول نرمال می‌شود.

قضیه **ACR** به ما کمک می‌کند تا بفهمیم چه چیزی از تخمین **2SLS** می‌آموزیم. برای مثال، متغیرهای ابزاری مشتق شده از سن حضور اجباری در مدرسه و قوانین کودکان کار اثر علّی افزایش تحصیل در محدوده کلاس ۱۲-۶ را نشان می‌دهد، اما نه ادامه تحصیل پس از دبیرستان. این موضوع در شکل ۶-۵-۴ که از آسمگلو و آنگریست (۲۰۰۰) برداشت شده نمایش داده شده است.

شکل اختلاف میان احتمال اینکه دستیابی به تحصیل برابر سطح کلاس بوده یا از آن پیشی می‌گیرد را روی محور **X** رسم می‌کند (یعنی یک واحد کمتر از **CDF**). اختلاف میان مردان تحت قوانین مختلف کار کودکان و قوانین حضور اجباری در مدرسه در نمونه‌ای از مردان سفیدپوست بین ۴۹-۴۰ سال که

از سرشماری ۱۹۶۰، ۱۹۷۰ و ۱۹۸۰ بیرون کشیده شده‌اند می‌باشد. ابزار به صورت تعداد سال‌های تحصیل مورد نیاز برای کار (پنل **A**) یا ترک مدرسه (پنل **B**) در سالی که شخص مورد نظر سنی برابر با ۱۴ سال داشت کدگذاری شده است. مردانی که تحت آخرین قوانین محدودکننده هستند گروه مرجع می‌باشند. از هر ابزار (برای مثال متغیر موهومی برای ۷ سال تحصیل قبل از اینکه اجازه کار داده شود مورد نیاز است) می‌توان برای ساخت برآوردگر والد با مقایسه با گروه مرجع استفاده کرد.

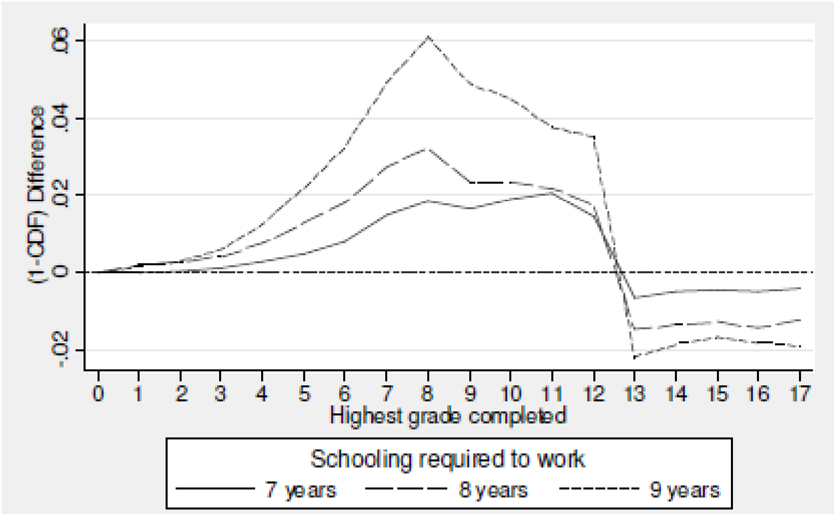
پنل **A** شکل ۱-۵-۴ نشان می‌دهد مردانی که در معرض قوانین کودکان کار محدودتری قرار داشتند با احتمال ۶-۱ درصد بیشتر امکان داشت کلاس‌های ۱۲-۸ را به پایان برسانند. شدت جابه‌جایی به این مسئله بستگی دارد که آیا قوانین به ۷، ۸ یا ۹ سال و بیشتر تحصیل قبل از اینکه اجازه کار داشته باشند نیاز دارد، اما در تمامی حالت‌ها، اختلاف‌های **CDF** در کلاس‌های پایین کاهش می‌یابند و با شیب زیادی بعد از کلاس ۱۲ افت می‌کنند. پنل **B** الگویی مشابه برای قوانین اجباری حضور در مدرسه نشان می‌دهد، با این حال آثار قدری کوچک‌تر هستند و رویدادها در اینجا در کلاس‌های بالاتر رخ می‌دهد، که مطابق با این حقیقت است معمولاً قوانین حضور اجباری در مدرسه در کلاس‌های بالاتر نسبت به قوانین کودکان کار از لزوم اجرای بیشتری برخوردارند.

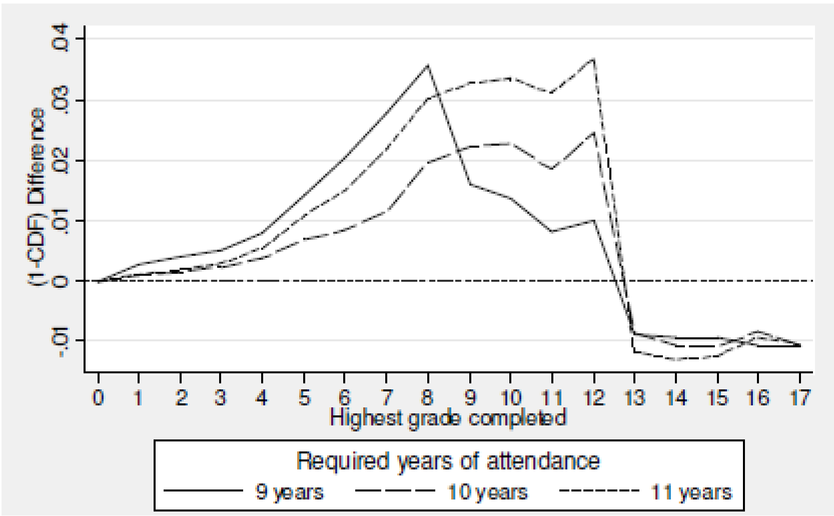
قبل از اینکه بحث درباره تعمیم **LATE** را جمع کنیم، بهتر است توجه داشته باشید بیشتر عناصری که در این کار پوشش دادیم به صورت ترکیبی کار می‌کنند. برای مثال، مدل‌ها با ابزارهای چندگانه و شدت رفتار متغیر میانگین وزنی **ACR** برای هر ابزار را ایجاد می‌نمایند. همچنین، قضیه‌های اشباع و وزن در مورد مدل‌هایی با شدت رفتار متغیر کاربرد دارند. از طرفی دیگر، ما هنوز بسط کاپای ابادی برای مدل‌هایی با شدت رفتار متغیر نداریم. بسط نهایی مهم در مورد سناریو است که در آن متغیر علّی مورد بررسی پیوسته می‌باشد و بنابراین می‌توانیم برای تابع پاسخ علّی مشتق در نظر بگیریم.

فرض کنید در مسئله تحصیل، مسائل غیرمرتبطی را تصور کنیم که توسط رابطه تابعی اساسی ایجاد شده‌اند. در هر حال، در این مورد متغیر علّی مورد بررسی می‌تواند هر مقدار غیر منفی اتخاذ نماید و فرض می‌شود رابطه تابعی دارای مشتق باشد. مثالی که این موضع در آن معنا پیدا کند منحنی تقاضا است، مقدار مورد نیاز بر حسب تابعی از قیمت می‌باشد. خصوصاً، بگذارید $q_i(p)$ مقدار مورد نیاز در بازار $\mathbf{i}$ در قیمت فرضی $\mathbf{p}$ را مشخص کند. این مسئله نتیجه‌ای بالقوه است، مثل $f_i(s)$، بجز اینکه به جای افراد واحد مشاهده زمان یا مکان یا هر دو است. برای مثال، آنگریست، گردی و ایمبنس (۲۰۰۰) قابلیت ارتجاعی مقدار مورد نیاز در بازار عمده‌فروشی ماهی فولتون در شهر نیویورک را تخمین می‌زنند. شیب منحنی این تقاضا برابر $\acute{q}_i(p)$ است؛ اگر مقدار و قیمت به صورت لگاریتمی محاسبه شوند، در این صورت برابر قابلیت ارتجاعی است.

شکل ۱-۵-۴. اثر ابزارهای تحصیل اجباری روی احتمال تحصیل

(برگرفته از آسمگلو و آنگریست ۲۰۰۰)





شکل فوق، اختلاف احتمال تحصیل در یا بیشتر از سطح کلاس را روی محور **X** نشان می‌دهد. گروه مرجع در پنل بالایی ۶ سال تحصیل مورد نیاز یا کمتر و در پنل پایینی ۸ سال یا کمتر است. پنل بالایی تفاضل **CDF** را با شدت قوانین کار کودکان نشان می‌دهد. پنل پایینی تفاضل **CDF** را با شدت قوانین اجباری حضور در مدرسه نمایش می‌دهد.

ابزار موجود در آنگریست، گردی و ایمبنس (۲۰۰۰) از اطلاعات وضعیت آب و هوا در ساحل لانگ آیلند، که چندان دور از زمین‌های تجاری ماهیگیری نیست، مشتق شده‌اند. آب و هوای طوفانی صید ماهی را دشوار کرده، باعث افزایش قیمت‌ها و کاهش تقاضا می‌گردد. آنگریست، گردی و ایمبنس از

متغیرهای موهومی مثل $stormy_i$ (طوفانی)، که نشان‌دهنده بازه‌های زمانی با موج‌های بلند و بادهای سریع است، جهت تخمین تقاضای ماهی استفاده می‌کنند. اطلاعات شامل مشاهدات روزانه از خریدهای کلی ماهی نرم‌باله به دست آمده است، ماهی ارزانی که برای پخت کیک ماهی از آن استفاده می‌شود.

برآوردگر والد با استفاده از ابزار $stormy_i$ را می‌توان به صورت:

$$\frac{E[q_i|stormy_i=1]-E[q_i|stormy_i=0]}{E[p_i|stormy_i=1]-E[p_i|stormy_i=0]} \tag{4.5.6}$$

$$= \frac{\int E[q_i'(t)|\ p_{1i}\geq t>p_{0i}]P[p_{1i}\geq t>p_{0i}]dt}{\int P[p_{1i}\geq t>p_{0i}]dt}, \tag{4.5.7}$$

نشان داد، که در آن p_i قیمت بازار (روز) $\mathbf{i}$ و p_{1i} و p_{0i} قیمت‌های بالقوه با اندیس $stormy_i$ هستند. این مشتق میانگین وزنی با تابع وزنی $P[p_{1i}\geq t>p_{0i}]=P[p_i\leq t|z_i=0]-P[p_i\leq t|z_i=1]$ در قیمت $\mathbf{t}$ است. به عبارت دیگر، تخمین IV با استفاده از $stormy_i$ مقدار میانگینی از مشتق $\acute{q}_i(t)$ ایجاد می‌کند، با وزن داده شده به هر قیمت ممکن (که دارای اندیس $\mathbf{t}$ است) نسبت به تغییرات ناشی از ابزار در تابع توزیع جمعی (CDF) قیمت‌ها در آن نقطه. این همان گروه از میانگین‌گیری است که در قضیه ACR به کار رفت با این تفاوت که اکنون پاسخ علّی اساسی به جای اینکه تفاضل یک واحدی باشد مشتق است.

فرمول پاسخ علّی میانگین، ۶-۵-۴، ناشی از این حقیقت است که:

$$E[q_i|stormy_i=1]-E[q_i|stormy_i=0]=E\int_{p_{0i}}^{p_{1i}} q_i'(t)dt, \tag{4.5.8}$$

که از قضیه بنیادی حسابان می‌آید. دو حالت به خصوص جالب از معادله ۸-۵-۴ به دست می‌آید. حالت اول موقعی است که تابع پاسخ علّی خطی باشد، یعنی $q_i(p)=\alpha_{0i}+\alpha_{1i}p$، برای بعضی ضرایب تصادفی، α_{0i} و α_{1i}. بنابراین خواهیم داشت:

$$\frac{E[q_i|stormy_i=1]-E[q_i|stormy_i=0]}{E[p_i|stormy_i=1]-E[p_i|stormy_i=0]}=\frac{E[\alpha_{1i}(p_{1i}-p_{0i})]}{E[p_{1i}-p_{0i}]}, \tag{4.5.9}$$

که میانگین وزنی ضریب تصادفی است، α_{1i}. وزن‌ها متناسب تغییرات قیمت القا شده به وسیله هوا در بازار $\mathbf{i}$ هستند.

حالت خاص دوم موقعی است که می‌توانیم تقاضا را به صورت:

$$q_i(p)=Q(p)+\eta_i, \tag{4.5.10}$$

بنویسیم، که در آن Q(p) تابع غیرتصادفی و η_i خطای تصادفی افزایشی است. در این حالت منظور ما این است که در هر روز یا هر بازار $\acute{q}_i(p)=\acute{Q}_i(p)$. در این حالت تابع پاسخ علّی به

$$\int Q'(t)\omega(t)dt, \text{where } \omega(t)=\frac{P[p_{1i}\geq t>p_{0i}]}{\int P[p_{1i}\geq r>p_{0i}]dr}.$$

تبدیل می‌گردد. این حالت‌های خاص دو مدل از میانگین‌گیری موجود در داخل قضیه **ACR** و نتیجه پیوسته آن را نشان می‌دهد، ۶-۵-۴ در ابتدا، میانگینی از بازار گرفته می‌شود، که وزن‌ها متناسب با اثر مرحله اول روی قیمت‌ها در هر بازار است. بازارهایی که در آن قیمت‌ها نسبت به وضعیت آب و هوا بسیار حساس است، بیشترین سهم را دارند. دوم، میانگینی در امتداد طول تابع پاسخ علّی در یک بازار داده شده گرفته می‌شود. **IV** مشتق میانگین گرفته شده روی محدوده‌ای از قیمت‌ها که در آن **CDF** قیمت‌ها بیشترین شیب تغییرات را دارد بازیابی می‌کند.

۶-۴. جزئیات IV

۱-۶-۴. اشتباهات 2SLS

محاسبه تخمین‌های **2SLS** ساده است، خصوصاً از زمانی که نرم‌افزارهایی مثل **SAS** و **Stata** آن را برای شما انجام می‌دهند. با این حال، ممکن است بعضی مواقع وسوسه شوید تا تخمین‌ها را خودتان انجام داده و ببینید آیا واقعاً ثمربخش است. یا شاید خود را روی سیاره کریکیت گرفتار ببینید در حالی که پروانه تمام نرم‌افزارهایتان منقضی شده است (کریک سیاره‌ای در ناحیه‌ای از فضاست که در آن زمان به کُندی می‌گذرد، بنابراین تجدید دوباره پروانه‌ها از شما زمان زیادی خواهد برد). «**2SLS** دستی» تنها برای چنین موارد اضطراری است. در روش **2SLS** دستی، مرحله اول را خودتان تخمین می‌زنید (که در هر حال باید خودتان آن را انجام دهید) و مقادیر متناسب را در معادله بزرگ مرحله دوم قرار دهید، که سپس با **OLS** تخمین زده می‌شود. با بازگشت به سیستم در ابتدای این بخش، مراحل اول و دوم عبارتند از:

$$\begin{aligned} s_i &= X_i'\pi_{10} + \pi_{11}' Z_i + \xi_{1i} \\ Y_i &= \alpha' X_i + \rho \hat{s}_i + [\eta_i + \rho(s_i - \hat{s}_i)] \end{aligned}$$

که در آن X_i مجموعه‌ای از متغیرهای کمی، z_i مجموعه‌ای از ابزارهای عدم شمول و مقادیر برازش شده مرحله اول برابر $\hat{s}_i = X_i'\hat{\pi}_{10} + \pi_{11}' Z_i$ هستند.

2SLS دستی بعضی از رازهای نهفته در **2SLS** را مشخص ساخته و در حالتی که به نرم‌افزار دسترسی نداریم مفید است، اما ممکن است باعث بروز اشتباهاتی گردد. همان طور که قبلاً بحث کردیم، خطاهای استاندارد **OLS** در مرحله دوم دستی تصحیح نخواهند شد (واریانس پسماند **OLS** واریانس $\eta_i + \rho(s_i + \hat{s}_i)$ است، در حالی که برای خطاهای مقتضی استاندارد **2SLS** شما تنها به واریانس η_i نیاز دارید). به همین ترتیب ریسک‌های ظریف بیشتری وجود دارند.

دوگانگی متغیرهای کمکی[1]

فرض کنید بردار متغیر کمی حاوی دو نوع متغیر باشد، بعضی، مثلاً (x_{0i}) که با آن راحتید و بقیه مثلاً (x_{1i}) که نسبت به آن احساسی دوگانه دارید. گریلیچس و میسون (۱۹۷۲) موقع ایجاد تخمین‌های **2SLS** از معادله دستمزد که با امتیازات **AFQT**[2] (آزمایش قابلیت مورد استفاده توسط نیروهای مسلح) به عنوان متغیر کنترل درون‌زای مجهز به ابزارهای اندازه‌گیری سروکار دارد با این سناریو موجه شدند. ابزارهای مورد استفاده در **AFQT** شامل متغیرهای تحصیل زودهنگام (قبل از خدمت سربازی به پایان رسیده است)، نژاد و پیشینه خانوادگی است. آنها سیستمی را تخمین زدند که می‌توان آن را به صورت:

$$\begin{aligned} s_i &= X_{0i}'\pi_{10} + \pi_{11}'Z_i + \xi_{1i} \\ Y_i &= \alpha_0'X_{0i} + \alpha_0'X_{1i} + \rho\hat{s}_i + [\eta_i + \rho(s_i - \hat{s}_i)]. \end{aligned}$$

توصیف کرد. این رابطه بسیار شبیه به **2SLS** دستی است.

با این حال نگاهی دقیق‌تر، از اختلافی مهم بین معادلات بالا و روش مرسوم **2SLS** پرده برمی‌دارد: متغیرهای کمی در مرحله‌های اول و دوم به یکدیگر شبیه نیستند. برای مثال گریلیچس و میسون سن را در مرحله دوم قرار دادند، اما در مرحله اول نه، حقیقتی که توسط کاردل و هاپکینز (۱۹۷۷) در اظهارنظری درباره این مقاله به آن اشاره شد. این امر یک اشتباه است. تخمین‌های مرحله دوم گریلیچس و میسون مشابه **2SLS** نیست. از آن بدتر، در جاهایی که **2SLS** ممکن است خوب باشد آنها متناقض هستند. برای بررسی علت، توجه داشته باشید که پسماند مرحله اول، $s_i - \hat{s}_i$، از اساس با x_{0i} نامتناسب است چون پسماندهای **OLS** همواره با رگرسورهای موجود نامتناسب هستند، اما چون x_{1i} در مرحله اول قرار نگرفته است احتمال دارد با پسماندهای مرحله اول متناسب باشد (برای مثال، سن احتمالاً متناسب با پسماند **AFQT** از مرحله اول گریلیچس و میسون (۱۹۷۲) است). تناقض حاصل از این معادله روی تمام ضرایب در مرحله دوم تأثیرگذار است. نتیجه اخلاقی داستان: متغیرهای کمی برون‌زای یکسانی را در مرحله اول و دوم قرار دهید. اگر یک متغیر کمی برای مرحله دوم مناسب باشد، برای مرحله اول هم مناسب است.

رگرسیون‌های ممنوع[3]

رگرسیون‌های ممنوع توسط پروفسور جری هاوسمن از دانشگاه **MIT** در سال ۱۹۷۵ ممنوع اعلام شد، و در حالی که گاهی اوقات در پایان‌نامه‌های کمتر نظارت شده ظاهر می‌شوند، همچنان از لحاظ فنی خارج از محدوده مجاز می‌باشند. رگرسیون ممنوع زمانی پدیدار می‌شود که محققان استدلال **2SLS**

1. Covariate Ambivalence
2. Armed Forces Qualification Test
3. Forbidden Regressions

را به شکل مستقیم برای مدل‌های غیرخطی به کار می‌برند. سناریوی متداول، متغیر درون‌زای موهومی است. برای مثال فرض کنید، مدل علّی مورد بررسی:

$$Y_i = \alpha' X_i + \rho D_i + \eta_i, \tag{4.6.1}$$

است، که در آن D_i متغیر موهومی برای وضعیت کهنه‌سرباز است. مرحله اول معمول **2SLS** عبارت است از:

$$D_i = \pi'_{10} X_i + \pi'_{11} Z_i + \xi_{1i}, \tag{4.6.2}$$

که رگرسیون خطی از D_i روی متغیرهای کمی و رگرسورها می‌باشد.

چون D_i متغیر موهومی است، **CEF** مربوط به مرحله اول، $E[D_i|X_i,Z_i]$، احتمالاً غیرخطی است. بنابراین مرحله اول **OLS** معمول تقریبی از **CEF** غیرخطی اساسی است. بنابراین ممکن است، از مرحله اول غیرخطی جهت نزدیک‌تر شدن به **CEF** بهره بجوییم. فرض کنید ما از پروبیت جهت مدلسازی $E[D_i|X_i,Z_i]$ استفاده کنیم. مرحله اول پروبیت برابر است با $\Phi[X_i'\pi_{p0} + \pi'_{p1}Z_i]$، که در آن π_{p0} و π_{p1} ضرایب پروبیت و مقادیر مناسب با $\hat{D}_{pi} = \Phi[X_i'\hat{\pi}_{p0} + \hat{\pi}'_{p1}Z_i]$ برابر هستند. رگرسیون ممنوع در این حالت معادله مرحله دوم ایجاد شده حاصل از جایگذاری $\hat{D}_{pi}$ به جای D_i است:

$$Y_i = \alpha' X_i + \rho \hat{D}_{pi} + [\eta_i + \rho(D_i - \hat{D}_{pi})]. \tag{4.6.3}$$

مشکل رابطه ۳-۶-۴ این است که تنها تخمین **OLS** ۲-۶-۴ این امکان را دارد تا پسماندهای مرحله اولی ایجاد کند که با مقادیر مناسب و متغیرهای کمی نامتناسب است. اگر $E[D_i|X_i,Z_i] = \Phi[X_i'\pi_{p0} + \pi'_{p1}Z_i]$، آنگاه پسماندهای حاصل از مدل غیرخطی به طور متناوب با x_i و $\hat{D}_{pi}$ نامتناسب هستند، اما چه کسی می‌تواند بگوید که **CEF** مرحله اول واقعاً پروبیت است؟ در مقابل، با وجود **2SLS** معمولی، نیازی نیست بابت اینکه آیا **CEF** مرحله اول واقعاً خطی است نگران باشیم.[۱]

جایگزینی ساده برای مرحله دوم ممنوعه، ۳-۶-۴، مانع بروز مشکلات ناشی از مرحله اول غیرخطی نادرست می‌گردد. به جای قرار دادن مقادیر مناسب غیرخطی، می‌توانیم از این مقادیر به عنوان ابزار استفاده کنیم. به عبارت دیگر، از $\hat{D}_{pi}$ به عنوان ابزاری برای ۱-۶-۴ در روش **2SLS** مرسوم استفاده نمایید (مثل همیشه، متغیرهای کمی برون‌زا، X_i، باید در لیست ابزار باشند). استفاده از مقادیر مناسب به عنوان ابزار همانند وصل کردن مقادیر مناسب هنگامی که اولین مرحله توسط **OLS** تخمین زده شده می‌باشد، اما همیشه این گونه نیست. به علاوه تناسب‌های غیرخطی به عنوان ابزار دارای این مزیت هستند که، اگر مدل غیرخطی تقریب بهتری از **CEF** مرحله اول نسبت به مدل خطی بدهد، تخمین‌های **2SLS** حاصل نسبت به آنهایی که از مرحله اول خطی استفاده می‌کنند مؤثرتر هستند (نیوی، ۱۹۹۰).

---

۱. این دیدگاه که پایداری تخمین‌های 2SLS در SEM مرسوم به مشخصات درست CEF مرحله اول بستگی ندارد به کلجیان (۱۹۷۱) بازمی‌گردد. استفاده از رابطه مرحله اول غیرخطی ممکن است در تمرین آسیب چندانی وارد نکند، پروبیت مرحله اول ممکن است بسیار نزدیک به حالت خطی باشد، اما چرا وقتی نیازی نیست دست به چنین اقدامی بزنیم؟

اما در اینجا هم، یک ایراد وجود دارد. تناسب‌های غیرخطی به عنوان ابزار به شکل ضمنی از غیرخطی بودن در مرحله اول به عنوان منبعی برای شناسایی اطلاعات بهره می‌جوید. برای دیدن این موضوع، فرض کنید مدل علّی مورد بررسی شامل ابزارهای، $\boldsymbol{Z_i}$، باشد:

$$Y_i = \alpha' X_i + \gamma' Z_i + \rho D_i + \eta_i. \qquad (4.6.4)$$

اکنون، با داشتن اولین مرحله که توسط ۲-۶-۴ داده شده، مدل ناشناس است و تخمین‌های **2SLS** متداول از ۴-۶-۴ وجود ندارند، اما تخمین‌های **2SLS** با استفاده از $\boldsymbol{Z_i}$، $\boldsymbol{X_i}$ و $\boldsymbol{\widehat{D}_{pi}}$ موجود است، چون $\boldsymbol{\widehat{D}_{pi}}$ تابع غیرخطی از $\boldsymbol{X_i}$ و $\boldsymbol{Z_i}$ است که از مرحله دوم حذف شده است. آیا باید از این ویژگی غیرخطی به عنوان منبعی برای شناسایی اطلاعات استفاده کرد؟ معمولاً ما سعی‌مان بر این است از چنین شناسایی‌هایی دوری کنیم چون واقعاً معلوم نیست آزمایش مبنا چه چیزی است.

به عنوان قانون، به‌کارگیری زودهنگام مقادیر مناسب مرحله اول در مدل‌های غیرخطی ایده بدی است. این مسئله شامل مدل‌هایی با مرحله دوم غیرخطی و همچنین آنهایی که **CEF** برای مرحله اول غیرخطی است، می‌باشد. برای مثال فرض کنید، بر این باورید رابطه علّی بین تحصیل و درآمد تقریباً درجه دوم است (همان طور که در مدل ساختاری کارد (۱۹۹۵) هم این گونه است). به عبارت دیگر، مدل مورد بررسی

$$Y_i = \alpha' X_i + \rho_1 s_i + \rho_2 s_i^2 + \eta_i. \qquad (4.6.5)$$

می‌باشد. با دو ابزار داده شده، تخمین ۵-۶-۴ که با $\boldsymbol{s_i}$ و $\boldsymbol{s_i^2}$ به صورت درون‌زا رفتار می‌کند آسان است. در این حالت، دو معادله مرحله اول وجود دارند، یکی برای $\boldsymbol{s_i}$ و دیگری برای $\boldsymbol{s_i^2}$. مسلماً نیاز به حداقل دو ابزار برای این حالت دارید تا به نتیجه برسید. طبیعی است از $\boldsymbol{Z_i}$ و مربع آن استفاده کنید (مگر اینکه $\boldsymbol{Z_i}$ موهومی باشد، که در این حالت نیاز به ایده بهتری دارید).

با این حال، ممکن است وسوسه شوید تا با یک مرحله اول کار کنید، برای مثال معادله ۲-۶-۴ و مرحله دوم پیش رو را دستی تخمین بزنید:

$$Y_i = \alpha' X_i + \rho_1 \hat{s}_i + \rho_2 \hat{s}_i^2 + [\eta_i + \rho_1(s_i - \hat{s}_i) + \rho_2(s_i^2 - \hat{s}_i^2)].$$

این کار اشتباه است چون $\boldsymbol{s_i}$ می‌تواند وابسته به $\boldsymbol{s_i^2 - \hat{s}_i^2}$ باشد در حالی که $\boldsymbol{\hat{s}_i^2}$ می‌تواند هم به $\boldsymbol{s_i - \hat{s}_i}$ و هم $\boldsymbol{s_i^2 - \hat{s}_i^2}$ وابسته باشد. از طرفی دیگر، تا زمانی که $\boldsymbol{X_i}$ و $\boldsymbol{Z_i}$ به $\boldsymbol{\eta_i}$ در ۵-۶-۴ وابسته نباشد و شما به اندازه کافی ابزار در $\boldsymbol{Z_i}$ داشته باشید، تخمین **2SLS** از ۵-۶-۴ درست است.

۲-۶-۴. آثار همتا[1]

متن‌های زیادی در علوم اجتماعی متأثر از اثر همتا هستند. این مسئله به معنای اثر علّی ویژگی‌های گروه روی نتایج فردی است. بعضی مواقع در اقدامی از رگرسیون برای آشکار ساختن این آثار استفاده

1. Peer Effects

می‌گردد. در عمل، استفاده از مدل‌های رگرسیون برای تخمین آثار همتا مملو از خطرات است. با این حال این مسئله به خودی خود مشکل **IV** نیست، زبان و جبر **2SLS** به ما در فهمیدن اینکه چرا شناسایی آثار همتا دشوار است، کمک می‌کند.

به طور کلی، دو دسته آثار همتا وجود دارند. اولین دسته شامل اثر ویژگی‌های گروه مثل میانگین تحصیلات در ایالت یا شهر روی متغیر خروجی که برای هر فرد اندازه‌گیری شده است. این اثر همتا میانگین یک متغیر را به نتایج اختصاصی که توسط متغیری دیگر توصیف شده مربوط می‌کند. برای مثال، آسمگلو و آنگریست (۲۰۰۰)، پرسیدند آیا درآمد فردی تحت تأثیر میانگین تحصیلات در محل سکونت شخص می‌باشد. قضیه اثر خارجی سرمایه انسانی اظهار می‌دارد زندگی در ایالتی با نیروی کار تحصیلکرده‌تر ممکن است تمام افراد در این ایالت را سودبخش‌تر کند، نه فقط کسانی که از تحصیلات بیشتری برخوردارند. گفته می‌شود این مسئله بازگشت اجتماعی تحصیل است: سرمایه انسانی که به همه سود می‌رساند، چه از تحصیلات بیشتری برخوردار باشند و چه نباشند.

مدل علّی که اجازه چنین اثر خارجی را می‌دهد می‌توان به صورت:

$$Y_{ijt} = \delta_j + \lambda_t + \gamma \overline{S}_{jt} + \rho s_i + u_{jt} + \eta_{ijt}, \qquad (4.6.6)$$

نوشت، که در آن $\boldsymbol{Y_{ijt}}$ سند دستمزد هفتگی فرد $\mathbf{i}$ در ایالت $\mathbf{j}$ در سال $\mathbf{t}$، $\boldsymbol{u_{jt}}$ مؤلفه خطای سالیانه ایالتی و $\boldsymbol{\eta_i}$ عبارت خطای فردی است. کنترل‌های $\boldsymbol{\delta_j}$ و $\boldsymbol{\lambda_t}$ آثار ایالت محل سکونت و سال است. ضریب $\boldsymbol{\rho}$ بازگشت به تحصیل برای هر فرد است، در حالی که ضریب $\boldsymbol{\gamma}$ به منظور گنجاندن اثر تحصیل میانگین، $\boldsymbol{\overline{S}_{jt}}$، در ایالت $\mathbf{j}$ و سال $\mathbf{t}$ می‌باشد.

علاوه بر نگرانی‌های معمول درباره $\boldsymbol{s_i}$، مهم‌ترین مسئله شناسایی ایجاد شده توسط معادله ۶-۶-۴ حذف متغیرهای اریب از رابطه بین تحصیل متوسط و سایر آثار سال ـ ایالتی مجسم شده در مؤلفه خطا $\boldsymbol{u_{jt}}$ است. برای مثال سیستم دانشگاه‌های دولتی ممکن است طی پیشرفت‌های دوره‌ای گسترش یابند، و روندی متداول در سطوح تحصیل میانگین ایالتی و درآمدهای میانگین ایالتی ایجاد کنند. آسمگلو و آنگریست (۲۰۰۰) اقدام به حل این مسئله با به‌کارگیری متغیرهای ابزاری مشتق شده از قوانین تاریخی حضور اجباری در مدرسه که با $\boldsymbol{\overline{S}_{jt}}$ مرتبط و با $\boldsymbol{u_{jt}}$ و $\boldsymbol{\eta_i}$ غیرمرتبط است کردند.

در حالی که آثار ایالتی ـ دولتی حذف شده جزء دلایل اصلی هستند که باعث محاسبه تخمین متغیرهای ابزاری توسط آسمگلو و آنگریست (۲۰۰۰) شدند، این حقیقت که یک رگرسور، $\boldsymbol{\overline{S}_{jt}}$، میانگین رگرسوری دیگر است، $\boldsymbol{s_i}$، مفهوم تخمین‌های **OLS** از معادله ۶-۶-۴ را کامل می‌کند. برای دیدن این مسئله، مدل ساده‌تری از ۶-۶-۴ با دیمانسیون سطح مقطع را در نظر بگیرید. می‌توان مدل را به صورت

$$Y_{ij} = \mu + \pi_0 s_i + \pi_1 \overline{S}_j + \nu_i; \text{ where } E[\nu_i s_i] = E[\nu_i \overline{S}_j] \equiv 0. \qquad (4.6.7)$$

نوشت، که در آن $\boldsymbol{Y_{ij}}$ سند دستمزد هفتگی فرد $\mathbf{i}$ در ایالت $\mathbf{j}$ و $\boldsymbol{\overline{S}_j}$ میانگین تحصیلات در آن ایالت است. اکنون بگذارید $\boldsymbol{\rho_0}$ ضریبی از رگرسیون دومتغیره $\boldsymbol{Y_{ij}}$ تنها روی $\boldsymbol{s_i}$ و $\boldsymbol{\rho_1}$ ضریبی از رگرسیون دومتغیره

$\boldsymbol{Y}_{\mathbf{ij}}$ تنها روی $\overline{\boldsymbol{S}}_{\boldsymbol{j}}$ مشخص کند. از بحث گروه‌بندی و **2SLS** که در ابتدای این فصل مطرح گشت، واضح است $\boldsymbol{\rho_1}$ تخمین **2SLS** از ضریب روی $\boldsymbol{s_i}$ در رگرسیون دومتغیره $\boldsymbol{Y}_{\mathbf{ij}}$ روی $\boldsymbol{s_i}$ با استفاده از مجموعه کامل متغیرهای موهومی ایالتی به عنوان ابزار است. پیوست از این واقعیت بهره می‌جوید تا نشان دهد پارامترهای معادله ۷-۶-۴ را می‌توان به صورت عبارت‌هایی از $\boldsymbol{\rho_0}$ و $\boldsymbol{\rho_1}$ به صورت:

$$\pi_0 = \rho_1 + \phi(\rho_0 - \rho_1) \qquad (4.6.8)$$

$$\pi_1 = \phi(\rho_1 - \rho_0)$$

نوشت که در آن $1 < \Phi=1/1-R^2 > 1$ است، و R^2 عبارت مجذور **R** مرحله اول است.

نتیجه رابطه ۸-۶-۴ این است که، اگر به هر دلیلی، تخمین‌های **OLS** از رگرسیون دومتغیره دستمزدها روی تحصیل فردی از تخمین‌های **2SLS** با استفاده از ابزارهای موهومی ایالتی متفاوت باشد، ضریب تحصیل متوسط در رابطه ۷-۶-۴ غیر صفر خواهد بود. برای مثال، اگر مجهز ساختن به متغیرهای موهومی ایالتی تقلیل اریب ناشی از خطای اندازه‌گیری در $\boldsymbol{s_i}$ را تصحیح کند، خواهیم داشت $\boldsymbol{\rho_1} > \boldsymbol{\rho_0}$ و ظاهر موهومی خروجی مثبت بازمی‌گردد. در مقابل، اگر مجهز ساختن با متغیرهای موهومی ایالتی اریب حاصل از رابطه مثبت میان $\boldsymbol{s_i}$ و درآمدهای بالقوه نادیده را حذف کند، خواهیم داشت $\boldsymbol{\rho_1} < \boldsymbol{\rho_0}$، و ظاهر اجتماعی منفی باز می‌گردد.[1] بنابراین، در عمل، بسیار مشکل است به آثار اجتماعی به دست آمده به وسیله تخمین **OLS** با یک معادله مانند ۶-۶-۴ تحقق بخشید، با این حال در عمل، از استراتژی‌های پیچیده‌تری جهت رفتار با میانگین‌های گروهی و فردی استفاده شده است.

اثر همتای دوم که آشکار ساختن آن سخت‌تر هم هست، اثر میانگین گروهی متغیر روی سطح فردی همین متغیر است. در واقع این مشکلی مربوط به **IV** نیست؛ این موضوع ما را به موضوعات اساسی رگرسیون باز می‌گرداند. برای درک این نکته، فرض کنید $\overline{\boldsymbol{S}}_{\boldsymbol{j}}$ نرخ فارغ‌التحصیلان مقطع دبیرستان در مدرسه **j** است و ما می‌خواهیم بدانیم آیا دانش‌آموزان زمانی که تمام همکلاسی‌های دوروبرشان میل به فارغ‌التحصیلی از دبیرستان دارند، احتمال بیشتری برای فارغ‌التحصیلی خواهند داشت یا خیر. برای آشکار ساختن اثر همتا روی نرخ فارغ‌التحصیلی از دبیرستان، ممکن است با مدل رگرسیونی به شکل:

$$s_{ij} = \mu + \pi_2 \overline{S}_j + \xi_{ij}, \qquad (4.6.9)$$

کار کنیم، که در آن $\boldsymbol{S_{ij}}$ وضعیت فارغ‌التحصیلی فرد **i** و $\overline{\boldsymbol{S}}_{\boldsymbol{j}}$ نرخ فارغ‌التحصیلی متوسط دبیرستان در مدرسه **j**، که فرد **i** به آن می‌رود است. در اولین نگاه، معادله ۹-۶-۴ فرمولی با معنا از سؤال علّی که به خوبی تعریف شده به نظر می‌رسد، اما در واقع بی‌معناست. رگرسیون $\boldsymbol{S_{ij}}$ روی $\overline{\boldsymbol{S}}_{\boldsymbol{j}}$ همیشه ضریبی برابر با یک دارد، نتیجه‌ای که با شناخت $\overline{\boldsymbol{S}}_{\boldsymbol{j}}$ به عنوان مقدار مناسب مرحله اول از رگرسیون $\boldsymbol{S_{ij}}$ روی مجموعه

۱. ضریب تحصیل میانگین در یک معادله با تحصیل فردی را می‌توان تفسیر کرد همانند هاوسمن (۱۹۷۸).

کاملی از متغیرهای موهومی مدرسه به سرعت قابل دستیابی است.[1] پس، معادله‌ای مثل ۹-۶-۴ احتمالاً حاوی اطلاعاتی درباره آثار علّی نمی‌باشد.

مدل ساده توسعه‌یافته رگرسیون همتای نامناسب معادله ۹-۶-۴ را به:

$$s_{ij} = \mu + \pi_4 \overline{S}_{(i)j} + \xi_{ij}, \qquad (4.6.10)$$

تغییر می‌دهد، که در آن $\overline{S}_{(i)j}$ مقدار میانگین S_{ij} در مدرسه j، با حذف دانش‌آموز i است. این کار قدمی در مسیر درست است ـ در تعریف، i در گروهی که برای ایجاد $\overline{S}_{(i)j}$ استفاده شده نیامده است ـ اما همچنان مشکل‌ساز است چون S_{ij} و $\overline{S}_{(i)j}$ هر دو تحت تأثیر شوک‌های تصادفی در سطح مدرسه هستند. حضور آثار تصادفی در عبارت خطا مسائل مهمی در نتیجه‌های آماری ایجاد می‌کند، مسائلی که به شکل کاملی در بخش ۸ در مورد آنها بحث شده است، اما در معادله‌ای مثل ۱۰-۶-۴، شوک‌های تصادفی در سطح گروهی بیشتر از یک مشکل برای خطاهای استاندارد هستند: هر شوک معمول برای گروه (مدرسه) آثار همتای موهومی ایجاد می‌نماید. برای مثال، مدیران به خصوص مؤثر مدارس ممکن است نرخ فارغ‌التحصیلی را در مدرسه محل کارشان برای همه افزایش دهند. این موضع به صورت اثر همتا به نظر می‌رسد خصوصاً اینکه همبستگی میان S_{ij} و $\overline{S}_{(i)j}$ ایجاد می‌کند حتی اگر ارتباط علّی میان روش‌های همتا و دستاوردهای فردی دانش‌آموز وجود نداشته باشد. بنابراین ترجیح بر این است رگرسیون‌هایی مثل ۱۰-۶-۴ را هم نبینیم.

بیشترین اقدامات در جستجوی علّی آثار همتا روی تغییرات در ویژگی‌های همتای پیش‌بینی شده تمرکز دارد، که عبارت است از، بعضی سنجش‌های کیفیت همتا که قبل از متغیر نتیجه رخ می‌دهد و بنابراین تحت تأثیر شوک‌های معمول قرار نمی‌گیرد. مثال اخیر آمرمولر و پیشکه (۲۰۰۶) است، که رابطه بین پیشینه خانوادگی همکلاسی را مورد مطالعه قرار می‌دهد و این پیشینه به وسیله تعداد کتاب‌های موجود در خانه و دستاوردهای دانش‌آموز در مدارس ابتدایی اروپا سنجش شده است. رگرسیون‌های آمرمولر و پیشکه مدل‌هایی از:

$$s_{ij} = \mu^* + \pi_4 \overline{B}_{(i)j} + \xi_{ij},$$

که در آن $\overline{B}_{(i)j}$ تعداد میانگین کتاب‌ها در خانه همتای دانش‌آموز i است. این رابطه مشابه ۱۰-۶-۴

آمار آزمایش برای کیفیت تخمین‌های OLS و تخمین‌های 2SLS برای بازگشت‌های خصوصی به تحصیل با استفاده از متغیرهای موهومی ایالتی به عنوان ابزار. بورخاس (۱۹۹۲) مسئله‌ای مشابه که روی تخمین آثار زمینه نژادی اثر می‌گذارد را مطرح می‌کند.

[1] در اینجا اثباتی مستقیم از اینکه رگرسیون s_{ij} روی $\overline{S}_j$ همراه یک است آورده‌ایم:

$$\frac{\sum_j \sum_i s_{ij}(\overline{S}_j - \overline{S})}{\sum_j n_j(\overline{S}_j - \overline{S})^2} = \frac{\sum_j (\overline{S}_j - \overline{S}) \sum_i s_{ij}}{\sum_j n_j(\overline{S}_j - \overline{S})^2} = \frac{\sum_j (\overline{S}_j - \overline{S})(n_j \overline{S}_j)}{\sum_j n_j(\overline{S}_j - \overline{S})^2} = 1.$$

است، اما با یک فرق مهم. متغیر $\overline{B}_{(i)j}$ ویژگی محیط خانه است که پیش از نمره‌های آزمون بررسی شده و بنابراین تحت تأثیر شوک‌های تصادفی در سطح مدرسه نمی‌باشد.

آنگریست و لنگ (۲۰۰۴) مثالی دیگر از اقدام برای مرتبط ساختن دستاورد دانش‌آموز با ویژگی‌های همتای پیش‌بینی شده فراهم آورده‌اند. مطالعه آنگریست و لنگ به مشاهده اثر دانش‌آموزان جدیدالورود با نمرات پایین روی نمرات دانش‌آموزان مستقر با نمرات بالا می‌پردازد. رگرسیون مورد استفاده در این حالت به شکل:

$$s_{ij} = \mu + \pi_3 \overline{m}_j + \xi_{ij}, \qquad (4.6.11)$$

است که در آن $\overline{m}_j$ تعداد دانش‌آموزان با نمرات دریافتی پایین در مدرسه j و S_{ij} نمره آزمون دانش‌آموز مستقر است. همبستگی کاذب ناشی از شوک‌های متداول به دو دلیل در این متن جای نگرانی ندارد. اول، $\overline{m}_j$ ویژگی جامعه مدرسه می‌باشد که توسط دانش‌آموزان خارج از نمونه که برای تخمین ۱۱-۶-۴ استفاده شده‌اند تعیین شده است. دوم، تعداد دانش‌آموزان با نمرات دریافتی پایین متغیر پیش‌بینی شده متمایل به اطلاعات اولیه درباره اینکه دانش‌آموزان اهل کجا هستند و نه متغیر خروجی، S_{ij}، می‌باشد. با این حال، آثار تصادفی در سطح مدرسه موضوعی مهم برای نتیجه باقی می‌مانند، چون $\overline{m}_j$ متغیری گروهی است.

۳-۶-۴. هزینه متغیرهای وابسته محدود[1]

در بخش ۲-۴-۳ به توصیف پیامدهای متغیرهای وابسته محدود برای مدل‌های رگرسیون پرداختیم. زمانی که متغیر وابسته دودویی یا غیر منفی باشد، برای مثال وضعیت اشتغال یا ساعات کاری، **CEF** معمولاً غیرخطی است. بسیاری از مدل‌های غیرخطی **LDV** حول تبدیل شاخص نهفته خطی ساخته شده‌اند. مثال‌ها شامل پروبیت، لوجیت و توبیت است. این مدل‌ها ویژگی‌های **CEF**های همبسته را نشان می‌دهند (برای مثال، مقادیر مناسب پروبیت بین صفر و یک هستند، در حالی که مقادیر مناسب توبیت غیر منفی می‌باشند). با این حال مشاهده نمودیم پیچیدگی‌های اضافه شده و کارهای اضافی مورد نیاز جهت تفسیر نتایج مدل‌های شاخص نهفته ارزش دردسر را ندارند.

نکته مهم قابل توجه به نفع **OLS** استحکام مفهومی است که مدل‌های ساختاری اغلب فاقد آنند. **OLS** همواره تقریب خطی **MMSE** برای **CEF** است. در واقع، می‌توانیم به **OLS** به عنوان طرحی برای محاسبه آثار حاشیه‌ای فکر کنیم، طرحی که دارای روح سادگی، اتوماسیون و مقایسه در میان مطالعات است. مدل‌های شاخص نهفته خطی بیشتر شبیه به **GLS** هستند، وقتی که عیناً به کار گرفته شوند باعث ایجاد بازده می‌گردند، اما نیاز به تعهد به فرم عملکردی و فرضیات توزیعی دارند که معمولاً

1. Limited Dependent Variables Reprise

احساس قوی نسبت به آن نداریم.[۱] مطلب دوم، تفاوت میان پارامترهای شاخص‌های نهفته در قلب مدل‌های غیرخطی و آثار علّی میانگین است که باور داریم باید در بسیاری از پروژه‌های تحقیقاتی به عنوان هدف در نظر گرفته شود. استدلال‌های به نفع **OLS** مرسوم با **LDV**ها با نیرویی یکسان در **2SLS** و مدل‌هایی با متغیرهای درون‌زای به کارگیری می‌شوند. روش‌های **IV** بدون توجه به اینکه آیا متغیر وابسته دودویی، غیر منفی، یا به شکل پیوسته روی خط واقعی توزیع شده‌اند آثار رفتار میانگین محلی را ضبط می‌کنند. با متغیرهای کمی، می‌توان **2SLS** را به عنوان تخمین میانگین **LATE** در امتداد سلول‌های متغیر کمی در نظر گرفت. در مدل‌های با شدت رفتار پیوسته یا متغیر، **2SLS** به ما پاسخ علّی میانگین یا مشتق میانگین می‌دهد. با وجود اینکه ابادی (۲۰۰۳) نشان داد **2SLS** به طور کلی تقریب **MMSE** برای تابع پاسخ علّی پذیرنده را فراهم نمی‌سازد، در عمل، تخمین‌های **2SLS** به طور قابل توجهی نزدیک به تخمین‌های با استفاده از روش دقیق‌تر مبتنی بر ابادی هستند (و با مدلی اشباع برای متغیرهای کمی، **2SLS** و ابادی عین هم هستند) و مسلماً، **2SLS** مستقیماً **LATE** را تخمین می‌زند؛ مرحله‌ای میانی شامل محاسبه آثار حاشیه‌ای وجود ندارد.

2SLS تنها راه پیش رو نیست. روش ماهرانه‌تر جایگزین با توصیف جزئیات روند ایجاد **LDV**ها سعی در ساخت داستانی علّی دارد. مثالی خوب پروبیت دومتغیره است، که می‌توان آن را برای مثال آنگریست و اوانز (۱۹۹۸) به کار برد. فرض کنید خانمی تصمیم می‌گیرد تا فرزند سومی با مقایسه هزینه‌ها و فواید توسط به‌کارگیری تابع سود خالص یا شاخص نهفته که در متغیرهای کمی و ابزارهای عدم شمول خطی است، با مؤلفه تصادفی یا عبارت خطا، v_i، داشته باشد. پروبیت دومتغیره مرحله اول را می‌توان به صورت:

$$D_i = 1[X_i'\gamma_0 + \gamma_1 z_i > v_i], \qquad (4.6.12)$$

نوشت که در آن z_i متغیر ابزاری است که فایده فرزند سوم را افزایش می‌دهد و مشروط روی متغیرهای کمی، X_i، است. برای مثال، به نظر می‌رسد والدین آمریکایی زمانی برای فرزند سوم ارزش بیشتری قائل هستند که دو دختر یا دو پسر داشته باشند، نوعی پدیده متنوع‌سازی انواع که می‌توان آن را به صورت افزایش سود فرزند سوم در خانواده‌هایی دارای فرزندان همجنس درک کرد.

---

۱. قیاس بین مدل‌های غیرخطی LDV و GLS بیشتر از حرف است. مدلی پروبیت با CEF غیرخطی $E[Y_i|X_i] = \Phi\left[\frac{X_i'\beta}{\sigma}\right] \equiv p_i$ را در نظر بگیرید. شرایط مرتبه اول جهت تخمین درستنمایی بیشینه این مدل

$$\sum \frac{(Y_i - p_i)X_i}{p_i(1-p_i)} = 0.$$

پس، بیشینه درستنمایی مشابه تخمین GLS از مدل غیرخطی است.

$$Y_i = \Phi\left[\frac{X_i'\beta}{\sigma}\right] + \xi_i.$$

پایداری برآوردگر احتمال بیشینه بر مبنای این فرض است که واریانس شرطی Y_i برابر $p_i(1-p_i)$ باشد. توجه کنید که می‌توان از این فرض رهایی یافت و به سادگی به کمک حداقل مربعات غیرخطی (NLLS) Y_i را با $\Phi\left[\frac{X_i'\beta}{\sigma}\right]$ برازش کرد. این گروه از NLLS خواص مستحکم OLS را به اشتراک می‌گذارد؛ و بهترین MMSE مناسب کلاسی از توابع تقریبی را می‌دهد.

نتیجه‌ای با اهمیت در این زمینه وضعیت اشتغال است، متغیر تصادفی برنولی با مقدار میانگین مشروط بین صفر و یک. جهت تکمیل مدل، فرض کنید وضعیت اشتغال، $\boldsymbol{Y_i}$، با شاخص نهفته

$$\mathrm{Y}_i = 1[\mathrm{X}_i'\beta_0 + \beta_1\mathrm{D}_i > \varepsilon_i], \tag{4.6.13}$$

تعیین شود، که در آن $\boldsymbol{\varepsilon_i}$ مؤلفه تصادفی یا عبارت خطای دوم است. این شاخص نهفته حاصل مقایسه بین هزینه‌ها و فواید کار کردن است.

منبع اریب‌های متغیر حذف شده در تنظیمات پروبیت دومتغیره رابطه بین $\boldsymbol{v_i}$ و $\boldsymbol{\varepsilon_i}$ می‌باشد. به عبارت دیگر، تعیین‌کننده‌های تصادفی اندازه‌گیری نشده مربوط به وضع حمل مرتبط با تعیین‌کننده‌های تصادفی اندازه‌گیری نشده اشتغال هستند. مدل با فرض اینکه $\boldsymbol{z_i}$ مستقل از این مؤلفه‌هاست شناخته شده و اینکه مؤلفه‌های تصادفی به صورت نرمال توزیع شده‌اند. با وجود نرمال بودن، پارامترهای ۱۲-۶-۴ و ۱۳-۶-۴ را می‌توان با بیشینه درستنمایی تقریب زد. تابع لگاریتمی درستنمایی عبارت است از:

$$\sum \mathrm{Y}_i \ln \Phi_b\left(\frac{\mathrm{X}_i'\beta_0 + \beta_1\mathrm{D}_i}{\sigma_\varepsilon}, \frac{\mathrm{X}_i'\gamma_0 + \gamma_1\mathrm{Z}_i}{\sigma_\nu}; \rho_{\varepsilon\nu}\right) \tag{4.6.14}$$
$$+(1-\mathrm{Y}_i)\ln\left[1 - \Phi_b\left(\frac{\mathrm{X}_i'\beta_0 + \beta_1\mathrm{D}_i}{\sigma_\varepsilon}, \frac{\mathrm{X}_i'\gamma_0 + \gamma_1\mathrm{Z}_i}{\sigma_\nu}; \rho_{\varepsilon\nu}\right)\right],$$

که در آن $\Phi_b(\cdot,\cdot;\rho_{\varepsilon\nu})$ تابع توزیع نرمال دومتغیره با ضریب تناسب $\boldsymbol{\rho_{ev}}$ است. با این حال توجه داشته باشید، می‌توان شاخص‌های نهفته را بدون تغییر در درستنمایی در ثابتی مثبت ضرب نمود. بنابراین هدف از تخمین، نسبت ضرایب شاخص به انحراف استاندارد از عبارت‌های خطاست (برای مثال $\boldsymbol{\beta_1/\sigma_\epsilon}$).

نتیجه‌های بالقوه که توسط مدل پروبیت دومتغیره تعریف شده‌اند برابر

$$\mathrm{Y}_{0i} = 1[\mathrm{X}_i'\beta_0 > \varepsilon_i] \text{ and } \mathrm{Y}_{1i} = 1[\mathrm{X}_i'\beta_0 + \beta_1 > \varepsilon_i],$$

هستند.

این در حالی است که تخصیص‌های تیمار بالقوه عبارتند از:

$$\mathrm{D}_{0i} = 1[\mathrm{X}_i'\gamma_0 > v_i] \text{ and } \mathrm{D}_{1i} = 1[\mathrm{X}_i'\gamma_0 + \gamma_1 > v_i].$$

مثل همیشه، تنها یک نتیجه بالقوه و یک انتخاب بالقوه برای هر نفر مشاهده شده است. همچنین از این مثال واضح است که رابطه بین $\boldsymbol{v_i}$ و $\boldsymbol{\varepsilon_i}$ مثل رابطه بین انتخاب‌های رفتار بالقوه و نتایج بالقوه است.

ضرایب شاخص نهفته به خودی خود به ما درباره اندازه اثر علّی وضع حمل روی اشتغال چیزی نمی‌گویند. برای مشاهده این موضوع، توجه داشته باشید که اثر علّی میانگین وضع حمل

$$E[\mathrm{Y}_{1i} - \mathrm{Y}_{0i}] = E\{1[\mathrm{X}_i'\beta_0 + \beta_1 > \varepsilon_i] - 1[\mathrm{X}_i'\beta_0 > \varepsilon_i]\}$$

است، در حالی که اثر میانگین روی افراد مورد مطالعه

$$E[\mathrm{Y}_{1i} - \mathrm{Y}_{0i}|\mathrm{D}_i = 1] = E\{1[\mathrm{X}_i'\beta_0 + \beta_1 > \varepsilon_i] - 1[\mathrm{X}_i'\beta_0 > \varepsilon_i]|\mathrm{X}_i'\gamma_0 + \gamma_1\mathrm{Z}_i > v_i\}.$$

است. با داشتن فرض‌های توزیعی متغیر برای $\boldsymbol{v_i}$ و $\boldsymbol{\varepsilon_i}$ اینها می‌توانند هر چیزی باشند (اگر عبارت‌های

خطا هم‌واریانس نباشند آنگاه در این صورت حتی علامت هم نامشخص است).

تحت شرایط نرمال، آثار علّی میانگین ایجاد شده به وسیله مدل پروبیت نرمال به سادگی قابل ارزیابی هستند. اثر علّی میانگین برابر:

$$\begin{aligned} & E\left\{1\left[X_i'\beta_0+\beta_1>\varepsilon_i\right]-1\left[X_i'\beta_0>\varepsilon_i\right]\right\} \\ = & E\left\{\Phi\left[\frac{X_i'\beta_0+\beta_1}{\sigma}\right]-\Phi\left[\frac{X_i'\beta_0}{\sigma}\right]\right\}, \end{aligned} \qquad (4.6.15)$$

است که در آن **Φ[0]**، **CDF** نرمال است. اثر روی افراد تحت مطالعه قدری پیچیده‌تر است چون شامل **CDF** نرمال دومتغیره می‌باشد.

$$\begin{aligned} & E\left[Y_{1i}-Y_{0i}\mid D_i=1\right] \\ = & E\left\{\frac{\Phi_b\left(\frac{X_i'\beta_0+\beta_1}{\sigma_\varepsilon},\frac{X_i'\gamma_0+\gamma_1 z_i}{\sigma_\nu};\rho_{\varepsilon\nu}\right)-\Phi_b\left(\frac{X_i'\beta_0}{\sigma_\varepsilon},\frac{X_i'\gamma_0+\gamma_1 z_i}{\sigma_\nu};\rho_{\varepsilon\nu}\right)}{\Phi\left(\frac{X_i'\gamma_0+\gamma_1 z_i}{\sigma_\nu}\right)}.\right\} \end{aligned} \qquad (4.6.16)$$

چون **CDF** نرمال دومتغیره در بسیاری از بسته‌های نرم‌افزاری تابعی سربسته است، در عمل محاسبه آن راحت است.

پروبیت دومتغیره از این نقطه‌نظر که چندان پیچیده نیست بی‌آزار به نظر می‌رسد و در صورت استفاده از روتین‌های بسته نرم‌افزاری به درستی قابل به‌کارگیری است، اما همچنان، دارای ضعف مدلسازی شاخص نهفته غیرخطی است که در بخش قبل درباره آن بحث شد. در ابتدا، بعضی محققان در اقدامی جهت شناسایی ضرایب شاخص به جای آثار علّی متوسط گمراه شدند. برای مثال، متنی بزرگ در حوزه اقتصادسنجی در رابطه با شناسایی ضرایب شاخص بدون نیاز به فرض‌های توزیعی است. محققان کاربردی که به آثار علّی علاقه‌مند هستند می‌توانند از این کار چشم‌پوشی کنند.[۱]

همچنین دومین ایراد در این زمینه هم یک مزیت است. می‌توان از پروبیت دومتغیره و سایر مدل‌های از این دست جهت شناسایی آثار علّی میانگین جامعه و/یا آثار روی افراد مورد مطالعه استفاده کرد. **2SLS** به شما قول آثار علّی میانگین را نمی‌دهد، فقط آثار علّی میانگین محلی، اما باید از رابطه ۱۵-۶-۴ واضح باشد که نرمال بودن فرضی عبارت‌های خطای شاخص نهفته برای این مسئله ضروری است. مثل همیشه، بهترین کاری که می‌توانید بدون فرض توزیعی انجام دهید **LATE** است، اثر علّی میانگین برای پذیرندگان. برای پروبیت دومتغیره، می‌توانیم **LATE** را به صورت

$$\begin{aligned} & E\left[Y_{1i}-Y_{0i}\mid D_{1i}>D_{0i}\right] \\ = & E\{1[X_i'\beta_0+\beta_1>\varepsilon_i]-1[X_i'\beta_0>\varepsilon_i]\mid X_i'\gamma_0+\gamma_1>v_i>X_i'\gamma_0\}, \end{aligned}$$

---

۱. فرض کنید عبارت خطای نهفته دارای توزیع ناشناخته‌ای باشد، با $\Lambda[\cdot]$ CDF. در این حالت اثر سببی میانگین برابر

$$E\left\{\Lambda\left[X_i'\beta_0+\beta_1\right]-\Lambda\left[X_i'\beta_0\right]\right\}=\Lambda'[X_i'\beta_0+\bar{\beta}_1]\beta_1,$$

است که در آن $\bar{\beta}_1$ در بازه $[0,\bar{\beta}_1]$ قرار دارد. این مسئله همواره به شکل $\Lambda[0]$ بستگی دارد.

بنویسیم، که مانند ۱۶-۶-۴، به کمک استفاده از نرمال بودن مشترک $\boldsymbol{v_i}$ و $\boldsymbol{\varepsilon_i}$ قابل محاسبه است، اما شما نیازی به استفاده از نرمال بودن جهت ارزیابی $E[Y_{1i}-Y_{0i}|D_{1i}>D_{0i}]$ ندارید، چون **LATE** را می‌توان به کمک **IV** برای هر $\boldsymbol{X_i}$ و میانگین با استفاده از هیستوگرام متغیرهای کمی تخمین زد. به عنوان جایگزین، **2SLS** را انجام داده و برای میانگین وزنی واریانس **LATE**های با متغیر کمی خاص حاضر شوید.

شاید شما بخواهید بدانید که آیا **LATE** کافی است. شاید شما دوست داشته باشید تا اثر درمان میانگین یا اثر درمان روی افراد تحت درمان را تخمین بزنید و بخواهید فرض‌های اضافی بدین منظور ارائه نمایید. همه اینها خوب است، اما تجربه به ما می‌گوید، شما نمی‌توانید از سنگ خون بگیرید، حتی با وجود فرض‌های بسیار خوب. چون اطلاعات محلی تنها چیزی است که در داده‌ها موجود می‌باشد، در عمل آثار علّی میانگین ایجاد شده به وسیله پروبیت دومتغیره احتمالاً مشابه تخمین‌های **2SLS** است و بیانگر این مسئله که مدل متغیرهای کمی از انعطاف‌پذیری کافی برخوردار است. این موضوع در جدول ۱-۶-۴ نشان داده شده است، که تخمین‌های **2SLS** و دومتغیره از آثار فرزند سوم روی نیروی کار زن با استفاده از ابزارهای همجنسیتی آنگریست ـ اوانز (۱۹۹۸) و همان نمونه سرشماری ۱۹۸۰ مربوط به زنان متأهل با دو فرزند یا بیشتر به‌کار رفته در این مقاله را گزارش می‌کند. متغیر وابسته، متغیری موهومی برای کار کردن در سال گذشته است؛ متغیر درون‌زای متغیر موهومی برای داشتن بچه سوم است. اثر مرحله اول داشتن خواهر یا برادر همجنسیت روی احتمال وضع حمل سوم در حدود ۷ درصد امتیاز است.

پلن **A** جدول ۱-۶-۴ تخمین‌هایی از مدلی بدون متغیرهای کمی گزارش می‌کند. تخمین **2SLS** برابر ۰/۱۳۸- در ستون یک از لحاظ عددی برابر اثر علّی تخمینی توسط ابادی با استفاده از مدل خطی در ستون دو است، همان گونه که در این حالت باید باشد. بدون متغیرهای کمی، ضریب شیب **2SLS** بهترین تقریب خطی برای تابع پاسخ پذیرنده همانند روش وزنی ـ کاپای ابادی فراهم می‌سازد. اثر حاشیه‌ای در صورتی به مقدار کم تغییر می‌کند که، به جای تقریب خطی، از حداقل مربعات غیرخطی با **CEF** پروبیت استفاده کنیم. اثر حاشیه‌ای تخمینی با کمینه‌سازی

$$E\left\{\kappa_i\left(Y_i-\Phi\left[\frac{\beta_0+\beta_1 D_i}{\sigma_\varepsilon}\right]\right)^2\right\}$$

برابر ۰/۱۳۷- است که در ستون سه گزارش شده است. این موضوع چندان جای تعجب ندارد چون مدل بدون متغیرهای کمی فرضیه‌های عملکردی را تحمیل نمی‌کند.

شاید این حقیقت جالب‌تر باشد که آثار حاشیه‌ای و آثار رفتار میانگین محاسبه شده به وسیله روابط ۱۵-۶-۴ و ۱۶-۶-۴ همانند تخمین‌های **2SLS** و ابادی هستند. این نتایج در ستون‌های ۶-۴ گزارش

شده‌اند. اثر حاشیه‌ای محاسبه شده با استفاده از مشتق برای تقریب به تفاضل محدود در رابطه ۱۵-۶-۴ برابر ۰/۱۳۸- است (در ستون ۴، **MFX** برچسب زده برای آثار حاشیه‌ای)، در حالی که آثار رفتار میانگین در ستون‌های ۵ و ۶ برابر ۰/۱۳۹- است. افزودن تعداد کمی متغیر کمی دارای اثر ناچیزی روی تخمین‌هاست، همان طور که در پنل **B** قابل رؤیت است. در این حالت، متغیرهای کمی همه متغیرهای موهومی هستند، سه تا برای نژاد (سیاهپوست، اسپانیایی و غیره) و دو تا نشانگر اولین و دومین پسرهای متولد شده (ابزار عدم شمول حاصل تعامل این دو است). پنل‌های **C** و **D** نشان می‌دهند اضافه کردن عبارت خطی به سن در اولین وضع حمل و متغیر موهومی برای سن مادری هم نتایج را بی‌تغییر باقی می‌گذارد.

ناوردایی (عدم نوسان) در متغیرهای کمی به نظر مطلوب است؛ چون ابزار همجنسیتی اساساً مستقل از متغیرهای کمی است، کنترل متغیرهای کمی برای حذف تمایل لازم نیست و باید در وهله اول روی دقت اثرگذار باشد. همچنان، همان طور که پنل **E** نشان می‌دهد، آثار حاشیه‌ای ایجاد شده توسط پروبیت دومتغیره به لیست متغیرهای کمی حساس است. تعویض متغیر موهومی که نشانگر مادران بالای ۳۰ سال است با عبارت خطی سن به شکل محسوسی تخمین‌های پروبیت دومتغیره را افزایش می‌دهد، به ۰/۱۷۱-، در حالی که تغییری در برآوردگرهای **2SLS** و ابادی ایجاد نمی‌کند. این مسئله احتمالاً منعکس‌کننده این حقیقت است که تغییرات خطی سن باعث القای برون‌زایی به سلول‌هایی می‌شود که در آنجا داده‌های کمی وجود دارد. با اینکه هیچ ضرری در گزارش نتایج در پنل **E** وجود ندارد، سخت است دریابیم چرا برآوردگرهای قوی‌تر **2SLS** و ابادی با اینکه از قابلیت اعتماد بیشتری برخوردارند نباید نشان داده شوند.[1]

۱. آنگریست (۲۰۰۱) هم به این نکته با استفاده از ابزار دوقلوها اشاره داشته و الگویی مشابه در مقایسه 2SLS گزارش می‌نماید.

جدول ۱-۶-۴. تخمین‌های 2SLS، ابادی و پروبیت دومتغیره از آثار فرزند سوم روی عرضه نیروی کار زنان

	2SLS	Abadie Estimates		Bivariate probit		
		Linear	Probit	MFX	ATE	TOT
	(1)	(2)	(3)	(4)	(5)	(6)
			A. No Covariates			
Employment	-0.138	-0.138	-0.137	-0.138	-0.139	-0.139
	(0.029)	(0.030)	(0.030)	(0.029)	(0.029)	(0.029)
			B. Some covariates (no age controls)			
Employment	-0.132	-0.132	-0.131	-0.135	-0.135	-0.135
	(0.029)	(0.029)	(0.028)	(0.028)	(0.028)	(0.028)
			C. Some covariates plus age at first birth			
Employment	-0.129	-0.129	-0.129	-0.133	-0.133	-0.133
	(0.028)	(0.028)	(0.028)	(0.026)	(0.026)	(0.026)
			D. Some covariates plus age at first birth and a dummy for age>30			
Employment	-0.124	-0.125	-0.125	-0.131	-0.131	-0.131
	(0.028)	(0.029)	(0.029)	(0.025)	(0.025)	(0.025)
			E. Some covariates plus age at first birth and age			
Employment	-0.120	-0.121	-0.121	-0.171	-0.171	-0.171
	(0.028)	(0.026)	(0.026)	(0.023)	(0.023)	(0.023)

نکته: اقتباس شده از آنگریست (۲۰۰۱). این جدول به مقایسه تخمین‌های 2SLS و تخمین‌های جایگزین نوع IV مربوط به اثر وضع حمل روی منبع نیروی کار به کمک مدل‌های غیرخطی می‌پردازد. خطاهای استاندارد برای تخمین‌های ابادی با استفاده از ۱۰۰ تکرار از زیرنمونه‌ها با ابعاد ۲۰۰۰۰ استخراج شده‌اند. MFX آثار حاشیه‌ای را مشخص می‌کند؛ ATE میانگین اثر رفتار است؛ TOT میانگین اثر رفتار روی نمونه مورد بررسی است.

۴-۶-۴. اریب 2SLS*

یک واقعیت مطلوب این است که برآوردگر **OLS** نه تنها سازگار، بلکه ناریب است. این بدان معناست که در یک نمونه با هر اندازه‌ای بردار ضریب **OLS** برآورد شده توزیعی دارد که مرکز آن در بردار ضریب جامعه قرار دارد.[۱] برآوردگر **2SLS** در مقایسه سازگار، اما اریب است. این بدین معناست که برآوردگر **2SLS** تنها در نمونه‌های بزرگ نزدیک به اثر علّی مورد نظر خواهد بود. در نمونه‌های کوچک **2SLS** می‌تواند به طور سیستماتیک از جامعه مورد برآورد متفاوت باشد.

برای سال‌های متمادی محققان کاربردی بدون تردید فکر می‌کردند که **2SLS** یک برآوردگر اریب است. هیچ‌کدام از ما چیز زیادی در مورد اریبی **2SLS** در کلاس‌های اقتصادسنجی خود در دانشگاه نشنیدیم. هرچند تعدادی مقاله در اوایل دهه ۱۹۹۰ میلادی این موضوع را تغییر دادند. این مقالات

ابادي و برآوردهای ساختاری غیرخطی مدل‌ها برای ساعات کار. Angrist (۱۹۹۱) 2SLS و برآورد پروبیت دومتغیره را در آزمایش‌های نمونه‌گیری مقایسه می‌کند.

۱. به عبارت دقیق‌تر OLS زمانی که (a) CEF خطی است یا (b) رگرسورها غیرتصادفی‌اند، ناریب است، مثال این مورد در نمونه‌های مکرر ثابت است. در عمل این شرایط خیلی مهم به نظر نمی‌رسند. به عنوان یک قاعده توزیع نمونه‌گیری $\hat{\beta} = [\sum_i X_i X_i']^{-1} \sum_i X_i Y_i$، تمایل دارد که بر روی نمونه نظیر جمعیت، $\beta = E[X_i X_i']^{-1} E[X_i Y_i]$ ، قرار گیرد، در نمونه‌هایی با هر اندازه چه CEF خطی باشد یا نه و چه رگرسورها غیرتصادفی باشند یا نه.

نشان دادند که برآوردهای **2SLS** می‌توانند در موارد تجربی کاملاً گمراه‌کننده باشند.[۱]

برآوردگر **2SLS** اکثراً زمانی که سنجه‌ها ضعیف هستند، یعنی همبستگی با رگرسورهای درون‌زا کم است و زمانی که تعداد زیادی قید بیش‌شناسایی شده وجود دارد، اریب است. زمانی که سنجه‌ها هم زیاد و هم ضعیف‌اند، برآوردگر **2SLS** نسبت به حد احتمال برآورد **OLS** متناظر اریب است. در بدترین سناریو ممکن برای تعداد زیادی سنجه و زمانی که سنجه‌ها خیلی ضعیف‌اند، به طوری که هیچ مرحله اولی[۲] در جامعه آماری وجود ندارد، توزیع نمونه‌گیری **2SLS** بر روی حد احتمال **OLS** قرار می‌گیرد. تئوری پشت این نتیجه‌گیری کمی فنی است، اما ایده اصلی آن را به راحتی می‌توان دید. منبع اریبی در برآوردهای **2SLS** تصادفی بودن برآوردهای مقادیر برازش شده در مرحله اول است. در عمل برآوردهای مرحله اول برخی از چیزهای تصادفی در متغیر درون‌زا را نشان می‌دهند. چراکه ضرایب مرحله اول از رگرسیون متغیر درون‌زا در سنجه‌ها به دست می‌آیند. اگر مرحله اول جامعه آماری صفر است، پس تمام موارد تصادفی در مرحله اول ناشی از متغیر درون‌زاست. این تصادفی بودن به همبستگی نمونه محدود بین مقادیر برازش شده در مرحله اول و خطاهای مرحله دوم تبدیل می‌شود، زیرا متغیر درون‌زا با خطاهای مرحله دوم همبستگی دارد (و در غیر این صورت شما نمی‌توانید از همان ابتدا سنجه را انتخاب کنید).

یک راه رسمی‌تر برای به دست آوردن اریبی **2SLS** به این صورت است. برای ساده کردن بحث از ماتریس‌ها و بردارها و یک مدل ساده اثرات ثابت[۳] (بحث درباره اریبی در دنیایی با اثرات ناهمگون دشوار است زیرا پارامتر هدف ممکن است متغیری از میان برآوردگرها باشد). با این فرض که شما به برآورد اثر یک رگرسور درون‌زای مفرد علاقه‌مندید که در یک بردار **x** در یک متغیر وابسته ذخیره شده است و در بردار **Y** بدون متغیر کمکی دیگری ذخیره شده است. مدل علّی مورد نظر می‌تواند به صورت زیر نوشته شود.

$$y = \beta x + \eta. \qquad (4.6.17)$$

ماتریس **N×Q** متغیر ابزاری **Z** است. معادله مرحله اول مرتبط با آن به صورت زیر است.

$$x = Z\pi + \xi. \qquad (4.6.18)$$

برآوردهای **OLS** در معادله (۱۷-۶-۴) اریبند، زیرا η_i با ξ_i همبسته است. سنجه‌های Z_i از نظر ساختار با ξ_i ناهمبسته هستند و طبق فرض با η_i ناهمبسته‌اند.

برآوردگر **2SLS** به صورت زیر است:

$$\widehat{\beta}_{2SLS} = (x'P_Z x)^{-1} x'P_Z y = \beta + (x'P_Z x)^{-1} x'P_Z \eta.$$

---

۱. منابع اصلی Nelson و Startz (۱۹۹۰، a و b)، Buse (۱۹۹۲)، Bekker (۱۹۹۴) و به خصوص Bound, Jaeger, و Baker (۱۹۹۵) هستند.

2. First-stage

3. Constant-effects

$P_Z = Z(Z'Z)^{-1}Z'$ ماتریس تصویر (ماتریس طرح‌ریزی)[1] است که مقادیر برازش شده را از رگرسیون $\mathbf{x}$ بر $\mathbf{Z}$ تهیه می‌کند و با جایگزین کردن $\mathbf{x}$ در $x'P_Z\eta$ رابطه زیر به دست می‌آید:

$$\widehat{\beta}_{2SLS} - \beta = (x'P_Zx)^{-1}(\pi'Z' + \xi')P_Z\eta = (x'P_Zx)^{-1}\pi'Z'\eta + (x'P_Zx)^{-1}\xi'P_Z\eta \qquad (4.6.19)$$

اریبی در **2SLS** از آنجا ناشی می‌شود که امید ریاضی (مقدار چشم‌داشتی) جملات سمت راست معادله صفر نباشد.

ارزیابی امید ریاضی معادله (۱۹-۶-۴) دشوار است، زیرا عملگر امید ریاضی از معکوس تابع غیرخطی $(x'P_Zx)^{-1}$ عبور نمی‌کند. هرچند می‌توان نشان داد که امید ریاضی ضرایب در طرف راست معادله (۱۹-۶-۴) می‌تواند به خوبی توسط ضرایب امیدهای ریاضی تقریب زده شود. به عبارت دیگر:

$$E[\widehat{\beta}_{2SLS} - \beta] \approx (E[x'P_Zx])^{-1}E[\pi'Z'\eta] + (E[x'P_Zx])^{-1}E[\xi'P_Z\eta].$$

این تقریب بسیار بهتر از تقریب مجانبی مرتبه یک معمول است که در نظریه نمونه بزرگ استفاده می‌شود. بنابراین ما فکر می‌کنیم این تقریب به ما مقیاس خوبی از رفتار نمونه متناهی برآوردگر **2SLS** می‌دهد.[2] به علاوه از آنجا که $E[\pi'Z'\xi] = 0$ و $E[\pi'Z'\eta] = 0$ بنابراین داریم:

$$E[\widehat{\beta}_{2SLS} - \beta] \approx \left[E(\pi'Z'Z\pi) + E(\xi'P_Z\xi)\right]^{-1}E(\xi'P_Z\eta). \qquad (4.6.20)$$

بنابراین اریبی تقریبی **2SLS** از آنجا می‌آید که $E(\xi'P_Z\eta)$ صفر نیست، مگر آنکه η_i و ξ_i ناهمبسته باشند، اما همبستگی η_i و ξ_i چیزی است که از همان ابتدا ما را به استفاده از **IV** هدایت کرد. دستکاری بیشتر (۲۰-۶-۴) جمله بسیار مفید دیگری را تولید می‌کند.

$$E[\widehat{\beta}_{2SLS} - \beta] \approx \frac{\sigma_{\eta\xi}}{\sigma_\xi^2}\left[\frac{E(\pi'Z'Z\pi)/Q}{\sigma_\xi^2} + 1\right]^{-1}$$

(می‌توانید چگونگی به دست آوردن رابطه را در ضمیمه مشاهده کنید). جمله $(1/\sigma_\xi^2)E(\pi'Z'Z\pi)/Q$ آماره **F** برای معناداری توأم تمام رگرسورها در رگرسیون مرحله اول است.[3] این آماره را **F** می‌نامیم. بنابراین می‌توانیم بنویسیم:

$$E[\widehat{\beta}_{2SLS} - \beta] \approx \frac{\sigma_{\eta\xi}}{\sigma_\xi^2}\frac{1}{F+1}. \qquad (4.6.21)$$

از این رابطه به این نتیجه می‌رسیم که هرچقدر آماره **F** در این رابطه کوچک‌تر شود اریبی **2SLS**

1. Projection Matrix

۲. مطالعه بکر (۱۹۹۴) و آنگریست و کروگر (۱۹۹۵) را ببینید. به این کار همچنین تقریب مجانبی گروهی نیز می‌گویند زیرا می‌تواند از یک دنباله مجانبی ناشی شود با ثابت نگه داشتن تعداد مشاهدات در هر سنجه اجازه می‌دهد که سنجه‌های عددی به همان اندازه که تعداد مشاهدات به بی‌نهایت می‌رسد به بی‌نهایت برسند.

۳. به نوعی اف(F)-آماره دقیق $(1/\hat{\sigma}_\xi^2)\hat{\pi}'Z'Z\hat{\pi}/Q$ است که در آن کلاه‌ها برآوردها را نشان می‌دهند. بنابراین گاهی اوقات $(1/\sigma_\xi^2)E(\pi'Z'Z\pi)/Q$ با نام اف ـ آماره جمعیت شناخته می‌شود، چراکه اف ـ آماره‌ای است که ما برای یک نمونه بزرگ بی‌نهایت به دست می‌آوریم. در عمل، تفاوت بین F جمعیت و نمونه در این زمینه بسیار کم است.

به $\frac{\sigma_{\eta\xi}}{\sigma_{\xi}^{2}}$ میل می‌کند. اریبی برآوردگر **OLS**، $\frac{\sigma_{\eta\xi}}{\sigma_{x}^{2}}$ است که اگر $\pi = 0$ برابر با $\frac{\sigma_{\eta\xi}}{\sigma_{\xi}^{2}}$ می‌شود. پس نشان دادیم زمانی که مرحله اول صفر است مرکز **2SLS** در همان نقطه‌ای قرار دارد که مرکز **OLS** قرار دارد. در حالت کلی‌تر می‌توان گفت زمانی که چیز زیادی در مرحله اول وجود ندارد، برآوردهای **2SLS** به طرف برآوردهای **OLS** اریب می‌شوند. از طرف دیگر زمانی که **F** بزرگ می‌شود اریبی **2SLS** به سمت صفر می‌رود. این اتفاق همچنین باید در نمونه‌های بزرگ، زمانی که $\pi \neq 0$ است بیفتد.

زمانی که سنجه‌ها ضعیف هستند خود آماره **F** به صورت معکوس با تعداد سنجه‌ها تغییر می‌کند. برای اینکه دلیل این مسئله را ببینیم، فرض می‌کنیم به مدل **2SLS** خود سنجه‌های به دردنخور اضافه کردیم. این سنجه‌ها تأثیری بر توان دوم مرحله اول ندارند. جمع مربعات مدل $E(\pi' Z' Z\pi)$ و واریانس مانده σ_{ξ}^{2} هر دو تا زمانی که **Q** زیاد می‌شود به همان صورت باقی می‌مانند. در نتیجه آماره **F** کوچک می‌شود. از این مسئله نتیجه می‌گیریم که اضافه کردن تعداد زیادی سنجه ضعیف اریبی را افزایش می‌دهد.

اریبی در **2SLS** مستقیماً نتیجه برآورد مرحله اول است. اگر ضرایب مرحله اول شناخته شده بودند می‌توانستیم از $\widehat{x}_{pop} = Z\pi$ برای مقادیر برازش شده در مرحله اول استفاده کنیم. این مقادیر برازش شده با خطاهای مرحله دوم ناهمبسته‌اند. هرچند در عمل ما از $\widehat{x} = P_Z x = Z\pi + P_Z\xi$ استفاده می‌کنیم که در جمله $P_Z\xi$ با $\widehat{x}_{pop}$ متفاوت است. اریبی در **2SLS** از این حقیقت که $P_Z\xi$ با η همبستگی دارد ناشی می‌شود. بنابراین برخی از همبستگی بین خطاها در مرحله اول و دوم از برآوردهای ما در **2SLS** از طریق تغییرپذیری نمونه‌گیری در $\hat{\pi}$ ناشی می‌شود. به طور موازی این همبستگی نیز ناچیز است، اما دنیای واقعی این گونه کار نمی‌کند.

فرمول اریبی (۲۱-۶-۴) نشان می‌دهد که اریبی در **2SLS** یک تابع افزایشی از تعداد سنجه‌هاست. بنابراین مشخصاً در موارد درست شناسایی شده زمانی که تعداد سنجه‌ها تا جایی که ممکن است کم باشند، اریبی حداقل است. هرچند معلوم می‌شود **2SLS** درست شناسایی شده (برآوردگر ساده والد) تقریباً ناریب است. نشان دادن این موضوع دشوار است، چراکه **2SLS** درست شناسایی شده گشتاوری ندارد (برای مثال توزیع نمونه‌گیری دم‌های کلفتی دارد). با وجود این حتی با سنجه‌های ضعیف **2SLS** درست شناسایی شده تقریباً در جایی که مرکزیت دارد (بنابراین ما می‌گوییم که **2SLS** درست شناسایی شده میانه ناریب است). این بدین معنا نیست که شما می‌توانید به راحتی از سنجه‌های ضعیف در مدل‌های درست شناسایی شده استفاده کنید. با یک سنجه ضعیف، برآوردهای درست شناسایی شده **IV**، بسیار ناپایدار و غیردقیق خواهند بود.

برآوردگر **LIML** برای مدل‌های اثر ثابت بیش‌شناسایی شده تقریباً میانه ناریب است به همین دلیل یک جایگزین جالب برای برآورد درست شناسایی شده با استفاده از یک سنجه در هر زمان تولید می‌کند (برای مثال می‌توانید مطالعه دیویدسون و مکینون ۱۹۹۳ و ماریانو ۲۰۰۱ را ببینید). مزیت

LIML این است که زمانی که کاهش اریبی نمونه متناهی را تولید می‌کند همان توزیع نمونه بزرگ **2SLS** را دارد (تحت عوامل ثابت). تعدادی از برآوردگرها در مدل‌های بیش‌شناسایی شده **2SLS** اریبی را کاهش می‌دهند، اما یک مطالعه گسترده مونته کارلو توسط فلورس ـ لاگونس (۲۰۰۷) به این نتیجه رسیده است که **LIML** در طیف وسیعی از موقعیت‌ها حداقل به خوبی جایگزین‌ها عمل می‌کند (در جمله‌های اریبی، میانگین خطای مطلق و نرخ رد کردن تجربی برای آزمون‌های **t**). یک مزیت دیگر برای **LIML** این است که بسته‌های نرم‌افزاری آماری فراوانی آن را محاسبه می‌کنند در حالی که برآوردگرهای دیگر معمولاً به مقداری برنامه‌نویسی نیاز دارند.[۱]

ما از یک آزمایش مونته کارلو کوچک برای نشان دادن تعدادی از نتایج تئوری مبحث بالا استفاده کردیم. داده‌های شبیه‌سازی شده از مدل زیر کشیده شده‌اند.

$$\begin{aligned} y_i &= \beta x_i + \eta_i \\ x_i &= \sum_{j=1}^{Q} \pi_j z_{ij} + \xi_i \end{aligned}$$

با $\beta = 1,\, \pi_1 = 0.1,\, \pi_j = 0 \ \forall j > 1$

$$\left(\begin{array}{c} \eta_i \\ \xi_i \end{array} \right) \Bigg| Z \sim N\left(\left(\begin{array}{c} 0 \\ 0 \end{array} \right), \left(\begin{array}{cc} 1 & 0.8 \\ 0.8 & 1 \end{array} \right) \right)$$

z_{ij} ها غیروابسته هستند. متغیرهای تصادفی با میانگین صفر و واریانس واحد به طور عادی توزیع شده‌اند. اندازه نمونه ۱۰۰۰ است.

شکل ۱-۶-۴ توزیع مونته کارلو با چهار برآوردگر را نشان می‌دهد: **OLS**، **IV** درست شناسایی شده (برای مثال **2SLS** با **Q=1** برچسب **IV** خورده است)، **2SLS** با دو سنجه (برای **Q=2** برچسب **2SLS** خورده است) و **LIML** با **Q=2**. برآوردگر **OLS** اریب است و اطراف مقدار ۱/۷۹ مرکزیت دارد. **IV** در اطراف ۱ مقدار β مرکزیت دارد. **2SLS** با یک سنجه ضعیف و یک سنجه مبهم نسبتاً به طرف **OLS** اریب شده است (میانه ۱/۰۷ است). تابع توزیع برای **LIML** با **Q=2** از تابع توزیع **IV** درست شناسایی شده غیرقابل تشخیص است، با اینکه برآوردگر **LIML** از یک سنجه کاملاً مبهم استفاده می‌کند.

شکل ۲-۶-۴ نتایج شبیه‌سازی را زمانی که **Q=20** است گزارش می‌کند. به این ترتیب علاوه بر

۱. LIML در SAS و STATA در دسترس است. با سنجه‌های ضعیف خطاهای استاندارد LIML کاملاً صحیح نخواهند بود، اما بکر (۱۹۹۴) یک راه‌حل ساده برای این موضوع ارائه می‌کند. چرا LIML نااریب است؟ عبارت (۲۱-۶-۴) نشان می‌دهد که اریبی تقریبی 2SLS متناسب با اریبی OLS است. از این موضوع نتیجه می‌گیریم که یک ترکیب خطی تقریباً نااریب بین 2SLS و OLS وجود دارد. معلوم می‌شود LIML چنین برآوردگر ترکیبی است. مانند اریبی 2SLS نااریب بودن تقریبی LIML را می‌توان نشان داد برای این کار از دنباله مجانب گروهی سبک بکر (Bekker-style) استفاده می‌شود که نسبت سنجه‌ها به اندازه نمونه را ثابت نگه می‌دارد. هرچند ذکر این نکته هم ارزشمند است که LIML در مدل‌ها با نوع خاصی از ناهمواریانسی اریب است؛ برای جزئیات بیشتر به مطالعه هاسمن نیووی و واترسون در سال ۲۰۰۶ مراجعه کنید.

یک سنجه مبهم، اما ضعیف ما ۱۹ سنجه بی‌ارزش دیگر اضافه کردیم. این بار نیز شکل توزیع **OLS**، **2SLS** و **LIML** را نشان می‌دهد. اکنون اریبی در **2SLS** بسیار بدتر است (میانه در ۱/۵۳ و نزدیک به میانه **OLS** است). توزیع نمونه برآوردگر **2SLS** نیز بسیار فشرده‌تر از مورد **Q=2** است. **LIML** همچنان عملکرد خوبی دارد و با کمی پراکندگی بیشتر نسبت به مورد **Q=2** در اطراف **β=1** مرکزیت دارد.

در نهایت شکل ۳-۶-۴ نتایج یک مدل کاملاً ناشناس را گزارش می‌کند. در این مورد $\pi_j=0$ **, j=1,2,3,…,20** است. عجیب نیست که تمام توزیع‌های نمونه در اطراف مقدار **OLS** قرار دارند. از طرف دیگر توزیع نمونه **2SLS** بسیار فشرده‌تر از توزیع **LIML** است. ما در این مورد از واژه مزیت ـ **LIML** استفاده می‌کنیم، زیرا پراکندگی گسترده توزیع نمونه **LIML** به خوبی این حقیقت را که نمونه در مورد پارامتر مورد نظر مبهم است را نشان می‌دهد.

این مسئله در عمل به چه معناست؟ علاوه بر حفظ حس مبهم نگرانی در مورد مرحله اول، توصیه‌های ما به شرح زیر است:

۱. مرحله اول را گزارش کنید و درباره معنادار بودن آن فکر کنید. آیا بزرگی و علامت اعداد محاسبه شده مطابق با انتظار شماست و یا اینکه اعداد به قدر کافی بزرگ ولی با علامت غلط هستند؟ در این صورت شاید سازوکار درستی در فرض مرحله اول شما وجود ندارد و تنها به صورت شانسی اعداد بزرگ به دست آمده‌اند.

۲. آماره **F** را گزارش کنید. هرچه مقدار این عدد بزرگ‌تر باشد بهتر است. استوک، رایت و یوگو (۲۰۰۲) نشان داده‌اند در صورتی که آماره **F** بزرگ‌تر از ۱۰ به دست آید شما را در منطقه امن قرار می‌دهد، اما بدیهی است که این یک قضیه ریاضی نیست.

۳. بهترین سنجه خود را انتخاب کنید و برآوردهای درست شناسایی شده را که تنها با استفاده از این سنجه به دست آمده‌اند را گزارش کنید. برآورد درست شناسایی شده برای **IV** میانه نااریب است. بنابراین بعید به نظر می‌رسد که نقد ضعیف بودن ابزار به آن وارد باشد.

۴. برآوردهای 2SLS بیش‌شناسایی شده را با LIML مقایسه کنید. LMIL دقیق‌تر از 2SLS است، اما اریبی کمتری دارد. اگر نتایج مشابه هم شد خوشحال شوید در غیر این صورت نگران باشید و به دنبال ابزار قوی‌تری بگردید.

۵. به ضرایب، آماره T و آماره F برای سنجه‌های حذف شده در شکل فروکاسته رگرسیون متغیرهای وابسته در سنجه‌ها توجه کنید. به یاد داشته باشید که فرم کاهش یافته متناسب با اثر علّی مورد نظر است. از همه مهم‌تر اینکه برآوردها در فرم کاهش‌یافته از آنجا که OLS هستند نااریب‌اند. همان طور که آنگریست و کروگر (۲۰۱۱) اشاره کرده‌اند اگر نمی‌توانید رابطه علّی مدنظر را در فرم کاهش‌یافته

ببینید احتمالاً چنین رابطه‌ای وجود ندارد.[1]

ما برخی از این نتایج را در یک بازتحلیل از مطالعه چارک تولد توسط آنگریست و کروگر (۱۹۹۱) نشان می‌دهیم. بوند جایگر و بیکر (۱۹۹۵) بیان کردند که اریبی در هنگام استفاده از چارک تولد به عنوان یک سنجه برای تحصیل از اهمیت ویژه‌ای برخوردار می‌شود، حتی با وجود این حقیقت که تعداد نمونه‌ها بیش از ۳۰۰۰۰۰ است. نمونه کوچک به وضوح نسبی است. قبلاً در این فصل مشاهده کردیم که الگو QOB در تحصیلات در شکل فروکاسته به وضوح بازتاب شده است که کمی نگران‌کننده است. از طرفی بوند جاگر و بیکر (۱۹۹۵) استدلال می‌کنند که مناسب‌ترین مدل‌های دارای کنترل اضافه مشمول این شکل‌های فروکاسته نیستند. جدول ۲-۶-۴ برخی از مشخصه‌ها را از آنگریست و کروگر(۱۹۹۱) و همچنین سایر مشخصه‌ها را از مطالعه بوند جگر و بیکر(۱۹۹۱) بازتولید می‌کند.

جدول ۲-۶-۴. برآوردهای جایگزین IV بازده اقتصادی برای تحصیل

	(1)	(2)	(3)	(4)	(5)	(6)
2SLS	0.105	0.435	0.089	0.076	0.093	0.091
	(0.020)	(0.450)	(0.016)	(0.029)	(0.009)	(0.011)
LIML	0.106	0.539	0.093	0.081	0.106	0.110
	(0.020)	(0.627)	(0.018)	(0.041)	(0.012)	(0.015)
F-statistic (excluded instruments)	32.27	0.42	4.91	1.61	2.58	1.97
Controls						
Year of birth	✓	✓	✓	✓	✓	✓
State of birth					✓	✓
Age, Age squared		✓		✓		✓
Excluded Instruments						
Quarter of birth	✓	✓				
Quarter of birth*year of birth			✓	✓	✓	✓
Quarter of birth*state of birth					✓	✓
Number of excluded instruments	3	2	30	28	180	178

نکته: جدول برآوردهای 2SLS و LIML را با استفاده از چند دسته سنجه‌ها و کنترل‌های جایگزین مقایسه می‌کند. برآورد OLS مربوط به گزارش مدل‌ها در ستون ۱ تا ۴، ۰/۷۱ است؛ برآورد OLS مربوط به گزارش مدل‌ها در ستون ۵ و ۶، ۰/۶۷ است. داده‌ها از مطالعه آنگریست و کروگر (۱۹۹۱) با ۱۹۸۰ نمونه سرشماری گرفته شده‌اند. اندازه نمونه ۳۲۹۵۰۹ است. خطاهای استاندارد داخل پرانتز گزارش شده‌اند.

ستون اول در جدول برآوردهای **2SLS** و **LIML** را در یک مدل با استفاده از متغیر ظاهری سه چارک تولد به عنوان سنجه و متغیر ظاهری سال تولد به عنوان متغیر کمکی گزارش می‌کند. برآورد **OLS** برای این مشخصات ۰/۰۷۱ است، در حالی که برآورد **2SLS** مقداری بیشتر و ۰/۱۰۵ است. اولین مرحله آماره **F** بیش از ۳۲ است، که بالاتر از منطقه خطر است. جای تعجب نیست که برآورد **LIML** در این مورد تقریباً برابر با **2SLS** است.

آنگریست و کروگر (۱۹۹۱) با مدل‌هایی آزمایش کردند که شامل سن و مربع سن می‌شوند که در

۱. مقاله اخیر توسط چرنوژوکف و هانسن (۲۰۰۷) به این قاعده کلی رسمیت می‌بخشد.

چارک‌ها به عنوان کنترل‌های اضافی اندازه‌گیری شده‌اند. این کنترل‌ها به منظور برداشتن اثرات حذف شده مربوط به سن هستند، چراکه این آثار ممکن است سنجه‌های چارک تولد را مختل کنند. از آنجا که سن در چارک‌ها، سال تولد و چارک تولد به صورت خطی وابسته‌اند، اضافه کردن سن و مربع سن تعداد سنجه‌ها را به دو تا کاهش می‌دهد. همان طور که در ستون دو نشان داده شده است زمانی که سن و مربع سن به عنوان کنترل اضافه شده‌اند مرحله اول آماره **F** تا ۰/۴ افت پیدا کرده است که نشانه محکمی برای وجود دردسر است، اما خطای استاندارد **2SLS** آنقدر زیاد است که ما نمی‌توانیم از این برآورد نتیجه مهمی بگیریم. برآورد **LIML** حتی نادقیق‌تر است. این مدل مشخصاً ناشناس است.

ستون ۳ و ۴ نتایج اضافه کردن اثر متقابل بین متغیرهای ظاهری چارک تولد و متغیرهای ظاهری سال تولد را به فهرست سنجه‌ها گزارش می‌کنند. بنابراین ۳۰ سنجه یا ۲۸ سنجه در زمانی که متغیرهای سن و مربع سن هم در نظر گرفته می‌شوند وجود دارد. مرحله اول آماره **F** ۴/۹ و با این دو مشخصه ۱/۶ است. برآوردهای **2SLS** کمی پایین‌تر از ستون یک هستند و ازاین‌رو به **OLS** نزدیک‌ترند، اما **LIML** خیلی از **2SLS** دور نیست. همچنین خطای استاندارد **LIML** در ستون ۴ تقریباً بزرگ است، اما آنقدر بزرگ نیست که برآورد بی‌معنا باشد. به طور خلاصه، به نظر می‌رسد دلایل کمی برای نگرانی در مورد سنجه ضعیف، حتی با اضافه کردن سن درجه دو وجود دارد.

نگران‌کننده‌ترین مشخصات آنهایی هستند که در ستون ۵ و ۶ آمده‌اند. این برآوردها با اضافه کردن ۱۵۰ اثر متقابل بین چارک تولد و مکان تولد به ۳۰ اثر متقابل بین چارک تولد و سال تولد تولید شده‌اند. منطق وارد کردن اثرهای متقابل مکان تولد در فهرست سنجه‌ها این است که تفاوت در قوانین اجباری تحصیل در کشورهای مختلف نیز در نظر گرفته شود، اما این مسئله موجب بیش‌شناسایی شدید در مدل‌ها با ۱۸۰ (یا ۱۷۸) سنجه که بسیاری از آنها ضعیف‌اند می‌شود. مرحله اول آماره **F** برای این مدل‌ها ۲/۶ و ۲ است که کاملاً در منطقه نامطمئن است. از طرف دیگر برآوردهای **LIML** بار دیگر بسیار شبیه به **2SLS** هستند. به علاوه در این مورد خطای استاندارد **LIML** با خطای استاندارد **2SLS** متفاوت است. این موضع نشان می‌دهد که نمی‌توانید همیشه با استفاده از یک قانون مکانیکی مانند **F** >۱۰ ارتباط سنجه‌ها را مشخص کنید. در برخی موارد **F** پایین ممکن است فاجعه نباشد.[1]

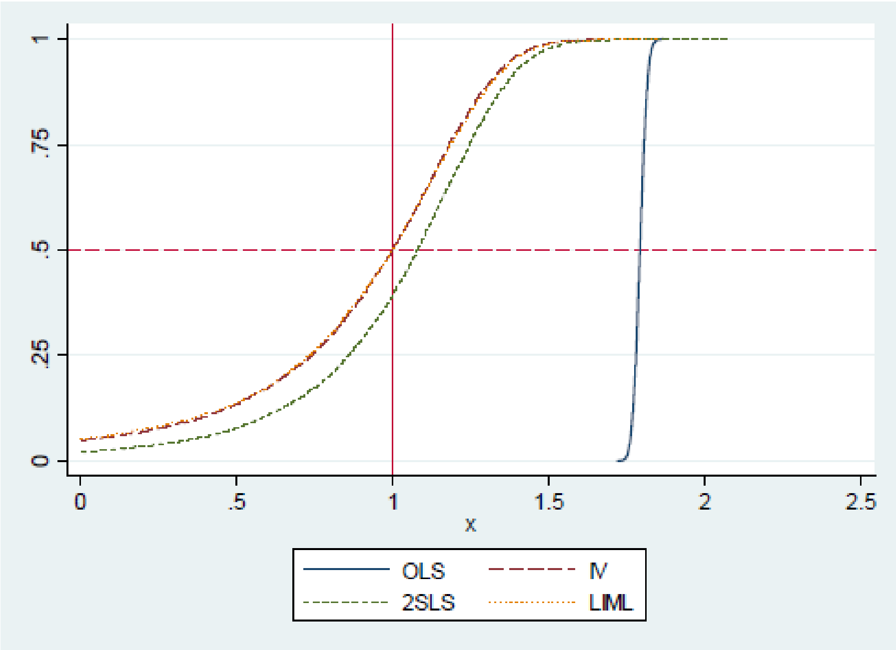
در نهایت، شایان ذکر است که در برنامه‌هایی با متغیرهای درون‌زا چندگانه، اولین مرحله متعارف **F** دیگر مناسب نیست. برای اینکه دلیل این موضوع را بدانیم فرض کنید دو سنجه برای دو متغیر درون‌زا وجود دارد و اولین سنجه قوی است و هر دو متغیر را به خوبی پیش‌بینی می‌کند در حالی که سنجه دوم ضعیف است. مرحله اول آماره **F** در هرکدام از معادلات مرحله اول به احتمال زیاد بالاست، اما مدل ضعیف شناسایی شده است، زیرا یک سنجه برای گرفتن دو اثر علّی کافی نیست. یک اصلاح ساده در

---

۱. کروز و موریرا (۲۰۰۵) به همین ترتیب نتیجه می‌گیرد که با وجود آماره F پایین اریبی کوچکی در مطالعه آنگریست و کروگر (۱۹۹۱)با مشخص‌سازی ۱۸۰ سنجه وجود دارد.

مرحله اول برای این مورد در ضمیمه داده شده است.

شکل ۱-۶-۴. توزیع برآوردگرهای OLS، IV، 2SLS و LIML. برآوردگر IV از یک سنجه استفاده می‌کند در حالی که 2SLS و LIML از دو سنجه استفاده می‌کنند



۷-۴. ضمیمه

به دست آوردن معادله (۸-۶-۴)

معادله ۷-۶-۴ را به صورت زیر بازنویسی می‌کنیم:

$$Y_{ij} = \mu^* + \pi_0 \tau_i + (\pi_0 + \pi_1)\overline{S}_j + \nu_i;$$

جایی که $\tau_i \equiv s_i - \overline{S}_j$ از آنجا که τ_i و $\overline{S}_j$ از نظر ساختاری ناهمبسته‌اند، داریم:

$$\begin{aligned} \rho_1 &= \pi_0 + \pi_1. \\ \pi_0 &= \frac{C(\tau_i, Y_{ij})}{V(\tau_i)}. \end{aligned}$$

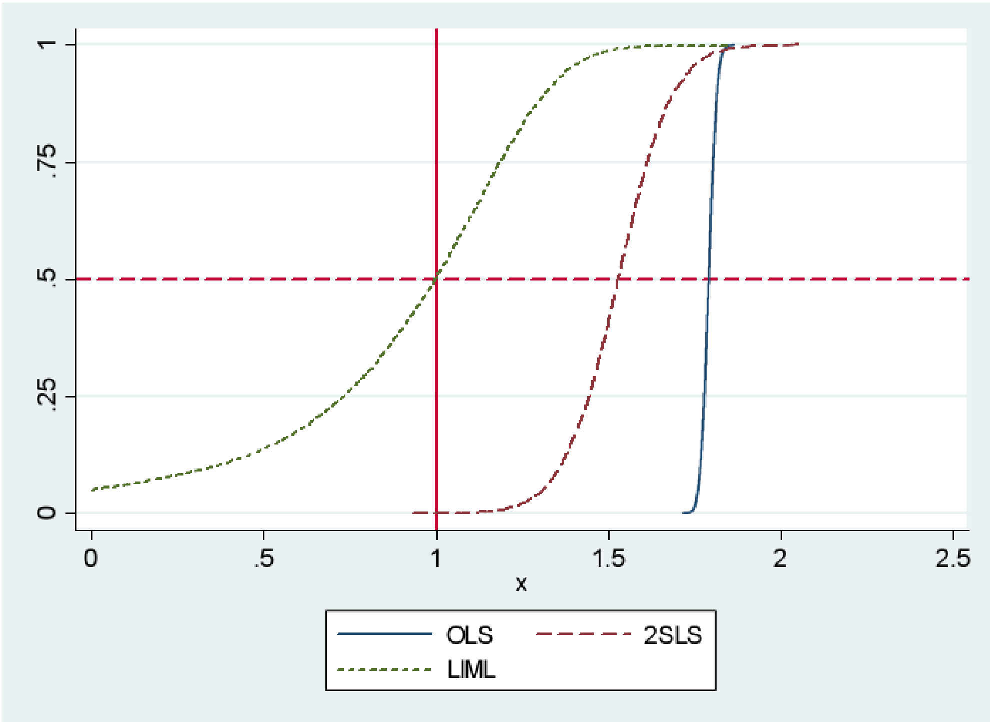
خط دوم را به صورت زیر ساده می‌کنیم:

$$\begin{aligned} \pi_0 &= \frac{C[(s_i - \overline{S}_j), Y_{ij}]}{[V(s_i) - V(\overline{S}_j)]} \\ &= \left[\frac{C(s_i, Y_{ij})}{V(s_i)}\right]\left[\frac{V(s_i)}{V(s_i) - V(\overline{S}_j)}\right] - \left[\frac{C(\overline{S}_j, Y_{ij})}{V(\overline{S}_j)}\right]\left[\frac{V(\overline{S}_j)}{V(s_i) - V(\overline{S}_j)}\right] \\ &= \rho_0\phi + \rho_1(1-\phi) = \rho_1 + \phi(\rho_0 - \rho_1) \end{aligned}$$

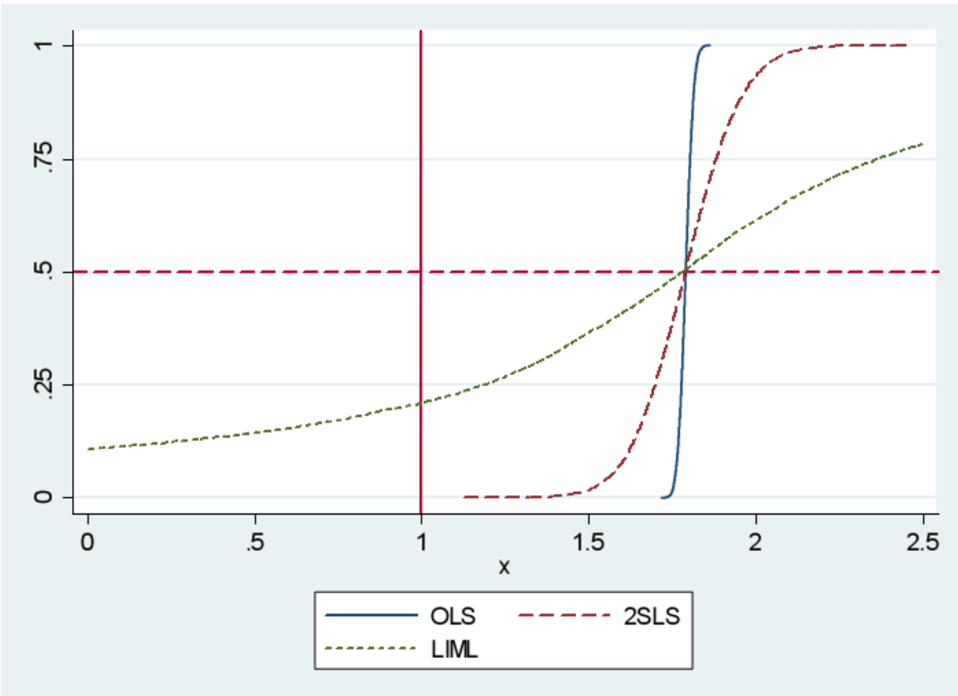
جایی که $\phi \equiv \frac{V(s_i)}{V(s_i)-V(\bar{S}_j)}$ با حل معادله برای π_1 داریم:

$$\pi_1 = \rho_1 - \pi_0 = \phi(\rho_1 - \rho_0).$$

شکل ۲-۶-۴. توزیع برآوردگرهای OLS، 2SLS و LIML با ۲۰ سنجه



شکل ۳-۶-۴. توزیع برآوردگرهای OLS، 2SLS و LIML با ۲۰ سنجه بی‌ارزش



چگونگی محاسبه اریب تقریبی 2SLS

از آخرین برابری در ۲۰-۶-۴ شروع می‌کنیم:

$$E[\widehat{\beta}_{2SLS} - \beta] \approx \left[E\left(\pi' Z' Z\pi\right) + E\left(\xi' P_Z \xi\right)\right]^{-1} E\left(\xi' P_Z \eta\right)$$

جادوی جبر خطی به ما کمک می‌کند این عبارت را ساده کنیم. جمله $\xi' P_Z \xi$ نرده‌ای است و در نتیجه با اثرش برابر است؛ اثر یک عملگر خطی است که از امید ریاضی عبور می‌کند و نسبت به جایگشت دوری ناوردا است. در نهایت اثر $\mathbf{P_Z}$ یک ماتریس خودتوان و با رتبه خود $\mathbf{Q}$ برابر است. با استفاده از این حقایق داریم:

$$\begin{aligned}
E\left(\xi' P_Z \xi\right) &= E\left[\mathrm{tr}\left(\xi' P_Z \xi\right)\right] \\
&= E\left[\mathrm{tr}\left(P_Z \xi \xi'\right)\right] \\
&= \mathrm{tr}\left(P_Z E\left[\xi \xi'\right]\right) \\
&= \mathrm{tr}\left(P_Z \sigma_\xi^2 I\right) \\
&= \sigma_\xi^2 \mathrm{tr}\left(P_Z\right) \\
&= \sigma_\xi^2 \mathrm{Q},
\end{aligned}$$

ما فرض کردیم ξ_i هم‌واریانس است. استفاده مشابه از ترفند اثر بر روی $\xi' P_Z \eta$ نشان می‌دهد که این جمله با $\sigma_{\eta\xi} \mathrm{Q}$ برابر است. بنابراین:

$$\begin{aligned}
E[\widehat{\beta}_{2SLS} - \beta] &\approx \left[E\left(\pi' Z' Z\pi\right) + \sigma_\xi^2 \mathrm{Q}\right]^{-1} E\left[\mathrm{tr}\left(\xi' P_Z \eta\right)\right] \\
&= \sigma_{\eta\xi} \mathrm{Q}\left[E\left(\pi' Z' Z\pi\right) + \sigma_\xi^2 \mathrm{Q}\right]^{-1} \\
&= \frac{\sigma_{\eta\xi}}{\sigma_\xi^2}\left[\frac{E\left(\pi' Z' Z\pi\right)/\mathrm{Q}}{\sigma_\xi^2} + 1\right]^{-1}.
\end{aligned}$$

مرحله اول آماره F چندمتغیره

فرض کنید هر متغیر کمکی بیرونی از فهرست سنجه‌ها خارج شده پس دو متغیر درون‌زا $\mathbf{x_1}$ و $\mathbf{x_2}$ با ضرایب δ_1 و δ_2 وجود دارند. ما به اریبی برآوردگر **2SLS** از δ_2 علاقه‌مندیم در حالی که با $\mathbf{x_1}$ نیز مانند یک متغیر درون‌زا رفتار می‌شود. مرحله دوم معادله به صورت زیر است:

$$y = P_Z x_1 \delta_1 + P_Z x_2 \delta_2 + [\eta + (x_1 - P_Z x_1)\delta_1 + (x_2 - P_Z x_2)\delta_2]. \tag{4.7.1}$$

$P_Z x_1$ و $P_Z x_2$ مقادیر برازش شده مرحله اول از رگرسیون $\mathbf{x_1}$ و $\mathbf{x_2}$ در $\mathbf{Z}$ هستند. با استفاده از فرمول آناتومی معمول برای رگرسیون چندمتغیره δ_2 در ۱-۷-۴ رگرسیون دومتغیره $\mathbf{y}$ در مانده رگرسیون $P_Z x_2$ در $P_Z x_1$ است. این مانده به صورت زیر است:

$$[I - P_Z x_1(x_1' P_Z x_1)^{-1} x_1' P_Z] P_Z x_2 = M_{1z} P_Z x_2$$

در اینجا $M_{1z} = [I - P_Z x_1(x_1' P_Z x_1)^{-1} x_1' P_Z]$ ماتریس مربوط به ساختن مانده است. به علاوه شایان ذکر است که $M_{1z} P_Z x_2 = P_Z[M_{1z} x_2]$

از اینجا ما نتیجه می‌گیریم که برآوردگر **2SLS** از δ_2 رگرسیون **OLS** بر $P_Z[M_{1z} x_2]$ است. به عبارت دیگر **OLS** از مقادیر برازش شده از یک رگرسیون $M_{1z} x_2$ در **Z** پس برآوردگر **2SLS** از δ_2 می‌تواند به صورت زیر نوشته شود:

$$[x_2' M_{1z} P_Z M_{1z} x_2]^{-1} x_2' M_{1z} P_Z y = \delta_2 + [x_2' M_{1z} P_Z M_{1z} x_2]^{-1} x_2' M_{1z} P_Z \eta$$

مجموع مربعات توضیح داده شده (شمارنده آماره **F**) که اریبی برآوردگر **2SLS** از δ_2 را تعیین می‌کند به همین دلیل امید ریاضی $[x_2' M_{1z} P_Z M_{1z} x_2]$است، در حالی که اریبی از این موضوع می‌آید که امید ریاضی $E[\xi' M_{1z} P_Z \eta]$زمانی که η و ξ همبسته‌اند غیرصفر است.

در اینجا بیان می‌کنیم که در عمل چگونه آماره **F** را محاسبه می‌کنیم: (**a**) مقادیر برازش مورد نظر مرحله اول را برای رگرسیون مورد نظر $P_Z x_2$ در مقادیر برازش مرحله اول دیگر و هر متغیر کمکی بیرونی رگرس کنید. مانده‌ها از این مرحله را ذخیره کنید. (**b**) آماره **F** را برای سنجه‌های حذف شده در مرحله اول رگرسیون مقادیر مانده از (**a**) در سنجه‌های حذف شده ایجاد کنید. توجه کنید که شما باید ضرایب مورد نظر **2SLS** را جایی که پسماندها از (**a**) با استفاده از **Z** سنجیده شده‌اند بدون داشتن متغیر کمکی یا متغیر درون‌زای دیگری با یک روش **2SLS** به دست بیاورید. از این حقیقت برای بررسی محاسبات خود استفاده کنید.

۵. جهان‌های موازی: اثرات ثابت،[1] تفاضلِ تفاضل[2] و داده‌های تابلویی[3]

اولین نکته برای فهم جهان‌های موازی این است که... آنها موازی نیستند.

داگلاس آدامز، تقریباً بی‌ضرر (۱۹۹۵)

کلید استنباط علّی در فصل ۳ کنترل عوامل اختلاطی[4و5] مشاهده شده است. اگر عوامل اختلاط‌ساز مهم مشاهده نشوند، باید با استفاده از **IV** که در فصل ۴ بحث کردیم، اثرات علّی را مشخص کنیم. ابزارهای خوب را به راحتی نمی‌توان پیدا کرد، هرچند دوست داریم برای بررسی متغیرهای اختلاطی

1. Fixed Effects
2. Differences-in-differences
3. Panel Data
4. Confounding Factors

۵. در آمار، اختلاط (Confounding) به پدیده‌ای گفته می‌شود که پیوستگی آماری مشاهده شده بین دو متغیر تصادفی حاصل از وجود متغیرهای تصادفی دیگر است و نه برهم‌کنش متغیرهای مشاهده شده با یکدیگر. به متغیرهایی که باعث ایجاد پیوستگی بین متغیرهای مطلوب می‌شوند متغیرهای اختلاطی یا متغیرهای نهان گفته می‌شود.

مشاهده نشده، ابزار دیگری نیز در اختیار داشته باشیم. این فصل تغییر موضوع کنترل را در نظر می‌گیرد: استراتژی‌هایی که از داده‌ها با زمان یا ابعاد گروهی برای کنترل متغیرهای حذف شده مشاهده نشده، اما ثابت استفاده می‌کنند. این استراتژی‌ها باید در سطوح مختلف مقایسه‌هایی انجام دهند و در عین حال رفتار روند شرطی خلاف واقع (جایگزین)[1] تیمار و گروه‌های کنترل یکسان باشد. ما همچنین ایده کنترل متغیرهای وابسته با تأخیر[2] را بحث می‌کنیم که استراتژی دیگری است که از زمان‌بندی استفاده می‌کند.

۵-۱. اثرات ثابت منفرد[3]

یکی از قدیمی‌ترین پرسش‌ها در اقتصاد کار، رابطه بین عضویت در اتحادیه و دستمزد است. آیا کارگرانی که دستمزدشان توسط چانه‌زنی گروهی تعیین می‌شود، درآمد بیشتری دارند یا در هر حالتی درآمد بیشتری دارند؟ برای مطرح کردن این پرسش، فرض کنیم Y_{it} مساوی درآمدهای کارگر **i** در زمان **t** است و فرض کنیم D_{it} نماینده وضعیت اتحادیه او باشد. Y_{it} مشاهده شده یا Y_{0it} است یا Y_{1it} که به وضعیت اتحادیه بستگی دارد. فرض کنید که:

$$E(Y_{0it}|A_i, X_{it}, t, D_{it}) = E(Y_{0it}|A_i, X_{it}, t),$$

یعنی وضعیت اتحادیه به خوبی شرط تصادفی توانایی مشاهده نشده کارگر A_i و سایر متغیرهای مشترک مشاهده شده X_{it} مانند سن و مدرسه رفتن است.

کلید تخمین اثرات ثابت این فرض است که A_i مشاهده نشده بدون زیروند[4] زمان در مدل خطی برای $E(Y_{0it}|A_i,X_{it}, t)$ قرار می‌گیرد:

$$E(Y_{0it}|A_i, X_{it}, t) = \alpha + \lambda_t + A_i'\gamma + X_{it}\delta, \qquad (5.1.1)$$

در نهایت، فرض می‌کنیم که تأثیر علّی عضویت اتحادیه، جمعی[5] و ثابت است:

$$E(Y_{1it}|A_i, X_{it}, t) = E(Y_{0it}|A_i, X_{it}, t) + \rho.$$

یعنی:

$$E(Y_{it}|A_i, X_{it}, t, D_{it}) = \alpha + \lambda_t + \rho D_{it} + A_i'\gamma + X_{it}\delta, \qquad (5.1.2)$$

که در آن ρ تأثیر علّی مدنظر است. مجموعه فرضیات منتهی به ۵-۱-۲ محدودکننده‌تر از چیزی است که به استفاده از رگرسیون در فصل ۳ منجر شود؛ ما به فرم تابعی خطی و جمعی احتیاج داریم تا با استفاده از داده‌های تابلویی بدون ابزار،[6] به مسئله متغیرهای اختلاطی مشاهده نشده برسیم.

1. Counterfactual Trend
2. Lagged Dependent Variables
3. Individual Fixed Effects
4. Subscript
5. Additive
6. Panel Data With No Instruments

معادله ۲-۱-۵ به معنای زیر است:

$$\mathrm{Y}_{it} = \alpha_i + \lambda_t + \rho \mathrm{D}_{it} + \mathrm{X}_{it}\delta + \varepsilon_{it}. \qquad (5.1.3)$$

که در آن

$$\alpha_i \equiv \alpha + A_i'\gamma.$$

این یک مدل اثرات ثابت است. با توجه به داده‌های تابلویی، مشاهدات تکراری بر روی افراد، تأثیر علّی وضعیت اتحادیه برای دستمزدها را می‌توان با تیمار α_i محاسبه کرد، اثر ثابت به عنوان پارامتری که قرار است تخمین زده شود. تأثیر سال، λ_t نیز به عنوان پارامتری در نظر گرفته می‌شود که قرار است تخمین زده شود. اثرات منفرد مشاهده نشده ضرایب متغیرهای ساختگی برای هر فرد هستند در حالی که اثرات سال، ضرایب متغیرهای موهومی زمان هستند.

به نظر می‌رسد که تعداد زیادی پارامتر برای تخمین در مدل اثرات ثابت وجود دارد مثلاً پیمایش تابلویی پویایی‌های درآمد، یک مجموعه داده تابلویی پُرکاربرد، شامل حدود ۵۰۰۰ مرد در سنین کاری مشاهده شده به مدت ۲۰ سال است. در نتیجه حدود ۵۰۰۰ اثر ثابت وجود دارد. در عمل اما این مسئله جدی نیست. در نظر گرفتن اثرات منفرد به عنوان پارامترهایی که قرار است تخمین زده شوند، از نظر جبری مانند تخمین در اریب از میانگین باشد. به عبارت دیگر، ما ابتدا میانگین منفرد را محاسبه می‌کنیم.

$$\overline{\mathrm{Y}}_i = \alpha_i + \overline{\lambda} + \rho\overline{\mathrm{D}}_i + \overline{\mathrm{X}}_i\delta + \overline{\varepsilon}_i.$$

این را از رابطه ۳-۱-۵ کم می‌کنیم که رابطه زیر به دست می‌آید:

$$\mathrm{Y}_{it} - \overline{\mathrm{Y}}_i = \lambda_t - \overline{\lambda} + \rho\left(\mathrm{D}_{it} - \overline{\mathrm{D}}_i\right) + \left(\mathrm{X}_{it} - \overline{\mathrm{X}}_i\right)\delta + \left(\varepsilon_{it} - \overline{\varepsilon}_i\right), \qquad (5.1.4)$$

در نتیجه اریب از میانگین، اثرات منفرد مشاهده نشده را از بین می‌برد.

جایگزین اریب از میانگین، دیفرانسیل‌گیری است. به عبارت دیگر، رابطه زیر را تخمین می‌زنیم:

$$\Delta\mathrm{Y}_{it} = \Delta\lambda_t + \rho\Delta\mathrm{D}_{it} + \Delta\mathrm{X}_{it}\delta + \Delta\varepsilon_{it}, \qquad (5.1.5)$$

که در آن پیشوند Δ نشان‌دهنده تغییر از یک سال به سال دیگر است. مثلاً، $\Delta\mathrm{Y}_{it} = \mathrm{Y}_{it} - \mathrm{Y}_{it-1}$. با دو دوره، دیفرانسیل‌گیری به صورت جبری مانند اریب از میانگین است. هر دو باید کار کنند هرچند با اریب ε_{it} همواریانس و ناهمبسته رشته‌ای[1] از میانگین که مؤثرتر است. اگر قرار باشد دستی آن را انجام دهید، ممکن است فکر کنید دیفرانسیل‌گیری راحت‌تر باشد، هرچند خطاهای استاندارد دیفرانسیل باید با این حقیقت جورشدگی داده شوند که پسماندهای دیفرانسیل‌گیری شده به صورت رشته‌ای به هم مرتبط هستند.

برخی بسته‌های نرم‌افزاری رگرسیون، تخمین‌گر اریب از میانگین[2] را به طور خودکار انجام داده و جورشدگی خطای استاندارد مناسب برای مقادیر آزادی از دست رفته در تخمین **N** میانگین منفرد را

1. Homoskedastic and Serially Uncorrelated
2. Deviations-from-means

انجام می‌دهند. این همه چیزی است که برای درست کردن خطاهای استاندارد با واریانس همسان لازم است. تخمین‌گر اریب از میانگین نام‌های زیادی دارد از جمله «تخمین‌گر درونی»[1] و «تحلیل کوواریانس». تخمین در فرم اریب از میانگین، جذب اثرات ثابت[2] نیز نامیده می‌شود.

فریمن (۱۹۸۴) از چهار مجموعه داده برای تخمین اثرات دستمزد اتحادیه، تحت این فرض استفاده می‌کند که انتخاب درون وضعیت اتحادیه بر اساس خصوصیات منفرد مشاهده نشده، اما ثابت[3] است. جدول ۱-۱-۵ نشان‌دهنده برخی از تخمین‌های او است. برای هر مجموعه داده، این جدول نشان‌دهنده نتایج تخمین‌گر اثرات ثابت و تخمین‌های سطح مقطع[4] متناظر است. تخمین‌های سطح مقطع معمولاً بالاتر (از ۱۵ تا ۲۵) از تخمین‌های اثرات ثابت هستند (از ۱۰ تا ۲۰). این نشان‌دهنده اریب انتخاب ثابت در تخمین‌های سطح مقطع است، هرچند اریب انتخاب تنها توضیح تخمین‌های پایین‌تر اثرات ثابت نیست.

جدول ۱-۱-۵. اثرات تخمین زده شده وضعیت اتحادیه برای لگاریتم دستمزدها

Survey	Cross section estimate	Fixed effects estimate
May CPS, 1974-75	0.19	0.09
National Longitudinal Survey of Young Men, 1970-78	0.28	0.19
Michigan PSID, 1970-79	0.23	0.14
QES, 1973-77	0.14	0.16

نکته: اقتباس شده از فریمن (۱۹۸۴). جدول فوق تخمین‌های اثر دستمزد نسبی اتحادیه را در مقطع و داده‌های تابلویی گزارش می‌کند. تخمین‌ها با استفاده از پیمایش‌های فهرست شده در سمت چپ محاسبه شدند. تخمین‌های سطح مقطع شامل کنترل متغیرهای مردم‌شناختی و سرمایه انسانی هستند.

هرچند، آنها برای نوع خاصی از متغیر حذف شده کنترل می‌شوند، اما تخمین‌های اثرات ثابت به اریب کاهشی[5] از خطای اندازه‌گیری حساس هستند. از یک سو، متغیرهای اقتصادی مانند وضعیت اتحادیه ثابت هستند و از سوی دیگر، خطای اندازه‌گیری اغلب از سالی به سال دیگر تغییر می‌کند. ازاین‌رو، در حالی که وضعیت اتحادیه ممکن است تنها برای چند کارگر در یک سال، به درستی گزارش یا کدگذاری نشود، تغییرات سالیانه مشاهده شده در وضعیت اتحادیه بیشتر نوفه[6] هستند. به عبارت دیگر، خطای اندازه‌گیری بیشتری در رگرسورها در معادله‌ای مانند ۵-۱-۵ یا ۴-۱-۵ وجود دارد تا در سطوح رگرسورها. این حقیقت توضیح‌دهنده تخمین‌های کوچک‌تر اثرات ثابت است.

بر اساس این حقیقت که دیفرانسیل‌گیری و تخمین‌گرهای اریب از میانگین مورد استفاده برای کنترل اثرات ثابت، معمولاً تغییر خوب و بد را از بین می‌برند، نوعی از مسئله خطای اندازه‌گیری ایجاد

1. Within Estimator
2. Absorbing the fixed effects
3. Unobserved-But-Fixed Individual Characteristics
4. Cross-Section Estimates
5. Attenuation Bias
6. Noise

می‌شود. به عبارت دیگر، این تغییرات ممکن است مقداری از اریب متغیرهای حذف شده را از بین ببرند، اما بیشتر اطلاعات سودمند ـ متغیر مدنظر ـ را نیز از بین می‌برند. یک نمونه استفاده از دوقلوها برای تخمین اثر علّی مدرسه رفتن بر دستمزد است. هرچند بُعد زمانی برای این مسئله وجود ندارد، اما ایده اصلی مانند مسئله اتحادیه است که در بالا بحث کردیم: دوقلوها پیشینه خانوادگی و ژنتیکی مشابه، اما مشاهده نشده دارند. ازاین‌رو می‌توان پیشینه خانوادگی مشترک آنها را با شمول اثر ثابت خانواده در نمونه جفت‌های دوقلو کنترل کرد.

آشنفلتر و کروگر (۱۹۹۴) و آشنفلتر و رز (۱۹۹۸) بازده مدرسه رفتن را با استفاده از نمونه‌های دوقلوها تخمین زدند و اثرات ثابت خانواده را کنترل کردند. چون دو دوقلو از هر خانواده حضور داشتند، این کار مانند رگرسیون‌گیری تفاوت‌ها در درآمدها درون جفت‌های دوقلو در تفاوت مدرسه رفتن آنهاست. جالب اینکه، تخمین‌های همراه خانواده، بزرگ‌تر از **OLS** هستند، اما چطور تفاوت در مدرسه رفتن بین افرادی ایجاد می‌شود که به شدت شبیه هم هستند؟ باند و سولون (۱۹۹۹) اشاره دارند که تفاوت‌های اندکی بین دوقلوها وجود دارد و آنهایی که معمولاً اول دنیا می‌آیند وزن کمتر و **IQ** بالاتری دارند. در حالی که این تفاوت‌های بین دوقلوها زیاد نیستند، تفاوت مدرسه رفتن آنها نیز زیاد نیست. ازاین رو، مقدار اندکی از تفاوت توانایی مشاهده نشده در میان دوقلوها ناشی از اریب جدی در تخمین‌های نهایی است.

در مورد خطای اندازه‌گیری و مسائل مرتبط در مدل‌هایی با اثرات ثابت، چه کار باید کرد؟ یک راه احتمالی برای از بین بردن خطای اندازه‌گیری، متغیرهای ابزاری است. آشنفلتر و کروگر (۱۹۹۴) از گزارشات متقاطع خواهر/برادری برای ساخت ابزار برای تفاوت‌های مدرسه رفتن بین دوقلوها استفاده کردند. مثلاً آنها از گزارش هریک از دوقلوها در مورد مدرسه رفتن برادرش به عنوان ابزاری برای گزارشات منفرد استفاده کردند. دومین رویکرد درج اطلاعات خارجی[1] در مورد میزان خطای اندازه‌گیری و جورشدگی تخمین‌های ساده است. در یک مطالعه بر روی اثرات دستمزد اتحادیه، کارد (۱۹۹۶) از اطلاعات خارجیِ پیمایش اعتبارسنجی جداگانه برای جورشدگی تخمین‌های داده‌های تابلویی برای خطای اندازه‌گیری در وضعیت گزارش شده اتحادیه استفاده می‌کند، اما داده‌های گزارشات مختلف و معیارهای تکراری مانند نوع استفاده شده توسط آشنفلتر و رز (۱۹۹۴) و کارد (۱۹۹۶) غیرمعمول هستند. ازاین‌رو، حداقل اجتناب از مطرح کردن ادعاها هنگام تعبیر تخمین‌های اثرات ثابت، مهم است.

۲-۵. تفاضلِ تفاضل: پیش و پس، تیمار و کنترل

استراتژی اثرات ثابت نیازمند داده‌های تابلویی است، این به معنای مشاهدات تکراری بر روی افراد یکسان است. اغلب، رگرسور مدنظر تنها در سطح کلی مانند وضعیت یا گروه تفاوت دارد. مثلاً سیاست‌های ایالتی

---

1. External Information

در رابطه با مزایای تیمار سلامت برای کارگران باردار یا حداقل دستمزد بین ایالت‌ها تغییر می‌کند، اما درون خود ایالت این گونه نیست. منبع اریب متغیرهای حذف شده هنگام ارزیابی این سیاست‌ها باید متغیرهای مشاهده نشده در ایالت یا سطح سال باشد.

برای مشخص شدن این موضوع، فرض کنید که ما به اثر حداقل دستمزد اشتغال که یک پرسش قدیمی در اقتصاد کار است، علاقه داریم. در بازار کار رقابتی، افزایش حداقل دستمزد ما را به بالای منحنی تقاضا با شیب نزولی می‌برد. حداقل‌های بالاتر اشتغال را کم می‌کنند و شاید به سیاست‌های حداقل دستمزد کارگران نیز آسیب بزنند. کارد و کروگر (۱۹۹۴) از تغییر جدی در حداقل دستمزد ایالت نیوجرسی استفاده می‌کنند تا ببینند آیا این مسئله درست یا خیر.

در اول آوریل ۱۹۹۲، ایالت نیوجرسی حداقل دستمزد را از ۴/۲۵ به ۵/۰۵ دلار افزایش داد. کارد و کروگر داده‌های اشتغال در رستوران‌های فست فود نیوجرسی در فوریه ۱۹۹۲ و سپس در نوامبر ۱۹۹۲ را جمع‌آوری کردند. این رستوران‌ها (برگر کینگ، وندیز و غیره) جزو بزرگ‌ترین رستوران‌هایی هستند که حداقل دستمزد را پرداخت می‌کنند. کارد و کروگر از همین نوع رستوران‌ها در شرق پنسیلوانیا درست در طرف دیگر رودخانه دلور داده‌ها را جمع‌آوری کردند. حداقل دستمزد در پنسیلوانیا در کل این دوره همچنان ۴/۲۵ دلار بود. آنها از مجموعه داده خود برای محاسبه تخمین‌های تفاضلِ تفاضل (**DD**) اثرات افزایش حداقل دستمزد نیوجرسی استفاده کردند. یعنی آنها تغییر اشتغال در نیوجرسی را با تغییر اشتغال در پنسیلوانیا حول زمان افزایش حداقل دستمزد نیوجرسی مقایسه کردند.

DD (تفاضلِ تفاضل) یک نسخه تخمین‌زن اثرات ثابت با استفاده از داده‌های کلی است. برای مشاهده این مسئله فرض می‌کنیم:

Y_{1ist} = اشتغال فست فودی در رستوران **i** و دوره **t** اگر حداقل دستمزد ایالت بالا باشد.

Y_{0ist} = اشتغال فست فودی در رستوران **i** و دوره **t** اگر حداقل دستمزد ایالت پایین باشد.

اینها خروجی‌های بالقوه هستند، اما در عمل، ما تنها یکی از آنها را مشاهده می‌کنیم. مثلاً Y_{1ist} را در نیوجرسی در نوامبر ۱۹۹۲ مشاهده می‌کنیم. قلب طرح **DD** یک ساختار جمع‌پذیر برای خروجی‌های بالقوه در وضعیت بدون تیمار است. به طور خاص، فرض می‌کنیم:

$$E(Y_{0ist}|s,t) = \gamma_s + \lambda_t \qquad (5.2.1)$$

که در آن **s** نماینده ایالت و **t** نماینده زمان است. این پرسش‌ها می‌گویند که در غیاب حداقل تغییر دستمزد، اشتغال توسط مجموع اثر زمان ایالت بدون توجه به زمان و **s** سال تعیین می‌شود که در بین ایالت‌ها مشترک است. تأثیر جمع‌پذیر ایالت نقش اثر منفرد مشاهده نشده در بخش فرعی قبلی را بازی می‌کند.

فرض کنیم D_{st} یک متغیر موهومی برای ایالت‌هایی باشد که حداقل دستمزد بالایی دارند، جایی که ایالت‌ها توسط نمایه **s** نشان داده شده و در دوره **t** مشاهده می‌شوند. فرض کنیم

$E(Y_{1ist} - Y_{0ist}|s,t)$ یک ثابت باشد که توسط β نشان داده می‌شود، داریم:

$$Y_{ist} = \gamma_s + \lambda_t + \beta D_{st} + \varepsilon_{ist} \qquad (5.2.2)$$

که در آن $E(\varepsilon_{ist}|s,t) = 0$. از اینجا داریم:

$$E[Y_{ist}|s = PA, t = Nov] - E(Y_{ist}|s = PA, t = Feb) = \lambda_{Nov} - \lambda_{Feb}$$

و

$$E(Y_{ist}|s = NJ, t = Nov) - E(Y_{ist}|s = NJ, t = Feb) = \lambda_{Nov} - \lambda_{Feb} + \beta.$$

تفاضلِ تفاضل‌های جامعه

$$[E(Y_{ist}|s = PA, t = Nov) - E(Y_{ist}|s = PA, t = Feb)]$$

$$- [E(Y_{ist}|s = NJ, t = Nov) - E(Y_{ist}|s = NJ, t = Feb)] = \beta,$$

تأثیر علّی مدنظر است. این را به راحتی می‌توان با استفاده از مشابه نمونه میانگین جامعه محاسبه کرد.

جدول ۲-۱-۵. میانگین اشتغال برای هر فروشگاه قبل و بعد از افزایش حداقل دستمزد نیوجرسی

Variable	PA (i)	NJ (ii)	Difference, NJ-PA (iii)
1. FTE employment before, all available observations	23.33 (1.35)	20.44 (0.51)	-2.89 (1.44)
2. FTE employment after, all available observations	21.17 (0.94)	21.03 (0.52)	-0.14 (1.07)
3. Change in mean FTE employment	-2.16 (1.25)	0.59 (0.54)	2.76 (1.36)

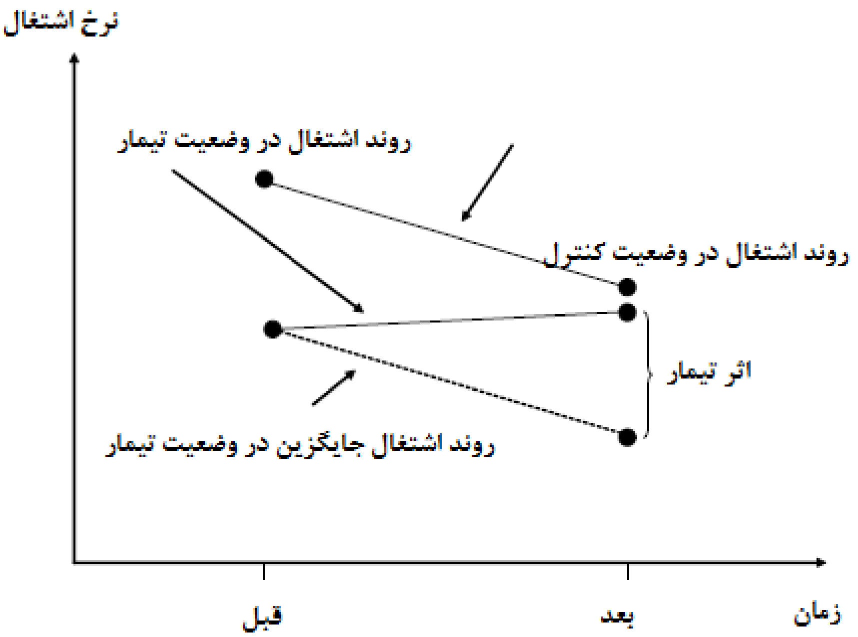
نکته: جدول ۱-۲-۵ (بر مبنای جدول ۳ کارد و کروگر، ۱۹۹۴) نشان‌دهنده میانگین اشتغال تمام‌وقت در رستوران‌های فست فود نیوجرسی و پنسیلوانیا قبل و بعد از تغییر در حداقل دستمزد نیوجرسی است. چهار سلول در دو ردیف و ستون اول و حاشیه‌ها که نشان‌دهنده تفاوت‌های ایالاتی در هر دوره، تغییرات در طول زمان در هر ایالت و تفاضلِ تفاضل‌ها هستند. اشتغال در رستوران‌های پنسیلوانیا به نوعی بیشتر از نیوجرسی در فوریه است، اما در نوامبر کمتر می‌شود. اشتغال در نیوجرسی، اما کمی بیشتر می‌شود. این دو تغییر تفاضلِ تفاضل‌های مثبت ایجاد می‌کنند در صورتی که اگر دستمزد بالاتر، کسب‌وکار را به بالای منحنی تقاضای کار ببرد، باید انتظار دیگری داشته باشیم.

این شاهد در برابر ماجرای تقاضای نیروی کار استاندارد چقدر متقاعدکننده است؟ فرضیه اصلی در اینجا این است که روندهای اشتغال در هر دو ایالت در غیاب این تیمار یکسان خواهند بود. تیمار اریب از این روند مشترک را منجر می‌شود که در شکل ۱-۲-۵ مشاهده می‌کنید. هرچند تیمار و کنترل با هم فرق دارند، اما این تفاوت توسط اثر ثابت ایالت مشخص می‌شود که همان نقش اثر منفرد مشاهده نشده در ۳-۱-۵ را بازی می‌کند.

فرضیه روندهای مشترک را می‌توان با استفاده از داده‌های دوره‌های مختلف بررسی کرد. کارد و

کروگر (۲۰۰۰) در به‌روزرسانی مطالعه قبلی خود یعنی حداقل دستمزد، داده‌های پرداخت اداری برای رستوران‌های نیوجرسی و پنسیلوانیا را برای مدت چند سال به دست آوردند. این داده‌ها را در شکل ۲-۲-۵ ملاحظه می‌کنید که مانند شکل ۲ در مطالعه بعدی آنهاست. خطوط عمومی نشان‌دهنده تاریخ‌هایی هستند که پیمایش اصلی آنها انجام شد و سومین خط عمومی نشان‌دهنده افزایش حداقل دستمزد فدرال به ۴/۷۵ دلار در اکتبر ۱۹۹۶ است که بر پنسیلوانیا تأثیر گذاشت، اما بر نیوجرسی تأثیری نداشت. این داده‌ها به ما این فرصت را می‌دهند تا نگاهی داشته باشیم به آزمایش جدید حداقل دستمزد.

شکل ۱-۲-۵. اثرات علّی در مدل تفاضلِ تفاضل‌ها



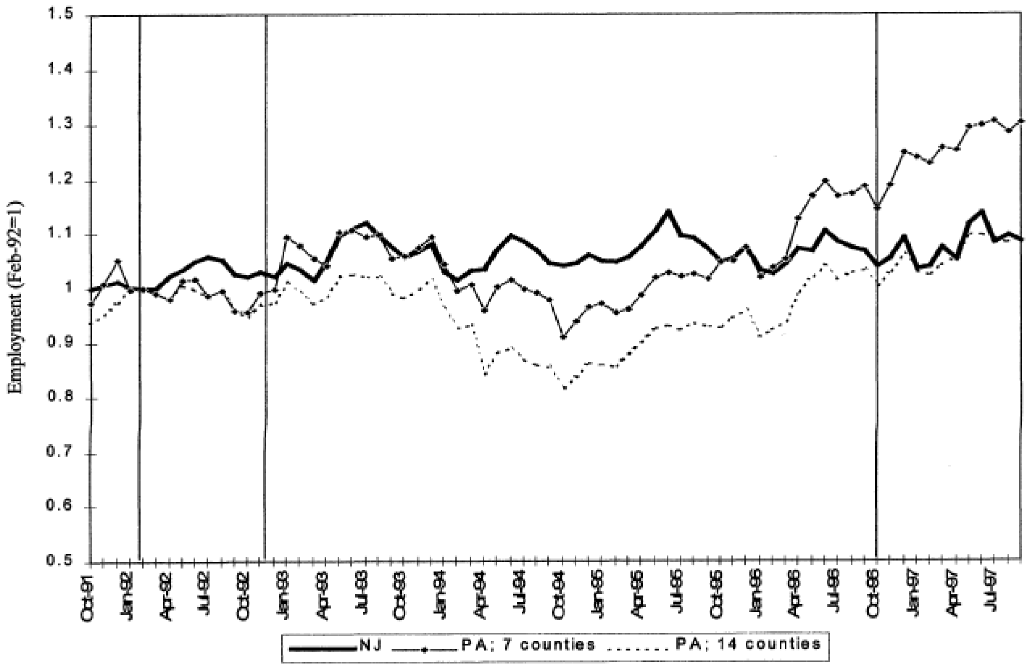
مانند پیمایش اصلی کارد و کروگر، داده‌های اجرایی نشان‌دهنده کاهش اندک اشتغال از فوریه تا نوامبر در پنسیلوانیا هستند و تغییر اندکی در نیوجرسی در همان دوره مشاهده می‌شود. هرچند، داده‌ها تغییر اشتغال سال به سال نسبتاً جدی را در دوره‌های دیگر نشان می‌دهند. این تغییرات اغلب در دو ایالت تفاوت زیادی دارند. به طور خاص، در حالی که سطوح اشتغال در نیوجرسی و پنسیلوانیا در انتهای ۱۹۹۱ یکسان بودند، اشتغال در پنسیلوانیا طی سه سال بعد، نسبت به اشتغال در نیوجرسی افت پیدا کرد که بیشتر قبل از تغییر حداقل دستمزد فدرال در سال ۱۹۹۶ بود. در نتیجه پنسیلوانیا ممکن است معیار خوبی برای نرخ اشتغال غیرحقیقی در نیوجرسی در غیاب تغییر خط‌مشی و بالعکس نباشد.

یک نمونه جذاب‌تر از سوی پیشکه (۲۰۰۷) مطرح می‌شود که به اثر طول ترم مدرسه بر عملکرد دانش‌آموز با استفاده از تغییر ایجاد شده توسط تغییر جدی سیاست در آلمان می‌پردازد. تا دهه ۱۹۶۰،

کودکان در ایالت‌های آلمان بجز باواریا مدرسه را در تابستان شروع می‌کردند. با شروع سال تحصیلی ۱۹۶۷ ـ ۱۹۶۶، آنهایی که در بهار شروع می‌کردند، مدرسه را در پاییز شروع کردند. گذار به شروع در پاییز دو سال تحصیلی کوتاه برای گروه‌های تحت تأثیر لازم داشت، یعنی طول ۲۴ هفته به جای ۳۷ هفته. دانش‌آموزان در این گروه‌ها نسبت به گروه‌های دو طرف دیگر و نسبت به دانش‌آموزان باواریا زمان خوبی داشتند که قبلاً در پاییز شروع می‌کردند.

شکل ۲-۲-۵ اشتغال در رستوران‌های فست فود نیوجرسی و پنسیلوانیا، اکتبر ۱۹۹۱ تا سپتامبر ۱۹۹۷ (از کارد و کروگر، ۲۰۰۰). خطوط عمودی نشان‌دهنده تاریخ‌های اصلی کارد و کروگر (۱۹۹۴) و افزایش حداقل دستمزد فدرال در اکتبر ۱۹۹۶ است.

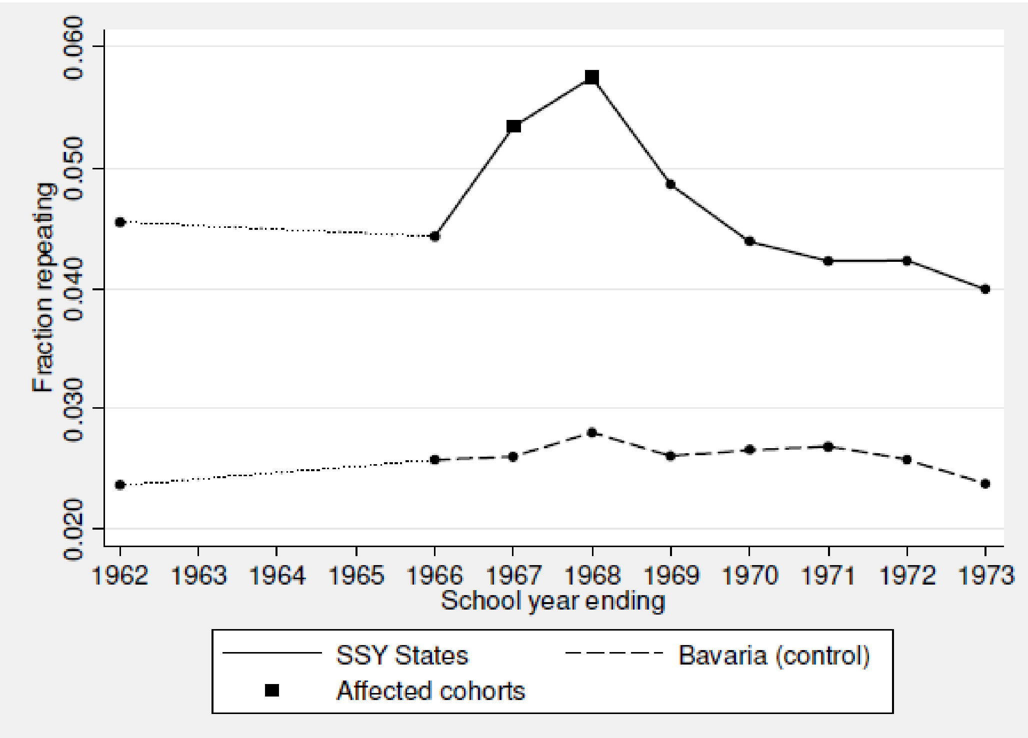
شکل ۲-۲-۵. اشتغال در رستوران‌های فست فود نیوجرسی و پنسیلوانیا، اکتبر ۱۹۹۱ تا سپتامبر ۱۹۹۷



شکل ۳-۲-۵ احتمال تکرار کلاس برای گروه‌های ۱۹۶۲ تا ۱۹۷۳ برای کلاس دومی‌های باواریا و ایالت‌های تحت تأثیر را نشان می‌دهد. نرخ تکرار در باواریا از ۱۹۶۶ به بعد حدود ۲/۵ درصد ثابت بود. نرخ تکرار در ایالت‌هایی که سال تحصیلی کوتاه‌تری داشتند، بالاتر بود حدود ۴ تا ۴/۵ درصد در سال ۱۹۶۲ و ۱۹۶۶ قبل از تغییر طول ترم، اما نرخ تکرار برای دو گروه تحت تأثیر در این ایالت‌ها تقریباً یک درصد بالاتر رفت، یک مقدار بیشتر از گروه دوم تا اول، که قبل از بازده سطح مبنا بود. این گراف شاهد تصویری قوی از وضعیت‌های تیمار و کنترل با روند یکسان مشترک و یک اثر تیمار است که انحراف تند، اما موقتی از این روند ایجاد می‌کند. به نظر می‌رسد، سال کوتاه‌تر مدرسه نرخ تکرار برای گروه‌های تحت تأثیر را افزایش داده است.

شکل ۳-۲-۵. نرخ میانگین تکرار پایه در کلاس دوم برای مدارس تیمار و کنترل در آلمان

(از پیشکه، ۲۰۰۷).



این داده‌ها برای دوره قبل و بعد تغییر طول ترم برای دانش‌آموزان خارجی باواریا هستند.

۱-۲-۵. تفاضلِ تفاضلِ رگرسیون

می‌توانیم مانند مدل اثرات ثابت، از رگرسیون برای تخمین مانند ۲-۲-۵ استفاده کنیم. فرض کنیم NJ_s متغیر موهومی برای رستوران‌ها در نیوجرسی و d_t متغیر موهومی زمان باشد که برای مشاهدات

به دست آمده در نوامبر فرض می‌شود (پس از تغییر حداقل دستمزد). پس،

$$\mathrm{Y}_{ist} = \alpha + \gamma NJ_s + \lambda d_t + \beta(NJ_s \cdot d_t) + \varepsilon_{ist} \qquad (5.2.3)$$

مانند ۲-۲-۵ است که در آن $NJ_s \cdot d_t = \mathrm{D}_{st}$. در زبان بخش ۴-۱-۳، این مدل شامل دو اثر اصلی برای ایالت و سال و عبارت تعامل است که نشان‌دهنده مشاهدات نیوجرسی در نوامبر است. این یک مدل اشباع شده است چون تابع میانگین شرطی $E(\mathrm{Y}_{ist}|s,t)$ ۴ مقدار احتمالی را می‌گیرد و چهار پارامتر وجود دارد. رابطه بین پارامترها در معادله رگرسیون ۳-۲-۵ و موارد موجود در مدل **DD** برای تابع میانگین شرطی ۲-۲-۵ به صورت زیر است:

$$\begin{aligned}
\alpha &= E(\mathrm{Y}_{ist}|s=PA, t=Feb) = \gamma_{PA} + \lambda_{Feb} \\
\gamma &= E(\mathrm{Y}_{ist}|s=NJ, t=Feb) - E(\mathrm{Y}_{ist}|s=PA, t=Feb) = \gamma_{NJ} - \gamma_{PA} \\
\lambda &= E(\mathrm{Y}_{ist}|s=PA, t=Nov) - E(\mathrm{Y}_{ist}|s=PA, t=Feb) = \lambda_{Nov} - \lambda_{Feb} \\
\beta &= \{E(\mathrm{Y}_{ist}|s=NJ, t=Nov) - E(\mathrm{Y}_{ist}|s=NJ, t=Feb)\} \\
&\quad -\{E(\mathrm{Y}_{ist}|s=PA, t=Nov) - E(\mathrm{Y}_{ist}|s=PA, t=Feb)\}.
\end{aligned}$$

فرمولاسیون رگرسیون مدل تفاضلِ تفاضل‌ها روش ساده‌ای برای ساخت تخمین‌های تفاضلِ تفاضل و خطاهای استاندارد ارائه می‌کند. همچنین افزودن ایالت‌ها (حالت‌ها) یا دوره‌های بیشتر به طرح رگرسیون ساده است. ممکن است به عنوان مثال، ایالت‌های کنترل یا دوره‌های پیش‌تیمار اضافی را به نمونه نیوجرسی/پنسیلوانیا اضافه کنیم. تعمیم نهایی ۳-۲-۵ شامل متغیر موهومی برای هر ایالت و دوره است، اما بدون تغییر می‌ماند.

کارد (۱۹۹۲) از تغییر منطقه‌ای در تأثیر حداقل دستمزد فدرال استفاده کرد. رویکرد او تحت تأثیر معادله‌ای مانند زیر ایجاد شد:

$$\mathrm{Y}_{ist} = \gamma_s + \lambda_t + \beta(\mathrm{FA}_s \cdot d_t) + \varepsilon_{ist} \qquad (5.2.4)$$

که در آن متغیر FA_sمعیار نسبتی از نوجوانان است که ممکن است تحت تأثیر افزایش حداقل دستمزد در هر ایالت قرار بگیرند و d_t متغیر موهومی برای مشاهدات پس از ۱۹۹۰، زمانی که حداقل دستمزد فدرال از ۳/۳۵ دلار به ۳/۸۰ دلار افزایش یافت، است. متغیر FA_s نسبت (پیش از افزایش) خط مبنای نیروی کار نوجوان هر ایالت را اندازه‌گیری می‌کند که کمتر از ۳/۸۰ دلار درآمد دارد.

مانند مطالعه نیوجرسی/پنسیلوانیا، کارد (۱۹۹۲) با داده‌های دو دوره، قبل و بعد، در این مورد ۱۹۸۹ و ۱۹۹۲ کار می‌کند، اما این مطالعه از ۵۱ ایالت (شامل منطقه کلمبیا) برای مجموع ۱۰۲ مشاهده ایالت ـ سال استفاده می‌کند. چون متغیرهای مشترک در سطح منفرد در ۴-۲-۵ وجود ندارد، این امر مانند تخمین با داده‌های خُرد است. توجه داشته باشید که $\mathrm{FA}_s \cdot d_t$ مانند $NJ_s \cdot d_t$ در ۳-۲-۵

عبارت تعامل[1] است، هرچند در اینجا عبارت تعامل مقدار متمایزی برای هر مشاهده در مجموعه داده می‌گیرد. در نهایت، چون کارد (۱۹۹۲) داده‌ها را تنها برای دو دوره تحلیل می‌کند، تخمین‌های گزارش شده از یک معادله در اولین تفاوت‌ها هستند:

$$\Delta\bar{Y}_s = \lambda^* + \beta FA_s + \Delta\bar{\varepsilon}_s,$$

که در آن $\Delta\bar{Y}_s$ تغییر میانگین اشتغال نوجوانان در ایالت **s** و $\Delta\bar{\varepsilon}_s$ عبارت خطا در معادله دیفرانسیل‌گیری شده است.

جدول ۲-۲-۵ بر اساس جدول ۳ در کارد (۱۹۲۲)، نشان می‌دهد که دستمزدها در ایالت‌هایی بیشتر شده‌اند که در آنها افزایش حداقل دستمزد به گزندگی بیشتر منجر شده است. این یک مرحله مهم در تحلیل کارد است ـ این انگاره را تأیید می‌کند که متغیر کسر تحت تأثیر پیش‌بینی‌کننده خوبی برای تغییرات دستمزد ایجاد شده توسط افزایش حداقل فدرال است. اشتغال اما از سوی دیگر، ارتباطی به کسر تحت تأثیر ندارد که در ستون ۳ مشاهده می‌کنید. ازاین‌رو، نتایج در کارد (۱۹۹۲) با نتایج مطالعه نیوجرسی/پنسیلوانیا هم‌راستا هستند.

جدول ۲-۲-۵. تخمین‌های رگرسیون ـ تفاضلِ تفاضل تأثیر حداقل دستمزد بر نوجوانان، ۱۹۸۹ تا ۱۹۹۲

Explanatory Variable	Equations for Change in Mean Log Wage: (1)	(2)	Equations for change in Teen Employment-Population Ratio: (3)	(4)
1. Fraction of Affected Teens	0.15	.14	0.02	-.01
	(0.03)	(0.04)	(0.03)	(0.03)
2. Change in Overall Emp./Pop. Ratio	–	0.46	–	1.24
		(0.60)		(0.60)
3. R-squared	0.30	0.31	0.01	0.09

نکته: برگرفته از کارد (۱۹۹۲). این جدول، برآوردهای رگرسیون تغییرات در متوسط اشتغال نوجوانان در هر ایالت بر روی نسبت نوجوانان متأثر از تغییرات حداقل دستمزد ایالتی در هر ایالت را گزارش می‌کند. داده‌ها از پیمایش **CPS** سال‌های ۱۹۸۹ و ۱۹۹۲ است. رگرسیون‌ها توسط حجم نمونه هر ایالت و هر سال وزن‌دهی شده‌اند.

تحلیل کارد (۱۹۹۲) نشان‌دهنده مزیت دیگر رگرسیون تفاضلِ تفاضل است: افزودن متغیرهای مشترک بیشتر در این چارچوب ساده است. مثلاً، ممکن است بخواهیم اشتغال بزرگسالان را به عنوان منبع روندهای ویژه و حذف شده ایالت کنترل کنیم. به عبارت دیگر، می‌توانیم اشتغال غیرحقیقی را در غیاب تغییر در حداقل دستمزد مدلسازی کنیم:

$$E[Y_{0ist}|s,t,X_{st}] = \gamma_s + \lambda_t + X'_{st}\delta.$$

که در آن X_{st} بردار متغیرهای مشترک است که بر اساس ایالت و زمان تغییر می‌کنند و شامل

1. Interaction Term

اشتغال بزرگسالان نیز هست. همان طور که مشخص است، افزودن کنترل اشتغال بزرگسالان تأثیر چندانی بر تخمین‌های کارد ندارد که در ستون‌های ۲ و ۴ جدول ۲-۲-۵ مشخص است.

شایان تأکید است که کارد (۱۹۹۲) میانگین‌های ایالت را به جای داده‌های منفرد تحلیل می‌کند. او ممکن است از نمونه چندساله تجمعی داده‌های خُرد از **CPS** برای تخمین یک معادله مانند زیر استفاده کرده باشد:

$$Y_{ist} = \gamma_s + \lambda_t + \beta(FA_s \cdot d_t) + X'_{ist}\delta + \varepsilon_{ist}, \qquad (5.2.5)$$

که در آن X_{ist} می‌تواند شامل خصوصیات سطح منفرد مانند نژاد باشد. بردار متغیر مشترک ممکن است شامل متغیرهایی باشد که با توجه به زمان تغییر می‌کنند و در سطح ایالت اندازه‌گیری می‌شوند. تنها مورد آخر می‌تواند منبع اریب متغیرهای حذف شده باشد، اما کنترل‌های سطح منفرد می‌توانند دقت را افزایش دهند که در بخش ۳-۲ نیز بدان اشاره کردیم. استنباط در چارچوبی که داده‌های خُرد متغیرهای وابسته را با رگرسورهای سطح گروه ترکیب می‌کند، کمی پیچیده‌تر است. مسئله کلیدی این است که چطور به بهترین شکل اثرات تصادفی سطح گروه را جورشدگی دهیم.

هنگامی که نمونه شامل سال‌های زیادی است، مدل رگرسیون **DD** در سایه گرنجر (۱۹۶۹) تحت تست قرار می‌گیرد. ایده گرنجر این است که ببیند آیا علل قبل از عواقب رخ می‌دهند و نه بالعکس. فرض کنید متغیر سیاست مدنظر D_{st}، در زمان‌های مختلف در ایالت‌های مختلف تغییر کند. در این بافت، تست علّیت گرنجر یعنی بررسی این نکته که آیا اثرات مشروط ایالت و سال، قبل از D_{st}، پیش‌بینی‌کننده Y_{ist} است در حالی که D_{st} آینده این گونه نیست. اگر D_{st} باعث Y_{ist} شود و بالعکس آن صحیح نباشد، نتیجه نباید در معادله‌ای مانند زیر مهم باشد:

$$Y_{ist} = \gamma_s + \lambda_t + \sum_{\tau=0}^{m} \beta_{-\tau} D_{s,t-\tau} + \sum_{\tau=1}^{q} \beta_{+\tau} D_{s,t+\tau} X_{ist}\delta + \varepsilon_{ist}, \qquad (5.2.6)$$

که در آن مجموع طرف سمت راست اجازه **m** تأخیر[1] $(\beta_{-1}, \beta_{-2}, \ldots, \beta_{-m})$ یا اثرات پس از تیمار و **q** تقدم $(\beta_{+1}, \beta_{+1}, \ldots, \beta_{+q})$ یا اثرات مقدم[2] یا پیش‌بینی‌کننده را ممکن می‌کند. الگوی اثرات تأخیر معمولاً جذاب است. مثلاً ممکن است باور کنیم که اثرات علّی باید با گذر زمان رشد کرده یا از بین بروند.

آتور (۲۰۰۳) تست گرنجر را در بررسی اثر حفاظت اشتغال بر استفاده از کمک موقتی شرکت اجرا کرد. حفاظت اشتغال نوعی قانون کار است که توسط قوه مقننه ایالت ایجاد و اخراج کارگران را دشوارتر می‌کند. به عنوان یک قانون، قانون کار ایالات متحده آمریکا «اشتغال با میل»[3] را اجازه می‌دهد که یعنی کارگران را می‌توان با دلیل یا بی‌دلیل اخراج کرد، اما برخی دادگاه‌های ایالتی تعداد استثنا را برای دکترین اشتغال با میل در نظر گرفته‌اند که به پیگرد قانونی برای اخراج ناعادلانه منجر می‌شود. آتور به

1. Lag
2. Lead
3. Employment at Will

این نکته علاقه‌مند است که آیا ترس از شکایت کارمند باعث می‌شود شرکت‌ها از کارگران موقتی برای کارهایی استفاده کنند که در غیر این صورت باید نیروی کار خود را افزایش دهند. کارگران موقتی در کنار شرکت برای فرد دیگری نیز کار می‌کنند. در نتیجه، شرکتی که از آنها استفاده می‌کند را نمی‌توان تحت پیگرد قرار داد.

استراتژی تجربی آتور، اشتغال کارگران موقتی در یک ایالت را به متغیرهای موهومی ربط می‌دهد که نشان می‌دهند قوانین دادگاه ایالتی استثنائاتی برای دکترین اشتغال با میل در نظر گرفته‌اند. مدل رگرسیون **DD** او شامل تقدم و تأخر است مانند معادله ۶-۲-۵. این تقدم و تأخر تخمین زده شده از دو سال جلوتر تا چهار سال عقب‌تر ادامه دارد که در شکل ۴-۲-۵ مشاهده می‌کنید که بازتولید شکل ۳ از آتور (۲۰۰۳) است. تخمین‌ها نشان‌دهنده اثرات در دو سال قبل از اتخاذ استثنا از سوی دادگاه نیستند و اثرات در اشتغال موقتی در چند سال اول این تصمیم رشد شدیدی دارد که سپس با نرخ بالای اشتغال موقتی در ایالت‌های تحت تأثیر، به شکل ثابت درمی‌آید. این الگو با تعبیر علّی نتایج آتور همخوانی دارد.

در یک طرح جایگزین، استراتژی شناسایی تفاضلِ تفاضل، روندهای زمانی ویژه ایالت را به رگرسورهای X_{ist} اضافه می‌کند. به عبارت دیگر، ما به شکل زیر تخمین می‌زنیم:

$$Y_{ist} = \gamma_{0s} + \gamma_{1st} + \lambda_t + \beta D_{st} + X'_{ist}\delta + \varepsilon_{ist}, \qquad (5.2.7)$$

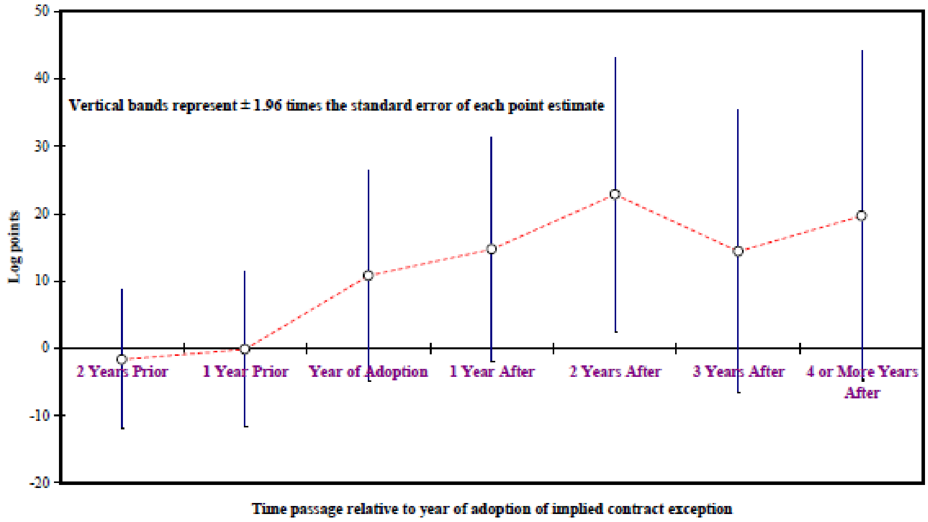
که در آن γ_{0s} عرض از مبدأ ویژه ایالت همانند قبل و γ_{1s} ضریب روند ویژه ایالت، ضرب در متغیر روند زمان **t** است. این امر، تیمار و کنترل برای دنبال کردن روندهای مختلف در حالتی محدود، اما آشکارکننده را ممکن می‌کند. درک اینکه آثار تخمینی مدنظر توسط شمول این روندها تغییری نمی‌کنند، ساده است. توجه داشته باشید که حداقل سه دوره را برای تخمین یک مدل با روندهای ویژه ایالت لازم داریم. علاوه بر این، در عمل، سه دوره برای درج هر دو روند و اثر تیمار نامناسب است. به عنوان یک قانون، تخمین تفاضلِ تفاضل با روندهای ویژه ایالت قوی‌تر و متقاعدکننده‌تر است زمانی که داده‌های پیش‌تیمار روند دقیقی را ایجاد کنند که بتوان آن را در دوره پس از تیمار نیز در اختیار داشت.

در مطالعه اثر مقررات کار بر کسب‌وکارها در ایالت‌های سرخپوستی، بسلی و برگس (۲۰۰۴) از روندهای ایالتی به عنوان چک قدرت استفاده کردند. ایالت‌های مختلف نظام‌های نظارتی را در زمان‌های مختلف تغییر دادند که باعث ایجاد طرح تحقیق تفاضلِ تفاضل (**DD**) شد. مانند کارد (۱۹۹۲)، واحد مشاهده بسلی و برگس (۲۰۰۴) میانگین ایالت ـ سال است. جدول ۳-۲-۵ نتایج کلیدی را نشان می‌دهد.

تخمین جدول ۱ از مدل رگرسیون **DD** بدون روندهای ویژه ایالت، نشان می‌دهد که مقررات کار به سرانه خروجی کمتر منجر می‌شود. این مدل‌ها برای ساخت تخمین‌ها در ستون‌های ۲ و ۳ استفاده شدند و متغیرهای مشترک ویژه ایالت که با زمان تغییر می‌کنند مانند سرانه هزینه دولت و جامعه ایالت را اضافه

می‌کنند. این امر در سایه افزودن نرخ اشتغال بزرگسالان در سطح ایالت به عنوان کنترل در مطالعه حداقل دستمزد است که توسط کارد (۱۹۹۲) انجام شد. افزودن کنترل‌ها بر تخمین‌های بسلی و برگس تأثیر اندکی داشت، اما افزودن روندهای ویژه ایالت مانند تأثیر مقررات کار را می‌توان در جدول ۴ مشاهده کرد. مشخصاً، مقررات نیروی کار در هندوستان در ایالت‌ها افزایش پیدا می‌کند که در آن خروجی در هر صورت کم می‌شود. کنترل برای این روند باعث می‌شود که تأثیر مقررات تخمینی صفر باشد.

شکل ۴-۲-۵. تأثیر تخمینی اتخاذ استثنای قرارداد از سوی دادگاه برای دکترین اشتغال با میل بر استفاده از کارگران موقتی (از آتور، ۲۰۰۳)



در این شکل، متغیر وابسته لگاریتم اشتغال کمک موقتی ایالتی در ۱۹۷۹ تا ۱۹۹۵ است. تخمین‌ها از مدلی هستند که اثرات قبل، طی و پس از اتخاذ تصمیم را ممکن می‌کند.

جدول ۳-۲-۵. تأثیر مقررات کار بر عملکرد شرکت‌ها در ایالت‌های هندوستان

	(1)	(2)	(3)	(4)
Labor regulation (lagged)	-0.186	-0.185	-0.104	0.0002
	(.0641)	(.0507)	(.039)	(.02)
Log development		0.240	0.184	0.241
expenditure per capita		(.1277)	(.1187)	(.1057)
Log installed electricity		0.089	0.082	0.023
capacity per capita		(.0605)	(.0543)	(.0333)
Log state population		0.720	0.310	-1.419
		(.96)	(1.1923)	(2.3262)
Congress majority			-0.0009	0.020
			(.01)	(.0096)
Hard left majority			-0.050	-0.007
			(.0168)	(.0091)
Janata majority			0.008	-0.020
			(.0235)	(.0333)
Regional majority			0.006	0.026
			(.0086)	(.0234)
State-specific trends	NO	NO	NO	YES
Adjusted R-squared	0.93	0.93	0.94	0.95

نکته: برگرفته از جدول ۴ بسلی و برگس (۲۰۰۴). این جدول برآوردهای رگرسیون تفاضل تفاضل تأثیرات قاعده‌گذاری نیروی کار بر بهره‌وری را نشان می‌دهد. متغیر وابسته، لگاریتم سرانه تولید کارخانه‌ای (Manufacturing Output) است. تمام مدل‌های شامل تأثیرات ایالت و سال هستند. خطاهای استاندارد استوار خوشه‌بندی شده در سطح ایالتی در پرانتز نوشته شده است. اصلاحات ایالتی به قانون مشاجرات صنعتی به این صورت کد شده‌اند که اگر به نفع کارگر باشند، عدد ۱، اگر خنثی باشند صفر و اگر به نفع کارفرما باشد، منفی یک می‌گیرند و در طول آن دوره جمع می‌شوند تا سنجه قاعده‌گذاری کار را حاصل کنند. لگاریتم ظرفیت الکتریسیته نصب شده با کیلووات سنجیده می‌شود و لگاریتم مصارف عمرانی شامل مخارج سرانه ایالتی در امور اقتصادی و خدماتی است. کنگره، چپ سخت، جاناتا و اکثریت محلی تعداد سال‌هایی هستند که این گروه‌های سیاسی اکثریت کرسی‌های پارلمان ایالتی را در اختیار داشتند. داده‌های ۱۶ ایالت اصلی دوره سال‌های ۱۹۵۸ تا ۱۹۹۲ را دربرمی‌گیرد که شامل ۵۵۲ مشاهده است.

انتخاب کنترل‌ها

ما دو بُعد را در طرح «ایالت‌ها» و «زمان» **DD** نامگذاری کردیم چون این یک نمونه **DD** اولیه در اقتصادسنجی کاربردی است، اما ایده **DD** کلی‌تر است. به جای ایالت‌ها، زیروند **s** نشان‌دهنده گروه‌های مردم‌شناختی است که برخی از آنها تحت تأثیر سیاست هستند و برخی نیستند. مثلاً، کاگلر، جیمنو و هرنانز (۲۰۰۵) تأثیرات سیاست‌های حفاظت اشتغال ویژه سن را در اسپانیا بررسی کردند. همچنین، به جای زمان، ممکن است داده‌ها را با انواع گروه یا سایر انواع خصوصیات گروه‌بندی کنند. یک نمونه آنگریست و ایوانز (۱۹۹۹) است که تأثیر تغییرات در قوانین سقط جنین ایالتی برای حاملگی نوجوانان با استفاده از تغییر توسط ایالت و سال تولد بررسی کردند. هرچند طرح‌های **DD** همیشه مقایسه تیمار ـ کنترل ضمنی دارند. پرسش اینکه آیا این مقایسه خوب است، ملاحظات جدی لازم دارد.

اصلاح دو به دو طرح **DD** از قراردادهای مرتبه بالاتر استفاده می‌کند تا به استنباط‌های علّی برسد. یک نمونه بسط پوشش برنامه بیمه درمانی مدیکید (**Medicaid**) در ایالات متحده است که توسط

یلوویتز (۱۹۹۵) مطالعه شد. شایستگی (صلاحیت) برای قرار گرفتن زیر پوشش مدیکید، برنامه بیمه گسترده در ایالات متحده برای فقرا، زمانی به صلاحیت **AFDC**[1] ارتباط داشت که برنامه رفاه نقدی گسترده بود. در زمان‌های مختلف در دهه ۱۹۸۰، برخی ایالت‌ها پوشش مدیکید را به کودکان در خانواده‌هایی بسط دادند که برای **AFDC** واجد شرایط نبودند. یلوویتز به این نکته علاقه داشت که چطور مشارکت و درآمدهای مادران کارگر تحت تأثیر قرار گرفته است.

علاوه بر ایالت و زمان، سن کودکان سومین بُعد را اضافه می‌کند که در آن سیاست مدیکید تغییر می‌کند. یلوویتز از این تغییر با تخمین رابطه زیر استفاده می‌کند:

$$\mathrm{Y}_{iast} = \gamma_{st} + \lambda_{at} + \theta_{as} + \beta \mathrm{D}_{ast} + \mathrm{X}_{iast}\delta + \varepsilon_{iast},$$

که در آن **s** نمایه ایالت‌هاست، **t** نمایه زمان و **a** سن کوچک‌ترین فرزند خانواده است. این مدل کنترل کامل غیرپارامتری را برای اثرات زمانی ویژه ایالت فراهم می‌کند که در همه گروه‌های سنی مشترک هستند (γ_{st})، اثرات سن که با زمان تغییر می‌کنند (λ_{at}) و اثرات سن ویژه ایالت (θ_{as}). رگرسور مدنظر، D_{ast}، نشان‌دهنده کودکان در گروه‌های سنی تحت تأثیر در ایالت‌ها و دوره‌هایی است که در آن پوشش فراهم شده است. این مدل تفاوت‌های سه‌گانه مجموعه متقاعدکننده‌تری از نتایج ارائه می‌کند تا تحلیل سنتی **DD** که از تفاوت‌های ایالت و زمان به تنهایی استفاده می‌کند.

۳-۵. اثرات ثابت در برابر متغیرهای وابسته با تأخیر[2]

تخمین‌گرهای اثرات ثابت و تفاضلِ تفاضل‌ها بر اساس پیش‌فرض متغیرهای حذف شده‌ای هستند که با زمان تغییر نمی‌کردند. فرض کنید ما به اثرات مشارکت در برنامه آموزشی یارانه‌ای علاقه داریم، مانند دهجیا و وهبا (۱۹۹۹) و لالوند (۱۹۸۶) که در بخش (۳-۳-۳) مطالعات آنها را بررسی کردیم. فرضیه اصلی که تخمین اثرات ثابت را منجر می‌شود، در این مورد به عبارت زیر است:

$$E(\mathrm{Y}_{0it}|\alpha_i, \mathrm{X}_{it}, \mathrm{D}_{it}) = E(\mathrm{Y}_{0it}|\alpha_i, \mathrm{X}_{it}), \qquad (5.3.1)$$

که در آن α_i خصوصیت منفرد مشاهده نشده است که در کنار متغیرهای مشترک، X_{it}، مشخص می‌کند که آیا فرد **i** آموزش را دریافت می‌کند یا خیر. یعنی، α_i ممکن است معیار مهارت‌های حرفه‌ای باشد هرچند ضربه در برابر طرح اثرات ثابت این حقیقت است که ماهیت دقیق متغیرهای مشاهده نشده معمولاً رازآلود باقی می‌ماند. در هر حالت، به همراه مدل خطی برای $E(\mathrm{Y}_{0it}|\alpha_i,\mathrm{X}_{it})$، فرضیه ۱-۳-۵ به استراتژی‌های تخمین ساده منجر می‌شود که شامل تفاوت‌ها یا اریب از میانگین است.

برای بسیاری از پرسش‌ها، این گمان که مهم‌ترین متغیرهای حذف شده ربطی به زمان ندارند، منطقی به نظر نمی‌رسد. ارزیابی برنامه‌های آموزشی موردی است که باید بررسی شود. به نظر می‌رسد

1. Aid to Families with Dependent Children
2. Fixed Effects Versus Lagged Dependent Variables

که مردم به دنبال بهبود گزینه‌های خود در بازار کار هستند و این را با شرکت در برنامه‌های آموزشی یارانه‌ای دولتی انجام می‌دهند که کمی ایراد دارند. بسیاری از برنامه‌های آموزشی افرادی را هدف قرار می‌دهند که از یک مشکل اخیر رنج می‌برند مثلاً مردانی که به تازگی کار خود را از دست داده‌اند. هم‌راستا با این مسئله، آشنفلتر (۱۹۷۸) و آشنفلتر و کارد (۱۹۸۵) متوجه شدند که شرکت‌کنندگان در برنامه آموزشی معمولاً تاریخچه درآمدی دارند که نشان‌دهنده افت پیش از برنامه است. درآمدهای گذشته محدودکننده وابسته به زمان هستند که نمی‌توان آنها را در متغیر مستقل از زمان مانند α_i مفروض داشت.

تاریخچه درآمد کارآموزان باعث تشکیل استراتژی تخمینی می‌شود که درآمدهای گذشته را مستقیماً کنترل می‌کند و با اثرات ثابت توزیع می‌شود. یعنی، به جای ۱-۳-۵ باید استنباط علّی خود را معطوف فرضیه استقلال شرطی کنیم:

$$E(Y_{0it}|Y_{it-h}, X_{it}, D_{it}) = E(Y_{0it}|Y_{it-h}, X_{it}). \qquad (5.3.2)$$

این مانند این است که بگوییم چیزی که کارآموزان را خاص می‌کند، درآمد آنها در h دوره قبلی است. سپس می‌توانیم از داده‌های تابلویی برای تخمین زیر استفاده کنیم:

$$Y_{it} = \alpha + \theta Y_{it-h} + \lambda_t + \beta D_{it} + X_{it}\delta + \varepsilon_{it}, \qquad (5.3.3)$$

که در آن اثر علّی آموزش β است. برای اینکه آن را کلی‌تر کنیم، Y_{it-h} می‌تواند برداری باشد که شامل درآمدهای با تأخیر برای دوره‌های مختلف است.

محققان کاربردی که از داده‌های تابلویی استفاده می‌کنند، اغلب با چالش انتخاب بین اثرات ثابت و مدل‌های متغیرهای وابسته به تأخیر مواجه می‌شوند یعنی بین استنباط‌های علّی بر اساس ۱-۳-۵ و ۲-۳-۵. یک راه‌حل برای این مسئله کار با مدلی است که شامل متغیرهای وابسته با تأخیر و اثرات منفرد مشاهده نشده باشد. به عبارت دیگر، شناخت ممکن است بر اساس فرضیه مستقل شرطی ضعیف‌تر باشد:

$$E(Y_{0it}|\alpha_i, Y_{it-h}, X_{it}, D_{it}) = E(Y_{0it}|\alpha_i, Y_{it-h}, X_{it}), \qquad (5.3.4)$$

که نیازمند مشروط‌سازی بر روی α_i و Y_{it-h} است. ما می‌توانیم اثرات علّی را با استفاده از یک تصریح مانند زیر تخمین بزنیم:

$$Y_{it} = \alpha_i + \theta Y_{it-h} + \lambda_t + \beta D_{it} + X_{it}\delta + \upsilon_{it}. \qquad (5.3.5)$$

متأسفانه، شرایط برای تخمین پیوسته β در رابطه ۵-۳-۵ دشوارتر از موارد مورد نیاز با اثرات ثابت یا تنها متغیرهای وابسته با تأخیر هستند. این را می‌توان در یک نمونه ساده دید که در آن متغیر وابسته با تأخیر Y_{it-1} است. ما اثر ثابت را با دیفرانسیل‌گیری از بین می‌بریم که رابطه زیر را ایجاد می‌کند:

$$\Delta Y_{it} = \theta \Delta Y_{it-1} + \Delta\lambda_t + \beta \Delta D_{it} + \Delta X_{it}\delta + \Delta\upsilon_{it}. \qquad (5.3.6)$$

مسئله در اینجا این است که پسماند دیفرانسیل‌گیری شده $\Delta\upsilon_{it}$ الزاماً با متغیر وابسته با تأخیر

ΔY_{it-1} رابطه دارد چون هر دو تابع v_{it-1} هستند. در نتیجه، تخمین‌های **OLS** برای ۶-۳-۵ برای پارامترهای ۵-۳-۵ همخوانی ندارند که این مسئله را برای اولین بار نیکل (۱۹۸۱) مطرح کرد. این مسئله را می‌توان حل کرد هرچند راه‌حل نیازمند فرضیات قوی است. ساده‌ترین راه‌حل استفاده از Y_{it-2} به عنوان ابزاری برای ΔY_{it-1} در ۶-۳-۵ است، اما این امر نیازمند آن است که Y_{it-2} به پسماندهای دیفرانسیل‌گیری شده ربطی نداشته باشد Δv_{it}. این امر به نظر محتمل نیست چون پسماندها بخشی از درآمد پسماند پس از محاسبه برای متغیرهای مشترک هستند. بیشتر درآمد مردم از یک سال تا سال بعدی همبستگی دارد. در نتیجه درآمدها پیش‌بینی‌کننده عالی درآمدهای آینده و رشد درآمدها هستند. اگر v_{it} همبستگی رشته‌ای داشته باشد، ممکن است تخمین‌گر پیوسته برای ۶-۳-۵ وجود نداشته باشد.

با توجه به دشواری‌هایی که هنگام تلاش برای تخمین ۶-۳-۵ ایجاد می‌شود، ممکن است بپرسید که آیا تمایز بین اثرات ثابت و متغیرهای وابسته با تأخیر اهمیت دارد یا خیر. متأسفانه پاسخ «آری» است. اثرات ثابت و مدل‌های متغیرهای وابسته با تأخیر لانه‌بندی شده نیستند، یعنی نمی‌توان امید داشت که یکی را تخمین زد و در صورت نیاز، دیگری را به عنوان یک مورد ویژه به دست آورد. تنها مدل کلی و مدلی که سخت‌تر شناسایی می‌شود ۵-۳-۵، اثرات ثابت و متغیرهای وابسته با تأخیر را آشیانه[1]بندی می‌کند.[2]

پس یک محقق کاربردی باید چه کار کند؟ یک پاسخ بررسی قدرت یافته‌ها با استفاده از فرضیات شناسایی جایگزین است. یعنی اینکه شما باید نتایج مشابه را با استفاده از هر دو مدل پیدا کنید. تخمین‌های اثرات ثابت و متغیرهای وابسته با تأخیر ویژگی براکت‌بندی سودمندی دارند. پیوست این فصل نشان می‌دهد که اگر ۲-۳-۵ صحیح باشد، اما اگر شما به اشتباه از اثرات ثابت استفاده کرده باشید، تخمین‌های اثر تیمار مثبت بیش از حد بزرگ خواهند بود. از سوی دیگر، اگر ۱-۳-۵ صحیح باشد و به اشتباه رابطه را با خروجی‌های با تأخیر مانند ۳-۳-۵ تخمین بزنید، تخمین‌های اثر تیمار مثبت بیش از حد کوچک خواهند بود. ازاین‌رو می‌توان اثرات ثابت و متغیرهای وابسته با تأخیر را به عنوان تأثیر علّی در نظر گرفت که به دنبالش هستید. گوریان (۲۰۰۴) نشان‌دهنده این نوع استدلال در مطالعه‌ای است که اثرات حمل با دستور دادگاه بر نرخ فارغ‌التحصیلی دبیرستان سیاهپوستان را تخمین می‌زند.

۴-۵. ضمیمه: موارد بیشتری در مورد اثرات ثابت و متغیر وابسته تأخیری

برای ساده‌سازی، متغیرهای کمکی و تأثیرات سالیانه را نادیده می‌گیریم و فرض می‌کنیم که تنها دو دوره وجود دارد که تیمار همه افراد در دور اول مساوی صفر است (اصل مطلب همانند تنظیمات عمومی

---

1. Nesting

۲. به ویژه اینکه تنظیمات $\theta = 1$ در رابطه ۳-۳-۵، الگوي با تأثيرات ثابت را به عنوان مورد ويژه‌اي از الگوي با متغير وابسته تأخيري به دست نمي‌دهد. در عوض داريم:

$$\Delta Y_{it} = \alpha + \lambda_t + \beta D_{it} + X_{it}\delta + \varepsilon_{it}$$

به معناي متغير وابسته تفاضلي با رگرسورهايي در سطح. اين الگوي با يك تفاضل در هر دو سمت راست و چپ، مورد نيز براي از بين بردن تأثيرات ثابت، نيست.

است). تأثیر علّی مورد نظر یعنی β مثبت است. فرض می‌کنیم که تیمار با یک اثر منفرد مشاهده نشده، a_i، همبسته است و خروجی می‌تواند به صورت زیر توضیح داده شود:

$$\mathrm{Y}_{it} = a_i + \beta \mathrm{D}_{it} + \varepsilon_{it}. \qquad (5.4.1)$$

که در آن ε_{it} به صورت رشته‌ای (سریالی) ناهمبسته بوده و با a_i و D_{it} ناهمبسته است. همچنین داریم:

$$\mathrm{Y}_{it-1} = a_i + \varepsilon_{it-1},$$

که در آن a_i و ε_{it-1} ناهمبسته هستند. برآورد اثر D_{it} در مدلی که Y_{it-1} را کنترل کرده، اما اثرات ثابت را نادیده می‌گیرد اشتباه است. حد احتمال برآوردگر حاصل $\frac{Cov(\mathrm{Y}_{it},\bar{\mathrm{D}}_{it})}{V(\bar{\mathrm{D}}_{it})}$ است، که در آن $\bar{\mathrm{D}}_{it} = \mathrm{D}_{it} - \gamma \mathrm{Y}_{it-1}$ پسماند رگرسیون D_{it} روی Y_{it-1} است.

اکنون با جایگذاری $a_i = \mathrm{Y}_{it-1} - \varepsilon_{it-1}$ در رابطه ۱-۴-۵ به رابطه زیر می‌رسیم:

$$\mathrm{Y}_{it} = \mathrm{Y}_{it-1} + \beta \mathrm{D}_{it} + \varepsilon_{it} - \varepsilon_{it-1}.$$

از اینجا داریم:

$$\frac{Cov(\mathrm{Y}_{it},\bar{\mathrm{D}}_{it})}{V(\bar{\mathrm{D}}_{it})} = \beta - \frac{Cov(\varepsilon_{it-1},\bar{\mathrm{D}}_{it})}{V(\bar{\mathrm{D}}_{it})} = \beta - \frac{Cov(\varepsilon_{it-1},\mathrm{D}_{it} - \gamma \mathrm{Y}_{it-1})}{V(\bar{\mathrm{D}}_{it})} = \beta + \frac{\gamma \sigma_\varepsilon^2}{V(\bar{\mathrm{D}}_{it})}.$$

که در آن σ_ε^2 واریانس ε_{it-1} است. به دلیل Y_{it-1} پایین آموزش‌گیرندگان، $\gamma < 0$ و برآورد منتج از β خیلی کم است.

فرض می‌کنیم که تیمار توسط Y_{it-1} پایین تعیین می‌شود. تشخیص (رابطه شناسایی و مشخص‌سازی) درست، نسخه ساده شده‌ای از رابطه ۳-۳-۵ است. به این صورت:

$$\mathrm{Y}_{it} = \alpha + \theta \mathrm{Y}_{it-1} + \beta \mathrm{D}_{it} + \varepsilon_{it}, \qquad (5.4.2)$$

که در آن ε_{it} به طور رشته‌ای[1] ناهمبسته است. برآورد یک معادله یک تفاضلی[2] برای از بین بردن اثرات ثابت اشتباه است، چراکه متغیرهای وابسته تأخیری را نادیده می‌گیرد. به عنوان یک مثال ساده، وقتی برای همه $\mathrm{D}_{it-1} = 0$ است، حد احتمال برآوردگر یک تفاضلی به صورت زیر خواهد بود:

$$\frac{Cov(\mathrm{Y}_{it} - \mathrm{Y}_{it-1}, \mathrm{D}_{it} - \mathrm{D}_{it-1})}{V(\mathrm{D}_{it} - \mathrm{D}_{it-1})} = \frac{Cov(\mathrm{Y}_{it} - \mathrm{Y}_{it-1}, \mathrm{D}_{it})}{V(\mathrm{D}_{it})}. \qquad (5.4.3)$$

با تفریق Y_{it-1} از هر دو سمت رابطه ۲-۴-۵ داریم:

$$\mathrm{Y}_{it} - \mathrm{Y}_{it-1} = \alpha + (\theta - 1)\mathrm{Y}_{it-1} + \beta \mathrm{D}_{it} + \varepsilon_{it}.$$

با جایگزینی این در رابطه ۲-۲-۴، الگویی که به طور نامناسب تفاضل‌گیری شده است خروجی زیر را به دست می‌دهد:

1. Serially
2. First-differenced (FD) Equation

$$\frac{Cov(Y_{it}-Y_{it-1},D_{it})}{V(D_{it})}=\beta+(\theta-1)\left[\frac{Cov(Y_{it-1},D_{it})}{V(D_{it})}\right].$$

به طور کلی، ما فکر می‌کنیم θ عددی بین صفر و یک است و در غیر این صورت Y_{it} نامانا[1] است (به عبارتی، یکسری زمانی انفجاری[2] است). بنابراین، از آنجایی که آموزش‌گیرندگان Y_{it-1} پایینی دارند، برآورد β با یک تفاضل خیلی بزرگ می‌شود.

ج) اضافات

۶. کمی تلاطم: طرح‌های ناپیوستگی رگرسیون[3]

اما هنگامی که شما شروع به استفاده از آن قواعد می‌کنید، انواع فرایندها شروع می‌شوند و شما شروع به پیدا کردن انواع چیزها درباره افراد می‌کنید... این فقط یک راه تفکر درباره یک مسئله است، که اجازه می‌دهد شکل مسئله پدیدار شود. در مورد بیشتر قواعد، هرچه نحیف‌تر و دلبخواه‌تر باشند، بهتر است.

داگلاس آدامز، تقریباً بی‌ضرر (۱۹۹۵)

ناپیوستگی رگرسیون از دانش دقیق قواعدی که تعیین‌کننده روش تیمار هستند بهره می‌برد. شناسایی ناپیوستگی رگرسیون بر اساس این ایده است که در دنیای مبتنی بر قاعده، برخی قواعد اختیاری هستند و ازاین‌رو آزمایشات خوبی را فراهم می‌کنند. ناپیوستگی رگرسیون در دو سبک، فازی و تیز[4] وجود دارد. طرح تیز می‌تواند به عنوان یک انتخاب از روی موارد قابل مشاهده[5] دیده شود. طرح فازی منجر به تنظیمی شامل نوعی ابزارهای متغیر می‌شود.

۶-۱. ناپیوستگی رگرسیون تیز (Sharp)

ناپیوستگی رگرسیون تیز هنگامی که وضعیت تیمار عملکرد یک متغیر، مثلاً x_i، قطعی و گسسته است استفاده می‌شود.

مثلاً فرض کنید:

$$D_i=\begin{cases}1 & \text{if } x_i\geq x_0\\ 0 & \text{if } x_i<x_0\end{cases}.$$

که در آن x_0 یک آستانه شناخته شده است. این مکانیسم تخصیص تابع جبری x_i است چون

1. Non-stationary
2. Explosive
3. Regression Discontinuity (RD)
4. Sharp
5. Selection-on-observables

وقتی که x_i را بشناسیم، D_i را هم خواهیم شناخت. این تابع ناپیوسته است چون مهم نیست x_i چقدر به x_0 نزدیک می‌شود و تیمار تا زمانی که $x_i = x_0$ بدون تغییر است.

آنچه گفته شد ممکن است زیادی انتزاعی به نظر برسد، بنابراین به یک مثال توجه کنید. به دانش‌آموزان دبیرستانی آمریکا جایزه بورسیه شایستگی بر اساس نمره **PSAT** اهدا می‌شود. این آزمون عمداً از کسانی گرفته می‌شود که بعداً در آزمون **SAT** نیز شرکت می‌کنند. سؤالی که اولین بحث ناپیوستگی رگرسیون را تحریک می‌کند این است که کدام دانش‌آموزان که موفق به دریافت این جوایز شدند بیشتر احتمال دارد کالج را به پایان برسانند. ناپیوستگی رگرسیونتیز نرخ به پایان بردن کالج توسط دانش‌آموزان را با نمرات **PSAT** بالاتر و پایین‌تر از آستانه جایزه بورسیه شایستگی مقایسه می‌کند. به طور کلی، انتظار می‌رود که دانش‌آموزانی که بالاترین نمره **PSAT** را دارند کالج را تمام کنند، اما این اثر می‌تواند با یک رگرسیون بجا توسط روابط بین تکمیل کالج و نمره **PSAT** ، حداقل در محدوده آستانه جایزه کنترل شود. به عنوان مثال، جهش در روابط بین نمره **PSAT** و حضور کالج در محدوده آستانه جایزه به عنوان شواهدی از یک اثر تیمار گرفته شده‌اند. این جهش در خط‌های رگرسیون، ناپیوستگی رگرسیون نامیده می‌شود.

یک ویژگی جالب و مهم از ناپیوستگی رگرسیون، که در پیمایش اخیر ناپیوستگی رگرسیون توسط ایمبنس و لمیکس برجسته شده، آن است که مقداری برای x_i وجود ندارد که در آن تیمار و مشاهدات کنترل را مشاهده کنیم. برخلاف استراتژی جورسازی کامل متغیر مشترک،[1] که بر مبنای مقایسه‌های تیمار ـ کنترل مشروط بر مقدارهای متغیر مشترک استوار هستند و در آنها همپوشانی وجود دارد، اعتبار[2] ناپیوستگی رگرسیون، تمایل ما را به استخراج از میان مقدار متغیرهای مشترک، حداقل در محدوده ناپیوستگی ایجاد می‌کند. این تنها یک دلیل ناپیوستگی رگرسیونتیز است که متمایز از دیگر استراتژی‌های کنترل دیده می‌شود. به همین دلیل، ما معمولاً نمی‌توانیم به عنوان اگنستیک برای شکل تابعی رگرسیون در جهان ناپیوستگی رگرسیون توانایی مالی داشته باشیم همانند جهان فصل سه.

شکل ۱-۱-۶ نشان‌دهنده سناریوی ناپیوستگی رگرسیون فرضی است که آنها با $x_i \geq 0.5$ تیمار می‌شوند. در پنل **A**، روند روابط بین Y_i و x_i خطی است، در حالی که در پنل **B**، خطی نیست. در هر دو مورد، یک ناپیوستگی در رابطه بین $E[Y_{0i}|x_i]$ و x_i در اطراف نقطه x_0 وجود دارد.

یک مدل ساده از ایده ناپیوستگی رگرسیون رسمی می‌شود. فرض کنید که علاوه بر مکانیسم تخصیص، (۱-۱-۶)، نتایج بالقوه می‌تواند توسط یک مدل خطی تأثیرات ثابت توضیح داده شود.

1. Full-covariate Matching Strategies
2. Validity

$$E[\mathrm{Y}_{0i}|x_i] = \alpha + \beta x_i$$

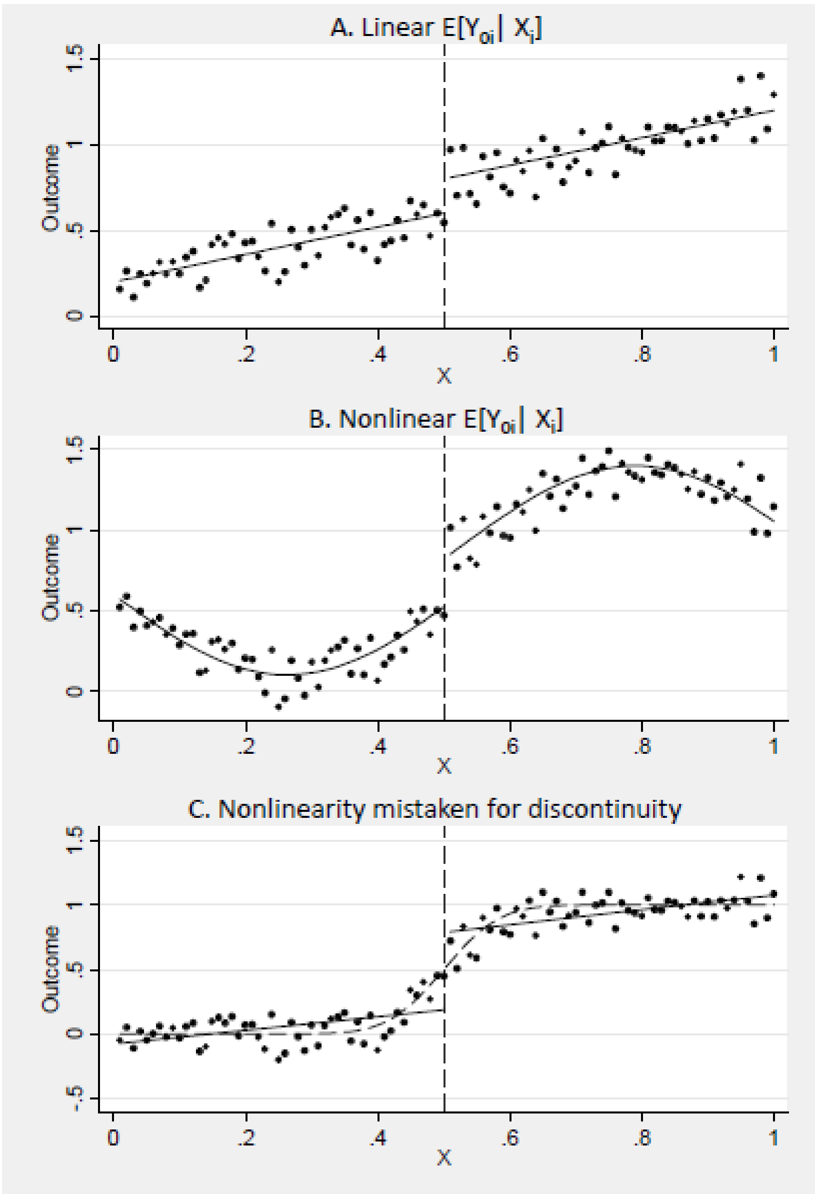
$$\mathrm{Y}_{1i} = \mathrm{Y}_{0i} + \rho$$

این به رگرسیون منجر می‌شود،

$$\mathrm{Y}_i = \alpha + \beta x_i + \rho \mathrm{D}_i + \eta_i, \quad (6.1.2)$$

که در آن ρ اثر علّی مدنظر است. تفاوت کلیدی بین این رگرسیون و دیگری را برای تخمین زدن آثار تیمار مورد استفاده قرار می‌دهیم (مثلاً در فصل سه) D_i، رگرسور مدنظر، با x_i نه تنها همبستگی دارد، بلکه تابع جبری x_i است. ناپیوستگی رگرسیون مشخص‌کننده اثر علّی با ایجاد تمایز غیرخطی و تابع ناپیوسته، $1(x_i \geq x_0)$، از نرم و (در این مورد) تابع خطی، x_iاست.

شکل ۱-۱-۶. طرح ناپیوستگی رگرسیونتیز



اما چه می‌شود اگر رابطه جریان، $E[Y_{0i}|x_i]$، غیرخطی باشد؟ برای اینکه دقیق باشیم، فرض کنید که $E[Y_{0i}|x_i] = f(x_i)$ برای تابع نرم منطقی، $f(x_i)$ است. پنل **B** در شکل ۱-۱-۶ حاکی از این است که هنوز حتی به این مورد عمومی‌تر امید وجود دارد. اکنون ما می‌توانیم ساختار ناپیوستگی رگرسیون را تخمین بزنیم با برازش:

$$Y_i = f(x_i) + \rho D_i + \eta_i, \qquad (6.1.3)$$

در جایی که دوباره، $D_i = 1(x_i \geq x_0)$، در x_i در x_0 ناپیوسته شده است. تا زمانی که $f(x_i)$ در محدوده x_0 پیوسته است، باید تخمین زدن یک مدل شبیه (۳-۱-۶)، حتی با یک فرم تابعی انعطاف‌پذیر برای $f(x_i)$ ممکن باشد. برای مثال، مدلسازی $f(x_i)$ با چند جمله‌ای درجه pام، تخمین‌های ناپیوستگی رگرسیون را می‌توان از رگرسیون‌های زیر ایجاد کرد:

$$Y_i = \alpha + \beta_1 x_i + \beta_2 x_i^2 + \dots + \beta_p x_i^p + \rho D_i + \eta_i. \qquad (6.1.4)$$

تعمیم دادن ناپیوستگی رگرسیون بر اساس (۴-۱-۶) تابع‌های جریان ناهمسان برای $E[Y_{0i}|x_i]$ و $E[Y_{1i}|x_i]$ اجازه می‌دهد. مدلسازی هر دو این **CEF**ها با چندجمله‌ای درجه pام، داریم:

$$E[Y_{0i}|x_i] = f_0(x_i) = \alpha + \beta_{01}\tilde{x}_i + \beta_{02}\tilde{x}_i^2 + \dots + \beta_{0p}\tilde{x}_i^p$$

$$E[Y_{1i}|x_i] = f_1(x_i) = \alpha + \rho + \beta_{11}\tilde{x}_i + \beta_{12}\tilde{x}_i{}^2 + \dots + \beta_{1p}\tilde{x}_i{}^p,$$

که در آن $\tilde{x}_i \equiv x_i - x_0$. قرار دادن x_i در مرکز x_0 یک عادی‌سازی است؛ اطمینان حاصل می‌کند که اثر تیمار در $x_i = x_0$ ضریب D_i در مدل رگرسیون با اثر متقابل است.

برای ایجاد یک مدل رگرسیون که می‌تواند برای تخمین اثرات مدنظر در این مورد استفاده شود، ما از این حقیقت که D_i یک تابع جبری x_i است استفاده می‌کنیم تا رابطه زیر را بنویسیم:

$$E[Y_i|x_i] = E[Y_{0i}|x_i] + E[Y_{1i} - Y_{0i}|x_i]D_i.$$

با جایگزینی چندجمله‌ای‌ها برای انتظارات شرطی، داریم:

$$\begin{aligned} Y_i &= \alpha + \beta_{01}\tilde{x}_i + \beta_{02}\tilde{x}_i^2 + \dots + \beta_{0p}\tilde{x}_i^p \\ &\quad + \rho D_i + \beta_1^* D_i\tilde{x}_i + \beta_2^* D_i\tilde{x}_i{}^2 + \dots + \beta_p^* D_i\tilde{x}_i{}^p + \eta_i, \end{aligned}$$

در آن $\beta_1^* = \beta_{11} - \beta_{01}$، $\beta_2^* = \beta_{12} - \beta_{02}$، و $\beta_p^* = \beta_{1p} - \beta_{0p}$ و عبارت خطا η_i پسماند CEF است.

معادله (۴-۱-۶) یک مورد خاص از (۶-۱-۶) که در آن $\beta_1^* = \beta_2^* = \beta_p^* = 0$. است. در مدل عمومی‌تر، اثر تیمار در $x_i - x_0 = c > 0$، $\rho + \beta_1^* c + \beta_2^* c^2 + \dots + \beta_p^* c^p$ است، در حالی که اثر تیمار در x_0، ρ است. مدل با اثر متقابل به این دلیل جاذبه دارد که هیچ محدودیتی بر شرط اساسی تابع‌ها اعمال نمی‌کند، اما در تجربه ما، تخمین‌های ρ ناپیوستگی رگرسیون بر اساس مدل ساده‌تر (۶-

۴-۱)، معمولاً مشابه آنها در (۶-۱-۶) مشخص می‌شود.

اگرچه مدل‌های چندجمله‌ای یک توصیف کافی از $E[Y_{0i}|X_i]$ را فراهم می‌کند، اعتبار تخمین‌های ناپیوستگی رگرسیون بر اساس (۴-۱-۶) یا (۶-۱-۶) مشخص می‌شود. اگر نه، بنابراین چیزی که مانند یک جهش به دلیل تیمار است ممکن است ناشی از غیرخطی بودن تابع میانگین شرطی غیرحقیقی باشد. این احتمال در پنل **C** در شکل ۱-۱-۶ نشان داده شده است، که نشان می‌دهد چگونه یک تیزی در $E[Y_{0i}|x_i]$ ممکن است برای یک جهش از یک خط رگرسیون به دیگری اشتباه گرفته شود. برای کاهش احتمال این گونه اشتباهات، ما می‌توانیم فقط به یک داده در محدوده ناپیوستگی نگاه کنیم، یعنی فاصله $[x_0-\delta, x_0+\delta]$ برای عدد بسیار کوچک δ . سپس ما داریم:

$$E[Y_i|x_0-\delta<x_i<x_0] \simeq E[Y_{0i}|x_i=x_0]$$

$$E[Y_i|x_0<x_i<x_0+\delta] \simeq E[Y_{1i}|x_i=x_0],$$

بنابراین:

$$\lim_{\delta\to 0} E[Y_i|x_0<x_i<x_0+\delta]-E[Y_i|x_0-\delta<x_i<x_0]=E[Y_{1i}-Y_{0i}|x_i=x_0]. \qquad (6.1.7)$$

به عبارت دیگر، مقایسه خروجی‌های میانگین در یک محدوده کافی کوچک در چپ و راست x_0 باید یک تخمین از اثر تیمار تهیه کند که به تعریف صحیح از یک مدل برای $E[Y_{0i}|x_i]$ بستگی ندارد.

علاوه بر این، روایی این استراتژی تخمین غیرپارامتری با فرضیه اثرات ثابت مشخص نمی‌شود $Y_{1i}-Y_{0i}=\rho$؛ مورد تخمینی در ۷-۱-۶ تأثیر علّی میانگین است، $E[Y_{1i}-Y_{0i}|x_i=x_0]$.

رویکرد ناپارامتری به ناپیوستگی رگرسیون نیازمند تخمین‌های خوب از میانگین Y_i در همسایگی کوچک راست و چپ x_0 است. به دست آوردن این تخمین‌ها دشوار است. اولین مسئله این است که کار در همسایگی کوچک مرز آستانه بدین معناست که شما داده زیادی در اختیار ندارید. همچنین، میانگین نمونه برای میانگین جامعه در محله یک مرز آستانه (در این مورد، x_0) انحراف دارد. راه‌حل این مسائل شامل استفاده از نسخه غیرپارامتری رگرسیون به نام رگرسیون خطی محلی و تخمین‌گرهای نسبتاً خطی و چندجمله‌ای محلی توسعه‌یافته توسط پورتر (۲۰۰۳) است. رگرسیون خطی محلی تخمین مجذورهای مربع وزنی معادله مانند (۶-۱-۶) را تنها با عبارت‌های خطی انجام می‌دهد و وزن بیشتری برای نقاط نزدیک به مرز قائل می‌شود.

روش‌های پیشرفته و غیرپارامتری ناپیوستگی رگرسیون هنوز کاربرد گسترده‌ای در کاربست تجربی پیدا نکرده‌اند. بیشتر کارهای کاربری ناپیوستگی رگرسیون هنوز پارامتری هستند، اما ایده تمرکز بر مشاهدات نزدیک مقدار مرزی آستانه، نشان‌دهنده ارزیابی استواری ارزشمندی است: هرچند تخمین‌های ناپیوستگی رگرسیون دقت‌شان کم می‌شود زمانی که پنجره مورد استفاده برای انتخاب نمونه ناپیوستگی

کوچک‌تر می‌شود، تعداد عبارت‌های چندجمله‌ای مورد نیاز برای مدلسازی $f(x_i)$ باید کمتر شوند. دومین ارزیابی مهم به رفتار متغیرهای قبل از تیمار نزدیک ناپیوستگی می‌پردازد. چون متغیرهای پیش‌تیمار تحت تأثیر تیمار قرار نمی‌گیرند، نباید جهشی در **CEF** این متغیرها در x_0 وجود داشته باشد.

مطالعه لی (۲۰۰۸) در مورد تأثیر جایگاه حزبی قبلی بر احتمال مجدد انتخاب شدن در انتخابات، نشان‌دهنده طرح **RD** نیز است. لی علاقه‌مند است بداند که آیا نامزدی حزب دموکرات در یک حوزه انتخابیه در انتخابات مجلس نمایندگان ایالات متحده، اگر این حزب آخرین بار در همان حوزه انتخابیه انتخابات را برده باشد، بر انتخاب مجدد نامزد آن حزب تأثیر دارد یا نه. موفقیت گسترده صاحب‌منصبان قبلی در انتخابات جدید، این سؤال را پیش می‌آورد که آیا نمایندگان از امتیازات و منابع خود برای پیروزی در انتخابات استفاده می‌کنند یا نه. اگرچه این حدس قابل اعتماد به نظر می‌رسد، اما الزاماً نشان‌دهنده وجود مزیت واقعی انتخاباتی در پیروزی صاحب‌منصبان قبلی نیست. چراکه این صاحب‌منصبان در مجلس ـ بنا به تعریف، افرادی که قبلاً نشان داده‌اند می‌توانند برنده باشند ـ ممکن است به سادگی در جلب رضایت رأی‌دهندگان موفق‌تر عمل کنند.

برای مشخص کردن تأثیر علّی قرار داشتن در یک منصب و احتمال پیروزی در انتخابات، لی (۲۰۰۸) به مطالعه احتمال پیروزی نامزد دموکرات به عنوان تابع سهم رأی نسبی در انتخابات قبلی پرداخت. به طور خاص، او از این حقیقت استفاده کرد که برنده انتخابات توسط $\mathrm{D}_i = 1(x_i \geq .0)$ مشخص می‌شود که در آن x_i حاشیه سهم رأی پیروزی[1] است. توجه داشته باشید که چون D_i تابع جبری x_i است، متغیر محدودکننده‌ای جز x_i وجود ندارد. این مشخصه سیگنال طرح ناپیوستگی رگرسیون است.

شکل ۲-۱-۶ الف، از سوی لی (۲۰۰۸) نشان‌دهنده طرح ناپیوستگی رگرسیونتیز در عمل است. این شکل ترسیم‌کننده احتمال پیروزی یک دموکرات در برابر تفاوت بین سهم رأی دموکرات و جمهوریخواه در انتخابات قبلی است. نقطه‌ها در شکل میانگین‌های محلی هستند؛ خطوط در شکل مقادیر برازش شده از مدل پارامتری با ناپیوستگی در صفر هستند. احتمال پیروزی دموکرات تابع صعودی سهم رأی گذشته است. مهم‌ترین شاخصه طرح جهش جدی در نرخ پیروزی در علامت صفر درصد است، نقطه‌ای که در آن نامزد دموکرات رأی بیشتری می‌آورد. بر اساس اندازه جهش، صاحب‌منصب بودن در مجلس، احتمال انتخابات مجدد حزب را تا ۴۰ درصد بیشتر می‌کند.

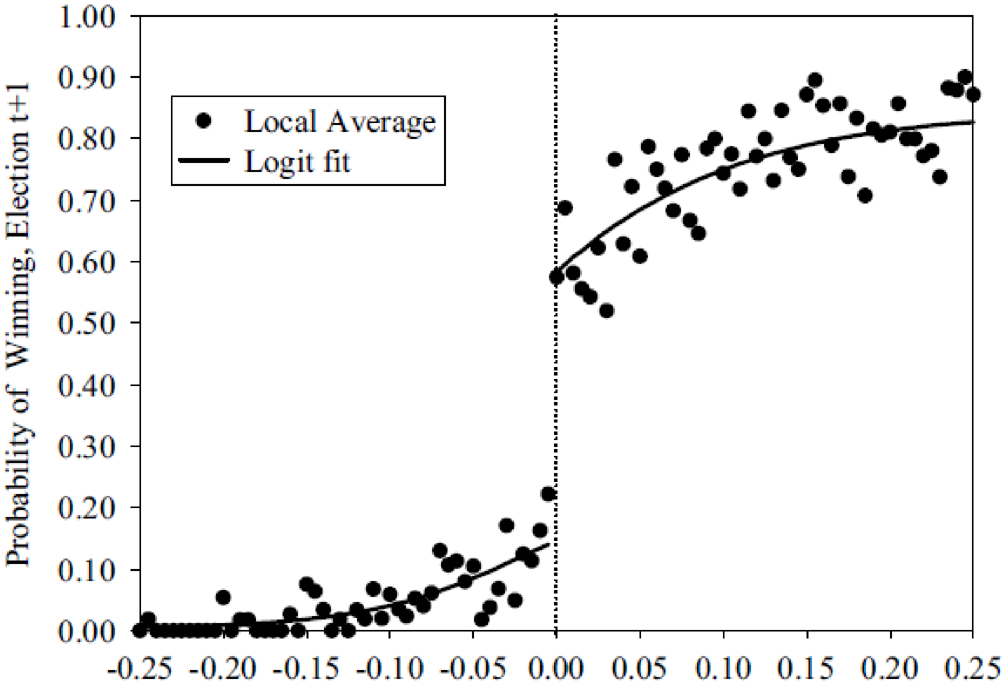
شکل ۲-۱-۶ ب فرضیات شناخت ناپیوستگی رگرسیونتیز را با نگاهی به پیروزی‌های دموکرات قبل از آخرین انتخابات بررسی می‌کند. نرخ پیروزی دموکرات در انتخابات قبلی نباید ارتباطی به آستانه مرزی

1. Vote Share Margin of Victory

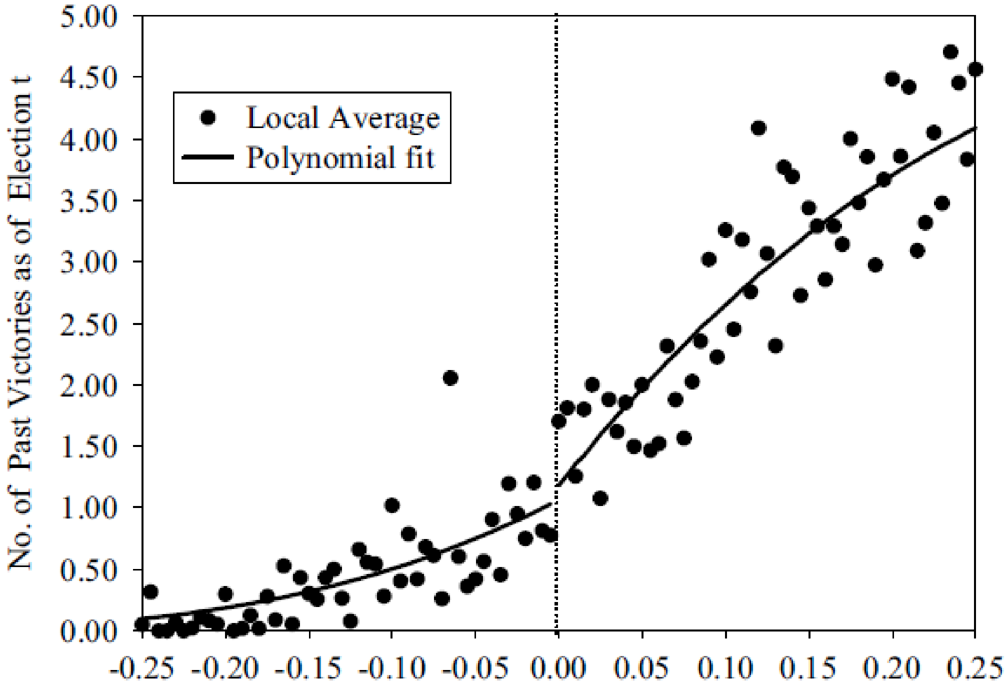
در آخرین انتخابات داشته باشد. بررسی پیروزی‌های پیش از تیمار توسط لی، یک نسخه از ایده‌ای است که می‌گوید متغیرهای مشترک باید توسط وضعیت تیمار در آزمایش تصادفی متعادل شوند. یک ارزیابی مرتبط تراکم x_i حول ناپیوستگی را بررسی می‌کند و نگاهی دارد به دسته‌بندی در توزیع x_i نزدیک x_0. مسئله در اینجا این است که افراد دارای سهم در D_i ممکن است برای تغییر x_i نزدیک آستانه مرزی تلاش کنند که در آن صورت، مشاهدات هر طرف نمی‌تواند قابل مقایسه باشد (مک کراری، ۲۰۰۸). تا همین اواخر، می‌گفتیم که این امر در مطالعات انتخابات مانند مطالعه لی، غیرممکن است، اما بازشماری در فلوریدا پس از انتخابات ریاست‌جمهوری سال ۲۰۰۰، نشان می‌دهد که وقتی انتخابات ایالات متحده تمام می‌شود، ما باید نگران سهم رأی قابل دستکاری باشیم.

شکل ۲-۱-۶ احتمال پیروزی یک انتخابات توسط سهم رأی گذشته و آینده (لی، ۲۰۰۸). الف) احتمال پیروزی انتخابات توسط نامزد **t+1**، با حاشیه پیروزی در انتخابات **t**: میانگین‌های محلی و برازش پارامتری. ب) تعداد انباشته پیروزی‌های انتخاباتی قبلی نامزد، توسط حاشیه پیروزی در انتخابات **t**: میانگین‌های محلی و برازش پارامتری.

a


Probability of Winning, Election t+1
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
Local Average
Logit fit
-0.25 -0.20 -0.15 -0.10 -0.05 0.00 0.05 0.10 0.15 0.20 0.25

b


No. of Past Victories as of Election t
5.00
4.50
4.00
3.50
3.00
2.50
2.00
1.50
1.00
0.50
0.00
Local Average
Polynomial fit
-0.25 -0.20 -0.15 -0.10 -0.05 0.00 0.05 0.10 0.15 0.20 0.25

۲-۶. ناپیوستگی رگرسیون فازی،[1] متغیر ابزاری (IV) است

ناپیوستگی رگرسیون فازی از ناپیوستگی‌ها در احتمال یا مقدار مورد انتظار تیمار مشروط به متغیر مشترک استفاده می‌کند. نتیجه، یک طرح تحقیقاتی است که در آن ناپیوستگی به متغیر ابزاری برای وضعیت تیمار به جای در نظر گرفتن یا در نظر نگرفتن جبری تیمار تبدیل می‌شود. برای اینکه ببینیم کارکرد این موضوع چطور است، فرض کنیم D_i نماینده تیمار باشد، هرچند در اینجا D_i دیگر به صورت جبری ربطی به قانون قطع آستانه $x_i \geq x_0$ ندارد. بلکه در احتمال تیمار در x_0یک جهش وجود دارد، در نتیجه:

$$P[D_i = 1|x_i] = \begin{cases} g_0(x_i) & \text{if } x_i \geq x_0 \\ g_1(x_i) & \text{if } x_i < x_0 \end{cases}, \text{ where } g_1(x_0) \neq g_0(x_0).$$

توابع $g_0(x_i)$ و $g_1(x_i)$ تا زمانی که در x_0 تفاوت داشته باشند، هر چیزی می‌توانند باشند. فرض می‌کنیم که $g_1(x_0) > g_0(x_0)$، در نتیجه $x_i \geq x_0$ تیمار را محتمل‌تر می‌کند. ما می‌توانیم رابطه بین احتمال تیمار و x_i را به صورت زیر بنویسیم:

$$E[D_i|x_i] = P[D_i = 1|x_i] = g_0(x_i) + [g_1(x_i) - g_0(x_i)]T_i,$$

که در آن،

$$T_i = 1(x_i \geq x_0)$$

متغیر موهومی T_i نشان‌دهنده نقطه ناپیوستگی در $E[D_i|x_i]$ است.

ناپیوستگی رگرسیون فازی طبیعتاً به استراتژی تخمین ساده **2SLS** منجر می‌شود. فرض کنیم $g_0(x_i)$ و $g_1(x_i)$ را بتوان توسط چندجمله‌ای‌های مرتبه **p**ام مانند ۴-۱-۶ تعریف کرد، سپس داریم:

$$\begin{aligned} E[D_i|x_i] &= \gamma_{00} + \gamma_{01}x_i + \gamma_{02}x_i^2 + \ldots + \gamma_{0p}x_i^p \\ &\quad + [\gamma_0^* + \gamma_1^* x_i + \gamma_2^* x_i^2 + \ldots + \gamma_p^* x_i^p]T_i \\ &= \gamma_{00} + \gamma_{01}x_i + \gamma_{02}x_i^2 + \ldots + \gamma_{0p}x_i^p \\ &\quad + \gamma_0^* T_i + \gamma_1^* x_i T_i + \gamma_2^* x_i^2 T_i + \ldots + \gamma_p^* x_i^p T_i. \end{aligned} \tag{6.2.1}$$

از این رابطه می‌بینیم که T_i و عبارت‌های تعامل $\{x_i T_i,\ x_i^2 T_i,\ .\ .\ .\ x_i^p T_i\}$ را می‌توان به عنوان ابزار برای D_i در (۴-۱-۶) استفاده کرد.

ساده‌ترین تخمین‌گر ناپیوستگی رگرسیون فازی، بدون عبارت‌های تعامل، تنها از T_iبه عنوان یک ابزار استفاده می‌کند. تخمین‌گر **IV** نهایی دارای ویژگی‌های شفافیت و نمونه نهایی خوب است. اولین مرحله در این مورد به عبارت زیر است:

1. Fuzzy RD

$$\mathrm{D}_i = \gamma_0 + \gamma_1 x_i + \gamma_2 x_i^2 + \ldots + \gamma_p x_i^p + \pi \mathrm{T}_i + \xi_{1i}, \qquad (6.2.2)$$

که در آن T_i ابزار حذف‌شده‌ای است که قدرت شناسایی را با تأثیر درجه اول ارائه شده توسط π فراهم می‌کند.

فرم تقلیل‌یافته ناپیوستگی رگرسیون فازی با جایگزینی (۶-۲-۲) در (۶-۱-۴) حاصل می‌شود:

$$\mathrm{Y}_i = \mu + \kappa_1 x_i + \kappa_2 x_i^2 + \ldots + \kappa_p x_i^p + \rho\pi \mathrm{T}_i + \xi_{2i}, \qquad (6.2.3)$$

که در آن $\mu = \alpha + \rho\gamma_0$ و $\kappa_j = \beta_1 + \rho\gamma_j$ برای $j = 1, \ldots, p$. مانند ناپیوستگی رگرسیونتیز، شناخت در مورد فازی، توانایی تمایز رابطه بین Y_i و تابع ناپیوسته $\mathrm{T}_i = 1(x_i \geq x_0)$ از تأثیر کنترل‌های چندجمله‌ای در اولین و دومین مرحله را ممکن می‌کند. در یکی از اولین مطالعات ناپیوستگی رگرسیون در اقتصادسنجی کاربردی، فن در کلاو (۲۰۰۲) از طرح فازی برای ارزیابی اثرات کمک مالی دانشگاه هنگام ثبت نام استفاده کرد. در مطالعه فن در کلاو، D_i اندازه کمک مالی ارائه شده و T_i متغیر موهومی است که نشان‌دهنده درخواست‌دهندگان با شاخص توانایی بالاتر از مرز آستانه پاداش از پیش تعیین شده است.

ناپیوستگی رگرسیون فازی با اثرات تیمار این گونه تخمین می‌زند که تغییرات به عنوان تابعی از x_i را می‌توان با تخمین **2SLS** یک معادله با تعاملات تیمار متغیر مشترک ایجاد کرد. در اینجا، مدل مرحله دوم با عبارت تعامل مانند (۶-۱-۶) است در حالی که مرحله دوم شبیه (۶-۲-۱) است جز اینکه برای جورشدگی پارامتری‌سازی مرحله دوم، عبارت چندجمله‌ای را در مرکز x_0 قرار می‌دهیم. در این مورد، ابزار حذف شده $\{\mathrm{T}_i, \bar{x}_i\mathrm{T}_i, \bar{x}_i^2\mathrm{T}_i, \ldots \bar{x}_i^p\mathrm{T}_i\}$ در حالی که متغیرها $\{\mathrm{D}_i, \bar{x}_i\mathrm{D}_i, \mathrm{D}_i\bar{x}_i{}^2, \ldots \mathrm{D}_i\bar{x}_i{}^p\}$ به عنوان موارد درونی در نظر گرفته می‌شوند. اولین مرحله برای D_i به شکل زیر است:

$$\begin{aligned}\mathrm{D}_i \;=\; & \gamma_{00} + \gamma_{01}\bar{x}_i + \gamma_{02}\bar{x}_i^2 + \ldots + \gamma_{0p}\bar{x}_i^p \\ & + \gamma_0^*\mathrm{T}_i + \gamma_1^*\bar{x}_i\mathrm{T}_i + \gamma_2^*\bar{x}_i^2\mathrm{T}_i + \ldots + \gamma_p^*\bar{x}_i^p\mathrm{T}_i + \xi_{1i}.\end{aligned} \qquad (6.2.4)$$

یک مرحله اول مشابه برای هریک از عبارت‌های تعامل چندجمله‌ای در مجموعه $\{\bar{x}_i\mathrm{D}_i, \mathrm{D}_i\bar{x}_i{}^2, \ldots \mathrm{D}_i\bar{x}_i{}^p\}$ ایجاد می‌شود.

نسخه غیرپارامتری ناپیوستگی رگرسیون فازی شامل تخمین **IV** در یک همسایگی کوچک حول ناپیوستگی است. انتظار شرطی فرم تقلیل‌یافته Y_i نزدیک x_0 به صورت زیر است:

$$E\left[\mathrm{Y}_i | x_0 < x_i < x_0 + \delta\right] - E\left[\mathrm{Y}_i | x_0 - \delta < x_i < x_0\right] \simeq \rho\gamma_0^*.$$

همچنین، برای مرحله اول برای D_i، داریم:

$$E\left[\mathrm{D}_i | x_0 < x_i < x_0 + \delta\right] - E\left[\mathrm{D}_i | x_0 - \delta < x_i < x_0\right] \simeq \gamma_0^*.$$

در نتیجه:

$$\lim_{\delta\to 0}\frac{E\left[\mathrm{Y}_i|x_0<x_i<x_0+\delta\right]-E\left[\mathrm{Y}_i|x_0-\delta<x_i<x_0\right]}{E\left[\mathrm{D}_i|x_0<x_i<x_0+\delta\right]-E\left[\mathrm{D}_i|x_0-\delta<x_i<x_0\right]}=\rho. \qquad (6.2.5)$$

مشابه نمونه (۵-۲-۶) تخمین‌گر والد از نوع بحث شده در بخش‌های قبلی است، در این مورد استفاده از T_i به عنوان یک ابزار برای D_i در همسایگی δ از x_0. مانند ابزار متغیر موهومی، نتیجه یک اثر تیمار میانگین محلی است. به طور خاص، تخمین والد برای ناپیوستگی رگرسیون فازی مشخص‌کننده اثر علّی بر کامپایلر تعریف شده به عنوان افرادی است که وضعیت تیمارشان با حرکت مقدار x_i از سمت چپ x_0 به سمت راست x_0 تغییر می‌کند. این تعبیر از ناپیوستگی رگرسیون فازی توسط هان، تاد و فن در کلاو (۲۰۰۱) معرفی شد. توجه داشته باشید که مورد دیگری وجود دارد که در آن این نسخه **LATE** محلی است: تخمین‌ها برای کامپایلرها با $x_i=x_0$ هستند، یک شاخصه از تخمین‌های غیرپارامتری تیز (تیز).

در نهایت توجه داشته باشید که نسخه غیرپارامتری ناپیوستگی رگرسیونتیز، رفتار نمونه متناهی مشابه نمونه (۵-۲-۶) چندان خوب نیست. هان، تاد و فن در کلاو (۲۰۰۱) یک رویه **IV** غیرپارامتری را با استفاده از رگرسیون خطی محلی ایجاد کردند تا بالا و پایین تخمین‌گر والد را با انحراف کمتر محاسبه کنند. این اما ما را به مدل **2SLS** با کنترل‌های خطی یا چندجمله‌ای برمی‌گرداند، اما این مدل در نمونه ناپیوستگی با استفاده از پهنای باند مبتنی بر داده برازش می‌شود. ایده استفاده از نمونه‌های ناپیوستگی به صورت غیررسمی در این بافت نیز کاربرد دارد: شروع با طرح پارامتری **2SLS** در یک نمونه کامل، بر اساس (۴-۱-۶). سپس محدود کردن نمونه به نقاط نزدیک ناپیوستگی و خلاص شدن از شر بیشتر یا همه کنترل‌های چندجمله‌ای. به صورت ایدئال، تخمین‌های **2SLS** در نمونه‌های ناپیوستگی با کنترل‌های اندک به صورت گسترده با تخمین‌های دقیق‌تر ایجاد شده توسط نمونه بزرگ‌تر همخوانی دارند.

آنگریست و لاوی (۱۹۹۹) از طرح تحقیق ناپیوستگی رگرسیون فازی برای تخمین اثرات اندازه کلاس بر نمرات آزمون کودکان استفاده کردند. ناپیوستگی رگرسیون فازی یک طرح تحقیقاتی قدرتمند و انعطاف‌پذیر است که ناپیوستگی رگرسیون فازی را به دو طریق نسبت به بحث بالا تعمیم می‌دهد. اول، متغیر علّی مدنظر، اندازه کلاس، مقادیر مختلفی می‌گیرد. در نتیجه اولین مرحله از جهش در میانگین اندازه کلاس به جای احتمالات استفاده می‌کند. دوم، طرح تحقیق آنگریست و لاوی (۱۹۹۹) از ناپیوستگی‌های چندگانه استفاده می‌کند.

مطالعه آنگریست و لاوی با این مشاهده شروع می‌شود که اندازه کلاس در مدارس حداکثر ۴۰ نفر است. انتظار می‌رود دانش‌آموزان در یک کلاس با حداکثر ۴۰ دانش‌آموز در یک کلاس به بزرگی ۴۰ قرار بگیرند، اما کلاس‌هایی با ۴۱ دانش‌آموز به دو کلاس تقسیم می‌شوند، کلاس‌هایی با ۸۱ دانش‌آموز

به سه کلاس تقسیم می‌شود و این روند ادامه دارد. آنگریست و لاوی این را «قانون میمونید»[1] می‌نامند، چراکه حداکثر اندازه ۴۰ نفری کلاس اولین بار توسط میمونید معرفی شد. برای تدوین (فرموله کردن) قانون میمونید، m_{sc} نماینده اندازه پیش‌بینی شده کلاس است که برای کلاس **c** در مدرسه **s** تخصیص یافته است که در آن ثبت نام در کلاس با e_s نشان داده می‌شود. با فرض اینکه گروه‌های کلاسی به کلاس‌هایی با اندازه مشابه تقسیم می‌شوند، اندازه پیش‌بینی شده کلاس که از کاربرد قانون میمونید به دست می‌آید به صورت زیر است:

$$m_{sc} = \frac{e_s}{int[\frac{(e_s - 1)}{40}] + 1}$$

که در آن $int(x)$ بخش عدد صحیح عدد واقعی **x** است. این تابع که با خطوط نقطه‌چین در شکل ۶-۲-۱ برای کلاس‌های چهارم و پنجم رسم شده است، الگوی دندانه‌اره‌ای با پیوستگی در عدد صحیح ۴۰ دارد. در عین حال، m_{sc} تابع صعودی ثبت نام است و e_s متغیر ثبت نام را به یک کنترل مهم تبدیل می‌کند.

آنگریست و لاوی از ناپیوستگی در قانون میمونید توسط ساخت تخمین‌های **2SLS** یک رابطه مانند زیر استفاده کردند:

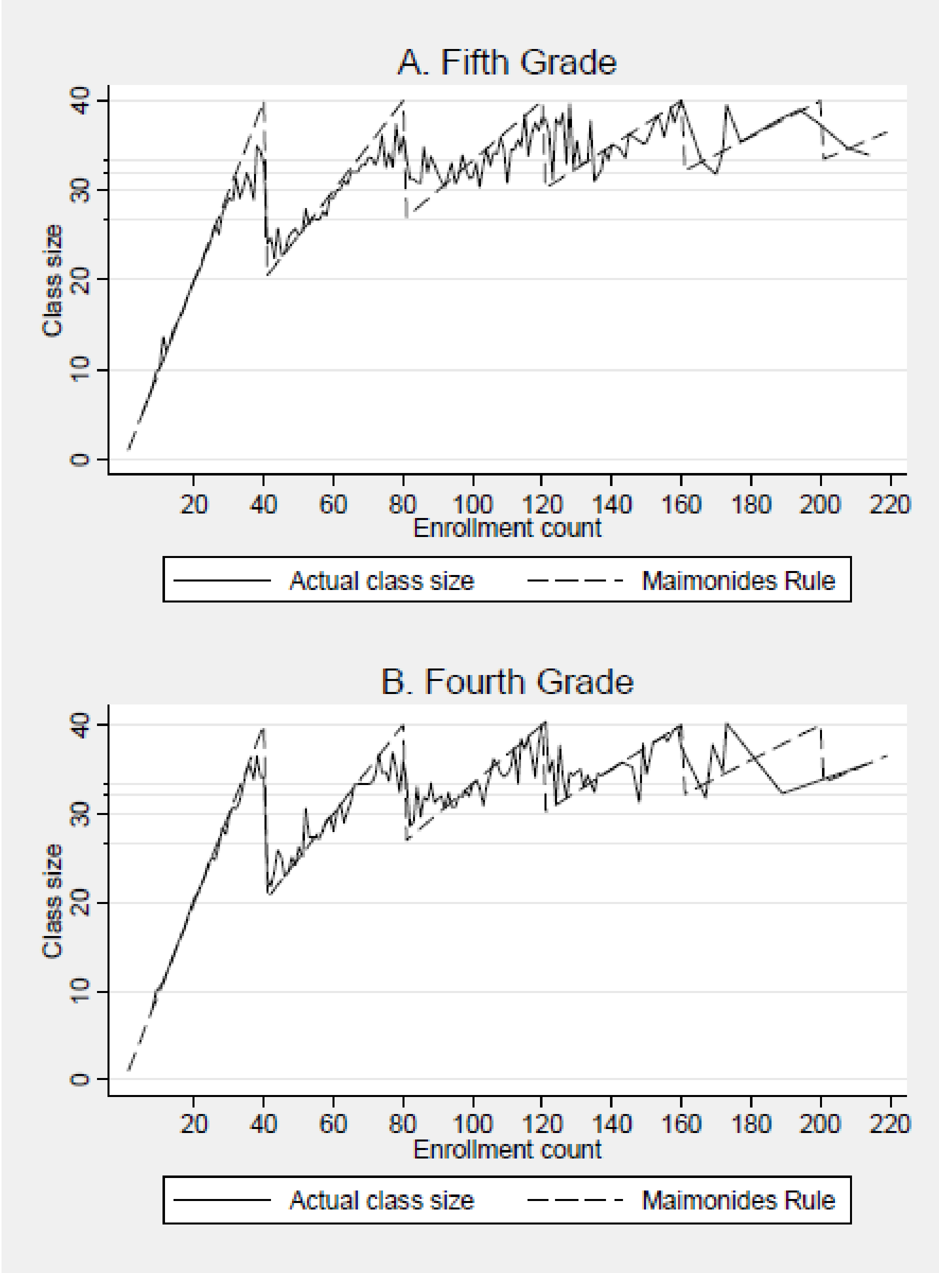
$$\mathrm{Y}_{isc} = \alpha_0 + \alpha_1 pd_s + \beta_1 e_s + \beta_2 e_s^2 + \dots + \beta_p e_s^p + \rho n_{sc} + \eta_{isc}, \qquad (6.2.6)$$

که در آن Y_{isc} نمره آزمون **i** در مدرسه **s** و کلاس **c** است، n_{sc} اندازه کلاس و e_s ثبت نام است. در این نسخه از ناپیوستگی رگرسیون فازی، m_{sc} نقش T_i، e_s نقش x_i و اندازه کلاس یعنی n_{sc} نقش D_i. را بازی می‌کنند. آنگریست و لاوی همچنین متغیر مشترک عدم ثبت نام pd_s را برای کنترل نسبت دانش‌آموزان در مدرسه با پیشینه عدم توانایی مالی در نظر می‌گیرند. این برای ناپیوستگی رگرسیون ضروری نیست، چون تنها منبع انحراف متغیرهای حذف شده در مدل ناپیوستگی رگرسیون، e_s است، اما این تصریح را با مدل استفاده شده برای ساخت مجموعه متناظر تخمین‌های **OLS** قابل قیاس می‌کند.

شکل ۱-۲-۶ آنگریست و لاوی (۱۹۹۹) میانگین اندازه واقعی و پیش‌بینی شده کلاس را در برابر ثبت نام در کلاس‌های چهارم و پنجم نشان می‌دهد. قانون میمونید اندازه کلاس را به خوبی پیش‌بینی نمی‌کند چون برخی مدارس کلاس‌ها را در ثبت نام کمتر از ۴۰ تقسیم می‌کنند. این چیزی است که طرح ناپیوستگی رگرسیون را فازی می‌کند. همچنان، افت مشخصی در اندازه کلاس در سطوح ثبت نام ۴۰، ۸۰ و ۱۲۰ وجود دارد. توجه داشته باشید که ابزار m_{sc} به خوبی تعاملات ناپیوستگی و ناپیوستگی شیب مانند $\tilde{x}_i \mathrm{T}_i$ را در (۴-۲-۶) در یک متغیر ترکیب می‌کند. این پارامتری‌سازی فشرده از شناخت ویژه نهادها و قوانینی است که اندازه کلاس را تعیین می‌کنند.

1. Maimonides Rule

شکل ۱-۲-۶. مرحله اول ناپیوستگی رگرسیون فازی برای تخمین‌های رگرسیون ناپیوستگی تأثیر اندازه کلاس بر نمرات آزمون دانش‌آموزان



برخلاف تخمین‌های **OLS** در ستون ۳، تخمین‌های **2SLS** با تصریح مشابه و استفاده از m_{sc} به عنوان ابزار برای n_{sc} نشان می‌دهند که کلاس‌های کوچک‌تر، نمرات تست را افزایش می‌دهند. این نتایج که در ستون ۴ برای مدل‌هایی گزارش شدند که شامل کنترل ثبت نام خطی و در ستون ۵ برای مدل‌هایی هستند که شامل کنترل ثبت نام چارکی هستند از ۰/۲۳- تا ۰/۲۶- با خطای استاندارد حول ۱ هستند. این نتایج نشان می‌دهند که کاهش هفت دانش‌آموز در اندازه کلاس، نمرات ریاضی را به

میزان ۱/۷۵ بالاتر می‌برد، برای اندازه اثر 18σ که در آن σ انحراف معیار میانگین نمرات کلاس است. این از تخمین‌های تنسی چندان دور نیست.

همچنین فرم تابعی کنترل ثبت نام به نظر چندان مهم نیست. ستون‌های ۶ و ۷ قدرت یافته‌های اصلی را با استفاده از نمونه ناپیوستگی +/-۵ چک می‌کنند. جای تعجب نیست که این نتایج دقت کمتری از موارد گزارش شده در ستون‌های ۵ و ۶ دارند چون تنها با یک‌چهارم داده‌های استفاده شده برای ساخت تخمین‌های نمونه کامل تخمین زده شده‌اند، اما هنوز، آنها حول عدد ۲۵ هستند. در نهایت، آخرین ستون نشان‌دهنده نتایج تخمین با استفاده از نمونه کوچک‌تر ناپیوستگی محدود به مدارس با ثبت نام مثبت و منفی سه دانش‌آموز حول ناپیوستگی در ۴۰، ۸۰ و ۱۲۰ است. اینها تخمین‌های والد در سایه هان، تاد و فن در کلاو (۲۰۰۱) و فرمول‌های ۵-۲-۶ هستند؛ ابزار برای ساخت این تخمین‌ها یک متغیر موهومی برای بودن در مدرسه با ثبت نام در سمت راست ناپیوستگی مرتبط است. نتایج دقت کافی را ندارد ۲۷۰- اما هنوز به شدت شبیه تخمین‌های دیگر جدول است. این مجموعه تخمین‌ها نشان‌دهنده قیمت بالایی است که از نظر دقت پرداخت می‌کنیم چون مجبوریم نمونه را حول ناپیوستگی‌ها کوچک کنیم. خوشبختانه، تصویری که از جدول ۱-۲-۶ ایجاد می‌شود، واضح است.

جدول ۱-۲-۶. برآوردهای OLS و RD فازی

در مورد تأثیر اندازه کلاس بر نمرات ریاضی کلاس پنجم

	OLS			2SLS					.S		
				Full sample		Discontinuity samples			iscontinuity samples		
						+/- 5		+/- 3	+/- 5		+/- 3
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	)	(7)	(8)
		67.3		67.3		67.0		67.0	67.0		67.0
		(9.6)		(9.6)		(10.2)		(10.6)	(10.2)		(10.6)
	.322	.076	.019	-.230	-.261	-.185	-.443	-.270	5	-.443	-.270
	(.039)	(.036)	(.044)	(.092)	(.113)	(.151)	(.236)	(.281)	1)	(.236)	(.281)
ged		-.340	-.332	-.350	-.350	-.459	-.435		9	-.435	
		(.018)	(.018)	(.019)	(.019)	(.049)	(.049)		9)	(.049)	
			.017	.041	.062		.079			.079	
			(.009)	(.012)	(.037		(.036)			(.036)	
/100					-.010						
					(.016)						
								-12.6			-12.6
								(3.80)			(3.80)
								-2.89			-2.89
								(2.41)			(2.41)
	9.36	8.32	8.30	8.40	8.42	8.79	9.10	10.2	9	9.10	10.2
	.048	.249	.252								
		2,018		2,018		471		302	471		302

نکته: برگرفته از آنگریست و لاوی (۱۹۹۹). این جدول برآوردهای معادله ۶-۲-۶ با استفاده از میانگین‌های کلاسی را گزارش می‌کند. خطاهای استاندارد گزارش شده در میان پرانتزها به خاطر همبستگی‌های بین کلاسی تصحیح شده‌اند.

۷. رگرسیون چندکی[1]

اینجا یک دعا برای تو هست. مداد دستت هست؟ ... محافظت کن از من در برابر دانستن چیزی که من نیازی به دانستنش ندارم. محافظت کن از من در برابر دانستن چیزهایی که دانستنی هستند و من نمی‌دانم. محافظت کن از من در برابر دانستن اینکه من تصمیم گرفتم چیزهایی در مورد چیزهایی که نمی‌دانم، ندانم. آمین. یک دعای دیگر همراه آن وجود دارد. پروردگارا، پروردگارا، پروردگارا، من را از عواقب دعای قبلی محافظت کن.

داگلاس آدامز، تقریباً بی‌ضرر (۱۹۹۵)

به درست یا غلط، ۹۵ درصد از اقتصادسنجی کاربردی مربوط به میانگین است. اگر به عنوان مثال، یک برنامه آموزشی، به اندازه کافی میانگین درآمدها را به گونه‌ای افزایش دهد که جبران هزینه‌ها را بکند، خوشحال هستیم. تمرکز بر میانگین تا اندازه‌ای به این دلیل است که به دست آوردن تخمین خوب از میانگین علّی آثار به اندازه کافی سخت است و اگر متغیر وابسته برای چیزهایی شبیه اشتغال، موهومی باشد، میانگین، کل توزیع را توصیف می‌کند، اما بسیاری از متغیرها مانند درآمد و نمرات آزمون دارای توزیع پیوسته هستند. این توزیع‌ها می‌توانند به شیوه‌ای که توسط معیارهای میانگین نشان داده نمی‌شوند تغییر کنند، به عنوان مثال، آنها می‌توانند گسترش پیدا کنند و یا فشرده‌تر شوند. به طور فزاینده اقتصاددانان کاربردی می‌خواهند بدانند که در کل یک توزیع، برای برندگان و بازنده‌های نسبی و همچنین در میانگین چه چیزهایی اتفاق می‌افتد.

به ویژه سیاستگذاران و کارشناسان اقتصادی در مورد تغییرات در توزیع دستمزد حساس بوده‌اند. مثلاً می‌دانیم که میانگین دستمزد به قیمت واقعی تنها بخش کوچکی از آنچه در ۲۵ سال گذشته در بازار کار جریان دارد، را توضیح می‌دهد. در حالی که چندک[2]‌های پایین‌تر سقوط کرده‌اند، چندک‌های بالای درآمدی افزایش پیدا کرده‌اند. به عبارت دیگر، ثروتمندان ثروتمندتر می‌شوند و فقرا فقیرتر، اما این تمام ماجرا نیست ـ به تازگی، نابرابری به طور نامتقارن رشد کرده است؛ برای مثال، در میان فارغ‌التحصیلان دانشگاهی، غالباً ثروتمندان ثروتمندتر می‌شوند و دستمزدها در دهک‌های پایین تغییر نمی‌کند. داستان توزیع دستمزد در حال تغییر بوده و نسبتاً پیچیده است و به نظر می‌رسد چندان قابل خلاصه کردن نباشد.

رگرسیون چندکی ابزاری قدرتمند است که باعث می‌شود حتی در صورتی که موضوعی مبنایی، پیچیده و چندبُعدی باشد، مدلسازی توزیع آسان شود. ما می‌توانیم از این ابزار استفاده کنیم که ببینیم آیا مشارکت در یک برنامه آموزشی یا عضویت در یک اتحادیه کارگری، نابرابری درآمد را به اندازه متوسط

1. Quantile Regression
2. Quantile

درآمد تغییر می‌دهد یا نه. ما همچنین می‌توانیم برخی تعاملات، مانند اینکه آیا رابطه بین تحصیل و نابرابری در طول زمان تغییر کرده است، را از این طریق بررسی کنیم. رگرسیون چندکی بسیار شبیه به رگرسیون معمولی است: عوامل اختلاطی[1] را می‌توان با استفاده از متغیر کمکی ثابت نگه داشت و شرایط تعامل همانند شرایط تعامل رگرسیون باشند. گاهی اوقات حتی می‌توانیم از روش‌های متغیرهای ابزاری مختلف برای برآورد مقادیر علیت در چندک، زمانی که یک موضوع انتخاب بر روی متغیرهای قابل مشاهده،[2] غیرممکن به نظر می‌رسد، استفاده کنیم.

۱-۷. الگوی رگرسیون چندکی

نقطه شروع رگرسیون چندکی تابع شرطی چندکی است (**CQF**[3]). فرض کنید ما علاقه‌مند به توزیع یک متغیر تصادفی توزیع شده Y_i با چگالی خوش‌رفتار[4] (بدون شکاف و یا لبه تیز[5]) هستیم. سپس سی کیو اف در چندک τ با داشتن یک بردار x_i شامل رگرسورها می‌تواند این چنین تعریف شود:

$$Q_\tau(\mathrm{Y}_i|\mathrm{X}_i) = F_Y^{-1}(\tau|\mathrm{X}_i)$$

جایی که $F_Y(y|\mathrm{X}_i)$ تابع توزیع Y_i مقید به X_i است. به عنوان مثال، وقتی $\tau = .10$، $Q_\tau(\mathrm{Y}_i|\mathrm{X}_i)$ دهک پایین Y_i را با توجه به X_i توصیف می‌کند، در حالی که $\tau = .5$ میانگین مقید را به دست می‌دهد.[6] با نگاهی به تغییرات سی کیو اف درآمدها به عنوان تابعی از تحصیلات، می‌توانیم در مورد اینکه پراکندگی درآمد بالا یا پایین می‌رود قضاوت کنیم. با نگاهی به تغییرات سی کیو اف درآمدها به عنوان تابعی از تحصیلات و زمان، می‌توانیم بفهمیم که آیا رابطه بین تحصیل و نابرابری در طول زمان تغییر می‌کند یا نه. **CQF** یک نسخه مقید ـ چندکی[7] **CEF** است. یادآوری می‌کنیم که **CEF** می‌تواند به عنوان راه‌حلی برای مسئله پیش‌بینی خطای میانگین مربعات استخراج شود،

$$E[\mathrm{Y}_i|\mathrm{X}_i] = \underset{m(\mathrm{X}_i)}{\arg\min} E\left[(\mathrm{Y}_i - m(\mathrm{X}_i))^2\right].$$

به همین ترتیب، **CQF** مشکل کمینه‌سازی زیر را حل می‌کند،

$$Q_\tau(\mathrm{Y}_i|\mathrm{X}_i) = \underset{q(X)}{\arg\min}\ E\left[\rho_\tau(\mathrm{Y}_i - q(\mathrm{X}_i))\right], \qquad (7.1.1)$$

به طوری که $\rho_\tau(u) = (\tau - 1(u \le 0))u$ تابع چک[8] نامیده می‌شود، چراکه زمانی که نمودار

1. Confounding Factors
2. Selection-on-observables
3. Conditional Quantile Function
4. Well-behaved Density
5. Spike
٦. عموماً، CQF متغیرهاي تصادفي گسسته و متغیرهاي تصادفي با چگالي‌هاي كمتر خوش‌رفتار را مي‌توانيم به اين صورت تعريف كنيم:

$Q_\tau(\mathrm{Y}_i|\mathrm{X}_i) = \inf\{y : F_Y(y|\mathrm{X}_i) \ge \tau\}.$

7. Conditional-quantile
8. Check Function

آن رسم می‌شود، شبیه علامت چک (تیک) می‌شود. اگر $\tau = .5$ باشد، حداقل انحراف مطلق[1] به دست می‌آید، چراکه $\rho_{.5}(u) = \frac{1}{2}(\text{sign}\, u)u = \frac{1}{2}|u|$. در این مورد $Q_\tau(Y_i|X_i)$ میانه شرطی (مقید) است، چراکه میانه شرطی، انحرافات مطلق را حداقل می‌کند. در غیر این صورت، تابع چک عبارات مثبت و منفی را به صورت نامتقارن وزن‌دهی می‌کند:

$$\rho_\tau(u) = 1(u > 0) \cdot \tau u + 1(u \leq 0) \cdot (1 - \tau)u.$$

این وزن‌بندی نامتقارن، مینیماندی را ایجاد می‌کند که چندک‌های شرطی را پیدا می‌کند (واقعیتی که به روشنی واضح نیست، اما می‌تواند با اندکی کار بر روی آن ثابت شود، کوئنکر را ببینید، ۲۰۰۵).

به عنوان یک ابزار عملی، سی کیو اف معایب سی ای اف را با X_i پیوسته و یا با ابعاد بزرگ دارد و بنابراین امکان دارد، برآورد و خلاصه کردن آن دشوار باشد. بنابراین ما دوست داریم این تابع را به مجموعه‌ای از اعداد، هر عدد برای هر عنصر X_i، بشکنیم. رگرسیون چندکی این کار را با جایگزینی یک مدل خطی برای $q(X_i)$ در (۱-۱-۷) انجام می‌دهد:

$$\beta_\tau \equiv \arg\min_{b \in \mathbb{R}^d} E\left[\rho_\tau(Y_i - X_i'b)\right]. \qquad (7.1.2)$$

برآوردگر رگرسیون چندکی، $\hat{\beta}_\tau$، نمونه مشابه (۲-۱-۷) است. متوجه می‌شویم که این یک مسئله برنامه‌نویسی خطی است که به راحتی توسط رایانه حل می‌شود. همان طور که **OLS** یک مدل خطی برای Y_i با به حداقل رساندن مربعات خطای انتظاری است، مینیماند رگرسیون چندکی یک مدل خطی را با استفاده از تابع زیان نامتقارن $\rho_\tau(\cdot)$ به Y_i برازش می‌کند. اگر در واقع $Q_\tau(Y_i|X_i)$ خطی باشد، مینیماند رگرسیون چندکی آن را پیدا خواهد کرد (همان طور که اگر **CEF** خطی باشد توسط **OLS** پیدا خواهد شد). مدل اصلی رگرسیون چندکی، که توسط کوئنکر و باست (۱۹۷۸) معرفی شده بود، بر این فرض استوار بود که سی کیو اف خطی است. همان طور که مشخص است، فرضیه یک سی کیو اف خطی غیرضروری است ـ رگرسیون چندکی مفید است چه آن را باور داشته باشیم یا نه.

قبل از بحث کلی نظری رگرسیون چندکی، استفاده از این ابزار برای مطالعه توزیع درآمد را نشان می‌دهیم. انگیزه استفاده از رگرسیون چندکی برای بررسی توزیع درآمد، از علاقه اقتصاددانان به سؤال این مسئله نشئت می‌گیرد که چگونه نابرابری در شرایط مختلف از جمله متغیرهایی مانند تحصیلات و تجربه تغییر می‌کند (نگاه کنید به بوچینسکی، ۱۹۹۴). شکاف کلی در درآمد بر اساس گروه تحصیلی (به عنوان مثال تفاوت میان دیپلم دانشگاه و دبیرستان) در دهه‌های ۱۹۸۰ و ۱۹۹۰ به طور قابل توجهی افزایش یافت. با این حال، اینکه توزیع درآمد در داخل گروه‌های آموزشی و تجربی تغییر کرده است، کمتر واضح است. بسیاری از اقتصاددانان کار، بر این باورند که افزایش به اصطلاح «درون‌گروهی»

1. Least Absolute Deviation

نابرابری، شواهد ویژه‌ای را مبنی بر تغییرات بنیادین در بازار کار ارائه می‌دهد که به آسانی نمی‌توان آن را به تغییرات در ویژگی‌های نهادی مانند درصد کارگران متعلق به اتحادیه‌های کارگری نسبت داد.

جدول ۱-۱-۷. رگرسیون چندکی برای ضرایب تحصیلات در سال‌های ۱۹۷۰، ۱۹۸۰ و ۲۰۰۰

		Desc. Stats.		Quantile Regression Estimates					OLS Estimates	
Census	Obs.	Mean	SD	0.1	0.25	0.5	0.75	0.9	Coeff.	Root MSE
				.074	.074	.068	.070	.079	.072	
1980	65023	6.4	0.67	(.002)	(.001)	(.001)	(.001)	(.001)	(.001)	0.63
				.112	.110	.106	.111	.137	.114	
1990	86785	6.46	0.06	(.003)	(.001)	(.001)	(.001)	(.003)	(.001)	0.64
				.092	.105	.111	.120	.157	.114	
2000	97397	6.5	0.75	(.002)	(.001)	(.001)	(.001)	(.004)	(.001)	0.69

نکته: برگرفته از آنگریست، چرنوژکو و فرناندزوال (۲۰۰۶). جداول برآوردهای رگرسیون چندکی درآمد و تحصیلات را نشان می‌دهد. برآوردهای OLS در سمت راست برای مقایسه آورده شده‌اند. نمونه مورد بررسی شامل مردان ۴۰ تا ۴۹ ساله سفید و سیاهپوست متولد آمریکاست. خطاهای استاندارد در پرانتز نمایش داده شده است. تمام مدل‌ها برای نژاد و تجربه بالقوه کنترل می‌شوند. وزن نمونه‌گیری برای برآوردهای سرشماری ۲۰۰۰ استفاده شده است.

جدول ۱-۱-۷ ضرایب تحصیلی رگرسیون‌های چندکی برآورد شده با استفاده از داده‌های سرشماری‌های ۱۹۸۰، ۱۹۹۰ و ۲۰۰۰ را گزارش می‌کند. مدل‌های مورد استفاده برای این برآوردها، نژاد و یک تابع درجه دوم تجربه بالقوه بازار کار را کنترل می‌کنند (تعریف شده تحت عنوان سن و آموزش). ضرایب چندک ۰/۵- برای میانگین مشروط ـ بسیار شبیه به همتای OLSشان در سمت راست بالای جدول هستند. به عنوان مثال، برآورد او ال اس از ۰/۷۲ در سرشماری سال ۱۹۸۰ چندان فرق زیادی با ضرایب برآورد چندک ۰/۵- حدود ۰/۶۸- با همان داده‌ها ندارد. اگر توزیع مشروط به متغیرهای کمکی[1] لگاریتم دستمزدها متقارن باشد، به طوری که میانه شرطی برابر با میانگین شرطی باشد، باید انتظار داشته باشیم که این دو ضریب یکسان باشند. همچنین این نکته قابل توجه است که ضرایب چندکی در چندک‌های سال ۱۹۸۰ مشابه هستند. افزایش یک سال به تحصیلات، میانه دستمزدها را ۶/۸ درصد اضافه می‌کند که این تأثیر در چارک‌های پایین‌تر و بالاتر از ۰/۷۴ و ۰/۷۰ اندکی بیشتر است. اگرچه درآمد متناسب با تحصیلات برآورد شده در سال‌های ۱۹۸۰ و ۱۹۹۰ به شدت افزایش یافت (تا ۰/۱۰۶ در میانه و با ۰/۱۱۴ درصد در OLS)، الگوی باثبات درآمدی در چندک‌ها در سرشماری ۱۹۹۰ وجود دارد. بیشترین تأثیر در دهک بالا، با ضریب ۰/۱۳۷ است، در حالی که ضرایب سایر دهک‌ها حدود ۰/۱۱ است.

اگر آثار تحصیلات بر دستمزد به چیزی که گاهی اوقات تغییر مکان نامیده می‌شود، نسبت داده شود، باید انتظار داشته باشیم که ضرایب ثابت[2] را در میان چندک‌ها مشاهده کنیم. در اینجا، این بدین

1. Conditional-on-covariates Distribution
2. Location Shift

معناست که هم‌زمان که افزایش سطح تحصیلات متوسط درآمد را افزایش می‌دهد، بخش‌های دیگر توزیع دستمزد در کنار یکدیگر حرکت می‌کنند (یعنی نابرابری درون‌گروهی تغییر نخواهد کرد). فرض کنید، برای مثال، لگاریتم دستمزدها را می‌توان با یک مدل رگرسیون خطی کلاسیک توصیف کرد:

$$Y_i \sim N(X_i'\beta, \sigma_\varepsilon^2), \qquad (7.1.3)$$

به طوری که $E[Y_i|X_i] = X_i'$ و $Y_i - X_i'\beta \equiv \varepsilon_i$ یک خطای توزیع نرمال با واریانس ثابت σ_ε^2است. هم‌واریانسی، به این معناست که توزیع مشروط لگاریتم دستمزدها برای فارغ‌التحصیلان دانشگاه بیشتر از فارغ‌التحصیلان دبیرستان نیست. پیامدهای مدل خطی هم‌واریانس برای چندک‌ها از این واقعیت آشکار است که:

$$P[Y_i - X_i'\beta < \sigma_\varepsilon \Phi^{-1}(\tau)|X_i] = \tau,$$

به طوری که $\Phi^{-1}(\tau)$ معکوسی از نرمال استاندارد CDF است. ازاین‌رو نتیجه می‌گیریم که $Q_\tau(Y_i|X_i) = X_i'\beta + \Phi^{-1}(\tau)$. به عبارت دیگر، صرفنظر از عرض از مبدأ در حال تغییر، تمام ضرایب رگرسیون‌های چندکی یکسان هستند. نتایج جدول ۱-۱-۷ برای سال‌های ۱۹۸۰ و ۱۹۹۰ چندان از نمایش آشکار شده مذکور دور نیستند.

در تمایز با الگوی ساده داده‌های سرشماری ۱۹۸۰ و ۱۹۹۰، رگرسیون چندکی بر روی داده‌های ۲۰۰۰، به طور قابل ملاحظه‌ای، به ویژه در دُم راست توزیع، در چندک‌ها تفاوت می‌کنند. اضافه کردن یک سال به تحصیلات، دستمزد دهک پایین را ۹/۲ درصد و دستمزدهای میانه را ۱۱/۱ درصد افزایش می‌دهد. در تمایز با این نتایج، یک سال اضافه کردن به تحصیلات، دستمزدهای دهک بالا را ۱۵/۷ درصد افزایش می‌دهد. بنابراین، علاوه بر افزایش نابرابری کلی در سال‌های ۱۹۸۰ و ۱۹۹۰ (حقیقتی که ما از آمارهای توصیفی ساده هم می‌فهمیم)، در سال ۲۰۰۰، نابرابری با آموزش هم شروع به افزایش کرد. این رویداد مورد توجه و بحث گسترده‌ای در میان اقتصاددانان کار قرار گرفت، چراکه بحث بر این بود که این مسئله مربوط به تغییرات بنیادی یا نهادی در بازار کار است یا نه (به عنوان مثال به اتور، کاتز و کرنی، ۲۰۰۵ و لمیکس ۲۰۰۸ مراجعه کنید).

باز هم یک مثال پارامتری به ما کمک می‌کند تا ببینیم یک الگوی افزایشی ضرایب رگرسیون چندکی از کجا ممکن است نشئت بگیرد. می‌توانیم با استفاده از مدل‌های رگرسیون نرمال کلاسیک (۳-۱-۷) با اضافه کردن ناهم‌واریانسی، ضرایب رگرسیون چندکی افزایشی تولید کنیم. فرض کنید که:

$$Y_i \sim N(X_i'\beta, \sigma^2(X_i)),$$

به طوری که $\sigma^2(X_i) = (\lambda' X_i)^2$ و λ یک بردار از ضرایب مثبت است، به طوری که

$\lambda' X_i > 0$ (شاید متناسب با β، به طوری که واریانس شرطی با میانگین شرطی رشد می‌کند).[۱]

بنابراین،

$$P[Y_i - X_i'\beta < (\lambda' X_i)\Phi^{-1}(\tau)|X_i] = \tau,$$

که پیامد آن به این صورت می‌شود:

$$Q_\tau(Y_i|X_i) = X_i'\beta + (\lambda' X_i)\Phi^{-1}(\tau) = X_i'[\beta + \lambda\Phi^{-1}(\tau)]. \qquad (7.1.4)$$

به طوری که ضرایب رگرسیون چندکی در میان چندک‌ها افزایش می‌یابد.

اگر قطعات مختلف این مسئله را کنار هم قرار دهیم، جدول ۱-۱-۷ دو بخش این ماجرا را خلاصه می‌کند که هر دو آنها به واریانس در نابرابری درون‌گروهی مربوط است. نخست، نتایج حاصل از سرشماری سال ۲۰۰۰ نشان‌دهنده نابرابری‌ای است که به شدت با آموزش افزایش می‌یابد. با این حال، این افزایش نامتقارن است و در دنباله بالا از توزیع دستمزد به مراتب واضح‌تر است. دوم، این افزایش یک پدیده جدید است. چراکه به طور متمایزی، در سال‌های ۱۹۸۰ و ۱۹۹۰، تحصیلات، توزیع دستمزد را به شیوه‌ای تقریباً منطبق با یک تغییر ساده در موقعیت مکانی تحت تأثیر قرار داد.[۲]

۱-۱-۷. رگرسیون چندکی سانسور شده

رگرسیون چندکی به ما اجازه می‌دهد تا به ویژگی‌های توزیع شرطی Y_i، هنگامی که بخشی از توزیع، پنهان است نگاه کنیم. فرض کنید شما داده‌ای با این فرم دارید:

$$Y_{i,obs} = Y_i \cdot 1[Y_i < c], \qquad (7.1.5)$$

به طوری که $Y_{i,obs}$ چیزی است که مشاهده می‌کنید و $Y_{i,obs}$ متغیری است که شما می‌خواهید آن را ببینید. متغیر $Y_{i,obs}$ سانسور شده است ـ اطلاعات Y_i در $Y_{i,obs}$ به دلایل محرمانگی یا سختی یا زمان‌بر بودن جمع‌آوری آن، محدود است. به عنوان مثال، در **CPS**، درآمد بالا برای حفاظت از حریم خصوصی پاسخ‌دهندگان بالا کدگذاری شده هستند. این به بدین معناست که مقدار درآمد در درآمدهای بالا عبارت کدگذاری شده نمایش داده می‌شود. داده‌های «طولِ مدت» نیز ممکن است سانسور شود: در مطالعه اثرات بیمه‌های بیکاری در طول مدت استخدام، ممکن است تقاضاهای جدید برای مقرری بیکاری را تا حداکثر ۴۰ هفته دنبال کنیم. هر کسی که از کار خارج می‌شود و برای مدت طولانی در دوره بیکاری

---

۱. براي يك مثال تجربي از مدل رگرسيون با اين نوع از ناهمواريانسي به كارد و لموكس (۱۹۹۶) مراجعه كنيد. كوئنكر و پورتنوي (۱۹۹۶) آن را مدل مكان – مقايس (Location – scale model) خطي مي‌نامند.

۲. نتايج جدول ۱-۱-۷ شامل دو مجموعه از خطاهاي استاندارد است. نخستينِ آنها خطاهاي استاندارد سنتي هستند كه توسط دستور qreg در استتا كه همچنين نشان‌دهنده استواري مدل است، نمايش داده مي‌شوند و فرض مي‌شود كه CQF به واقع خطي است. فرمول آن عبارتند از:

$$\tau(1-\tau)\{E[f_{u_\tau}(0|X_i)X_iX_i']^{-1}E[X_iX_i']E[f_{u_\tau}(0|X_i)X_iX_i']^{-1},$$

به طوري كه $f_{u_\tau}(0|X_i)$ ، چگالي سنتي پسماند رگرسيون چندكي در صفر است. اگر اين پسماندها هم واريانس باشند، به سادگي مي‌توان نوشت: $\frac{\tau(1-\tau)}{f_{u_\tau}^2(0)}E[X_iX_i']^{-1}$. دومين مجموعه در قبال تشخيص نادرست، استوار هستند كه اين مسئله با فرمول‌هاي انگريست، كرنوژوكو و فرناندز-وال (۲۰۰۶) محاسبه مي‌شود. در اين مثال، تأثير خطي نبودن بر خطاهاي استاندار اندك است.

است در مرز ۴۰ سال سانسور می‌شود. توجه داشته باشید که متغیرهای وابسته محدود مانند ساعات کار یا هزینه‌های پزشکی که در بخش ۲-۴-۳ مورد بحث قرار گرفتند سانسور شده نیستند، چراکه آنها معمولاً بر اساس ماهیت خود صفر می‌شوند، درست مانند متغیرهای موهومی مثل وضعیت اشتغال.

برای مواجهه با متغیرهای مستقل سانسور شده، می‌توان از رگرسیون چندکی برای ارزیابی تأثیرات متغیرهای چندک شرطی زیر نقطه سانسور شده (با فرض سانسور از بالا) استفاده کرد. این منتج از این واقعیت است که تغییر کدینگ درآمدهای بالاتر از دهک بالا به مقادیر دهک بالا تأثیری در میانه ندارد. بنابراین اگر کدگذاری مقادیر بالای پیمایش CPS، تعداد کمی از افراد را متأثر می‌کند (همان طور که اغلب چنین است)، سانسور کردن بر برآورد میانگین مشروط و یا حتی β_τ برای $\tau = .75$ تأثیری ندارد. به همین ترتیب، اگر کمتر از ۱۰ درصد نمونه مستقل مشروط به تمام مقادیر X_i سانسور شده باشد، هنگام برآورد β_τ برای τ تا ۰/۹ می‌توانید به سادگی آن را نادیده بگیرید. همچنین می‌توانید نمونه را به مقادیری از X_i محدود کنید که در آن $Q_\tau(Y_i|X_i)$ زیر c باشد (یا اگر سانسور شده از پایین با $Y_{i,obs} = Y_i \cdot 1[Y_i > c]$ باشد، بالای c است).

پاول (۱۹۸۶) این ایده را با برآوردگر رگرسیون چندکی سانسور شده فرموله می‌کند. به دلیل اینکه ممکن است ما ندانیم که کدام چندک شرطی زیر نقطه سانسور شده (به مثال کدگذاری فکر کنید) است، پاول پیشنهاد می‌کند که ما با عبارت زیر کار کنیم:

$$Q_\tau(Y_i|X_i) = \min(c, X_i'\beta_\tau^c).$$

بردار پارامتری β_τ^c رابطه زیر را حل می‌کند:

$$\beta_\tau^c \equiv \arg\min_{b\in\mathbb{R}^d} E\{1\left[X_i'\beta_\tau^c < c\right] \cdot \rho_\tau(Y_i - X_i'b)\}. \tag{7.1.6}$$

به عبارت دیگر، در اینجا مسئله بهینه‌سازی رگرسیون چندکی را برای مقادیر X_i حل می‌کنیم به طوری که $X_i'\beta_\tau^c < c$. به این ترتیب، ما شباهت نمونه‌ای (آنالوگ نمونه) را به حداقل می‌رسانیم (۷-۱-۶). تا زمانی که داده‌های سانسور نشده به اندازه کافی وجود دارد، برآوردهای حاصله تابع رگرسیون چندکی را به ما می‌دهند که گویی داده سانسور شده وجود ندارد (با این فرض که تابع چندکی شرطی در واقع خطی است) و اگر معلوم شود که چندک‌های شرطی شما زیر نقطه سانسور شده هستند، بنابراین به رگرسیون چندکی متعارف بازگشته‌اید.

شباهت نمونه‌ای از (۶-۱-۷) دیگر یک مشکل برنامه‌نویسی خطی نیست، اما بوچینسکی (۱۹۹۴) یک الگوریتم برنامه‌ریزی تکراری خطی را پیشنهاد می‌کند که ظاهراً به خوبی کار می‌کند. تکرارها به این صورت است: ابتدا c را نادیده می‌گیرید. سپس سلول‌های با $X_i'\beta_\tau^c < c$ خاتمه می‌یابند. سپس با استفاده از این سلول‌ها، رگرسیون چندکی را دوباره تخمین می‌زنید و به همین ترتیب عمل می‌کنید. این الگوریتم تضمین نمی‌کند که به پاسخ برسد، اما به نظر می‌رسد که در عمل جواب می‌دهد. خطاهای

استاندارد می‌توانند بوت استرپ (خودگردان) شوند. بوچینسکی (۱۹۹۴) و چمبرلین (۱۹۹۴) از این رویکرد برای برآورد رابطه درآمد به تحصیل برای کارگران با تجربه که ممکن است درآمد بالاتر از کد مربوط به افراد با درآمد بالا در پیمایش CPS را داشته باشند، استفاده می‌کنند. این تعدیلات در سانسور، به افزایش درآمد مرتبط با تحصیل در این گروه کمک می‌کند.

۲-۱-۷. ویژگی تقریب رگرسیون چندکی*

بعید به نظر می‌رسد **CQF** لگاریتم دستمزدها با توجه به تحصیلات، دقیقاً خطی باشد، بنابراین فرضیه اصلی مدل رگرسیون چندکی نمی‌تواند در این مثال حفظ شود. خوشبختانه، رگرسیون چندکی همچنین می‌تواند به عنوان یک تقریب خطی **MMSE** به **CQF** فهمیده شود، اگرچه در این مورد، مسئله **MMSE** کمی پیچیده‌تر و سخت‌تر از قضیه رگرسیون **CEF** است. برای هر شاخص چندکی $\tau \in (0,1)$، خطای تشخیص رگرسیون چندکی به صورت زیر تعریف می‌شود:

$$\Delta_\tau(\mathrm{X}_i, \beta_\tau) \equiv \mathrm{X}_i'\beta_\tau - Q_\tau(\mathrm{Y}_i|\mathrm{X}_i).$$

به منظور حداقل کردن میانگین موزون مربعات خطای تشخیص، $\Delta_\tau^2(\mathrm{X}_i,\beta)$، می‌توان بردار تجمعی رگرسیون چندکی را نشان داد، همان طور که در قضیه زیر از آنگریست، چرنوژوکوف و فرناندز وال (۲۰۰۶) نشان داده شده است:

قضیه ۱-۱-۷ (تقریب رگرسیون چندکی). فرض کنید که اولاً چگالی شرطی $f_Y(y|\mathrm{X}_i)$ تقریباً مطمئناً وجود دارد، ثانیاً $E[\mathrm{Y}_i]$، $E[Q_\tau(\mathrm{Y}_i|\mathrm{X}_i)]$ و $E\|\mathrm{X}_i\|$ متناهی هستند و ثالثاً β_τ به طور منحصربه‌فردی معادله ۲-۱-۷ را حل می‌کند، بنابراین:

$$\beta_\tau = \arg\min_{b\in\mathbb{R}^d} E\left[w_\tau(\mathrm{X}_i,b)\cdot\Delta_\tau^2(\mathrm{X}_i,b)\right], \qquad (7.1.7)$$

جایی که

$$\begin{aligned} w_\tau(\mathrm{X}_i,b) &= \int_0^1 (1-u)\cdot f_{\epsilon(\tau)}\left(u\Delta_\tau(\mathrm{X}_i,b)|\mathrm{X}_i\right)\, du \\ &= \int_0^1 (1-u)\cdot f_Y\left(u\cdot \mathrm{X}_i'b + (1-u)\cdot Q_\tau(\mathrm{Y}_i|\mathrm{X}_i)|\mathrm{X}_i\right) du \geq 0 \end{aligned}$$

و $\epsilon_i(\tau)$ یک پسماند خاص چندک است،

$$\epsilon_i(\tau) \equiv \mathrm{Y}_i - Q_\tau(\mathrm{Y}_i|\mathrm{X}_i),$$

با چگالی شرطی $f_{\epsilon(\tau)}(e|\mathrm{X}_i)$ در $\epsilon_i(\tau)=e$.

علاوه بر این، هنگامی که Y_i دارای چگالی شرطی هموار است، برای β در همسایگی β_τ داریم:

$$w_\tau(\mathrm{X}_i,\beta) \approx 1/2\cdot f_Y\left(Q_\tau(\mathrm{Y}_i|\mathrm{X}_i)|\mathrm{X}_i\right). \qquad (7.1.8)$$

اگرچه قضیه تقریبی رگرسیون چندکی پیچیده به نظر می‌رسد، اما تصویر کلان، ساده است. ما می‌توانیم به رگرسیون چندکی به عنوان تقریبی از $Q_\tau(Y_i|X_i)$ فکر کنیم، همان طور که OLS تقریبی از $E[Y_i|X_i]$ است. تابع وزنی OLS هیستوگرام X_iاست؛ که ما آن را با $\pi(X_i)$نشان می‌دهیم. تابع وزن رگرسیون چندکی، به صورت ضمنی توسط $w_\tau(X_i,\beta_\tau)\cdot\pi(X_i)$ استنتاج دقیق‌تر از $\pi(X_i)$ است (از آنجایی که انتظار در (۷-۱-۷) بیش از توزیع X_i است، هیستوگرام به طور ضمنی بخشی از تابع وزنی رگرسیون چندکی است). عبارت $w_\tau(X_i,\beta_\tau)$ شامل بردار رگرسیون چندکی است، β_τاما می‌توان آن را با β_τبازنویسی کرد تا آن را به طور کامل حذف کند، به طوری که فقط تابع X_i باشد (برای جزئیات این بحث نگاه کنید به آنگریست، چرنوژکوف و فرناندز وال، ۲۰۰۶). در هر صورت، وزن رگرسیون چندکی تقریباً متناسب با چگالی Y_i در محدوده CQF است.

ویژگی تقریب رگرسیون چندکی در شکل ۱-۱-۷ نشان داده شده است که در آن، تابع چندکی لگاریتم دستمزدهای دارای بیشترین رتبه را با استفاده از داده‌های سرشماری نشان می‌دهد. در اینجا، ما از مزیت گسسته بودن تحصیلات و نمونه‌های بزرگ سرشماری برای برآورد غیرپارامتری CQF با استفاده از محاسبه چندک دستمزدها برای هر سطح تحصیلی استفاده می‌کنیم. پنل‌های A-C یک برآورد غیرپارامتری $Q_\tau(Y_i|X_i)$ همراه با رگرسیون خطی t برای مقادیر ۱۰، ۰، ۵۰ و ۰/۹۰ را نشان می‌دهد که X_i شامل تنها متغیر تحصیلات و یک ثابت می‌باشد. برآورد غیرپارامتری سلول به سلول CQF با استفاده از حلقه‌هایی ترسیم شده است، در حالی که خط رگرسیون چندکی ثابت است. این شکل نشان می‌دهد که چگونه رگرسیون چندکی خطی، تقریب CQF را به دست می‌دهد.

همچنین مقایسه رگرسیون چندکی به یک برازش موزون با هیستوگرام CQF، همانند آنچه برای مقایسه OLS و CEF انجام شد، جالب است. تخمین‌زن فاصله حداقل چمبرلینی[1] (MD) یک مشابه نمونه‌ای بردار $\bar\beta_\tau$ است که با حل کردن رابطه زیر به دست می‌آید:

$$\bar{\beta}_\tau = \arg\min_{b\in\mathbb{R}^d} E\left[(Q_\tau(Y_i|X_i)-X_i'b)^2\right] = \arg\min_{b\in\mathbb{R}^d} E\left[\Delta_\tau^2(X_i,b)\right].$$

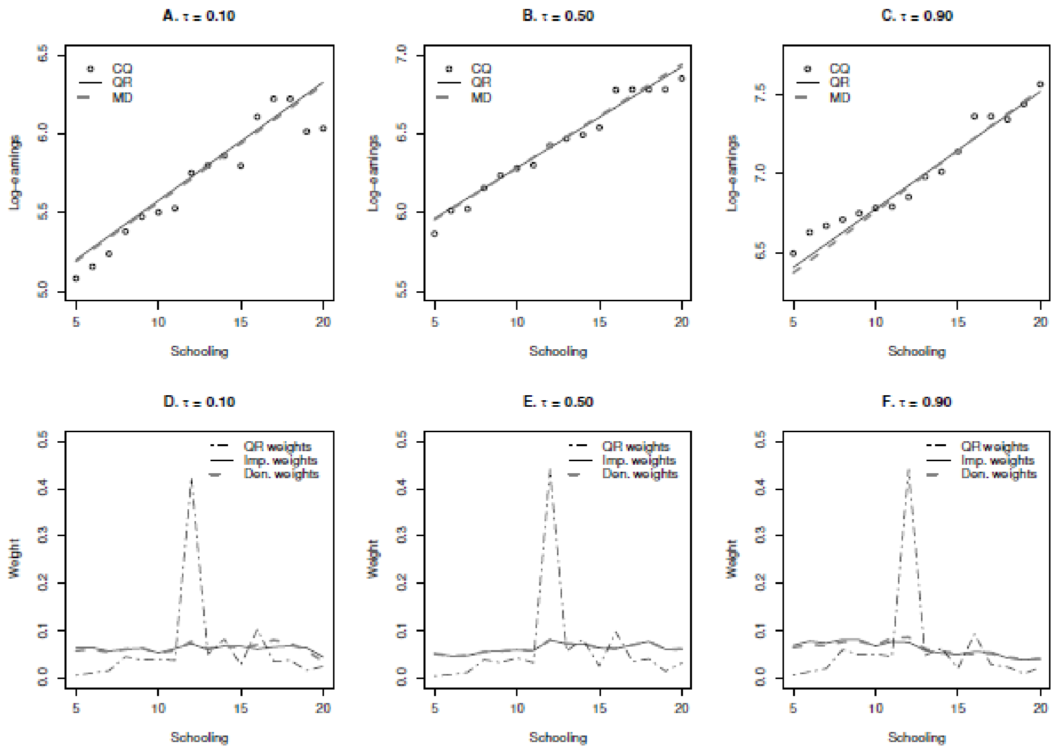
به عبارت دیگر، $\bar\beta_\tau$ شیب رگرسیون خطی $Q_\tau(Y_i|X_i)$ در X_iاست که با هیستوگرام X_i وزن‌دهی می‌شود: در مقایسه با رگرسیون چندکی، که تنها به یک مسیر داده نیاز دارد، **MD** مداوماً در مرحله اول بر توانایی برآورد $Q_\tau(Y_i|X_i)$ به شکل غیرپارامتری تکیه می‌کند.

شکل ۱ مقادیر برازش شده MD را با خط نقطه‌چین نشان می‌دهد. رگرسیون چندکی و خطوط MD اگرچه به هم نزدیک هستند، اما به دلیل وزن داده شدن رگرسیون چندکی با $w_\tau(X_i,\beta_\tau)$، یکسان نیستند. این وزن دادن، کیفیت برازش در مقادیر X_i، جایی که Y_i در نزدیکی CQF چگالی

1. Minimum Distance

بیشتری دارد را برجسته‌تر می‌کند. پنل‌های D-F در شکل ۱-۱-۷ وزن چندکی کلی $w_\tau(X_i, \beta_\tau) \cdot \pi(X_i)$ را در برابر X_i تصویر می‌کند. این پنل‌ها همچنین برآوردهای $w_\tau(X_i, \beta_\tau)$ را با برچسب «وزن‌های اهمیت» و تقریب چگالی آنها، $1/2 \cdot f_Y\left(Q_\tau(Y_i|X_i)|X_i\right)$ نشان می‌دهد. اهمیت وزن‌ها و چگالی وزن‌ها مشابه هم و تقریباً مسطح هستند. تابع وزن‌دهی کلی بسیار شبیه هیستوگرام تحصیلی است و بنابراین، بالاترین وزن را به ۱۲ و ۱۶ سال تحصیل می‌دهد:

شکل ۱-۱-۷. ویژگی تقریب رگرسیون چندکی (برگرفته از آنگریست، چرنوژوکوف و فرناندز وال، ۲۰۰۶)



این شکل، برآوردهای جایگزین تابع چندکی شرطی لگاریتم دستمزدها را بر اساس بیشترین نمره محاسبه شده در داده‌های سرشماری ۱۹۸۰، در کنار تابع ضمنی وزن‌دهی نشان می‌دهد. پنل‌های A تا C، ناپارامتری (CQ)، رگرسیون چندکی (QR) و برآورد حداقل فاصله (MD) را برای $\tau = .1, .5, .9$ نشان می‌دهد. پنل‌های D تا F، توابع وزن‌دهی مربوط به QR و MD را همان طور که در متن توضیح داده شده است نشان می‌دهد.

۳-۱-۷. نکات ظریف

زبان چندک‌های شرطی ظریف است. گاهی اوقات در مورد ضرایب «رگرسیون چندکی در میانه» و یا

«تأثیرات آنهایی که در دهک پایین هستند» صحبت می‌کنیم، اما مهم این است که به خاطر بیاوریم که ضرایب چندکی به ما در مورد تأثیرات توزیع‌ها و نه افراد اطلاعات می‌دهند. اگر برای مثال کشف کنیم که یک برنامه آموزشی باعث افزایش کمینه توزیع دستمزد می‌شود، لزوماً این بدان معنا نیست که کسی که فقیر بوده (به عنوان مثال در دهک پایین بدون آموزش) در حال حاضر کمتر فقیر است. این فقط به این معناست که کسانی که در ساختار با آموزش هستند، نسبت به افراد یک ساختار بدون آموزش، کمتر فقیر هستند.

تمایز قائل شدن بین ثروتمند کردن مجموعه‌ای از افراد فقیر و تغییر آنچه به معنای فقیر بودن است، ظریف است. این تمایز به این بستگی دارد که آیا ما فکر می‌کنیم مداخله، رتبه فرد را در توزیع دستمزد (یا سایر متغیر وابسته) حفظ می‌کند یا نه. اگر یک مداخله حافظ رتبه باشد، افزایش دهک پایین‌تر در واقع باعث می‌شود که کسانی که فقیر بوده‌اند، ثروتمند شوند، چراکه حفظ رتبه به این معناست که وضعیت نسبی تغییر نمی‌کند. در غیر این صورت تنها می‌توانیم بگوییم نیازمندان به عنوان گروهی که در ۱۰ درصد پایین توزیع دستمزد هستند، هر کس که می‌خواهد باشد، وضعیت بهتری پیدا کرده‌اند. ما در بخش ۲-۷ در ادامه توضیح بیشتری در این مورد خواهیم داد.

نکته ظریف دیگر انتقال از چندک شرطی به چندک حاشیه‌ای[1] است. برقراری یک پیوند از چندک‌های شرطی به حاشیه به ما اجازه می‌دهد تا به بررسی تأثیر تغییرات رگرسیون چندکی بر روی کل نابرابری بپردازیم. به عنوان مثال، فرض کنید که ضرایب چندکی با در نظر گرفتن تحصیلات، بیش از گذشته و فراتر از آنچه در داده‌های سرشماری ۲۰۰۰ مشاهده شد، پراکنده می‌شوند. این در ارتباط با نسبت دستمزدهای دهک بالا به دهک پایین متضمن چه معنایی است؟ در عوض می‌توانیم بپرسیم که چه مقدار از افزایش کلی در نابرابری (به عنوان مثال، نابرابری اندازه‌گیری شده توسط نسبت دهک بالا به پایین) به وسیله افزایش پراکندگی ضرایب رگرسیون چندکی توضیح داده می‌شود؟ پاسخ به این نوع سؤالات به طور شگفت‌آوری مشکل به نظر می‌رسد. مشکل این است که با این واقعیت که تمام چندک‌های شرطی برای تعیین کردن یک چندک حاشیه‌ای خاص مورد نیاز است (ماچادو و ماتا، ۲۰۰۵). به طور خاص، $Q_\tau(Y_i|X_i) = X_i'\beta_\tau$ به معنای $Q_\tau(Y_i) = Q_\tau(X_i)'\beta_\tau$ نیست. این تضاد با عملگر انتظاراتِ به مراتب قابل مدیریت‌تر، بیشتر است؛ جایی که اگر $E(Y_i|X_i) = X_i'\beta$ پس، با تکرار انتظارات، داریم $E(Y_i) = E(X_i)'\beta$.

استخراج چندک‌های حاشیه‌ای*

برای نشان دادن رابطه بین چندک‌های شرطی و توزیع‌های حاشیه‌ای فرض می‌کنیم **CQF** در واقع خطی است، به طوری که $Q_\tau(Y_i|X_i) = X_i'\beta_\tau$. اگر $F_Y(y|X_i) \equiv P[Y_i < y|X_i]$ با توزیع

1. Marginal Quantile

حاشیه‌ای $F_Y(y) = P[Y_i < y]$، بنابر تعریف یک چندک شرطی خواهیم داشت:

$$\int_0^1 1[F_Y^{-1}(\tau|X_i) < y]d\tau = F_Y(y|X_i). \quad (7.1.9)$$

به بیان دیگر، نسبت جامعه زیر y مشروط به X_i برابر است با نسبت چندک شرطی که پایین y قرار دارد.[1] جایگزینی CQF داخل انتگرال،

$$F_Y(y|X_i) = \int_0^1 1[X_i'\beta_\tau < y]d\tau.$$

در گام بعد، از CDF از X_i ، $F_X(x)$ استفاده می‌کنیم تا انتگرال گرفته شده و تابع توزیع حاشیه‌ای $F_Y(y)$ به دست آید:

$$F_Y(y) = \int\int_0^1 1[X_i'\beta_\tau < y]d\tau dF_X(x). \quad (7.1.10)$$

در نهایت، چندک حاشیه‌ای، $Q_\tau(Y_i)$، برای $\tau \in (0,1)$، از انحراف $F_Y(y)$ می‌آید:

$$Q_\tau(Y_i) = \inf\ \{y : F_Y(y) \geq \tau\}.$$

تخمین‌زننده توزیع حاشیه‌ای به جای انتگرال‌ها در رابطه ۱۰-۱-۷ جمع می‌گذارد، جایی که انتگرال روی چندک‌ها از برآوردهای رگرسیون چندکی، به عنوان مثال هر ۰/۰۱ چندک به دست می‌آید. در نمونه‌ای با اندازه n این یعنی:

$$\hat{F}_Y(y) = n^{-1}\sum_i (1/100)\sum_{\tau=0}^{\tau=1} 1[X_i'\hat{\beta}_\tau < y].$$

برآوردگر چندکی حاشیه‌ای مرتبط، $\hat{F}_Y(y)$ را معکوس می‌سازد.

این رویکرد در عمل با برخی مشکلات روبه‌رو است. به عنوان نمونه برای به دست آوردن یک چیز، باید تعداد زیادی رگرسیون چندکی انجام دهید. مشکل دیگر این است که نظریه توزیع، آشفته است (اگرچه غیرقابل تحمل نیست؛ نگاه کنید به، مانند ملی، ۲۰۰۵). ساده‌سازی انتقال معادله شرطی به سمت حاشیه‌ای در حال حاضر یک حوزه مطالعاتی فعال است. گوسلین، مچین و مغیر (۲۰۰۰) و ماچادو و ماتا (۲۰۰۵) در میان نخستین مطالعات تجربی هستند که از معادلات شرطی به حاشیه‌ای می‌رسند. هنگامی که متغیر مورد نظر در یک مدل رگرسیون چندکی یک متغیر موهومی است و سایر رگرسیون‌ها به عنوان کنترل دیده می‌شوند، می‌توان از یک طرح وزن‌دهی نوع نمره ـ گرایشی[2] برای به دست آوردن تفاوت‌ها در چندک‌ها در شرایط موهومی استفاده کرد. فیرپو (۲۰۰۷) برای مورد برون‌زا به فیرپو (۲۰۰۷)

---

۱. به عنوان مثال، اگر y میانه شرطی باشد، بنابراین $F_Y(y|X_i) = .5$ و نیمی از چندک‌های شرطی زیر y قرار می‌گیرند. رابطه ۹-۱-۷ با استفاده از تغییر در فرمول متغیرها قابل اثبات است.

2. Propensity-score Type Weighting

و برای یک طرح حاشیه‌ای‌سازی که مناسب تأثیرات تیماری که در بخش بعدی بحث خواهد شد، است. به وفرولیش و ملی (۲۰۰۷) مراجعه کنید.

۲-۷. اثرات تیمارهای چندکی

سؤال با ارزش در رابطه با هر مجموعه برآورد رگرسیون این است که آیا آنها یک تفسیر علّی دارند یا نه. این سؤال برای رگرسیون چندکی از OLS کم‌اهمیت‌تر نیست. فرض کنید ما علاقه‌مند به تخمین آثار یک برنامه آموزشی در مورد درآمد هستیم. برآوردهای رگرسیونی OLS، تأثیر برنامه بر میانگین درآمدها را در حالی که برآورد رگرسیون چندکی می‌تواند برای اندازه‌گیری تأثیر برنامه بر میانه درآمدها استفاده شود محاسبه می‌کنند.

در هر دو مورد، ما باید نگران این باشیم که آیا اثرات برآورد شده برنامه، به وسیله اریب متغیرهای حذف شده مشکل‌دار می‌شوند یا نه.

در اینجا نیز مشکلات متغیرهای حذف شده را می‌توان با استفاده از متغیرهای ابزاری حل کرد، هر چند روش IV برای مدل‌های چندکی یک توسعه نسبتاً جدید هستند و هنوز مثل 2SLS متداول نیستند. ما در اینجا رویکردی که ضرایب علیت یک متغیر باینری بر روی یک چندک (به عنوان مثال، اثر تیمار) با استفاده از یک ابزار باینری را بررسی می‌کند، بحث می‌کنیم. برآوردگر اثرات تیمار چندکی (QTE[1]) که در آبادیه، آنگریست وایمبنز (۲۰۰۲) معرفی شده است، اساساً بر همان مفروضاتی که چارچوب LATE بر میانگین اثرات علّی دارند، تکیه می‌کند. نتیجه، یک برآوردگر وزن‌دهی از نوع آبادیه برای اثرات علّی تیمار چندکی برای موافقت‌کنندگان (پذیرندگان)[2] است.[3]

بحث ما در مورد برآوردگر QTE بر اساس یک مدل افزایشی برای چندک‌های شرطی، به طوری که یک اثر تیماری منفرد تخمین زده می‌شود، است. اگر هیچ گونه ابزارسازی‌ای در کار نباشد، این برآوردگر به رگرسیون چندکی خطی، ساده می‌شود (کوئنکر و باست، ۱۹۷۸). رابطه بین QTE و رگرسیون چندکی مشابه رابطه میان 2SLS و OLS، زمانی که رگرسور مورد نظر موهومی باشد است: پارامترهای مورد نظر به شرح زیر تعریف می‌شوند. برای $\tau \in (0,1)$ فرض می‌کنیم که $\alpha_\tau \in \mathbb{R}$ و $\beta_\tau \in \mathbb{R}^r$ باشد به گونه‌ای که:

$$Q_\tau(Y_i|X_i, D_i, D_{1i}>D_{0i}) = \alpha_\tau D_i + X_i'\beta_\tau, \qquad (7.2.1)$$

به طوری که $Q_\tau(Y_i|X_i,D_i,D_{1i}>D_{0i})$ مقدار τ- چندکی از Y_i ، با توجه به X_i و D_i برای

1. Quantile Treatment Effects
2. Compliers
۳. براي يك رويكرد جايگزين، نگاه كنيد به: چرنوژكو و هانسن (۲۰۰۵) كه به رگرسورهايي از اين نوع (نه فقط موهومي‌ها) اجازه مي‌دهند، اما فرض عدم تغيير رتبه را نيز در نظر مي‌گيرند كه براي چارچوب QTE ضرورتي ندارد.

موافقت‌کنندگان را نشان می‌دهد. α_τو β_τ ضرایب رگرسیون چندکی برای تهیه‌کننده‌ها هستند.

یادآوری می‌کنیم همان طور که در (۲-۵-۴) مورد بحث قرار گرفت، D_i مستقل از خروجی بالقوه، مشروط به X_i و $D_{1i}>D_{0i}$است. پارامتر α_τ در این مدل، تفاوت چندک‌های مشروط به X_i برای Y_{1i} و Y_{0i} برای موافقان را به دست می‌دهد. به عبارت دیگر،

$$Q_\tau(Y_{1i}|X_i, D_{1i}>D_{0i}) - Q_\tau(Y_{0i}|X_i, D_{1i}>D_{0i}) = \alpha_\tau \qquad (7.2.2)$$

این رابطه می‌گوید که برای مثال آیا یک برنامه آموزشی، میانه شرطی یا دهک پایین درآمد موافقان را تغییر می‌دهد یا نه. توجه داشته باشید که α_τ پارامتر به ما نمی‌گوید که آیا تیمار، نتایج چندک توزیع غیرشرطی Y_{1i}و Y_{0i} را تغییر داد یا نه. برای این کار، ما باید خانواده‌های نتایج رگرسیون چندکی را با استفاده از رویه‌هایی که در بخش ۳-۱-۷ شرح داده شده است، ادغام کنیم.

این رابطه همچنین تأکید می‌کند که α_τ چندک شرطی اثرات منفرد تیمار $(Y_{1i} - Y_{0i})$ نیست. برای مثال، ممکن است بخواهید بدانید که آیا میانه اثرات تیمار مثبت است یا نه. متأسفانه، پاسخ دادن به چنین سؤالاتی بدون مفروضات استنباط علّی بسیار دشوار است.[1] حتی یک آزمایش تصادفیده با تکمین[2] نیز نمی‌تواند توزیع $(Y_{1i}-Y_{0i})$ را آشکار کند. از آنجا که میانگین تفاضل‌ها، تفاضل میانگین‌هاست، این به معنای متوسط اثرات تیمار نیست، اما همه ویژگی‌های دیگر توزیع $Y_{1i}-Y_{0i}$ پنهان هستند، چراکه ما هرگز نمی‌توانیم برای هریک از افراد Y_{1i}و Y_{0i}را با هم ببینیم. خبر خوب برای اقتصاددانان کاربردی این است که تفاوت در توزیع‌های حاشیه‌ای (۲-۲-۷) معمولاً بیشتر از توزیع تیماری است، چراکه مقایسه رفاه کل در اقتصاد معمولاً نیاز به توزیع حاشیه‌ای Y_{1i} و Y_{0i} و نه توزیع تفاضل آنها دارد (به عنوان مثال، برای دیدگاه سنتی در این موضوع نگاه کنید به آتکینسون (۱۹۷۰)). این نکته را می‌توان با مثال بدون اشاره به چندک‌ها بیان کرد. در هنگام ارزیابی یک برنامه اشتغال، اگر آن برنامه میزان اشتغال را افزایش دهد، ما تمایل داریم که این برنامه را مطلوب ارزیابی کنیم. به عبارت دیگر، اگر میانگین Y_{1i} بزرگ‌تر از میانگین Y_{0i} باشد، ما خوشحال می‌شویم و از آنجا که تعداد برندگان نسبت به بازندگان در یک برنامه خوب ضرورتاً بیشتر است، بنابراین، به نظر می‌رسد، تعداد افرادی که شغل به دست می‌آورند $(Y_{1i}-Y_{0i} = 1)$ یا شغلشان را از دست می‌دهند $(Y_{1i}-Y_{0i} = 0)$ در وهله دوم اهمیت قرار داشته باشد.

۱-۲-۷. برآوردگر QTE

انگیزه پشت برآوردگر QTE این مشاهده است که از آنجایی که پارامترهای مورد نظر، ضرایب رگرسیون چندکی هستند، بنابراین آنها به لحاظ نظری می‌توانند مداوماً از طریق اجرای رگرسیون‌های چندکی در

۱. برای مثال، نگاه کنید به: هکمن، اسمیت و کلمنتز (۱۹۹۷).

2. Compliance

جامعه موافقان برآورد شوند. هرچند، همچون همیشه، جامعه موافقان قابل شناسایی نیست و ما نمی‌توانیم موافقان را در یک مجموعه داده فهرست کنیم. با وجود این، همانند بخش ۲-۵-۴، مینیماند اقتصادسنجی مربوطه را می‌توان با استفاده از قضیه آبادی کاپا[1] به دست آورد. به ویژه:

$$(\alpha_\tau, \beta_\tau) = \arg\min_{a,b} E\{\rho_\tau(Y_i - aD_i - X_i'b)|D_{1i} > D_{0i}\} = \arg\min_{a,b} E\{\kappa_i\rho_\tau(Y_i - aD_i - X_i'b)\}, \qquad (7.2.3)$$

که در آن،

$$\kappa_i = 1 - \frac{D_i(1-Z_i)}{1-P(Z_i=1|X_i)} - \frac{(1-D_i)Z_i}{P(Z_i=1|X_i)},$$

که در آن، برآوردگر QTE نظیر نمونه‌ای ۳-۲-۷ است.

هنگام پیاده‌سازی QTE، برخی مسائل کاربردی به وجود می‌آیند. ابتدا باید κ_i تخمین زده شود و سپس واریانس نمونه‌گیری ناشی از تخمین اول مرحله در نظریه توزیع مجانبی مربوطه منعکس شود. آبادی، آنگریست و ایمبرز (۲۰۰۲) توزیع محدودی از نظیر نمونه‌ای (۳-۲-۷)، زمانی که κ_i به طور غیرپارامتری ارزیابی می‌شود را استخراج می‌کنند. با این حال، در عمل، خودگردان‌سازی (بوت استرپ) کل رویه (به عنوان مثال، با شروع کردن از ایجاد کاپاهای برآورد شده) نسبت به استفاده از فرمول‌های مجانبی ساده‌تر است.

دوم، هنگامی که $D_i \neq Z_i$ است، κ_i منفی است. بنابراین مینیماند (حداقل‌شوند) رگرسیون چندکی موزون با کاپا[2] کوژ (محدب) نیست و دیگر نمایش برنامه‌نویسی خطی ندارد. این مشکل می‌تواند با حداقل کردن مسئله زیر حل شود:

$$\min_{a,b} E\{E[\kappa_i|Y_i, D_i, X_i]\rho_\tau(Y_i - aD_i - X_i'b)\} \qquad (7.2.4)$$

این مینیماند از تکرار انتظارات (۳-۲-۷) به دست می‌آید. تفاوت‌های عملی بین (۳-۲-۷) و (۲-۷-۴) در این عبارت است:

$$E[\kappa_i|Y_i, D_i, X_i] = P[D_{1i} > D_{0i}|Y_i, D_i, X_i]$$

که یک احتمال است و بنابراین بین صفر و یک است. ساده‌سازی بیشتر از این واقعیت به دست می‌آید که:

$$E[\kappa_i|Y_i, D_i, X_i] = 1 - \frac{D_i(1-E[Z_i|Y_i, D_i=1, X_i])}{1-P(Z_i=1|X_i)} - \frac{(1-D_i)E[Z_i|Y_i, D_i=0, X_i]}{P(Z_i=1|X_i)}. \qquad (7.2.5)$$

آنگریست (۲۰۰۱) از این رابطه برای پیاده‌سازی QTE از طریق پروبیت مرحله اول برای تخمین جداگانه $D_i = 0$ و $D_i = 1$ $E[Z_i|Y_i, D_i, X_i]$ استفاده می‌کند، که این پیاده‌سازی از طریق برآورد

1. Abadie Kappa theorem
2. Kappa-weighted Quantile Regression

$E[\kappa_i|Y_i,D_i,X_i]$ با استفاده از (۵-۲-۷)، و سپس پیرایش[1] هریک از برآوردهای $E[\kappa_i|Y_i,D_i,X_i]$ که خارج از فاصله واحد هستند حاصل می‌شود. برآورد اولیه مرحله اول $E[\kappa_i|Y_i,D_i,X_i]$ می‌تواند به سادگی به عنوان وزن در هنگام ساخت برآوردهای رگرسیون چندکی در مرحله دوم با استفاده از دستور qreg در استتا استفاده شود.

محاسبه تأثیر آموزش در چندک‌های درآمدی کارآموزان

قانون مشارکت در آموزش مشاغل (JTPA)[2] یک برنامه بزرگ فدرال بود که آموزش‌های یارانه‌ای را برای کارگران محروم آمریکایی در دهه ۱۹۸۰ ارائه داد. خدمات JTPA در ۶۴۹ مکان مختلف، که نواحی ارائه خدمت[3] (SDAs) نیز نامیده می‌شدند، در سراسر کشور ارائه شد. مطالعه اصلی در مورد تأثیر JTPA بر بازار کار، بر اساس نمونه ۱۵,۹۸۱ نفری از افرادی بود که اطلاعات مداوم درآمدی آنها (از پرونده‌های بیمه بیکاری دولتی یا دو پیمایش پیگیری) حداقل ۳۰ ماه پس از انتساب تصادفی موجود بود. از این بین اطلاعات ۶,۱۰۲ زن و ۵,۱۰۲ مرد با درآمد ۳۰ماهه وجود دارد.

در نشانه‌گذاری ما، Y_i درآمد ۳۰ماهه، D_i نشان‌دهنده ثبت نام برای خدمات JTPA، و Z_i نشانه دریافت پیشنهاد تصادفی ارائه خدمات توسط JTPA است. یکی از ویژگی‌های کلیدی بسیاری از آزمایش‌های اجتماعی، همانند بسیاری از آزمایشات تصادفیده داروها و درمان‌های جدید، این است که برخی از شرکت‌کنندگان، مداخله ارائه شده را رد می‌کنند. در JTPA، این افرادی که به آنها پیشنهاد ارائه خدمت می‌شد، مجبور نبودند در برنامه آموزشی شرکت کنند. در نتیجه، اگرچه آموزش‌های یارانه‌ای به طور تصادفی اختصاص داده شد، تنها ۶۰ درصد از این افراد که به آنها خدمات آموزشی ارائه شد، JTPA را دریافت کردند. بنابراین تیمار دریافت شده تا حدودی انتخاب شده توسط خود افراد و احتمالاً با نتایج احتمالاً همبستگی دارد. از سوی دیگر، ارائه تصادفیده آموزش ابزاری خوبی برای خدمات آموزشی دریافت شده است، چراکه این متغیر همبسته بوده و پیشنهاد ارائه تیمار مستقل از خروجی بالقوه است. علاوه بر این، به دلیل تعداد بسیار کم (درصد پایین) افراد دریافت‌کننده خدمات JTPA در گروه کنترل (کمتر از ۲ درصد)، تأثیرات موافقان (پذیرندگان) در این مورد می‌تواند به عنوان تأثیرات آنهایی که مورد تیمار قرار گرفتند (همواره ـ گیرندگان[4] کمی وجود دارند) تفسیر شود.

از آنجا که آموزش‌های مطالعه ملی JTPA به صورت تصادفی ارائه شدند، نیازی به برآورد مداوم تأثیرات بر روی موافقان (پذیرندگان) توسط متغیرهای کمکی (X_i) نیست. با این حال، حتی در آزمایش‌هایی مانند این، کنترل متغیرهای کمکی برای تصحیح ارتباط شانسی بین وضعیت تیمار و

1. Trimming
2. Job Training Partnership Act
3. Service Delivery Areas
4. Always-takers

ویژگی‌های متقاضی به منظور افزایش دقت مطالعه متعارف است (نگاه کنید به فصل ۲). متغیرهایی که در اینجا استفاده می‌شود، اندازه‌گیری‌های اولیه از فرایند دریافت خدمات JTPA است. آنها شامل متغیرهای موهومی برای متقاضیان سیاهپوست و اسپانیایی، یک متغیر موهومی برای فارغ‌التحصیلان دبیرستانی (از جمله دارندگان GED)، متغیرهای موهومی برای متقاضیان متأهل، پنج متغیر موهومی برای گروه‌های سنی و یک متغیر موهومی برای اینکه بدانیم آیا متقاضی پیش از سال تخصیص تصادفی ۱۲ هفته کار کرده است یا نه، بودند. همچنین متغیرهای موهومی برای راهبرد اولیه توصیه شده برای خدمات، (کلاس، آموزش در محل کار (OJT)[1]، کمک به جستجوی شغلی (JSA)[2] و غیره) و یک متغیر موهومی برای اینکه آیا داده‌های درآمدی متعلق به دومین پیمایشی که برای پیگیری مطالعه انجام شد هست یا نه، استفاده شد. از آنجایی که این متغیرها مشخصات جامعه‌شناختی را خلاصه می‌کنند، عموماً جامعه‌شناسی پیچیده و سخت را خلاصه می‌کند، می‌توانیم از تحلیل چندکی برای این منظور استفاده کنیم که به ما بگوید چگونه آزمایش JTPA بر توزیع درآمد در گروه‌های با مشخصات مختلف جامعه‌شناختی اثر گذاشت.

به عنوان یک معیار، برآوردهای OLS و متغیرهای ابزاری سنتی (2SLS) تأثیر آموزش در ستون اول جدول ۱-۲-۷ گزارش شده است. ضریب OLS آموزش دقیقاً ۳,۷۵۴ دلار تخمین زده می‌شود که این ضریب D_i در رگرسیون Y_i روی D_i و X_i است. این تخمین‌ها این واقعیت که دانشجویان خودشان، خودشان را انتخاب کرده‌اند نادیده می‌گیرد. برآوردهای 2SLS در جدول ۱-۲-۷ از ارائه تصادفی تیمار Z_i به عنوان ابزار D_i استفاده می‌کنند. برآورد 2SLS برای مردان ۱۵۹۳ دلار با خطای استاندارد ۸۹۵ دلار، ۱۵۹۳ دلار است که کمتر از نصف اندازه تخمینی OLS مربوطه است.

برآورد رگرسیون چندکی نشان می‌دهد که شکاف در چندک‌ها در شرایط مناسب با توجه به وضعیت تحصیلات، به مراتب، بیشتر از اینکه بالای میانه باشد، کمتر از آن است. این نکته را می‌توان در ستون‌های سمت راست جدول ۱-۲-۷ مشاهده کرد که برآوردهای رگرسیون چندکی برای چندک‌های ۰/۱۵، ۰/۲۵، ۰/۵ یا ۰/۸۵ را نشان می‌دهد. به طور خاص، چندک ۰/۸۵ درآمدی آموزش‌دیدگان، ۱۳ درصد بیشتر از افرادی است که آموزش ندیده‌اند و این در حالی است که چندک ۰/۱۵، کارآموزان حدود ۱۳ درصد بیشتر از افرادی که آموزش ندیده‌اند. همانند برآوردهای OLS در جدول، این ضرایب رگرسیون چندکی، لزوماً یک تفسیر علّی ندارند، بلکه آنها یک مقایسه توصیفی از توزیع درآمد کارآموزان و غیرکارآموزان را ارائه می‌دهند.

برآوردهای QTE اثرات آموزش بر میانه درآمدها، مشابه معیارهای تخمین زده شده توسط 2SLS

1. on-the-job Training
2. Job Search Assistance

است اگرچه دقت کمتری دارند. از سوی دیگر، برآوردهای QTE یک الگوی بسیار متفاوت از برآوردهای رگرسیون چندکی را نشان می‌دهد و هیچ مدرکی مبنی بر تأثیرگذاری بر روی چندک ۰/۱۵ یا ۰/۲۵ وجود ندارد. برآوردهای چندکی در چندک‌های پایین، به طور قابل ملاحظه‌ای کوچک‌تر از برآوردهای رگرسیون چندکی مربوطه هستند و مقادیر مطلق آنها نیز اندک است. به عنوان مثال، برآورد QTE (و خطای استاندارد) تأثیر در چندک ۰/۱۵، ۱۲۱ دلار (۴۷۵) است، در حالی که برآورد رگرسیون چندکی مربوطه ۱,۱۸۷ دلار (۲۰۵) است. به همین ترتیب، تخمین اثر QTE (خطای استاندارد) در چندک ۰/۲۵ برای مردان ۷۰۲ دلار (۶۷۰) است، در حالی که برآورد رگرسیون چندکی مربوطه معادل ۲,۵۱۰ دلار (۳۵۶) است. با این حال، برخلاف نتیجه چندک‌های پایین، برآوردهای QTE از درآمد مردان که بالاتر از میانه قرار دارند، بزرگ و از لحاظ آماری معنادار هستند (گرچه هنوز هم کوچک‌تر از برآوردهای رگرسیون چندکی مربوطه پایین‌ترند).

این نکته که کارآموزی JTPA درآمد چندک‌های پایین‌تر درآمدی مردان را افزایش نمی‌دهد، جالب‌ترین نتیجه ناشی از این تحلیل است. این تحلیل نشان می‌دهد که برآوردهای رگرسیون چندکی در نیمه بالایی جدول ۱-۲-۷ توسط اریب گزینش مثبت مشکل‌دار شده‌اند. یک پاسخ به این یافته ممکن است این باشد که درآمد تعداد اندکی از متقاضیان JTPA بسیار خوب بوده و در نتیجه، اثرات توزیعی میان متقاضیان اهمیت کمتری نسبت به این نکته داشته است که برنامه مذکور به عده زیادی از متقاضیان، روی هم رفته، کمک کرده است. هرچند، چندک‌های بالایی درآمد برای افرادی که در مطالعه JTPA مشارکت داشتند، منطقاً بالا بود. در نتیجه، بعید است این افزایش درآمد در دنباله بالایی اولویت زیادی داشته باشد.

جدول ۱-۲-۷. برآوردهای رگرسیونی چندکی و اثرات تیمار JTPA

الف) برآوردهای OLS و رگرسیون چندکی

	OLS	Quantile				
		0.15	0.25	0.50	0.75	0.85
Training	3,754	1,187	2,510	4,420	4,678	4,806
	(536)	(205)	(356)	(651)	(937)	(1,055)
% Impact of Training	21.20	135.56	75.20	34.50	17.24	13.43
High school or GED	4,015	339	1,280	3,665	6,045	6,224
	(571)	(186)	(305)	(618)	(1,029)	(1,170)
Black	-2,354	-134	-500	-2,084	-3,576	-3,609
	(626)	(194)	(324)	(684)	(1087)	(1,331)
Hispanic	251	91	278	925	-877	-85
	(883)	(315)	(512)	(1,066)	(1,769)	(2,047)
Married	6,546	587	1,964	7,113	10,073	11,062
	(629)	(222)	(427)	(839)	(1,046)	(1,093)
Worked less than 13	-6,582	-1,090	-3,097	-7,610	-9,834	-9,951
weeks in past year	(566)	(190)	(339)	(665)	(1,000)	(1,099)
Constant	9,811	-216	365	6,110	14,874	21,527
	(1,541)	(468)	(765)	(1,403)	(2,134)	(3,896)

ب) برآوردهای 2SLS و QTE

	2SLS	Quantile 0.15	0.25	0.50	0.75	0.85
Training	1,593	121	702	1,544	3,131	3,378
	(895)	(475)	(670)	(1,073)	(1,376)	(1,811)
% Impact of Training	8.55	5.19	11.99	9.64	10.69	9.02
High school or GED	4,075	714	1,752	4,024	5,392	5,954
	(573)	(429)	(644)	(940)	(1,441)	(1,783)
Black	-2,349	-171	-377	-2,656	-4,182	-3,523
	(625)	(439)	(626)	(1,136)	(1,587)	(1,867)
Hispanic	335	328	1,476	1,499	379	1,023
	(888)	(757)	(1,128)	(1,390)	(2,294)	(2,427)
Married	6,647	1,564	3,190	7,683	9,509	10,185
	(627)	(596)	(865)	(1,202)	(1,430)	(1,525)
Worked less than 13	-6,575	-1,932	-4,195	-7,009	-9,289	-9,078
weeks in past year	(567)	(442)	(664)	(1,040)	(1,420)	(1,596)
Constant	10,641	-134	1,049	7,689	14,901	22,412
	(1,569)	(1,116)	(1,655)	(2,361)	(3,292)	(7,655)

نکته: جدول، برآوردهای OLS، رگرسیون چندکی، 2SLS و QTE اثرات کارآموزی بر درآمد را نشان می‌دهد. (اقتباس از آبادی، آنگریست و ایمبنز، ۲۰۰۲) وضعیت تخصیص به عنوان ابزاری برای وضعیت آموزش در پنل ب مورد استفاده قرار می‌گیرد. تمام مدل‌های شامل متغیرهای کمکی موهومی برای راهبرد خدمت توصیه شده و گروه سنی، همچنین موهومیِ نشانگر داده پیمایش دوم هستند. خطاهای استاندارد استوار در پرانتز گزارش شده است.

۸. مسائل غیراستاندارد خطای استاندارد[1]

شرایط عادی است، تکرار می‌کنم شرایط عادی است.

اما اگر هنوز مورد غیرقابل حلی برای شما وجود دارد این مشکل خودتان است.

داگلاس آدامز، راهنمای کهکشانگردی برای مسافران پیاده (۱۹۷۹)

امروزه، بسته‌های نرم‌افزاری به طور معمول، خطاهای استاندارد مجانبی را به گونه‌ای محاسبه می‌کنند که تحت فرضیات ضعیفی در مورد فرایند نمونه‌گیری یا مدل پایه به دست آمده‌اند. برای مثال، خطاهای استاندارد رگرسیون بر پایه فرمول (۷-۱-۳) با استفاده از گزینه «استواری» Stata به دست می‌آید. از آنجایی که استنباط‌های به دست آمده زمانی به صورت مجانبی معتبر هستند که پسماندهای رگرسیون ناهم‌واریانس هستند و این مورد اغلب زمانی اتفاق می‌افتد که رگرسیون یک CEF غیرخطی را تقریب می‌زند، خطاهای استاندارد استوار بر پایه اصلاح خطاهای استاندارد قدیمی شکل گرفته‌اند، در حالی که، خطاهای استاندارد قدیمی با فرض هم‌واریانسی پسماندهای رگرسیون به دست آمده‌اند. توقف ناگهانی (در این برنامه) به این دلیل است که خطاهای استاندارد استوار وقتی که تقریب مجانبی مناسب نباشد، می‌توانند گمراه‌کننده باشند. بخش اول این فصل به بررسی خطای استنباط مجانبی با خطاهای استاندارد استوار و با برخی کاهنده‌های ساده می‌پردازد.

1. Nonstandard Standard Error Issues

یک پایه استنباط مقطعی ـ و بحثی که در بخش ۳-۱-۳ ذکر شد ـ بر این فرض استوار است که داده‌ها مستقل از هم هستند. هر مشاهده‌ای به عنوان یک انتخاب تصادفی از یک جامعه یکسان تلقی می‌شود که نسبت به مشاهدات قبل و بعد از خود ناهمبسته است. ما امروزه می‌دانیم که این مدل نمونه‌گیری غیرواقعی است و بالقوه پُر از احتمال خطا (ریسک) است، (بلکه) بیشتر شبیه آنچه در مطالعات سری‌های زمانی در اقتصاد کلان رایج است، تحلیلگران مقطع داده‌ها می‌باید به ارتباط بین مشاهدات توجه نمایند. مهم‌ترین شکل وابستگی در داده‌هایی با ساختار گروهی ظاهر می‌شود ـ برای مثال، نمره‌های آزمون مشاهداتی که بر روی بچه‌ها در کلاس‌های درس یا مدارس انجام می‌شود. نمره‌های به دست آمده از بچه‌هایی که در یک مدرسه و یا یک کلاس هستند به هم وابسته‌اند، چراکه آنها آزمودنی‌هایی از یک محیط یکسان و یا با اثرات زمینه‌ای خانوادگی یکسانی هستند. ما این نوع همبستگی را مسئله خوشه‌بندی، یا مسئله مولتون[1] می‌نامیم، چراکه در سال (۱۹۸۶) مولتون بود که این نوع از همبستگی را معرفی کرد. مسئله‌ای که بسیار حائز اهمیت است همبستگی طی زمان‌های مختلف در مجموعه داده‌هایی است که معمولاً برای اعمال تفاوت‌هایی در راهکارهای متفاوت تخمین استفاده می‌شوند. برای مثال، تحقیق در مورد سطح حداقل دستمزدهای دولتی می‌باید با توجه به این حقیقت صورت گیرد که میزان متوسط استخدام دولتی طی زمان‌های مختلف به هم وابسته هستند. ما این را مسئله همبستگی سری می‌نامیم که بسیار به مسئله مولتون مربوط است، اما در عین حال از آن متمایز است.

محققانی که به خاطر همبستگی سریالی و خوشه‌بندی دچار مشکل می‌شوند، مجبورند که با این حقیقت روبه‌رو شوند که ساده‌ترین راه‌حل‌ها برای این دست مسائل، مانند انتخاب «خوشه» Stata، لزوماً مناسب‌ترین راهکار نخواهد بود. تقریب مجانبی که به داده‌های همبسته سریالی یا خوشه‌ای مربوط باشد، به تعداد زیاد خوشه‌ها و یا مشاهداتی که در سری‌های زمانی صورت گرفته باشند، وابسته است. متأسفانه ما به ندرت از خوشه‌های فراوان و یا از سری‌هایی در طول زمان طولانی برخوردار می‌شویم. مسائل استنباط به وجود آمده همواره هم غیرقابل حل نیستند، اگرچه گاهی بهترین راه‌حل برای دستیابی به داده‌های بیشتر نیاز دارد. بحث در مورد راه‌حل‌های اقتصادسنجی برای همبستگی‌های سریالی و خوشه‌ای در بخش دو این فصل ارائه شده است. به‌کارگیری برخی از مواد این فصل بدون جبر ماتریس‌ها بسیار دشوار است، در نتیجه ما دل را به دریا زده‌ایم و موضوع را به شکل ماتریسی آن پی می‌گیریم.

1. Moulton Problem

۱-۸. اریب خطاهای استاندارد استوار[1]*

در نمایش ماتریسی،

$$\hat{\beta} = \left[\sum_i X_i X_i'\right]^{-1} \sum_i X_i Y_i = (X'X)^{-1}X'y,$$

که در آن X یک ماتریس N*K با سطرهای X_i' و y برداری است N*1 از Y_iها. در بخش ۳-۱-۳ دیدیم که $\hat{\beta}$ دارای یک توزیع نرمال مجانبی است و می‌توانیم بنویسیم:

$$\sqrt{N}(\hat{\beta} - \beta) \sim N(0, \Omega)$$

که در آن Ω ماتریس کوواریانس مجانب است. با تکرار (۷-۱-۳)، فرمول Ω عبارت است از:

$$\Omega_r = E[X_i X_i']^{-1} E[X_i X_i' e_i^2] E[X_i X_i']^{-1}, \qquad (8.1.1)$$

که در آن $e_i = Y_i - X_i'\beta$ و هنگامی که پسماندها مجانبی هستند، Ω را می‌توان به شکل ساده $\Omega_c = \sigma^2 E[X_i X_i']^{-1}$ که در آن $\sigma^2 = E[e_i^2]$.

ما در اینجا به دنبال اریب خطاهای استاندارد استوار در نمونه‌های مستقل از هم (یعنی بدون همبستگی سریالی یا خوشه‌ای) هستیم. برای ساده‌سازی مشتق اریب، فرض می‌کنیم که بردار رگرسور را می‌توان به مثابه یک مقدار ثابت تکرار شده در همه نمونه‌ها در نظر گرفت و این مسئله شبیه وقتی است که بر نمونه‌های محدود X_i نمونه‌گیری را انجام داده باشیم. رگرسورهای غیرمجانب مدل پایه‌ای را برای نمونه‌گیری فراهم می‌کنند که اغلب برای مشاهده توزیع‌هایی با نمونه‌های محدود به کار می‌رود. می‌توان نتیجه گرفت که با این فرض که به ساده‌سازی قابل توجه مشتق‌ها می‌انجامد تنها میزان کمی (اطلاعات) را از دست می‌دهیم.

با فرض رگرسورهای ثابت خواهیم داشت:

$$\Omega_r = \left(\frac{X'X}{N}\right)^{-1} \left(\frac{X'\Psi X}{N}\right) \left(\frac{X'X}{N}\right)^{-1} \qquad (8.1.2)$$

که در آن $\Psi = E[ee'] = diag(\psi_i)$ ماتریس واریانس پسماندهاست. طبق هم‌واریانسی برای همه **i**ها داریم، $\psi_i = \sigma^2$ و در نتیجه خواهیم داشت:

$$\Omega_c = \sigma^2 \left(\frac{X'X}{N}\right)^{-1}$$

خطاهای استاندارد مجانبی پس از حذف نرمال‌سازی به وسیله تقسیم بر N، با ریشه دوم عنصرهای قطر اصلی Ω_r و Ω_c به دست آمده‌اند.

---

1. The Bias of Robust Standard Errors

در عمل، ماتریس کوواریانس مجانب را باید تخمین (تقریب) زد. ماتریس واریانس تخمین قراردادی عبارت است از:

$$\hat{\Omega}_c = (X'X)^{-1}\hat{\sigma}^2 = (X'X)^{-1}\left(\sum \frac{\hat{e}_i^2}{N}\right),$$

که در آن $\hat{e}_i = Y_i - X_i'\hat{\beta}$ عبارت است از مانده رگرسیون تخمین زده شده، و

$$\hat{\sigma}^2 = \sum \frac{\hat{e}_i^2}{N}$$

واریانس مانده را تخمین می‌زند. ماتریس واریانس تخمین متناظر آن عبارت است از:

$$\hat{\Omega}_r = (X'X)^{-1}\left(\sum \frac{X_i X_i' \hat{e}_i^2}{N}\right)(X'X)^{-1}. \qquad (8.1.3)$$

می‌توانیم به یک جمله مشترک به عنوان تخمین‌زننده شکل $\sum \frac{X_i X_i' \hat{\psi}_i}{N}$ برسیم که در آن $\hat{\psi}_i = \hat{e}_i^2$، ψ_i را تخمین می‌زند.

با توجه به قاعده اعداد بزرگ و اصول اسلاتسکی، $N\hat{\Omega}_c$ احتمالاً به سمت Ω_c میل می‌کند در حالی که $N\hat{\Omega}_r$ به سمت Ω_r میل می‌کند، اما در مورد نمونه‌های متن‌ها، هر دو تخمین‌زننده‌های واریانس اریب هستند. اریب در $\hat{\Omega}_c$ از تئوری قدیمی کوچک‌ترین مربعات به این سو معرفی شده و اصلاح آن ساده است. آنچه کمتر مورد توجه است این مسئله است که اگر پسماندهای (رگرسیون) مجانبی هستند، برآوردگر استوار اریب‌دارتر از حالت پیش‌فرض آن خواهد بود و این (اختلاف) اریب ممکن است خیلی زیادتر هم بشود. می‌توان از اینجا نتیجه گرفت که خطاهای استاندارد استوار بالقوه هستند از خطاهای استاندارد در شرایطی که ناهمواریانسی نسبتاً کم است. می‌توان با یک حساب سرانگشتی هم با استفاده از حداکثر خطاهای استاندارد استوار و قدیمی از بروز اشتباهات فاحش در دقت کار جلوگیری کرد.

با وجود رگرسورهای غیرتصادفی، داریم:

$$E[\hat{\Omega}_c] = (X'X)^{-1}\hat{\sigma}^2 = (X'X)^{-1}\left(\sum \frac{E(\hat{e}_i^2)}{N}\right).$$

برای تحلیل $E[\hat{e}_i^2]$ ، با بسط $\hat{e} = y - X\hat{\beta}$ شروع می‌کنیم:

$$\hat{e} = y - X(X'X)^{-1}X'y = \left[I - X(X'X)^{-1}X'\right](X\beta + e) = Me$$

که در آن e بردار پسماند جامعه است، $M = I_N - X(X'X)^{-1}X'$ یک ماتریس پسماندساز غیرتصادفی است با سطر iام m_i' ، و I_N ماتریس مشابه N*N است. سپس $\hat{e}_i = m_i'e$ ، و

$$\begin{aligned} E\left(\hat{e}_i^2\right) &= E\left(m_i'ee'm_i\right) \\ &= m_i'\Psi m_i \end{aligned}$$

برای ساده‌سازی بیشتر، می‌نویسیم $m_i = \ell_i - h_i$ که در آن ℓ_i ستون iام ماتریس I_N است و $h_i = X(X'X)^{-1}\mathrm{X}_i$، ستون iام ماتریس تصویر عبارت است از $H = X(X'X)^{-1}X'$.

سپس داریم:

$$\begin{aligned} E\left(\hat{\epsilon}_i^2\right) &= \left(\ell_i - h_i\right)' \Psi \left(\ell_i - h_i\right) \\ &= \psi_i - 2\psi_i h_{ii} + h_i' \Psi h_i \end{aligned} \qquad (8.1.4)$$

که در آن h_{ii}، عنصر iام قطری Hاست که در این معادله صدق می‌کند:

$$h_{ii} = h_i' h_i = \mathrm{X}_i'(X'X)^{-1}\mathrm{X}_i. \qquad (8.1.5)$$

همچنین h_{ii} /اهرم/[1] مشاهده‌گر iام نیز خوانده می‌شود. اهرم‌ها به ما می‌گویند که تا چه میزان می‌باید یک مقدار مشخص X_i را بر روی خط رگرسیون اعمال کنیم. به یاد داشته باشیم که مقدار مناسب iام (عنصر iام Hy) عبارت است از:

$$\hat{\mathrm{Y}}_i = h_i' y = h_{ii}\mathrm{Y}_i + \sum_{j \neq i} h_{ij}\mathrm{Y}_j. \qquad (8.1.6)$$

مقدار زیاد h_{ii} به این معناست که مشاهده iام تأثیر زیادی بر مقدار پیش‌بینی شده iام دارد. در یک رگرسیون دومتغیره با یک رگرسور (واحد)، x_i،

$$h_{ii} = \frac{1}{N} + \frac{\left(x_i - \overline{x}\right)^2}{\sum \left(x_j - \overline{x}\right)^2}. \qquad (8.1.7)$$

این نشان می‌دهد وقتی که x_i از مقدار متوسط فاصله دارد، اهرم‌ها افزایش پیدا می‌کنند. علاوه بر ۸-۱-۶، ما می‌دانیم که h_{ii} عددی است که در بازه [۰/۱] قرار دارد و $\sum_{j=1}^{N} h_{ij} = \mathrm{K}$ مقدار رگرسورها است (نگاه کنید به هاوگلین و ولش، ۱۹۷۸).[2]

فرض کنیم پسماندها همواریانس هستند، بنابراین $\psi_i = \sigma^2$. سپس ۸-۱-۴ به صورت زیر ساده می‌شود:

$$E\left(\hat{\epsilon}_i^2\right) = \sigma^2[1 - 2h_{ii} + h_i' h_i] = \sigma^2\left(1 - h_{ii}\right) < \sigma^2.$$

بنابراین $\hat{\Omega}_c$ کوچک‌تر خواهد شد. با استفاده از ویژگی‌های h_{ii}می‌توانیم یک گام پیش‌تر برویم:

---

1. Leverage

۲. ویژگی $\sum_{j=1}^{N} h_{ij} = \mathrm{K}$، از این حقیقت ناشی می‌شود که H عنصری آزاد (خودتوان) است. همچنین می‌توانید از (۷-۱-۸) هم برای اثبات یک رگرسیون دومتغیره $\sum_{j=1}^{N} h_{ij} = 2$ استفاده کنید.

$$\sum \frac{E(\hat{e}_i^2)}{N} = \sigma^2 \sum \frac{1-h_{ii}}{N} = \sigma^2 \left(\frac{N-\mathrm{K}}{N}\right).$$

بنابراین، اریب $\hat{\Omega}_c$ را می‌توان با یک اصلاح درجه آزادی به میزان کم ثابت (تنظیم) کرد: به جای تقسیم بر N که در بیشتر محاسبات واریانس تجربی معمول است با تقسیم بر $\mathrm{N}-\mathrm{K}$ در فرمولی که برای $\hat{\sigma}^2$ است، این اصلاح انجام می‌گیرد.

حالا می‌خواهیم نشان دهیم که تحت هم‌واریانسی، اریب در $\hat{\Omega}_r$ احتمالاً بسیار بیشتر از اریب در $\hat{\Omega}_c$است. اریب در ماتریس کوواریانس استوار تخمین‌زننده عبارت است از:

$$E[\hat{\Omega}_r] = N(X'X)^{-1}\left(\sum \frac{\mathrm{X}_i \mathrm{X}_i' E(\hat{e}_i^2)}{N}\right)(X'X)^{-1}, \qquad (8.1.8)$$

که در آن $E\left(\hat{e}_i^2\right)$ از رابطه (۸-۱-۴) به دست آمده است. تحت هم‌واریانسی، $\psi_i = \sigma^2$ و ما $E\left(\hat{e}_i^2\right) = \sigma^2\left(1-h_{ii}\right)$ را، همانند آنچه در $\hat{\Omega}_c$ داشتیم، داریم.

بنابراین روشن است که اریب در $\hat{e}_i^2$ تمایل به کاهش خطاهای استاندارد استوار دارد. اگرچه ارزیابی توصیف کامل (۸-۱-۸) دشوار است، چشر و جویت (۱۹۸۷) نشان داده‌اند تا وقتی که هم‌واریانسی زیادی وجود ندارد، خطاهای استاندارد که بر مبنای $\hat{\Omega}_r$ هستند در واقع اریب نزولی دارند.[۱]

چگونه می‌دانیم که $\hat{\Omega}_r$ ممکن است اریب‌دارتر از $\hat{\Omega}_c$ باشد؟ تا حدی این نتیجه‌گیری ناشی از شواهد مونت کارلو به دست آمده است (یعنی مک کینون و وایت، ۱۹۸۵ و مطالعه محدود ما که به شرح آن می‌پردازیم). همچنین ما این نتیجه را برای یک نمونه دومتغیره هم اثبات می‌کنیم، که در آن یک رگرسور واحد، $\tilde{x}_i$، به حالت انحرافی از صورت میانگین فرض می‌شود و بنابراین تنها یک ضریب وجود دارد. در این مورد، برآوردگر مورد نظر $\hat{\beta}_1 = \frac{\sum \tilde{x}_i \mathrm{Y}_i}{\sum \tilde{x}_i^2}$ و اهرم‌ها عبارتند از: $h_{ii} = \frac{\tilde{x}_i^2}{\sum \tilde{x}_i^2}$ (با حذف کردن مقدار ثابت جمله $\frac{1}{N}$ را از دست می‌دهیم). فرض کنیم $s_x^2 = \frac{\sum \tilde{x}_i^2}{N}$ برای برآوردگر کوواریانس قراردادی داریم:

$$E[\hat{\Omega}_c] = \frac{\sigma^2}{N s_x^2}\left[\frac{\sum(1-h_{ii})}{N}\right] = \frac{\sigma^2}{N s_x^2}\left[1-\frac{1}{N}\right],$$

بنابراین اریب در اینجا کم است. محاسبه ساده‌ای با استفاده از (۸-۱-۸) نشان می‌دهد که تحت (شرایط) هم‌واریانسی، برآوردگر استوار دارای امید ریاضی زیر است:

$$E[\hat{\Omega}_r] = \frac{\sigma^2}{N s_x^2} \sum \frac{(1-h_{ii})}{N}\left(\frac{\tilde{x}_i^2}{s_x^2}\right) = \frac{\sigma^2}{N s_x^2} \sum (1-h_{ii}) h_{ii} = \frac{\sigma^2}{N s_x^2}\left[1-\sum h_{ii}^2\right].$$

---

۱. به ویژه، تا وقتی که میزان (نرخ) بیشترین ψ_i نسبت به کوچک‌ترین ψ_i کمتر از ۲ باشد، خطاهای استاندارد استوار دارای شیب نزولی خواهند بود.

بنابراین اگر $\sum h_{ii}^2 > \frac{1}{N}$ طبق آنچه در نامعادله ینسن[1] ارائه شده، اریب $\hat{\Omega}_r$ خیلی بیشتر (شدیدتر) از اریب $\hat{\Omega}_c$ خواهد بود در غیر این صورت رگرسور اهرم‌ها (اهرم) ثابتی برای همه **i**ها در $h_{ii}=\frac{1}{N}$ خواهد داشت.[2]

ما می‌توانیم با تلاش برای دستیابی به یک برآوردگر بهتر ψ_i، مثلاً $\widehat{\psi}_i$، اریب را کاهش دهیم. برآوردگر $\hat{\Omega}_r$ منجر به $\widehat{\psi}_i = \widehat{e}_i^2$ می‌شود، برآوردگری که وایت (۱۹۸۰a) آن را مفروض گرفته همان چیزی است که ما این بخش را با آن آغاز می‌کنیم. در اینجا خلاصه‌ای از طرح‌هایی که توسط مک‌کینون و وایت (۱۹۸۵) پیگیری شده‌اند ارائه می‌شود:

$$HC_0 \quad : \quad \widehat{\psi}_i = \widehat{e}_i^2$$

$$HC_1 \quad : \quad \widehat{\psi}_i = \frac{N}{N-\mathrm{K}}\widehat{e}_i^2$$

$$HC_2 \quad : \quad \widehat{\psi}_i = \frac{1}{1-h_{ii}}\widehat{e}_i^2$$

$$HC_3 \quad : \quad \widehat{\psi}_i = \frac{1}{(1-h_{ii})^2}\widehat{e}_i^2.$$

HC_1 اصلاح درجه آزادی ساده‌ای است که برای $\hat{\Omega}_c$ به‌کار می‌رود. HC_2 از اهرم‌ها برای ارائه یک تخمین بدون اریب از تخمین واریانس مانده iام وقتی که پسماندها هم‌واریانس هستند استفاده می‌کند، در حالی که HC_3 یک برآوردگر به روش جک‌نایف را تقریب می‌زند.[3] در کاربردها دیدیم که هرچه به پایین لیست می‌رویم خطاهای استاندارد تخمین‌زده شده متمایل به بزرگ‌تر شدن هستند، اما این یک اصل (قاعده) نیست.

زمان خودگردان‌سازی (بوت استرپ)[4]

خودگردان‌سازی طرحی برای نمونه‌گیری دو بار است که جایگزینی برای استنباط بر پایه فرمول‌های

---

1. Jensen.s Inequality

۲. h_{ii} را به عنوان یک متغیر تصادفی با توزیع یکسان در نمونه‌ها مفروض می‌گیریم. در نتیجه داریم:

$$E[h_{ii}] = \frac{\sum h_{ii}}{N} = \frac{1}{N}$$

و

$$E[h_{ii}^2] = \frac{\sum h_{ii}^2}{N} > (E[h_{ii}])^2 = \left(\frac{1}{N}\right)^2$$

با نامعادله ینسن، در غیر این صورت hii را ثابت فرض می‌کنیم. بنابراین Σ h2ii > 1 / N. اهرم ثابت زمانی اتفاق می‌افتد که برای برخی ثابت‌ها مثل κ داشته باشیم: x~i = +/- κ.

۳. تخمین‌گر واریانس جک‌نایف واریانس نمونه‌گیری را از توزیع تجربی حاصل از حذف یک مشاهده‌گر در لحظه را تقریب می‌زند. استتا HC1, HC2 و HC3 را محاسبه می‌کند. همچنین می‌توانید از راه‌حل پیشنهادی توسط مسر و وایت (۱۹۸٤) استفاده کنید: تقسیم Yiو xi بر √φ^i و برآورد مدل به دست آمده با X / √ φ^i برای هر انتخابی از φi^.

4. Bootstrap

مجانبی ارائه می‌دهد. یک نمونه خودگردان‌ساز، نمونه‌ای است که از داده‌های خودمان به دست آمده است. به بیان دیگر، اگر نمونه‌ای با اندازه **N** داشته باشیم، این نمونه به مثابه جامعه است که ما مرتب از آن (با فرایند جابه‌جایی) نمونه می‌گیریم. خطای استاندارد خودگردان‌سازی انحراف استانداردی است از یک برآوردگر که طی نمونه‌گیری‌های بسیار به دست آمده است. به طور شهودی، انتظار داریم که توزیع نمونه‌گیری که با نمونه‌گیری از داده‌های خودمان شکل گرفته است تقریب مناسبی برای توزیع نمونه‌گیری فراهم آورد.

راه‌های بسیاری برای تخمین رگرسیون خودگردان‌سازی وجود دارد. ساده‌ترین راه، استخراج جفت‌هایی با مقادیر $\{Y_i, X_i\}$ وجود دارد، که گاهی اوقات «جفت‌های خودگردان‌ساز»[1] یا «خودگردان‌ساز غیرپارامتری»[2] نامیده می‌شود. به عنوان یک راه جایگزین می‌توانیم مقادیر X_iها را که از توزیع پسماندهای $(\hat{e}_i)$ استخراج کرده‌ایم، ثابت فرض کنیم و یک تخمین جدیدی از متغیرهای وابسته بر پایه مقادیر پیش‌بینی شده و (مقادیر) مانده به دست آمده برای مشاهده‌گر خاص به وجود می‌آید. این روند، که یک نوع از «خودگردان‌سازی پارامتری» است، از نمونه‌ای تقلید می‌کند که با رگرسورهای غیرتصادفی استخراج شده و تضمین می‌کند که X_i و پسماندهای رگرسیون از هم مستقل هستند. از سوی دیگر، این استقلال را وقتی که به دنبال خطاهای استاندارد تحت شرایط ناهم‌واریانس هستیم، نیاز نداریم. یک خودگردان‌سازی مانده جایگزین، که «خودگردان‌سازی (مهارنشدنی) شدید» نامیده می‌شود، $X_i'\hat{\beta}+\hat{e}_i$ (که البته، همان Y_i اصلی است) با احتمال ۰/۵ استخراج می‌کند، در غیر این صورت $X_i'\hat{\beta}-\hat{e}_i$را استخراج می‌کند (نگاه کنید به مامن ۱۹۹۳ و هورویتس ۱۹۹۷). این مسئله ارتباط بین واریانس‌های مانده و X_i مشاهده شده در نمونه اصلی را حفظ می‌کند.

خودگردان‌سازی به دو دلیل مفید است. یک اینکه، در برخی موارد، توزیع مجانبی یک برآوردگر را می‌توان به سختی محاسبه کرد (یعنی، توزیع‌های مجانبی رگرسیون چارکی شامل چگالی‌های ناشناخته‌ای است). خودگردان‌سازی یک راهکار محاسباتی رایانه‌ای سنگین ایجاد می‌کند که در غیر این صورت یک راهکار محاسباتی مستقیم خواهد بود. همه توزیع‌های مجانبی هم به شیوه خودگردان‌سازی تخمین زده نمی‌شوند، اما به نظر می‌رسد که این شیوه برای برآوردگرهای ساده که ما به آن می‌پردازیم به خوبی کار می‌کند. دوم آنکه، تحت شرایطی خاص، توزیع نمونه‌گیری به دست آمده با خودگردان‌سازی می‌تواند به توزیع نمونه‌های محدود بهره نزدیک‌تر از تقریب مجانبی باشد که آمارданان این ویژگی را تظریف مجانبی می‌نامند.

در اینجا ما بیشتر به خودگردان‌سازی می‌پردازیم که به خاطر تظریف مجانبی است. محاسبه توزیع مجانبی تخمین‌های رگرسیون ساده است، اما نگرانی ما در مورد اریب‌دار بودن برآوردگرهای

1. Pairs Bootstrap
2. Nonparametric Bootstrap.

HC_0 - HC_3 است. به عنوان یک قاعده، خودگردان‌سازی وقتی که برای آمار آزمون به کار می‌رود دارای توزیع‌های مجانبی هستند و به هیچ پارامتر ناشناخته‌ای وابسته نیستند، یک تظریف مجانبی را ایجاد می‌کند (نگاه کنید به هورویتس، ۲۰۰۱). یک چنین آمارهای آزمونی را مجانب محوری می‌نامند. یک مثال آن عبارت است از آماره t که عبارت است از یک نرمال استاندارد مجانبی. ضرایب رگرسیون، مجانبی محوری نیستند، بلکه دارای یک توزیع مجانبی است که وابسته به واریانس مانده ناشناخته‌ای است.

خلاصه بحث این است که اگر شما یک استنباط بهتر با نمونه‌های محدود برای ضرایب رگرسیون می‌خواهید، می‌باید آماره t را خودگردان‌سازی نمایید و این یعنی اینکه شما آماره t را برای هر نمونه خودگردان‌ساز محاسبه کنید و آماره t مشابه را در نمونه اصلی خود با این توزیع «t» خودگردان‌ساز مقایسه کنید. در شرایطی که مقدار مطلق آماره t اصلی بالاتر باشد؛ مثلاً در مورد صدک ۹۵ام مقادیر اصلی از توزیع خودگردان‌سازی فرض ما رد می‌شود.

با وجود جذابیت نظری این مسئله برای محققان، ما به ایده خودگردان‌سازی استاتیک‌محوری نمی‌پردازیم. چراکه توجه ما نباید به آزمون‌های نظری صوری منحصر شود: ما علاقه‌مند هستیم که خطاهای استاندارد را در پرانتزهایی تحت ضرایب رگرسیون مطالعه کنیم. یک معیار دقت را که می‌توان برای ایجاد فواصل اطمینان‌بخش، مقایسه برآوردگرها و آزمون هر فرضیه‌ای که با آن در هر زمانی مواجه می‌شویم، به کار گرفت. ما به طور حتم می‌توانیم خطاهای استاندارد را از طریق خودگردان‌سازی نمونه‌ها محاسبه کنیم، اما این محاسبه هیچ تظریف مجانبی را تضمین نمی‌کند. بنابراین به نظر ما، افراد متخصص که رفتار نمونه‌های محدود خطاهای استاندارد استوار را تحت نظر دارند، می‌باید به اطلاعات اریب HC_1-HC_3 تمرکز کنند. همچنین به ایده بررسی خطای استاندارد قراردادی (با درجه آزادی اصلاح شده) و یکی از سه (اریب ذکر شده) می‌پردازیم.

مثال:

برای درک بیشتر تفاوت‌های بین برآوردگرهای کوواریانس استوار، می‌باید یک مثال ساده، اما بسیار مهم را که در فصل‌های پیش‌تر این کتاب ویژگی‌هایش بیان شده، بررسی کنیم. فرض کنیم می‌خواهید β_1را در مدل زیر تخمین بزنید:

$$Y_i = \beta_0 + \beta_1 D_i + \varepsilon_i, \tag{8.1.9}$$

که در آن D_i یک متغیر موهومی است. تخمین **OLS**، β_1 تفاوت میانگینی است بین آنچه با D_i فعال و یا D_i غیرفعال بوده است.

با علامتگذاری این زیرنمونه‌ها با اندیس‌های ۱و ۰ داریم:

$$\hat{\beta}_1 = \overline{Y}_1 - \overline{Y}_0.$$

برای این نسبت اختلاف‌ها، D_iرا یک متغیر غیرتصادفی در نظر می‌گیریم در نتیجه با فرض

$\sum \mathrm{D}_i = N_1$ و $\sum(1-\mathrm{D}_i) = N_0$ ثابت خواهند بود. **r** را برابر N_1/N بگیرید.

ما چیزهایی در مورد رفتار نمونه محدود $\widehat{\beta}_1$از طریق تئوری آماری می‌دانیم. اگر Y_iبا واریانس ناشناخته، اما معادلی در هر دو جامعه $\mathrm{D}_i = 1$و $\mathrm{D}_i = 0$نرمال باشد.

در نتیجه آماره **t** قراردادی برای $\widehat{\beta}_1$ دارای توزیع **t** است. اینها دو نمونه مرسوم آزمون **t** هستند، ناهم‌واریانسی در این مبحث به این معناست که واریانس‌ها در جامعه‌های $\mathrm{D}_i = 1$و $\mathrm{D}_i = 0$متفاوت هستند. در این مورد، مسئله آزمون در نمونه‌های کوچک حل‌نشدنی است: توزیع دقیق نمونه کوچک برای همین مسئله ساده هم ناشناخته است.[۱] برآوردگرهای کوواریانس استوار HC_0 - HC_3تقریب‌های مجانبی برای توزیع ناشناخته نمونه محدود برای واریانس‌های نابرابر ارائه می‌دهند.

تفاوت بین HC_0 - HC_3، تفاوت‌هایی در مورد چگونگی پردازش واریانس‌های نمونه است در دو گروهی که با $\mathbf{D_i}$ تعریف شده‌اند. برای **j = 0, 1**، $S_j^2 = \sum_{\mathrm{D}_i=j} (\mathrm{Y}_i - \overline{\mathrm{Y}}_j)^2$تعریف می‌شود. اهرم‌ها در این مثال عبارتند از:

$$h_{ii} = \begin{cases} 1/N_0 & \text{if } \mathrm{D}_i = 0 \\ 1/N_1 & \text{if } \mathrm{D}_i = 1 \end{cases}$$

با استفاده از این معادله، پنج برآوردگر واریانسی که در مورد آنها بحث کرده‌ایم را در اینجا به طور ساده و مستقیم نشان می‌دهیم:

$$\begin{aligned}
\mathit{Conventional} &: \quad \frac{N}{N_0 N_1}\left(\frac{S_0^2 + S_1^2}{N-2}\right) = \frac{1}{Nr(1-r)}\left(\frac{S_0^2 + S_1^2}{N-2}\right) \\
HC_0 \text{ (White, 1980)} &: \quad \frac{S_0^2}{N_0^2} + \frac{S_1^2}{N_1^2} \\
HC_1 &: \quad \frac{N}{N-2}\left(\frac{S_0^2}{N_0^2} + \frac{S_1^2}{N_1^2}\right) \\
HC_2 &: \quad \frac{S_0^2}{N_0(N_0-1)} + \frac{S_1^2}{N_1(N_1-1)} \\
HC_3 &: \quad \frac{S_0^2}{(N_0-1)^2} + \frac{S_1^2}{(N_1-1)^2}.
\end{aligned}$$

زیرنمونه‌های ادغام برآوردگر قراردادی: وقتی دو واریانس مشابه هم باشند (این کار) کفایت می‌کند.

برآوردگر وایت (۱۹۸۰**a**)، (HC_0) با استفاده از برآوردگرهای واریانس ثابت (اما اریب‌دار) $\frac{S_j^2}{N_j}$؛ تخمین‌های مجزایی از واریانس‌های نمونه‌گیری میانگین افزوده‌اند. برآوردگر HC_2 از برآوردگرهای بدون اریب واریانس نمونه برای هر گروه استفاده می‌کند، زیرا باعث اصلاح درجه آزادی درستی می‌شود. HC_1اصلاح درجاتی از آزادی را خارج از مقدار مجموع ایجاد می‌کند که کمک‌کننده است، اما کاملاً

۱. به عنوان مسئله بهرنس ـ فیشر شناخته شده است (نگاه کنید به: دی گروت و چرویش، ۲۰۰۱، فصل ۸).

دقیق نیست. از آنجایی که می‌دانیم HC_2 تخمین بدون اریب واریانس نمونه‌گیری تحت شرایط هم‌واریانسی است، HC_3 می‌باید بسیار بزرگ‌تر باشد. به یاد داشته باشیم که با **r =0.5**، موردی که در آن طرح رگرسیون، متعادل خوانده می‌شود، برآوردگر قراردادی برابر HC_1 است و همه پنج برآوردگر تنها اختلاف‌های جزئی (با همدیگر) دارند.

مطالعه مختصر مونت کارلو بر پایه (۹-۱-۸) جمع‌ها و تفریق‌های برآوردگرها و گستره‌ای را که یک حساب سرانگشتی برای بهبود اریب کلاس **HC** دربرمی‌گیرد، نشان می‌دهد. ما **N =30** را برای روشن شدن نتایج نمونه کوچک انتخاب می‌کنیم و نیز **r =0.9** را در نظر می‌گیریم که با شرط $h_{ii} = 10/N = 1/3 \text{ if } \mathrm{D}_i = 1$ را به کار می‌برد. این یک طرح بسیار نامتعادل است، نتیجه می‌گیریم که:

$$\varepsilon_i \sim \begin{cases} N(0,\sigma^2) & \text{if } \mathrm{D}_i = 0 \\ N(0,1) & \text{if } \mathrm{D}_i = 1 \end{cases}$$

و نتایج در این سه مورد به ترتیب زیر است: اولی هم‌واریانسی زیادی با **σ =0.5** دارد، در حالی که دومی به نسبت هم‌واریانسی کمتری با **σ = 0.85** دارد. حالت معیار (مبنا) هیچ هم‌واریانسی ندارد.

جدول ۱-۱-۸ نتایج را نشان می‌دهد، ستون‌های ۱ و ۲ مقدار میانگین و انحراف معیار برآوردگرهای خطای استاندارد مختلفی را در تعداد ۲۵۰۰۰ تکرار عمل نمونه‌گیری گزارش می‌دهند. انحراف معیار $\widehat{\beta_1}$ واریانس نمونه‌گیری است که ما در تلاش برای اندازه‌گیری آن بوده‌ایم. با وجود هم‌واریانسی زیاد، در بخش بالایی جدول؛ خطاهای استاندارد قراردادی به طور نامناسبی اریب‌دار هستند که به طور میانگین تنها برابر با نیمی از اندازه واریانس نمونه‌گیری مونت کارلو است که هدف ما را شکل می‌دهد. از سوی دیگر، در حالی که خطاهای استاندارد استوار کارایی بهتری جز در مورد **HC3** دارند، اما همچنان کوچک هستند.[1]

خطاهای استاندارد خود، تخمینی هستند و تفاوت‌های قابل توجهی در نمونه‌گیری با هم دارند، به ویژه همان طور که در ستون ۲ می‌توان مشاهده کرد نکته قابل توجه این واقعیت است که خطاهای استاندارد استوار گوناگونی بیشتری در نمونه‌گیری دارند تا خطاهای استاندارد **OLS**.[2] حتی وقتی که تلاش می‌کنیم تا اریب را با تقسیم پسماندها بر $1 - h_{ii}$ یا $(1 - h_{ii})^2$ کاهش دهیم؛ بیشتر هم می‌شود. بدترین حالت مربوط به HC_3 است که دارای انحراف معیاری حدود ۵۰ درصد بیش از آن چیزی

۱. توجه داشته باشید که HC2 یک تخمین‌گر بدون اریب برای واریانس نمونه‌گیری است، در حالی که میانگین خطاهای استاندارد HC2 در طول تکرار عمل نمونه‌گیری (۰/۵۲) همچنان پایین‌تر از انحراف معیار (0.59) β است، و از این حقیقت ناشی می‌شود که خطای استاندارد ریشه دوم واریانس نمونه‌گیری است، واریانس نمونه‌گیری خودش، تخمینی است و بنابراین دارای گوناگونی نمونه‌گیری است و ریشه دوم یک تابع واگرا (کاو) است.

۲. واریانس نمونه‌گیری بزرگ تخمین‌گرهای خطای استاندارد استوار توسط چشر و آوستین (۱۹۹۱) مورد توجه قرار گرفته‌اند. کوئرمان و کارول (۲۰۰۱) تعدیل بازه‌های اطمینان را برای اصلاح این مورد پیشنهاد کرده‌اند.

است که خطای استاندارد وایت (۱۹۸۰a)، HC_0 داشته است.

دو ستون آخر در این جدول نرخ‌های رد تجربی در یک آزمون ۵ درصد اسمی را برای فرض‌های $\widehat{\beta}_1 = \beta^*$ نشان می‌دهند، که در آن β^* پارامتر جامعه است (که برابر صفر است در این مورد). آمارهای آزمون با یک توزیع نرمال و یک توزیع-t با درجات آزادی $N - 2$ مقایسه می‌شوند. نرخ‌های رد برای همه آزمون‌ها بسیار بالا هستند حتی برای HC_3. استفاده از توزیع-t به نسبت استفاده از یک توزیع نرمال تنها به شکلی حاشیه‌ای کمک‌کننده است.

نتایج با هم‌واریانسی کم، در بخش دوم جدول ارائه شده‌اند و نشان می‌دهند که خطاهای استاندارد قراردادی همچنان بسیار پایین هستند و این اریب از مرتبه ۱۵ درصد خواهد بود. HC_0 و HC_1 نیز همچنان بسیار کوچک هستند چیزی در حدود همان چه که در جملات اصلی وجود داشت هرچند که اکنون دارای وابستگی بیشتری به خطاهای استاندارد قراردادی هستند. خطاهای استاندارد HC_2 و HC_3 همچنان به طور میانگین بزرگ‌تر از خطاهای استاندارد قراردادی هستند، اما نرخ‌های رد تجربی برای این دو بیشتر از نرخی است که برای خطاهای استاندارد قراردادی به دست آمده است و به این معناست که خطاهای استاندارد استوار گاهی اوقات «به طور تصادفی» بسیار کوچک هستند، تصادفی که هر از گاهی برای تورم (آماس) نرخ‌های رد رخ می‌دهد که باعث می‌شود از نرخ‌های رد قراردادی پیشی بگیرند.

درسی که ما می‌توانیم از این (موضوع) بگیریم این است که خطاهای استاندارد استوار در همه موارد جواب نمی‌دهند، و ممکن است به دو دلیل کوچک‌تر از خطاهای استاندارد قراردادی باشند: یکی برای اریب نمونه کوچکی که ما در موردش بحث کردیم و دیگری به خاطر واریانس نمونه‌گیری بالاتر این خطاهای استاندارد. بنابراین نتایج تجربی به دست می‌آوریم که در آنها خطاهای استاندارد استوار پایین‌تر از خطاهای استاندارد قراردادی، به عنوان یک علامت (پرچم) قرمز می‌افتند، که به احتمال زیاد بر مبنای اریب یا احتمال وقوع به شکل بهتری کاهش می‌یابد.

به همین خاطر، ایده در نظرگرفتن بیشترین خطای استاندارد قراردادی و یک خطای استاندارد استوار به عنوان بهترین معیار دقت؛ مورد توجه ماست. قاعده حساب سرانگشتی در دو شمارش کمک‌کننده هستند: مقادیر کم را از تخمین‌گرهای استوار جدا می‌کند (می‌برد)، اریب و گوناگونی (واریانس) را هم کاهش می‌دهد. جدول ۱-۱-۸ نرخ‌های رد تجربی که با استفاده از $Max(HC_j, Conventional)$ به دست آمده را نشان می‌دهد. نرخ‌های رد تجربی با استفاده از قاعده حساب سرانگشتی در دو بخش اول جدول به خوبی به دست آمده‌اند و به تنهایی برآوردگرهای استوار را ارتقا می‌دهند.[1]

از آنجایی که هیچ گنجی بدون رنج به دست نمی‌آید، می‌باید هزینه‌ای بابت استفاده از

۱ یانگ، هسو و ژاو (۲۰۰۵) مفهوم روال‌های آزمون را بر پایه بیشترین، یک مجموعه آماره‌های آزمون با کارایی و ویژگی‌های استواری متفاوت به صورت فرمولی درآورده‌اند.

$Max(HC_j, Conventional)$ پرداخت شود. این هزینه عبارت است از اینکه بهترین خطای استاندارد در هنگام عدم وجود هم‌واریانسی، تخمین **OLS** قراردادی باشد. این مورد در بخش انتهایی جدول ثبت شده است. استفاده از بیشترین خطای استاندارد تورم‌یافته بدون نیاز به شرایط هم‌واریانسی، نرخ‌های رد را دچار رکود می‌کند. با وجود این؛ جدول نشان می‌دهد که حتی در این شرایط هم نرخ‌های رد خیلی کاهش نمی‌یابند. همچنین نگاهی خواهیم داشت بر برآورد دست پایین دقت که هزینه‌ای کمتر از برآورد دست بالای آن دارد. برآورد دست پایین دقت، به نظر می‌رسد ناشی از نبود اطلاعات کافی در داده‌ها باشد و ما بهتر است که داده‌های بیشتری جمع‌آوری کنیم، در حالی که در مورد اخیر ما ممکن است به اشتباه نتایجی که حاوی اطلاعات درست و مهمی باشند را نادیده بگیریم.

نظر آخر در مورد بررسی مونت کارلو راجع به اندازه نمونه است. اقتصاددانان کارگری مانند ما، به کار کردن با ده‌ها هزار مشاهده و یا حتی بیشتر عادت دارند، اما گاهی اوقات هم این کار را انجام نمی‌دهیم. در مطالعه‌ای مربوط به اثرات جابه‌جایی با اتوبوس بر دانش‌آموزان مدارس دولتی، آنگریست و لانگ (۲۰۰۴) با تعداد نمونه‌ای حدود ۳۰۰۰ دانش‌آموز که از ۵۶ مدرسه هستند، کار می‌کنند. رگرسور بهره در این مطالعه تنها بین سطوح مختلف مدرسه ایجاد تفاوت می‌کند، در نتیجه برخی از تحلیل‌ها در این مقاله از میانگین ۵۶ مدرسه استفاده می‌کند. بنابراین جای تعجب نیست که آنگریست و لانگ (۲۰۰۴) در هنگامی که با داده‌های سطوح مدارس کار می‌کردند، به خطاهای استاندارد **HC_1** پایین‌تر از خطاهای استاندارد **OLS** قراردادی دست یافتند. به عنوان یک قاعده، حتی اگر شما با داده‌های بسیار کوچک بر روی افراد (به صورت تکی) کار می‌کنید، رگرسور بهره در یک سطح بالاتر یک جمع ـ یک مدرسه، دولت یا دیگر گروه‌ها یا خوشه‌ها متفاوت است ـ اندازه‌های نمونه‌ها بر تعداد خوشه‌ها بیشتر اثرگذار است تا بر تعداد افراد. روال‌های استنباط برای داده‌های خوشه‌بندی شده در بخش بعدی به صورت مفصل به بحث گذارده می‌شوند.

۸-۲. خوشه‌بندی[1] و همبستگی رشته‌ای[2] در داده‌های تابلویی

۸-۲-۱. خوشه‌بندی و عامل مولتون

جدای از مسائل مربوط به اریب؛ هم‌واریانسی به ندرت منجر به تغییرات اساسی در استنباط می‌شود. در نمونه‌های بزرگ که در آنها اریب مسئله‌ساز نیست، احتمالاً هنگام حرکت از برآوردگر قراردادی به **HC_1**، شاهد افزایش خطاهای استاندارد در حدود ۲۵ درصد خواهیم بود. در مقابل، خوشه‌بندی می‌تواند باعث تفاوت‌های کلی شود.

مسئله خوشه‌بندی را می‌توان با استفاده از رگرسیون دومتغیره ساده که در داده‌هایی با ساختار

1. Clustering
2. Serial Correlation

گروهی تخمین زده شده است، نشان داد.

فرض کنید ما رگرسیونی دومتغیره در نظر گرفته‌ایم:

$$Y_{ig} = \beta_0 + \beta_1 x_g + e_{ig}, \qquad (8.2.1)$$

که در آن $\mathbf{Y_{ig}}$ متغیر وابسته برای هر فرد $\mathbf{i}$ در خوشه یا گروه $\mathbf{g}$ با تعداد $\mathbf{G}$ گروه است. بسیار مهم است که رگرسور بهره، $\mathbf{x_g}$ تنها به واسطه سطح گروه تفاوت ایجاد می‌کند. برای مثال، داده‌های به دست آمده از آزمایش **STAR** که توسط کروجر (۱۹۹۹) تحلیل شده‌اند به صورت $\mathbf{y_{ig}}$، نمره آزمون هر دانش‌آموز $\mathbf{i}$ در کلاس $\mathbf{g}$، و اندازه خوشه $\mathbf{x_g}$؛ ظاهر می‌شوند.

اگرچه در آزمایش **STAR** دانش‌آموزان به طور تصادفی در این خوشه‌ها جا گرفته‌اند، نامحتمل است که داده‌ها طی مشاهدات، از هم مستقل باشند. نمره‌های آزمون برای دانش‌آموزان در یک گروه مشابه به یکدیگر همبسته به نظر می‌رسند، چراکه دانش‌آموزان در یک گروه مشابه دارای ویژگی‌های زمینه‌ای اشتراکی هستند و در معرض محیط کلاسی یکسان و دارای یک معلم یکسان هستند. بنابراین با احتیاط می‌توان فرض کرد که برای دانش‌آموزان $\mathbf{i}$ و $\mathbf{j}$ در یک کلاس یکسان، $\mathbf{g}$، داریم:

$$E[e_{ig}e_{jg}] = \rho\sigma_e^2 > 0, \qquad (8.2.2)$$

که در آن $\boldsymbol{\rho}$ ضریب همبستگی درون کلاسی است و σ_e^2 واریانس پسماند است.[۱]

همبستگی درون‌گروه‌ها اغلب با استفاده از یک مدل اثرات تصادفی اضافی مدلسازی می‌شود. به ویژه، ما فرض می‌گیریم که مانده، $\mathbf{e_{ig}}$ یک ساختار گروهی دارد:

$$e_{ig} = v_g + \eta_{ig}. \qquad (8.2.3)$$

که در آن $\mathbf{v_g}$ یک عنصر تصادفی مخصوص کلاس $\mathbf{g}$ است و $\boldsymbol{\eta_{ig}}$ عنصر سطح دانش‌آموز میانگین ـ صفر است که مانده محسوب می‌شود. اینجا ما به مسئله همبستگی می‌پردازیم، در نتیجه هر دو این عناصر خطا ناهم‌واریانس فرض می‌شوند.

هنگامی که رگرسور بهره تنها در سطح گروه‌ها تفاوت دارد، یک ساختار خطا مانند (۳-۲-۸) می‌تواند به سرعت افزایش پیدا کند. این حقیقت ناخوشایند مسئله تازه‌ای نیست ـ کلوئک (۱۹۸۱) و مولتون (۱۹۸۶) هر دو به این مسئله اشاره کرده‌اند ـ اما به نظر می‌رسد درست‌تر این است که بگوییم خوشه‌بندی در واقع تا پیش از ۱۵ سال پیش بخشی از اقتصادسنجی کاربردی نبوده است.

ساختار خطای داده شده، (۳-۲-۸)، ضریب همبستگی درون‌گروهی می‌شود:

$$\rho = \frac{\sigma_v^2}{\sigma_v^2 + \sigma_\eta^2}.$$

که در آن $\boldsymbol{\sigma^2_e}$ واریانس $\mathbf{v_g}$ است و $\boldsymbol{\sigma^2_\eta}$ واریانس $\boldsymbol{\eta_{ig}}$. مطلبی در مورد نامگذاری هست و آن این

---

۱. به علاوه این نوع از ساختار اصلاح مانده پیامد نمونه‌گیری طبقه‌بندی شده است (نگاه کنید به: وولدریج، ۲۰۰۳). اکثر نمونه‌هایی که با آنها کار می‌کنیم به میزان کافی تصادفی هستند و ما نوعاً به وابستگی بر مبنای ساختار گروه توجه می‌کنیم تا به خوشه‌بندی بر اساس طبقه‌بندی.

است که **ρ** ضریب همبستگی درون‌گروهی نامیده می‌شود حتی در زمانی که گروه‌های بهره، کلاس درسی نباشند.

فرض کنیم $V_c(\widehat{\beta}_1)$ فرمول واریانس **OLS** قراردادی برای اریب رگرسیون باشد (که با استفاده از **Ω_c** در بخش قبل، ایجاد شده است) و $V(\widehat{\beta}_1)$ واریانس نمونه‌گیری صحیح ارائه شده در ساختار خطا را نشان می‌دهد (۳-۲-۸). با رگرسورهای ثابت در سطح گروه و گروه‌های با اندازه یکسان، **n**، داریم:

$$\frac{V(\widehat{\beta}_1)}{V_c(\widehat{\beta}_1)} = 1 + (n-1)\rho, \qquad (8.2.4)$$

(و این) فرمولی (است) که از پیوست این فصل به دست آمده است. ریشه دوم این نسبت را به خاطر مطالعات بسزای مولتون (۱۹۸۶) عامل مولتون می‌نامیم. معادله (۴-۲-۸) به ما می‌گوید که تا چه میزان دقت دست بالا تخمین زده شده است. با نادیده گرفتن همبستگی درون‌گروهی خطاهای استاندارد قراردادی به طور فزاینده‌ای با افزایش **η**و **ρ** گمراه‌کننده‌تر می‌شوند. برای مثال، فرض کنید که **ρ = 1**، در این مورد، تمام خطاها در یک گروه، مشابه هم خواهند بود، در نتیجه **Y_{ig}**ها هم مشابه خواهند بود. بزرگ‌تر کردن یک مجموعه داده با **n** بار کپی کردن یک مجموعه داده کوچک‌تر، هیچ اطلاعات جدیدی خلق نمی‌کند. بنابراین واریانس $V_c(\widehat{\beta}_1)$ می‌باید با یک فاکتور از **n** به مقیاس بزرگ‌تری برده شود. عامل مولتون با (افزایش) اندازه گروه افزایش می‌یابد، زیرا با یک اندازه نمونه ثابت کلی، گروه‌های بزرگ‌تر مستلزم خوشه‌های کمتری هستند که در این مورد اطلاعات مستقل کمتری در نمونه وجود دارد (زیرا داده‌ها در ازای خوشه‌ها مستقل هستند و نه درون خوشه‌های آنها).[1]

حتی ضرایب اصلاح درون‌گروهی کوچک می‌توانند یک عامل مولتون بزرگ ایجاد کنند. برای مثال، در مطالعه آنگریست و لاوی (۲۰۰۷)؛ ۴۰۰۰ دانش‌آموزی که در ۴۰ مدرسه گروه‌بندی شده‌اند؛ بنابراین مقدار میانگین **n**؛ ۱۰۰ است. رگرسور بهره وضعیت نحوه عمل (تیمار) سطح مدرسه است ـ تمام دانش‌آموزان این مدارس که برای گذراندن آزمون‌های ورودی نام‌نویسی کرده‌اند، واجد دریافت جوایز نقدی بودند. همبستگی درون‌گروهی در این مطالعه حول‌وحوش ۰.۱. در نوسان هستند. با به‌کارگیری فرمول (۴-۲-۸)، عامل بولتون بیش از ۳ خواهد شد: خطاهای استاندارد گزارش شده متعارف تنها یک‌سوم چیزی است که باید باشند.

معادله (۴-۲-۸) یک مورد مهم را پوشش می‌دهد که در آن رگرسورها در بین گروه‌ها ثابت هستند و اندازه گروه هم ثابت خواهد بود. فرمول کلی به رگرسور، **x_{ig}**، اجازه می‌دهد تا در سطح فردی و برای اندازه‌های مختلف گروه‌ها، **n_g**، متفاوت باشد. در این مورد عامل مولتون ریشه دوم معادله زیر است:

۱. با رگرسورهای غیرتصادفی و مانده‌های همواریانس، عامل مولتون یک نتیجه نمونه محدود است. محققان آمار عامل مولتون را اثر طراحی می‌خوانند، زیرا این عامل به ما می‌گوید که تا چه حد خطاهای استاندارد را در نمونه‌های طبقه‌بندی شده برای انحراف‌ها از نمونه‌گیری تصادفی ساده تعدیل کنیم.

$$\frac{V(\widehat{\beta}_1)}{V_c(\widehat{\beta}_1)} = 1 + \left[\frac{V(n_g)}{\overline{n}} + \overline{n} - 1\right]\rho_x \rho, \qquad (8.2.5)$$

که در آن $\overline{n}$ میانگین اندازه گروه است و ρ_x همبستگی درون‌گروهی (کلاسی) x_{ig} است:

$$\rho_x = \frac{\sum_g \sum_{i \neq k} (x_{ig} - \overline{x})(x_{kg} - \overline{x})}{V(x_{ig}) \sum_g n_g (n_g - 1)}.$$

به یاد داشته باشید که ρ_x یک ساختار واریانس عناصر را مانند (۳-۲-۸) اعمال نمی‌کند ـ در اینجا، ρ_x یک اندازه کلی همبستگی رگرسورها در بین گروه‌هاست. فرمول کلی مولتون به ما می‌گوید که خوشه‌بندی با اندازه‌های متغیر گروه‌ها و وقتی که ρ_x بزرگ است، اثر بزرگ‌تری بر خطاهای استاندارد خواهد داشت. این اثر وقتی که $\rho_x = 0$ باشد، محو می‌شود. به بیان دیگر، اگر x_{ig} ها در بین گروه‌ها ناهمبسته هستند، ساختار خطای گروهی برای تخمین خطاهای استاندارد مسئله‌ای ایجاد نمی‌کند. به همین دلیل است که ما هنگامی که رگرسور بهره بین گروه‌ها ثابت است، بیشتر به خوشه‌بندی توجه می‌کنیم.

ما فرمول (۱-۲-۸) را با استفاده از مثال **STAR** تنسی نمایش دادیم. رگرسور نمره درصدی کودکستان‌ها در مورد اندازه گروه (کلاس) تخمینی حدود ۰/۶۲- را با خطای استاندارد (HC_1) با استواری ۰/۰۹ را به وجود می‌آورد. در این مورد $\rho_x=1$ است، زیرا اندازه گروه در بین گروه‌ها (کلاس‌ها) ثابت است در حالی که $V(n_g)$ مثبت است، زیرا کلاس‌ها از جهت اندازه با هم متفاوت هستند (در این مورد $V(n_g) = 17.1$). ضریب همبستگی درون‌کلاسی برای افراد ۰/۳۱ است و اندازه متوسط کلاس ۱۹/۴ است. درج این اعداد در (۱-۲-۸) مقداری برابر با ۷ را برای $\frac{V(\widehat{\beta}_1)}{V_c(\widehat{\beta}_1)}$ به دست می‌دهد، در نتیجه خطاهای استاندارد قراردادی می‌باید در یک عامل $\sqrt{۷}$ = ۲/۶۵ ضرب شوند. بنابراین خطای استاندارد اصلاح شده چیزی حدود ۰/۲۴ خواهد بود.

عامل مولتون به طور مشابهی با **2SLS** کار می‌کند جز اینکه ρ_x می‌باید برای متغیر ابزاری محاسبه شود و نه برای رگرسور. به ویژه، استفاده از (۵-۲-۸) به جابه‌جایی ρ_x با ρ_z منجر می‌شود، که در آن ρ_z ضریب همبستگی درون‌کلاسی متغیر ابزاری است (شور ـ شپرد، ۱۹۹۶) و ρ همبستگی درون‌کلاسی مانده مرتبه دوم است. برای درک چرایی این کارکرد، به یاد آورید که خطاهای استاندارد قراردادی برای **2SLS** از تقسیم واریانس مانده معادله مرتبه دوم بر واریانس مقادیر مناسب مرتبه اول به دست آمده‌اند. این هم یک فرمول واریانس مجانبی مشابه برای **OLS** است که مقادیر مناسب مرتبه اول نقش رگرسور را در آن ایفا می‌کنند.[۱]

در اینجا راه‌حل‌هایی برای مسئله مولتون آورده شده است:

۱. (راه‌حل) پارامتری: خطاهای استاندارد قراردادی ثابت از (۵-۲-۸) استفاده می‌کند. محاسبه

---

۱. همچنین خوشه‌بندی می‌تواند در طراحی ناپیوستگی ـ رگرسیونی مسئله‌ساز شود اگر متغیری که تخصیص نحوه عمل را مشخص می‌کند تنها در سطح گروه تفاوت ایجاد نماید (برای مطالعه دقیق‌تر نگاه کنید به: کارد و لی، ۲۰۰۳).

همبستگی‌های درون‌کلاسی $\boldsymbol{\rho}_\mathbf{x}$ و $\boldsymbol{\rho}$ بسیار ساده است و به عنوان آمارهای توصیفی در برخی بسته‌های نرم‌افزاری تولید می‌شوند.[1]

۲. خطاهای استاندارد خوشه‌ای: لیانگ و زگر (۱۹۸۶) ماتریس کوواریانس استوار (**1980a**) را برای خوشه‌بندی همانند ناهم‌واریانسی کلیت می‌بخشد:

$$\hat{V}_c(\hat{\beta}) = (X'X)^{-1}\left(\sum_g X_g \hat{\Psi}_g X_g\right)(X'X)^{-1}, \text{ where} \tag{8.2.6}$$

$$\hat{\Psi}_g = a\hat{e}_g\hat{e}_g' = a\begin{bmatrix} \hat{e}_{1g}^2 & \hat{e}_{1g}\hat{e}_{2g} & \cdots & \hat{e}_{1g}\hat{e}_{n_g g} \\ \hat{e}_{1g}\hat{e}_{2g} & \hat{e}_{2g}^2 & & \vdots \\ \vdots & & \ddots & \hat{e}_{(n_g-1)g}\hat{e}_{n_g g} \\ \hat{e}_{1g}\hat{e}_{n_g g} & \cdots & \hat{e}_{(n_g-1)g}\hat{e}_{n_g g} & \hat{e}_{n_g g}^2 \end{bmatrix}.$$

در اینجا $\mathbf{x_g}$ ماتریس رگرسورهای گروه **g** است و **a** عامل تعدیل درجه‌های آزادی است و مشابه آن چیزی که در $\mathbf{HC_1}$ بروز می‌کند. برآوردگر واریانس خوشه‌بندی شده $\hat{V}_c(\hat{\beta})$ با افزایش نه تنها مدل پارامتری، بلکه با افزایش تعداد گروه‌ها تحت هر ساختار همبستگی درون‌گروهی، ثابت می‌ماند. $\hat{V}_c(\hat{\beta})$ با وجود ثابت بودن تعداد گروه‌ها؛ ثابت نیست، حتی وقتی که اندازه گروه متمایل به بی‌نهایت باشد. برای درک چرایی این مسئله، به یاد آورید که مجموع‌ها در $\hat{V}_c(\hat{\beta})$ بیش از **g** است و نه بیش از **i**. ثبات با قاعده تعداد بزرگ مشخص می‌شود، که مبین این مسئله است که ما می‌توانیم بر گشتاورهای نمونه برای همگرایی به سوی گشتاورهای جامعه تکیه کنیم (بخش ۳-۱-۳)، اما در اینجا مجموع‌ها در سطح گروه هستند و بر افراد به صورت تکی اعمال نمی‌شوند. بنابراین نامحتمل است که خطاهای استاندارد خوشه‌بندی شده با تعداد خوشه‌های کم قابل اتکا باشد، در ادامه به این نکته باز می‌گردیم.

۳. استفاده از میانگین‌های گروه به جای داده‌های میکرو (بسیار ریز): فرض کنیم $\overline{\mathrm{Y}}_g$ میانگین Y_{ig} در گروه **g** باشد. تخمین به وسیله کمترین توان‌های دوم موزون با به‌کارگیری اندازه گروه به عنوان اوزان صورت می‌گیرد.

$$\overline{\mathrm{Y}}_g = \beta_0 + \beta_1 x_g + \overline{e}_g$$

این فرمول معادل **OLS** با استفاده از داده‌های بسیار ریز است، اما خطاهای استاندارد به دست آمده از ساختار گروه (۳-۲-۸) به طور مجانبی درست هستند.

دوباره مجانب‌ها در اینجا بر پایه تعداد گروه‌ها هستند و نه بر پایه اندازه گروه‌ها. به هر حال مهم این است که از آنجایی که میانگین‌های گروه بسیار به توزیع نرمال به دست آمده از اندازه‌های متعادل گروه نزدیک هستند، می‌توانیم انتظار داشته باشیم که ویژگی‌های نمونه‌های محدود مناسب رگرسیون

---

۱. برای مثال می‌توان از دستور Loneway نرم‌افزار استتا استفاده کرد.

با خطاهای نرمال در این مسئله اثرگذار هستند ـ بنابراین خطاهای استانداردی که از تخمین گروهی به دست آمده‌اند احتمالاً قابل اتکاتر از خطاهای استاندارد خوشه‌بندی شده در نمونه‌های با تعداد کم خوشه‌ها هستند.

تخمین داده‌های گروهی را می‌توان به مدل‌هایی با متغیرهای کمی بسیار ریز با استفاده از روال دو مرحله‌ای، تعمیم داد.

فرض کنیم معادله بهره عبارت است از:

$$\mathrm{Y}_{ig} = \beta_0 + \beta_1 x_g + \mathrm{W}'_{ig}\delta + e_{ig}, \qquad (8.2.7)$$

که در آن W'_{ig} برداری از متغیرهای کمّی است که در بین گروه‌ها متفاوت هستند. در مرحله ۱، متغیر کمّی تعدیل یافته اثرات گروه، μ_g با تخمین زدن $\mathbf{Y_{ig}}$ ساخته می‌شود:

$$\mathrm{Y}_{ig} = \mu_g + \mathrm{W}'_{ig}\delta + \eta_{ig}.$$

μ_g را که اثرات گروه می‌نامیم؛ ضرایبی بر یک مجموعه کامل از ظواهر گروه است. μ_gهای تخمین‌زده شده میانگین‌های تعدیل شده گروه برای اثر متغیرهای سطح فردی W'_{ig} هستند. به یاد داشته باشید که بنابر (۷-۲-۸) و (۳-۲-۸)،

$$\hat{\mu}_g = \beta_0 + \beta_1 x_g + \{\nu_g + (\hat{\mu}_g - \mu_g)\}. \qquad (8.2.8)$$

$\mu_g = \beta_0 + \beta_1 x_g + \upsilon_g$ (خواهد بود). بنابراین در مرحله ۲، رگرسیون اثرات گروه تخمین‌زده شده بر روی متغیرهای سطح گروه را محاسبه می‌کنیم.

برآوردگر **GLS** مؤثر برای (۸-۲-۸) با استفاده از واریانس تخمین یافته دوسویه مانده سطح گروه $\{\nu_g + (\hat{\mu}_g - \mu_g)\}$ به عنوان اوزان، کمترین توان‌های دوم موزون است.

از آن جایی که واریانس υ_g با تعداد کم گروه‌ها به خوبی تخمین زده نمی‌شود، می‌تواند مسئله‌ساز شود. بنابراین ممکن است به استفاده از وزن دوسویه واریانس اثرات گروه تخمین‌زده شده، اندازه گروه، یا به عدم استفاده از هیچ‌گونه وزنی روی آوریم.[1] در تلاشی برای تقریب بهتر توزیع نمونه محدود مرتبط، دونالد و لانگ (۲۰۰۷) پیشنهاد کردند که استنباط‌ها در روال‌های گروهی بر پایه توزیع-**t** با درجات **G-K** آزادی باشند.

به یاد داشته باشیم که رویکرد گروه‌بندی هنگامی که x_{ig} در بین گروه‌ها تغییر می‌کند، قابل اعمال نیست. همان طور که در بخش ۴ مشاهده کردیم میانگین x_{ig} به صورت $\bar{x}_g$ نسخه‌ای از **IV** است. بنابراین با تغییرات بسیار کوچک در رگرسور بهره، گروه‌بندی پارامترهایی را تخمین می‌زند که متفاوت از پارامترهای هدف در مدل مشابه (۷-۲-۸) هستند.

۱. برای مثالی در مورد دو طرح وزنی اخیر، نگاه کنید به: آنگریست و لاوی (۲۰۰۷).

۴. خودگردان‌سازی بلوکی:[1] در کل، استنباط خودگردان‌سازی. از توزیع تجربی داده‌ها به وسیله نمونه‌گیری مجدد بهره می‌برد، اما نمونه‌گیری مجدد تصادفی ساده در این مورد انجام نمی‌شود. نکته ظریف در مورد داده‌های خوشه‌بندی شده، محافظت از ساختار وابستگی در جامعه هدف است. ما این کار را با خودگردان‌سازی بلوکی انجام می‌دهیم که عبارت است از استخراج بلوک‌های داده‌هایی که توسط گروه‌های **g** تعریف شده‌اند. برای مثال، در داده‌های **STAR** تنسی ما خودگردان‌سازی بلوکی را با استفاده از نمونه‌گیری مجدد کلی کلاس‌ها جدای از فرد فرد دانش‌آموزان انجام می‌دهیم.

۵. تخمین زدن یک **GLS** یا ماکسیمم پارامتری شبیه مدلی است که بر مبنای نسخه‌ای از (۸-۲-۱) می‌باشد. این کار مسئله خوشه‌بندی را حل می‌کند، اما علاوه بر آن مورد برآورد (آزمون) را هم تغییر می‌دهد، مگر اینکه **CEF** خطی باشد، جزئیات این مسئله در بخش ۱-۴-۳ آورده شده است، بنابراین ما رویکردهای دیگر را ترجیح می‌دهیم.

جدول ۱-۲-۸ اصلاحات خطای استاندارد را با هم در نمونه **STAR** مقایسه می‌کند. این جدول ۶ تخمین از خطاهای استاندارد را ارائه می‌دهد: خطاهای استاندارد استوار قراردادی (با استفاده از $\mathbf{HC_1}$)؛ دو نسخه از خطاهای استاندارد اصلاح شده پارامتری از فرمول مولتون (۸-۲-۵) استفاده می‌کنند، اولین کاربرد این فرمول برای همبستگی درون‌کلاسی است که با فرمول مولتون ارائه می‌شود و کاربرد دوم آن استفاده برآوردگر **Stata** از دستور **Loneway**؛ خطاهای استاندارد خوشه‌بندی شده؛ خطاهای استاندارد خودگردان‌سازی شده بلوکی؛ و خطاهای استاندارد ناشی از تخمین موزون در سطح گروه است. ضریب تخمین ۰/۶۲- است. در این مورد، تمام تغییرات (تعدیل‌ها) نتایج یکسانی را به بار می‌آورند، یعنی خطای استاندارد حدود ۰/۲۳. این خروجی مناسب در بخش عمده‌ای بر مبنای حقیقتی است که با ۳۱۸ کلاس درسی، ما خوشه‌هایی کافی برای مجانب بودن مناسب سطح گروه در اختیار داریم. هرچند با خوشه‌های کمتر، (نتایج) پیش‌بینی‌ناپذیرتر هستند و این نکته‌ای است که در انتهای فصل به آن می‌پردازیم.

۸-۲-۲. همبستگی پیاپی در داده‌های تابلویی و در مدل‌های تفاضلِ تفاضل[2]

همبستگی پیاپی ـ تمایل به اینکه یک مشاهده به آنچه قبلاً مشاهده شده بستگی داشته باشد ـ به نظر می‌رسد که مسئله دیگری است؛ به ویژه، در مورد ارواح نگون‌بختی که زندگی‌شان را ورای داده‌های خط سیر زمان می‌گذرانند (مثل متخصصان اقتصاد کلان)؛ بنابراین متخصصان اقتصاد خُرد کاربردی زمانی طولانی است که از این کار صرفنظر کرده‌اند.[3] اما داده‌های ما معمولاً یک بُعد زمانی هم دارند. به خصوص

1. Block Bootstrap
2. Serial Correlation in Panels and Difference-in-difference Models
۳. حوزه «مسئله فرد دیگر» (The Somebody Else.s Problem - SEP)؛ در ابتدا به عنوان یک پدیده طبیعی در اثر آدام، با عنوان جهان و همه چیز تبیین شد، طبق آنچه در دانشنامه (ویکی‌پدیا) آمده است، «یك حوزه انرژی تولید شده که بر ادراک

در مدل‌های تفاضل در تفاضل. این حقیقت با خوشه‌بندی‌ای ترکیب شده که می‌تواند یک اثر عمده بر استنباط آماری بگذارد.

فرض کنید، همان طور که در بخش ۲-۵ نشان داده شد به اثرات یک دستمزد حداقل دولتی بپردازیم. در این بافت، نسخه رگرسیون تفاوت در اختلافات اثر اضافی زمان و دولت را هم دربرمی‌گیرد. بنابراین معادله‌ای مشابه (۳-۲-۵) خواهیم داشت که در پایین تکرار می‌شود:

$$Y_{ist} = \gamma_s + \lambda_t + \beta D_{st} + \varepsilon_{ist}, \qquad (8.2.9)$$

همان طور که قبلاً گفته شد، $\mathbf{Y_{ist}}$ خروجی برای هر فرد $\mathbf{i}$ در دولت $\mathbf{s}$ در سال $\mathbf{t}$ است و $\mathbf{D_{st}}$ متغیر ظاهری است که حالت‌های تیمار را در دوره‌های پس از تیمار هم نشان می‌دهد. اصطلاح خطا در (۸-۲-۹) تنوع ویژه را در خروجی‌های بالقوه‌ای نمایان می‌کند که از فردی به فرد دیگر از دولتی به دولت دیگر و از زمانی به زمان دیگر متفاوت است. برخی از این تنوع‌ها، برای مثال، یک چرخه اقتصاد منطقه‌ای، احتمالاً برای افراد در یک دولت و سال یکسان مشابه است. ما می‌توانیم این عنصر مشترک را با اندیشه در مورد $\boldsymbol{\varepsilon_{ist}}$ به عنوان مجموع یک تکان (شوک) دولت ـ سال، $\boldsymbol{\upsilon_{st}}$ و یک عنصر فردی ویژه $\boldsymbol{\eta_{ist}}$ مدلسازی کنیم.

در نتیجه داریم:

$$Y_{ist} = \gamma_s + \lambda_t + \beta D_{st} + \upsilon_{st} + \eta_{ist}. \qquad (8.2.10)$$

ما فرض می‌گیریم که در استخراج‌های تکرار شده طی دولت‌ها و در طول زمان $\mathbf{E\,[\upsilon_{st}] = 0}$ است در حالی که طبق تعریف $\mathbf{E\,[\eta_{ist}] = 0}$ است.

شوک‌های دولت ـ سال؛ خبرهای ناخوشایندی برای مدل‌های تفاوت‌ها در اختلافات است. حتی با مسئله مولتون، اثرات تصادفی دولت ـ زمان ویژه یک مسئله خوشه‌بندی به وجود می‌آورند که بر استنباط آماری اثر می‌گذارد، اما این حداقل مشکلات ما در این مورد است. برای درک چرایی آن، فرض کنید همان طور که در مطالعه کارد و کروئگر (۱۹۹۴) نیوجرسی/ پنسیلوانیا آمده است، دو دوره زمانی و دو دولت داریم. تفاوت‌ها در اختلاف تجربی عبارت است از:

$$\hat{\beta}_{CK} = (\overline{Y}_{s=NJ,t=Nov} - \overline{Y}_{s=NJ,t=Feb}) - (\overline{Y}_{s=PA,t=Nov} - \overline{Y}_{s=PA,t=Feb}).$$

این برآوردگر به دلیل این معادله $\mathbf{E\,[\upsilon_{st}] = E\,[\eta_{ist}] = 0}$، بدون اریب است. از سوی دیگر، فرض کنید محدودیت‌های احتمال را به عنوان اندازه قابل افزایش گروه در نظر بگیریم در حالی که موارد دولت و دوره زمانی را ثابت در نظر می‌گیریم، شوک‌های دولت ـ سال $\boldsymbol{\beta_{ck}}^{\wedge}$ را به صورت غیرثابت ارائه می‌دهند.

$$plim\hat{\beta}_{CK} = \beta + \{(\upsilon_{s=NJ,t=Nov} - \upsilon_{s=NJ,t=Feb}) - (\upsilon_{s=PA,t=Nov} - \upsilon_{s=PA,t=Feb})\}.$$

میانگین نمونه‌های بزرگ و بزرگ‌تر در نیوجرسی یا پنسیلوانیا در یک دوره زمانی مشخص هیچ

اثر می‌گذارد... موجودات در این وادی توسط یک مشاهده‌گر بیرونی درک می‌شوند و این یعنی همان مسئله فرد دیگر؛ و بنابراین تا حد زیاد و تاثیرگذاری نامحسوس است مگر آنکه مشاهده‌گر به وضوح به دنبال آن موجودیتی باشد که درصدد مشاهده آن است».

تأثیری بر حذف شوک‌های منطقه‌ای مربوط به یک موقعیت مکانی یا زمانی خاص ندارد. تنها با وجود دو دوره زمانی و دو دولت، ما هیچ راهی برای تشخیص تفاوت‌های موجود در اختلافات که با تغییر سیاست‌ها به وجود آمده‌اند از تفاوت‌های موجود در اختلافاتی نداریم که بنا بر واقعیتی به وجود آمده‌اند که بیانگر ثبات اقتصادی نیوجرسی در ۱۹۹۲ بوده است در حالی که پنسیلوانیا در آن زمان یک چرخه نزولی خفیفی را در اقتصادش تجربه می‌کرده است. می‌توانیم حضور **υ_{st}** را به عنوان نقض روال عادی فرضیاتی در نظر بگیریم که در بخش ۲-۵ به بحث گذاشته شد.

راه‌حل عدم ثباتی که به خاطر شوک‌های تصادفی در مدل‌های تفاوت‌ها در اختلافات القا شده این است که یا دوره‌های زمانی چندگانه داشته باشیم یا تعداد دولت‌های (ایالت‌های) بیشتری را در نظر بگیریم (و یا هر دو این عامل‌ها را با هم در نظر بگیریم). برای مثال، کارد (۱۹۹۲) برای مطالعه تغییرات حداقل دستمزد دولتی (ایالتی) ۵۱ ایالت را در نظر گرفته در حالی که کارد و کروئگر (۲۰۰۰) از زاویه دیگری تجربه نیوجرسی/پنسیلوانیا را با فرض توالی‌های زمانی گسترده‌تری برای بررسی داده‌های دستمزد ماهیانه در نظر گرفتند. می‌توانیم امیدوار باشیم که میانگین **υ_{st}** به سمت صفر است. همان طور که در بخش اول این فصل در مورد مسئله مولتون گفته شد، چارچوب استنباط در این بافت متکی به تئوری توزیع مجانبی با گروه‌های بسیار است و نه بر اندازه گروه (یا حداقل تنها بر اندازه گروه متکی نیست). در نتیجه مهم‌ترین مسئله‌ای که استنباط می‌شود رفتار **υ_{st}**. به ویژه، اگر ما فرض کنیم که شوک‌ها در ایالت‌ها و زمان‌ها از هم مستقل هستند، -یعنی، آنها به طور متوالی ناهمبسته هستند ـ دوباره به سر مسئله **plain-vanilla** مولتون برمی‌گردیم که در بخش ۱-۲-۸ به بحث گذاشته شد و آنجا دیدیم که خوشه‌بندی را بر اساس ضرب دولت (ایالت) در سال انجام می‌دهیم، اما در اکثر موارد دیگر، فرض اینکه **υ_{st}** به طور متوالی ناهمبسته است به سختی قابل دفاع است. تقریباً با اطمینان می‌توان گفت، برای مثال شرکت‌های منطقه‌ای به طور پیاپی بسیار به هم بسته هستند: اگر شرایط برای یک ماه در پنسیلوانیا نامناسب است، احتمالاً به همین میزان ناخوشایندی در ماه بعد هم ادامه می‌یابد.

توالی همبستگی پیاپی داده‌های تابلویی خوشه‌بندی شده توسط برتراند، دوفلو و مالینتان (۲۰۰۴) و کزدی (۲۰۰۴) مورد توجه قرار گرفته است. هر طرح تحقیقی با یک ساختار گروهی که در آن میانگین گروه‌ها همبسته هستند را می‌توان مسئله همبستگی پیاپی در نظر گرفت. خلاصه تحقیقات اخیر بر همبستگی متوالی در داده‌ها با یک ساختار گروهی این است که تنها هنگامی که ما می‌باید خطاهای استاندارد را برای همبستگی بین گروه‌های ایجاد شده به خاطر حضور **υ_{st}**؛ تعدیل نماییم، می‌باید تعدیل‌های بیشتری را برای همبستگی پیاپی در خود **υ_{st}**ها هم انجام دهیم. راه‌های متعددی برای این کار وجود دارد که البته همگی به یک اندازه در همه شرایط کارایی ندارند. به نظر منصفانه است که بگوییم پرسش در مورد اینکه بهترین رویکرد تعامل با مسئله همبستگی پیاپی چه می‌تواند باشد همچنان

تحت مطالعه و بررسی است و هنوز اجماعی در مورد این مسئله حاصل نشده است. ما در اینجا تلاش می‌کنیم تا هریک از این رویکردها را به اختصار معرفی کنیم و نتایج به دست آمده را به صورت خلاصه ارائه دهیم.

ساده‌ترین و پُرکاربردترین رویکرد این است که خوشه‌بندی را یک پله بالاتر ببریم. بنابراین در مورد مثال دولت (ایالت) ـ سال، می‌توانیم خوشه‌بندی خطاهای استاندارد لاینگ و زگر (۱۹۸۶) را بر مبنای دولت (ایالت) به جای ایالت و سال بپذیریم (با استفاده از خوشه‌بندی **Stata**). شاید در وهله اول عجیب به نظر برسد چراکه مدل کنترل‌کننده‌ای برای اثرات عامل دولت (ایالت) است. اثر دولت، γ_s در فرمول (۱۰-۲-۸) میانگین زمان υ_{st} را که با $\overline{v}_s$ نشان می‌دهیم، حذف می‌کند. با وجود این $v_{st} - \overline{v}_s$ احتمالاً هنوز به صورت پیاپی همبسته هستند. خوشه‌بندی در سطح دولت از این مسئله بهره می‌برد، چراکه برآوردگر کوواریانس خوشه‌بندی در یک سطح بالاتر برای همبستگی مانده غیرپارامتری بین خوشه‌هاست که شامل همبستگی سری‌های زمانی در $v_{st} - \overline{v}_s$ است. این یک اصلاح ساده و سریع است. مسئله‌ای که در اینجا مطرح است، همان طور که ممکن است حدس زده باشید، این است که گذر به سمت سطوح بالاتر تعداد خوشه‌ها را کاهش می‌دهد و استنباط مجانبی این فرض را فراهم می‌کند که تعداد زیادی خوشه داریم، چراکه به طور منطقی تعداد زیادی دولت و بازه زمانی برای تخمین همبستگی بین $v_{st} - \overline{v}_s$ و $v_{st-1} - \overline{v}_s$ احتیاج داریم. تعداد کم خوشه‌ها به معنای خطاهای استاندارد اریب‌دار و استنباط‌های گمراه‌کننده است.

۳-۲-۸. کمتر از ۴۲ خوشه

اریب ناشی از خوشه‌های کم؛ هم در مسئله مولتون و در بافت همبستگی پیاپی محتمل خطاست، زیرا هر دو این موارد استنباط به خوشه وابسته است. با تعداد کم خوشه‌ها، ما به تخمین پایین‌دستی هم در همبستگی پیاپی در شوک تصادفی مانند υ_{st} می‌رسیم و هم به همبستگی درون‌کلاسی، ρ در مسئله مولتون. بُعد مرتبط برای شمارش خوشه‌ها در مسئله مولتون تعداد گروه‌ها، G است. در یک سناریوی تفاوت‌ها در اختلافات که در آن خوشه‌بندی بر حسب دولت (یا بر اساس دیگر ابعاد مقطعی) انجام می‌شود، بعد مرتبط برای شمارش خوشه‌ها، تعداد دولت‌ها یا گروه‌های مقطعی است. بنابراین، حکم داگلاس آدامز که جواب نهایی به زندگی، جهان و هر چیزی را ۴۲ می‌داند، سؤال قابل طرح به نظر ما این خواهد بود که: چه تعداد خوشه برای استنباط قابل اتکا با استفاده از تعدیل‌های خوشه استاندارد به دست آمده از (۶-۲-۸) کفایت می‌کند؟

اگر ۴۲ برای قابل اتکا بودن تعدیل خوشه استاندارد کفایت کند ـ و تعداد کمتر مناسب این کار نباشد ـ بنابراین وقتی تعداد خوشه‌ها کم باشد چه باید کرد؟ اولین و بهترین راه‌حل این است که با جمع‌آوری داده‌های بیشتر خوشه‌های بیشتری را فراهم نماییم، اما گاهی اوقات توان این کار وجود ندارد،

بنابراین ایده‌های دیگری برای انجام این کار در ادامه بیان می‌شود. شایسته است که در ابتدا به یاد داشته باشیم که همه این ایده‌ها به یک اندازه برای مسائل مولتون و همبستگی پیاپی مناسب نیستند.

۱. اصلاح اریب خطاهای استاندارد خوشه‌بندی شده. خطاهای استاندارد خوشه‌بندی شده در نمونه‌های کوچک اریب‌دار هستند، زیرا همان طور که در بخش ۱-۸ بیان شد $E\left(\widehat{e}_g\widehat{e}'_g\right) \neq E\left(e_g e'_g\right) = \Psi_g$. معمولاً $E\left(\widehat{e}_g\widehat{e}'_g\right)$ بسیار کوچک است، یک راه‌حل ایجاد تورم در پسماندها به امید کاهش اریب است. بل و مک‌کافری (۲۰۰۲) روالی را (خطی‌سازی کاهش اریب یا **BRL** می‌نامند) پیشنهاد می‌کنند که پسماندها را با فرمول‌های زیر تعدیل می‌کند:

$$\hat{\Psi}_g = a\widetilde{e}_g\widetilde{e}'_g$$

$$\widetilde{e}_g = A\widehat{e}_g$$

که در آن **A** فرمول‌های زیر را حل می‌کند.

$$A'_g A_g = (I - H_g)^{-1}$$

$$H_g = X_g(X'X)^{-1}X'_g.$$

این یک نسخه از **HC2** برای حالت خوشه‌بندی است. **BRL** برای درستی مسئله مولتون با تعداد کم خوشه‌ها می‌باشد، اما به دلایل فنی برای مسئله همبستگی پیاپی تفاوت‌ها در اختلافات معمولی به کار نمی‌رود.[1]

۲. برای تشخیص این مسئله که واحد اصلی مشاهده یک خوشه است و نه یک واحد فردی بین گروهی (خوشه‌ای)، بل و مک‌کافری (۲۰۰۲) و دونالد و لانگ (۲۰۰۷) پیشنهاد نمودند که استنباط بر پایه یک توزیع-**t** با درجات آزادی **G-K** باشد تا اینکه تنها بر پایه توزیع نرمال استاندارد باشد. برای **G** کوچک یک تفاوت اساسی به وجود می‌آید ـ به منظور اجتناب از برخی اشتباهات بازه‌های اطمینان بسیار بزرگ‌تر می‌شوند. کامرون، گلباخ و میلر (۲۰۰۸) مثال‌های مونت‌کارلو را گزارش دادند که در آنها ترکیب تعدیل **BRL** و استفاده از جدول‌های **t** به خوبی جواب می‌دهند.

۳. بحث دونالد و لانگ (۲۰۰۷) این است که تخمین با استفاده از میانگین گروه برای **G** کوچک در مسئله مولتون به خوبی کار می‌کند و حتی وقتی که استنباط بر اساس توزیع **t** با درجات آزادی **G-**

---

۱. ماتریس A_g یکتا نیست، بلکه تعداد زیادی از این تجزیه‌ها وجود دارد. بل و مک‌کافری (۲۰۰۲) از ریشه دوم متقارن $(I - H_g)^{-1}$ یا $A_g = P\Lambda^{1/2}$ استفاده کرده‌اند که در آن P ماتریس ویژه بردار $(I - H_g)^{-1}$ است، Λ ماتریس قطری ویژه مقدار معادل است و $\Lambda^{1/2}$ ماتریس قطری ریشه دوم ویژه مقدار است. مشکلی که با تعدیل بل و مک‌کافری وجود دارد این است که $(I - H_g)$ ممکن است از رتبه کامل نباشد و بنابراین معکوس آن شاید برای همه طراحی‌ها وجود نداشته باشد. برای مثال این اتفاق زمانی رخ می‌دهد که یکی از رگرسورها یک متغیر ظاهری باشد که دقیقاً برای یکی از خوشه‌ها مقدار یک و برای بقیه خوشه‌ها مقدار صفر می‌گیرد، این مسئله شامل مدل DD پانل که توسط برتراند و همکارانش (۲۰۰٤) بررسی شده، نیز می‌شود، که در آن شما یک مجموعه کامل از متغیرهای ظاهری و خوشه‌ها برای دولت خواهید داشت، به علاوه تجزیه ویژه مقدار برای ماتریس‌هایی که به اندازه گروه هستند به کار گرفته می‌شود. در بسیاری از این کاربردها، اندازه‌های گروه به اندازه کافی بزرگ هستند که غیرقابل محاسبه می‌شوند.

K است بهتر هم نتیجه می‌دهد، اما همان طور که پیش از این بحث شد، رگرسیون باید در بین گروه‌ها ثابت باشد. سطح تجمع، سطحی است که خوشه‌بندی را در آن انجام می‌دهیم، مانند (واحد) مدرسه در مطالعه آنگریست و لاوی (۲۰۰۷)، برای همبستگی پیاپی، این سطح تجمع، دولت (ایالت) است، اما میانگین‌های دولت را نمی‌توان برای تخمین یک مدل با مجموعه کامل از اثرات دولت به کار برد. همچنین، از آنجایی که وضعیت تیمار (نحوه عمل) از دولتی به دولت دیگر متفاوت است، ارتقای میانگین به سطح میانگین‌های دولتی موجب می‌شود که رگرسور بهره هم قواعد بازی را به گونه‌ای تغییر دهد که خوشایند ما نباشد (برآوردگر به واسطه تورم‌های گروه به عنوان ابزار به متغیرهای ابزاری تبدیل می‌شود). بنابراین رویکرد مسئله همبستگی پیاپی؛ میانگین گروه، خارج از مرزها خواهد بود.[1] همچنین به یاد داشته باشید که اگر پسماندهای گروه‌بندی شده ناهم‌واریانس هستند، در نتیجه شما باید از خطاهای استاندارد استوار استفاده کنید و می‌باید حواستان به اریب فرم (ساختار) باشد، که در بخش ۱-۸ به بحث گذاشته شد. اگر هم اثر تصادفی و هم مانده بسیار ریز پس‌زمینه، هم‌واریانس باشند، می‌توانید ناهم‌واریانسی را در میانگین گروه با استفاده از وزن (توزین) مربوط به اندازه گروه اصلاح نمایید، اما وقتی که CEF غیرخطی توزین، مورد براورد را تغییر می‌دهد ـ در نتیجه اصلاً یک مسئله ساده پیش‌رو نداریم (آنگریست و لاوی، ۱۹۹؛ عدم توزین میانگین‌های سطح مدرسه را انتخاب کردند، زیرا تنوع مطالعات آنها عمدتاً ناشی از مدارس کوچک بود). توزین یا عدم توزین، امن‌ترین دوره هنگام کار کردن با میانگین‌های سطح گروه استفاده از قاعده حساب سرانگشتی است که در بخش ۱-۸ به آن پرداختیم. بیشترین خطاهای استاندارد قراردادی یا استوار را به عنوان بهترین مقیاس (معیار) دقت انتخاب نمایید.

۴. کامرون، گلباخ و میلر (۲۰۰۸) این طور گزارش دادند که برخی اشکال خودگردان‌سازی بلوکی با تعداد کمتر گروه‌ها بهتر کار می‌کنند و اینکه خودگردان‌سازی بلوکی معمولاً عملکرد بهتری در خطاهای استاندارد خوشه‌بندی شده استات را بدون اصلاح اریب داراست. به نظر می‌رسد که برای هر دو مسئله مولتون و همبستگی پیاپی درست باشد، اما کامرون، گلباخ و میلر (۲۰۰۸) بر استفاده از میزان رد آمارهای آزمون (محوری) تمرکز می‌کنند در حالی که ما به بررسی خطاهای استاندارد می‌پردازیم.

۵. اصلاحات پارامتری: برای مسئله مولتون، این مسئله منجر به استفاده از عامل مولتون می‌شود. با وجود همبستگی پیاپی، خطاهای استاندارد برای همبستگی پیاپی مرتبه اول در سطح گروه اصلاح می‌شوند. بر اساس تجربه نمونه‌گیری و با توجه به مسئله مولتون و یکی دیگر از منابع مطالعاتی، رویکردهای پارامتری ممکن است به خوبی پاسخگو باشند و بهتر از تخمین‌گر غیرپارامتری عمل کند (۶-۲-۸)، به ویژه اگر مدل پارامتری خیلی دور (پرت) نباشد (نگاه کنید به هانسن، ۲۰۰۷، که همچنین یک اصلاح اریب را برای تخمین پارامترهای همبستگی پیاپی پیشنهاد می‌کند). متأسفانه، اگرچه خارج

---

۱. دونالد و لانگ (۲۰۰۷) نمونه‌های همبستگی پیاپی را بررسی می‌کنند که در آن رگرسور بین ابعاد خوشه‌بندی ثابت است، اما یک حالت (وضعیت) تفاوت‌ها در اختلافات محسوب نمی‌شود.

از جهان گلخانه‌ای (تحت کنترل و حفاظت) مطالعات مونت‌کارلو، نامحتمل است که بدانیم آیا فرضیات پارامتری مناسب هستند یا خیر.

متأسفانه، سیر اصلی در اینجا کاملاً مشخص نیست، به همین خاطر مسئله اصلی هنگامی است که تعداد کم خوشه‌ها برای استنباط خطای مهلک محسوب می‌شوند. شدت اریب به دست آمده به نظر وابسته به طبیعت مسئله مورد بررسی است، به خصوص وابسته است به اینکه آیا در آن با مسائل مولتون مواجه می‌شوید و یا با مسائل همبستگی پیاپی. تجمع در سطح گروه طبق تحقیق دونالد و لانگ (۲۰۰۷) به نظر در مورد مسئله مولتون تا جایی که رگرسور بهره در گروه‌ها ثابت باشند و ناهم‌واریانسی زمینه‌ای زیادی وجود نداشته باشد، به خوبی کار می‌کند. حداقل از آنجایی که این رویکرد شفاف و محافظه‌کارانه است، شاید بخواهید این گونه نشان دهید که نتایج به دست آمده با استفاده از استنباط‌های ناشی از تحلیل میانگین‌های گروه ثابت هستند. آنگریست و لاوی (۲۰۰۷) برای تعدیل خوشه‌بندی در سطح مدرسه از خطاهای استاندارد **BRL** استفاده می‌کنند، اما اعتبار آنها را با بیان اینکه نتایج اصلی منجر به استفاده مشابهی از میانگین‌های گروه تعدیل شده دومتغیره می‌شوند، نشان می‌دهند.

تا آنجا که همبستگی پیاپی پیش می‌رود، بیشتر شواهد مبنی بر این است که اگر شانس بیاورید و تحقیقاتتان در مورد ایالت‌های آمریکا باشد، ۵۱ خوشه خواهید داشت و شما در محیطی اصلی (واقعی) قرار دارید. می‌توانید کاربرد ساده و اولیه‌ای از دستور خوشه‌بندی **Stata** در سطح ایالت داشته باشید، و یا شاید بخواهید به مطالعه کانادا بپردازید که در این صورت ۱۰ خوشه به شکل استان خواهید داشت و در این صورت این تعداد کمتر از ۴۲ خواهد بود. هانسن (**2007 b**) دریافت که خطاهای استاندارد [خوشه‌بندی **Stata**] لیانگ و زگر (۱۹۸۶) منطقاً برای تصحیح همبستگی پیاپی در داده‌های تابلویی حتی در مورد مسئله‌ای با موضوعیت کانادا مناسب هستند. همچنین هانسن استفاده از توزیع **t** با درجات آزادی **G-K** را برای مقادیر بحرانی پیشنهاد می‌کند.

مسئله خوشه‌بندی متخصصان اقتصاد خُرد کاربردی را واداشته تا به اشتباه خود اعتراف کنند. ما با افتخار به خاطر کار بر روی مجموعه‌های بسیار بزرگ داده‌های بسیار ریز، می‌خواهیم کمی متخصصان اقتصاد کلان را با نمونه‌های کوچک توالی زمانی به چالش بکشیم، اما جوجه را آخر پاییز می‌شمارند: اگر رگرسور بهره تنها در یک سطح گروه دوره‌ای تغییر کند ـ مانند گذر زمان یا در گذر دولت‌ها و یا از کشوری به کشور دیگر ـ این متخصصان اقتصاد کلان هستند که وضعیت واقعی استنباط را در اختیار دارند.

۳-۸. پیوست: استخراج عامل ساده مولتون

می‌نویسیم:

$$y_g = \begin{bmatrix} y_{1g} \\ y_{2g} \\ \vdots \\ y_{n_g g} \end{bmatrix} \quad e_g = \begin{bmatrix} e_{1g} \\ e_{2g} \\ \vdots \\ e_{n_g g} \end{bmatrix}$$

و

$$y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_G \end{bmatrix} \quad x = \begin{bmatrix} \iota_1 x_1 \\ \iota_2 x_2 \\ \vdots \\ \iota_G x_G \end{bmatrix} \quad e = \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_G \end{bmatrix}$$

که در آن $\mathbf{\iota_g}$ بردار ستونی $\mathbf{n_g}$ هاست و $\mathbf{G}$ تعداد گروه‌هاست. به یاد داشته باشید که:

$$\begin{aligned} E(ee') &= \Psi = \begin{bmatrix} \Psi_1 & 0 & \cdots & 0 \\ 0 & \Psi_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \Psi_G \end{bmatrix} \\ \Psi_g &= \sigma_e^2 \begin{bmatrix} 1 & \rho & \cdots & \rho \\ \rho & 1 & & \vdots \\ \vdots & & \ddots & \rho \\ \rho & \cdots & \rho & 1 \end{bmatrix} = \sigma_e^2 \left[(1-\rho) I + \rho \iota_g \iota_g' \right], \end{aligned}$$

در آن

$$\rho = \frac{\sigma_v^2}{\sigma_v^2 + \sigma_\eta^2}.$$

اکنون

$$\begin{aligned} X'X &= \sum_g n_g x_g x_g' \\ X'\Psi X &= \sum_g x_g \iota_g' \Psi_g \iota_g x_g'. \end{aligned}$$

اما

$$
\begin{aligned}
x_g \iota_g' \Psi_g \iota_g x_g' &= \sigma_\epsilon^2 x_g \iota_g' \begin{bmatrix} 1+(n_g-1)\rho \\ 1+(n_g-1)\rho \\ \cdots \\ 1+(n_g-1)\rho \end{bmatrix} x_g' \\
&= \sigma_\epsilon^2 n_g \left[1+(n_g-1)\rho\right] x_g x_g'.
\end{aligned}
$$

فرض کنیم $\tau_g = 1+(n_g-1)\rho$، در نتیجه خواهیم داشت:

$$
\begin{aligned}
x_g \iota_g' \Psi_g \iota_g x_g' &= \sigma_\epsilon^2 n_g \tau_g x_g x_g' \\
X' \Psi X &= \sigma_\epsilon^2 \sum_g n_g \tau_g x_g x_g'.
\end{aligned}
$$

با در دست داشتن این فرمول می‌توانیم بنویسیم:

$$
\begin{aligned}
V(\widehat{\beta}) &= (X'X)^{-1} X'\Psi X (X'X)^{-1} \\
&= \sigma_\epsilon^2 \left(\sum_g n_g x_g x_g'\right)^{-1} \sum_g n_g \tau_g x_g x_g' \left(\sum_g n_g x_g x_g'\right)^{-1}.
\end{aligned}
$$

ما می‌خواهیم این فرمول را با برآوردگر کوواریانس **OLS** استاندارد مقایسه کنیم:

$$
V_c(\widehat{\beta}) = \sigma_\epsilon^2 \left(\sum_g n_g x_g x_g'\right)^{-1}.
$$

اگر اندازه‌های گروه با هم برابر باشند، $\mathbf{n_g = n}$ و $\boldsymbol{\rho}$ $\mathbf{\tau_g = \tau = 1 + (n-1)\,\rho}$، بنابراین،

$$
\begin{aligned}
V(\widehat{\beta}) &= \sigma_\epsilon^2 \tau \left(\sum_g n x_g x_g'\right)^{-1} \sum_g n x_g x_g' \left(\sum_g n x_g x_g'\right)^{-1} \\
&= \sigma_\epsilon^2 \tau \left(\sum_g n x_g x_g'\right)^{-1} \\
&= \tau V_c(\widehat{\beta}),
\end{aligned}
$$

که فرمول (۴-۲-۸) را به کار می‌گیرد.

جدول ۱-۱-۸. نتایج مونت کارلو برای خطاهای استاندارد استوار

	Mean	Standard Deviation	Empirical 5% Rejection Rates: Normal	Empirical 5% Rejection Rates: t
	(1)	(2)	(3)	(4)
A. Lots of Heteroskedasticity				
$\hat{\beta}_1$	-0.001	0.586		
Standard Errors:				
Conventional	0.331	0.052	0.278	0.257
HC0	0.417	0.203	0.247	0.231
HC1	0.447	0.218	0.223	0.208
HC2	0.523	0.26	0.177	0.164
HC3	0.636	0.321	0.13	0.12
max(Conventional, HC0)	0.448	0.172	0.188	0.171
max(Conventional, HC1)	0.473	0.19	0.173	0.157
max(Conventional, HC2)	0.542	0.238	0.141	0.128
max(Conventional, HC3)	0.649	0.305	0.107	0.097
B. Little Heteroskedasticity				
$\hat{\beta}_1$	0.004	0.6		
Standard Errors:				
Conventional	0.52	0.07	0.098	0.084
HC0	0.441	0.193	0.217	0.202
HC1	0.473	0.207	0.194	0.179
HC2	0.546	0.25	0.156	0.143
HC3	0.657	0.312	0.114	0.104
max(Conventional, HC0)	0.562	0.121	0.083	0.07
max(Conventional, HC1)	0.578	0.138	0.078	0.067
max(Conventional, HC2)	0.627	0.186	0.067	0.057
max(Conventional, HC3)	0.713	0.259	0.053	0.045
C. No Heteroskedasticity				
$\hat{\beta}_1$	-0.003	0.611		
Standard Errors:				
Conventional	0.604	0.081	0.061	0.05
HC0	0.453	0.19	0.209	0.193
HC1	0.486	0.203	0.185	0.171
HC2	0.557	0.247	0.15	0.136
HC3	0.667	0.309	0.11	0.1
max(Conventional, HC0)	0.629	0.109	0.055	0.045
max(Conventional, HC1)	0.64	0.122	0.053	0.044
max(Conventional, HC2)	0.679	0.166	0.047	0.039
max(Conventional, HC3)	0.754	0.237	0.039	0.031

نکته: این جدول نتایج یک نمونه‌گیری با ۲۵۰۰۰ بار تکرار را نشان می‌دهد.

جدول ۱-۲-۸. خطاهای استاندارد برای اثرات اندازه کلاس در داده‌های STAR

	Standard Error
Robust (HC1)	0.09
Parametric Moulton Correction (using Moulton intraclass coefficient)	0.222
Parametric Moulton Correction (using ANOVA intraclass coefficient)	0.23
Clustered	0.232
Block Bootstrap	0.231
Estimation using group means (weighted by class size)	0.226

نکته: این جدول تخمین‌های یک رگرسیون نمرات درصدی میانگین را بر اندازه کلاس برای کودکستان‌ها با استفاده از کاربرد عمودی مجموعه داده‌های پروژه **STAR** گزارش می‌دهد. ضریب اعمال شده بر اندازه کلاس ۰/۶۲- است. سطح گروه برای خوشه‌بندی، کلاس درسی است. تعداد مشاهدات ۵۷۴۳ بار است. تخمین خودگردان‌سازی از ۱۰۰۰ بار تکرار استفاده نموده است.

ختم کلام

اگر اقتصادسنجی کاربردی آسان بود، [اقتصادسنجی‌دان‌های] نظری انجامش می‌دادند، اما این حوزه آنچنان هم که صفحات ثقیل مجله اکونومتریکا ممکن است شما را به آن تصور برسانند، سخت نیست. اقتصادسنجی کاربردی اگر به دقت به سؤالات سازگار علّی، رگرسیون و 2SLS اعمال شود، معنا و مفهوم پیدا می‌کند. خطاهای استاندارد شما اگرچه به ندرت درست از آب درمی‌آیند، اما احتمالاً کاملاً درست نباشند. با شک و تردید زیاد به کار خود، خودتان را خجالت ندهید و مخصوصاً اینکه وحشت نکنید!

اختصارات

عبارات فنی (توضیح تفصیلی هرکدام از این عبارات در متن آمده است):

- ۲SLS - Two Stage Least Squares
- ACR - Average Causal Response
- ANOVA - Analysis of Variance
- BRL - Biased Reduced Linearization estimator
- CDF - Cumulative Distribution Function
- CEF - Conditional Expectation Function
- CIA - Conditional Independence Assumption
- COP - Conditional on Positive effect
- CQF - Conditional Quantile Function

- DD - Differences in Differences estimator
- GLS - Generalized Least Squares estimator
- GMM - Generalized Method of Moments
- HC0 - HC3 Heteroskedasticity Consistent variance estimators
- ILS - Indirect Least Squares estimator
- ITT - Intention to Treat effect
- IV - Instrumental Variables estimator
- LATE - Local Average Treatment Effect
- LDVs - Limited Dependent Variables
- LIML - Limited Information Maximum Likelihood estimator
- LM - Lagrange Multiplier test
- LPM - Linear Probability Model
- MFX - Marginal Effects
- MMSE - Minimum Mean Squared Error
- OLS - Ordinary Least Squares
- OVB - Omitted Variables Bias
- QTE - Quantile Treatment Effect
- RD - Regression Discontinuity design
- SEM - Simultaneous Equations Models
- SSIV - Split-Sample Instrumental Variables estimator
- TSIV - Two-Sample Instrumental Variables estimator
- VIV - Visual Instrumental Variables

مجموعه داده (دیتاست) و نام متغیر

- AFDC - Aid to Families with Dependent Children
- AFQT - Armed Forces Qualification Test
- CPS - Current Population Survey
- GED - General Educational Development certificate
- IPUMS - Integrated Public Use Microdata Series
- NHIS - National Health Interview Survey
- NLSV - National Longitudinal Survey of Youth
- PSAT - Preliminary SAT
- PSID - Panel Study of Income Dynamics
- QOB - Quarter of Birth
- RSN - Random Sequence Numbers
- SDA - Service Delivery Area
- SSA - Social Security Administration

مطالعات

- HIE - Health Insurance Experiment
- HRT - Hormone Replacement Therapy

- JSA - Job Search Assistance
- JTPA - Job Training Partnership Act
- MDVE - Minneapolis Domestic Violence Experiment
- NSW - National Supported Work demonstration
- OJT - On the Job Training
- STAR - the Tennessee Student/Teacher Achievement Ratio experiment
- WHI - Women's Health Initiative

فهرست مطالعات تجربی

- Abadie, Amgrist, amd Imbems (2002)
- Acemoglu amd Amgrist (2000)
- Amgrist (1990)
- Amgrist (1998)
- Amgrist (2001)
- Amgrist amd Evams (1998)
- Amgrist amd Imbems (1995)
- Amgrist amd Krueger (1991)
- Amgrist amd Lavy (1999)
- Amgrist, Chermozhukov, Fermamdez-Val (2006)
- Autor (2003)
- Besley amd Burgess (2004)
- Card (1992)
- Card amd Krueger (1994, 2000)
- Dehejia amd Wahba (1999)
- Freemam (1984)
- Krueger (1999)
- Lee (2008)
- Mammimg, et al (198Y)
- Pischke (200Y)

منابع و مآخذ

1. Abadie, Alberto (2003): “Semiparametric Instrumental Variable Estimation of Treatment Response Mod-els,”Journal of Econometrics, 113, 231–263.
2. Abadie, Alberto, Joshua D. Angrist, and Guido Imbens (2002): “Instrumental Variables Estimates of the Effect of Subsidized Training on the Quantiles of Trainee Earnings,”Econometrica, 70, 91–117.
3. Abadie, Alberto, Alexis Diamond, and Jens Hainmueller (2007): “Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California’s Tobacco Control Program,” National Bureau of Economic Research, Working Paper No. 12831.

4. Abadie, Alberto, and Guido Imbens (2006): “Large Sample Properties of Matching Estimators for Average Treatment Effects,”Econometrica, 74, 235–67.
5. _____ (2008) Bias-Corrected Matching Estimators for Average Treatment Effects,”Harvard University, Department of Economics, mimeo.
6. Acemoglu, Daron, and Joshua Angrist (2000): “How Large are the Social Returns to Education? Evidence from Compulsory Schooling Laws,” in National Bureau of Economics Macroeconomics Annual 2000, ed. by Ben S. Bernanke, and Kenneth S. Rogoff, pp. 9–58. The MIT Press, Cambridge.
7. Acemoglu, Daron, Simon Johnson, and James A. Robinson (2001): “The Colonial Origins of Comparative Development: An Empirical Investigation,”The American Economic Review, 91, 1369–1401.
8. Adams, Douglas (1979): The Hitchhiker’s Guide to the Galaxy. Pocket Books, New York.
9. _____ (1990) Dirk Gently’s Holistic Detective Agency. Simon & Schuster, New York.
10. _____ (1995) Mostly Harmless. Harmony Books, New York.
11. Altonji, Joseph G., and Lewis M. Segal (1996): “Small-Sample Bias in GMM Estimation of Covariance Structures,”Journal of Business and Economic Statistics, 14, 353–366.
12. Ammermueller, Andreas, and Jorn-Steffan Pischke (2006): “Peer Effects in European Primary
13. Schools: Evidence from PIRLS,” Institute for the Study of Labor (IZA), Discussion Paper No. 2077.
14. Ananat, Elizabeth, and Guy Michaels (2008): “The Effect of Marital Breakup on the Income Distri-bution of Women with Children,”Journal of Human Resources, forthcoming.
15. Anderson, Michael (2008): “Multiple Inference and Gender Differences in the Effect of Early Interven-tion: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects,” Journal of the American Statistical Association, forthcoming.
16. Angrist, Joshua, Eric Bettinger, Erik Bloom, Elizabeth King, and Michael Kremer (2002): “Vouchers for Private Schooling in Colombia: Evidence from a Randomized Natural Experiment,” The American Economic Review, 92, 1535–1558.
17. Angrist, Joshua D. (1988): “Grouped Data Estimation and Testing in Simple Labor Supply Models,” Princeton University, Industrial Section, Working Paper No. 234.
18. _____ (1990) Lifetime Earnings and the Vietnam Era Draft Lottery: Evidence from Social Security Administrative Records,”American Economic Review, 80, 313–335.
19. _____ (1991) Grouped Data Estimation and Testing in Simple Labor Supply Models,” Journal of Econometrics, 47, 243–266.
20. _____ (1998) Estimating the Labor Market Impact on Voluntary Military Service Using Social Security Data on Military Applicants,”Econometrica, 66,

249–288.

21. _____ (2001) Estimations of Limited Dependent Variable Models with Dummy Endogenous Regressors:
22. Simple Strategies for Empirical Practice,"Journal of Business and Economic Statistics, 19, 2–16.
23. _____ (2004) American Education Research Changes Track," Oxford Review of Economic Policy, 20, 198–212.
24. _____ (2006) Instrumental Variables Methods in Experimental Criminological Research: What, Why and How,"Journal of Experimental Criminology, 2, 22–44.
25. Angrist, Joshua D., Victor Chernozhukov, and Ivan Fernandez-Val (2006): "Quantile Regression Under Misspeci…cation, with an Application to the U.S. Wage Structure,"Econometrica, 74, 539–563.
26. Angrist, Joshua D., and William N. Evans (1998): "Children and Their Parents' Labor Supply:
27. Evidence from Exogenous Variation in Family Size,"American Economic Review, 88, 450–477.
28. _____ (1999) Schooling and Labor Market Consequences of the 1970 State Abortion Reforms," in Re-search in Labor Economics, ed. by Solomon W. Polachek, vol. 18, pp. 75–113. Elsevier Science, Amsterdam.
29. Angrist, Joshua D., Kathryn Graddy, and Guido W. Imbens (2000): "The Interpretation of Instru-mental Variables Estimators in Simultaneous Equations Models with an Application to the Demand for Fish,"Review of Economic Studies, 67, 499–527.
30. Angrist, Joshua D., and Jinyong Hahn (2004): "When to Control for Covariates? Panel Asymptotics for Estimates of Treatment Effects,"Review of Economics and Statistics, 86, 58–72.
31. Angrist, Joshua D., Guido Imbens, and Donald B. Rubin (1996): "Identi…cation of Causal Effects Using Instrumental Variables,"Journal of the American Statistical Association, 91, 444–472.
32. Angrist, Joshua D., and Guido W. Imbens (1995): "Two-Stage Least Squares Estimation of Aver-age Causal Effects in Models with Variable Treatment Intensity," Journal of the American Statistical Association, 90, 430–442.
33. Angrist, Joshua D., and Alan B. Krueger (1991): "Does Compulsory Schooling Attendance Affect Schooling and Earnings?,"Quarterly Journal of Economics, 106, 976–1014.
34. _____ (1992) The Effect of Age at School Entry on Educational Attainment: An Application of Instru-mental Variables with Moments from Two Samples,"Journal of the American Statistical Association, 418, 328–36.
35. ___ (1995) Split-Sample Instrumental Variables Estimates of the Return to Schooling," Journal of Business and Economic Statistics, 13, 225–35.
36. _____ (1999) Empirical Strategies in Labor Economics," in Handbook of

Labor Economics, ed. by Orley C. Ashenfelter, and David Card, vol. 3. North Holland, Amsterdam.

37. _____ (2001) Instrumental Variables and the Search for Identi…cation: From Supply and Demand to Natural Experiments,"Journal of Economic Perspectives, 15, 69–85.

38. Angrist, Joshua D., and Guido Kuersteiner (2004): "Semiparametric Causality Tests Using the Policy Propensity Score," National Bureau of Economic Research, Working Paper No. 10975.

39. Angrist, Joshua D., and Kevin Lang (2004): "Does School Integration Generate Peer Effects? Evidence from Boston' Metco Program,"The American Economic Review, 94, 1613–1634.

40. Angrist, Joshua D., and Victor Lavy (1999): "Using Maimonides'Rule to Estimate the Effect of Class Size on Scholastic Achievement,"Quarterly Journal of Economics, 114, 533–575.

41. _____ (2007) The Effects of High Stakes High School Achievement Awards: Evidence from a Group-Randomized Trial,"unpublished paper, Department of Economics, Massachusetts Institute of Technology.

42. Angrist, Joshua D., Victor Lavy, and Analia Schlosser (2006): "Multiple Experiments for the Causal Link Between the Quantity and Quality of Children," MIT Department of Economics Working Paper No. 06-26.

43. Arellano, Manuel, and Stephen Bond (1991): "Some Tests of Speci…cation for Panel Data: Monte Carlo Evidence and an Application to Employment Equations," The Review of Economic Studies, 58, 277–297.

44. Ashenfelter, Orley A. (1978): "Estimating the Effect of Training Programs on Earnings," Review of Economics and Statistics, 60, 47–57.

45. _____ (1991) How Convincing is the Evidence Linking Education and Income?," Princeton University, Industrial Relations Section, Working Paper No. 292.

46. Ashenfelter, Orley A., and David Card (1985): "Using the Longitudinal Structure of Earnings to Estimate the Effect of Training Programs,"The Review of Economics and Statistics, 67, 648–660.

47. Ashenfelter, Orley A., and Alan B. Krueger (1994): "Estimates of the Economic Return to Schooling from a New Sample of Twins,"American Economic Review, 84, 1157–1173.

48. Ashenfelter, Orley A., and Cecilia Rouse (1998): "Income, Schooling, and Ability: Evidence from a New Sample of Identical Twins,"Quarterly Journal of Economics, 113, 253–284.

49. Athey, Susan, and Guido Imbens (2006): "Identi…cation and Inference in Nonlinear Difference-in-Difference Models,"Econometrica, 74, 431–497.

50. Atkinson, Anthony B. (1970): "On the Measurement of Inequality," Journal of Economic Theory, 2, 244–263.

51. Autor, David (2003): "Outsourcing at Will: The Contribution of Unjust Dismissal Doctrine to the Growth of Employment Outsourcing,"Journal of

Labor Economics, 21, 1–42.

52. Autor, David, Lawrence F. Katz, and Melissa S. Kearney (2005): “Rising Wage Inequality: The Role of Composition and Prices,” National Bureau of Economic Research, Working Paper No. 11628.

53. Barnett, Steven W. (1992): “Bene…ts of Compensatory Preschool Education,” Journal of Human Re-sources, 27, 279–312.

54. Barnow, Burt S., Glen G. Cain, and Arthur Goldberger (1981): “Selection on Observables,” Evaluation Studies Review Annual, 5, 43–59.

55. Bekker, Paul A. (1994): “Alternative Approximations to the Distributions of Instrumental Variable Estimators,”Econometrica, 62, 657–681.

56. Bell, Robert M., and Daniel F. McCaffrey (2002): “Bias Reduction in Standard Errors for Linear Regression with Multistage Samples,”Survey Methodology, 28, 169–181.

57. Bennedsen, Morten, Kasper M. Nielsen, Francisco Pérez-González, and Daniel Wolfenzon (2007): “Inside the Family Firm: The Role of Families in Succession Decisions and Performance,” The Quarterly Journal of Economics, 122, 647–692.

58. Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan (2004): “How Much Should We Trust Differences-in-Differences Estimates?,”Quarterly Journal of Economics, 119, 249–275.

59. Bertrand, Marianne, and Sendhil Mullainathan (2004): “Are Emily and Greg More Employable than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination,”The American Economic Review, 94, 991–1013.

60. Besley, Timothy, and Robin Burgess (2004): “Can Labour Market Regulation Hinder Economic Per-formance? Evidence from India,”Quarterly Journal of Economics, 113, 91–134.

61. Bjorklund, Anders, and Markus Jantti (1997): “Intergenerational Income Mobility in Sweden Com-pared to the United States,”The American Economic Review, 87, 1009–1018.

62. Black, Dan A., Jeffrey A. Smith, Mark C. Berger, and Brett J. Noel (2003): “Is the Threat of Reemployment Services More Effective than the Services Themselves? Evidence from Random Assignment in the UI System,”The American Economic Review, 93, 1313–1327.

63. Bloom, Howard S. (1984): “Estimating the Effect of Job-Training Programs, Using Longitudinal Data:

64. Ashenfelter’ Findings Reconsidered,”The Journal of Human Resources, 19, 544–556.

65. Bloom, Howard S., Larry L. Orr, Stephen H. Bell, George Cave, Fred Doolittle, Winston Lin, and Johannes M. Bos (1997): “The Bene…ts and Costs of JTPA Title II-A Programs: Key Findings from the National Job Training Partnership Act Study,”The Journal of Human Resources, 32, 549–576.

66.Blundell, Richard, and Stephen Bond (1998): “Initial Conditions and Moment

Restrictions in Dy-namic Panel Data Models,"Journal of Econometrics, 87, 115–143.

67. Borjas, George (1992): "Ethnic Capital and Intergenerational Mobility,"Quarterly Journal of Economics, 107, 123–150.
68. _____ (2005) Labor Economics, 3rd edn. McGraw-Hill/Irwin, New York.
69. Bound, John, David Jaeger, and Regina Baker (1995): "Problems with Instrumental Variables Esti-mation when the Correlation between the Instruments and the Endogenous Variables is Weak," Journal of American Statistical Association, 90, 443–450.
70. Bound, John, and Gary Solon (1999): "Double Trouble: On the Value of Twins-based Estimation of the Returns of Schooling,"Economics of Education Review, 18, 169–182.
71. Bronars, Stephen G., and Jeff Grogger (1994): "The Economic Consequences of Unwed Motherhood:
72. Using Twin Births as a Natural Experiment,"American Economic Review, 84, 1141–1156.
73. Buchinsky, Moshe (1994): "Changes in the U.S. Wage Structure 1963-1987: Application of Quantile Regression,"Econometrica, 62, 405–458.
74. Buse, A. (1992): "The Bias of Instrumental Variable Estimators,"Econometrica, 60, 173–180.
75. Cameron, Colin, Jonah Gelbach, and Douglas L. Miller (2008): "Bootstrap-Based Improvements for Inference with Clustered Errors," The Review of Economics and Statistics, forthcoming, unpublished paper, Department of Economics, The University of California at Davis.
76. Campbell, Donald Thomas (1969): "Reforms as Experiments,"American Psychologist, 24, 409–429.
77. Campbell, Donald Thomas, and Julian C. Stanley (1963): Experimental and Quasi-experimental Designs for Research. Rand McNally, Chicago.
78. Card, David (1992): "Using Regional Variation to Measure the Effect of the Federal Minimum Wage," Industrial and Labor Relations Review, 46, 22–37.
79. _____ (1995) Earnings, Schooling and Ability Revisited," in Research in Labor Economics, ed. by Solomon W. Polachek, vol. 14, pp. 23–48. JAI Press, Greenwich, Connecticut.
80. _____ (1999) The Causal Effect of Education on Earnings," in Handbook of Labor Economics, ed. by Orley C. Ashenfelter, and David Card, vol. 3. North Holland, Amsterdam.
81. Card, David, and Alan Krueger (1994): "Minimum Wages and Employment: A Case Study of the Fast Food Industry in New Jersey and Pennsylvania,"American Economic Review, 84, 772–784.
82. _____ (2000) Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania: Reply,"American Economic Review, 90, 1397–420.
83. Card, David, and David S. Lee (2008): "Regression Discontinuity Inference with Speci…cation Error," Journal of Econometrics, 142, 655–674.

84. Card, David, and Thomas Lemieux (1996): “Wage Dispersion, Returns to Skill, and Black-White Dif-ferentials,”Journal of Econometrics, 74, 316–361.

85. Card, David E., and Daniel Sullivan (1988): “Measuring the Effect of Subsidized Training on Move-ments In and Out of Employment,”Econometrica, 56, 497–530.

86. Cardell, Nicholas Scott, and Mark Myron Hopkins (1977): “Education, Income, and Ability: A Comment,”Journal of Political Economy, 85, 211–215.

87. Chamberlain, Gary (1977): “Education, Income, and Ability Revisited,” Journal of Econometrics, 5, 241–57.

88. _____ (1978) Omitted Variables Bias in Panel Data: Estimating the Returns to Schooling,” Annales De L’INSEE, 30-31, 49–82.

89. _____ (1984) Panel Data,”in Handbook of Econometrics, ed. by Zvi Griliches, and Michael D. Intriligator, vol. 2, pp. 1247–1318. North Holland, Amsterdam.

90. _____ (1994) Quantile Regression, Censoring and the Structure of Wages,” in Proceedings of the Sixth World Cogress of the Econometrics Society, Barcelona, Spain, ed. by Christopher A. Sims, and Jean-Jacques Laffont, pp. 179–209. Cambridge University Press, New York.

91. Chamberlain, Gary, and Edward E. Leamer (1976): “Matrix Weighted Averages and Posterior Bounds,”Journal of the Royal Statistical Society, Series B, 38, 73–84.

92. Chernozhukov, Victor, and Christian Hansen (2005): “An IV Model of Quantile Treatment Effects,” Econometrica, 73, 245–261.

93. _____ (2007) A Simple Approach to Heteroskedasticity and Autocorrelation Robust Inference with Weak Instruments,” unpublished paper, Department of Economics, Massachusetts Institute of Technology.

94. Chesher, Andrew, and Gerald Austin (1991): “The Finite-Sample Distributions of Heteroskedasticity Robust Wald Statistics,”Journal of Econometrics, 47, 153–173.

95. Chesher, Andrew, and Ian Jewitt (1987): “The Bias of the Heteroskedasticity Consistent Covariance Estimator,”Econometrica, 55, 1217–1222.

96. Cochran, William G. (1965): “The Planning of Observational Studies of Human Populations,” Journal of the Royal Statistical Society, Series A, 128, 234–65.

97. Cook, Thomas D. (2008): “Waiting for Life to Arrive: A History of the Regression-Discontinuity Design in Psychology, Statistics, and Economics,”Journal of Econometrics, 142, 636–654, forthcoming.

98. Cook, Thomas D., and Vivian C. Wong (2008): “Empirical Tests of the Validity of the Regression-Discontinuity Design,”Annales d’Economie et de Statistique, forthcoming.

99. Crump, Richard K., V. Joseph Hotz, Guido W. Imbens, and Oscar A. Mitnik (2006): “Moving the Goalposts: Addressing Limited Overlap in the Estimation of Average Treatment Effects by Changing the Estimand,” National Bureau of

Economic Research, Technical Working Paper No. 330.

100. Cruz, Luiz M., and Marcelo J. Moreira (2005): “On the Validity of Econometric Techniques with Weak Instruments: Inference on Returns to Education Using Compulsory School Attendance Laws,”Journal of Human Resources, 40, 393–410.

101. Currie, Janet, and Aaron Yelowitz (2000): “Are Public Housing Projects Good for Kids?,” Journal of Public Economics, 75, 99–124.

102. Davidon, Russell, and James G. MacKinnon (1993): Estimation and Inference in Econometrics. Oxford University Press, New York.

103. Dearden, Lorraine, Sue Middleton, Sue Maguire, Karl Ashworth, Kate Legge, Tracey Allen, Kim Perrin, Erich Battistin, Carl Emmerson, Emla Fitzsimons, and Costas Meghir (2004): “The Evaluation of Education Maintenance Allowance Pilots: Three Years’Evidence. A Quanti-tative Evaluation,”Department for Education and Skills, Research Report No. 499.

104. Deaton, Angus (1997): The Analysis of Household Surveys: A Microeconometric Approach to Development Policy. Johns Hopkins University Press for the World Bank, Baltimore, MD.

105. Dee, Thomas S., and William N. Evans (2003): “Teen Drinking and Educational Attainment: Evidence from Two-Sample Instrumental Variables Estimates,”Journal of Labor Economics, 21, 178–209.

106. DeGroot, Morris H., and Mark J. Schervish (2001): Probability and Statistics, 3rd edn. Addison-Wesley, Boston.

107. Dehejia, Rajeev H. (2005): “Practical Propensity Score Matching: A Reply to Smith and Todd,”Journal of Econometrics, 125, 355–364.

108. Dehejia, Rajeev H., and Sadek Wahba (1999): “Causal Effects in Nonexperimental Studies: Reevalu-ating the Evaluation of Training Programs,”Journal of the American Statistical Association, 94, 1053–62.

109. Donald, Stephen G., and Kevin Lang (2007): “Inference with Difference-in-Differences and Other Panel Data,”Review of Economics and Statistics, 89, 221–233.

110. Duan, Naihua, Willard D. Manning, Jr., Carl N. Morris, and Joseph P. Newhouse (1983): “A Comparison of Alternative Models for the Models for the Demand for Medical Care,”Journal of Business & Economic Statistics, 1, 115–126.

111. _____ (1984) Choosing Between the Sample-Selection Model and the Multi-Part Model,” Journal of Business & Economic Statistics, 2, 283–289.

112. Durbin, James (1954): “Errors in Variables,”Review of the International Statistical Institute, 22, 23–32.

113. Eicker, Friedhelm (1967): “Limit Theorems for Regressions with Unequal and Dependent Errors,” in Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, vol. 1, pp. 59–82. University of California Press, Berkeley.

114. Elder, Todd E., and Darren H. Lubotsky (2008): “Kindergarten Entrance Age and Children’s Achieve-ment: Impacts of State Policies, Family Background,

and Peers,"Journal of Human Resources, forthcom-ing, forthcoming.

115. Finn, Jeremy D., and Charles M. Achilles (1990): "Answers and Questions About Class Size: A Statewide Experiment,"American Educational Research Journal, 28, 557–77.
116. Firpo, Sergio (2007): "Efficient Semiparametric Estimation of Quantile Treatment Effects,"Econometrica, 75, 259–276.
117. Flores-Lagunes, Alfonso (2007): "Finite Sample Evidence of IV Estimators under Weak Instruments," Journal of Applied Econometrics, 22, 677–694.
118. Freedman, David (2005): "Linear Statistical Models for Causation: A Critical Review," in The Wiley Encyclopedia of Statistics in Behavioral Science, ed. by B. Everitt, and D. Howell. John Wiley, Chichester, UK.
119. Freeman, Richard (1984): "Longitudinal Analyses of the Effect of Trade Unions," Journal of Labor Economics, 3, 1–26.
120. Frisch, Ragnar, and Frederick V. Waugh (1933): "Partial Time Regression as Compared with Indi-vidual Trends,"Econometrica, 1, 387–401.
121. Frolich, Markus, and Blaise Melly (2007): "Unconditional Quantile Treatment Effects Under Endo-geneity," Centre for Microdata Methods and Practice, Working Paper No. CWP32/07.
122. Fryer, Roland G., and Steven D. Levitt (2004): "The Causes and Consequences of Distinctively Black Names,"The Quarterly Journal of Economics, 119, 767–805.
123. Goldberger, Arthur S. (1972): "Selection Bias in Evaluating Treatment Effects: Some Formal Illustra-tions," University of Wisconsin, Department of Economics, Working Paper.
124. _____ (1991) A Course in Econometrics. Harvard University Press, Cambridge, MA.
125. Gosling, Amanda, Stephen Machin, and Costas Meghir (2000): "The Changing Distribution of Male Wages in the U.K.,"Review of Economic Studies, 67, 635–66.
126. Granger, Clive W. J. (1969): "Investigating Causal Relation by Econometric and Cross-Sectional Method,"Econometrica, 37, 424–438.
127. Griliches, Zvi (1977): "Estimating the Returns to Schooling: Some Econometric Problems,"Econometrica, 45, 1–22.
128. Griliches, Zvi, and Jerry A. Hausman (1986): "Errors in variables in panel data," Journal of Econo-metrics, 31, 93–118.
129. Griliches, Zvi, and William M. Mason (1972): "Education, Income, and Ability," Journal of Political Economy, 80, S74–S103.
130. Grumbach, Kevin, Dennis Keane, and Andrew Bindman (1993): "Primary Care and Public Emer-gency Department Overcrowding,"American Journal of Public Health, 83, 372–378.
131. Guryan, Jonathan (2004): "Desegregation and Black Dropout Rates," American Economic Review, 94, 919–943.

132. Haavelmo, Trygve (1944): “The Probability Approach in Econometrics,”Econometrica, 12, S1–S115.

133. Hahn, Jinyong (1998): “On the Role of the Propensity Score in E¢ cient Semiparametric Estimation of Average Treatment Effects,”Econometrica, 66, 315–31.

134. Hahn, Jinyong, Petra Todd, and Wilbur van der Klaauw (2001): “Identi…cation and Estimation of Treatment Effects with a Regression-Discontinuity Design,”Econometrica, 69, 201–209.

135. Hansen, Christian B. (2007a): “Asymptotic Properties of a Robust Variance Matrix Estimator for Panel Data when T is Large,”Journal of Econometrics, 141, 597–620.

136. (2007) b): “Generalized Least Squares Inference in Panel and Multilevel Models with Serial Corre-lation and Fixed Effects,”Journal of Econometrics, 140, 670–694.

137. Hansen, Lars Peter (1982): “Large Sample Properties of Generalized Method of Moments Estimators,” Econometrica, 50, 1029–1054.

138. Hausman, Jerry (1978): “Speci…cation Tests in Econometrics,”Econometrica, 46, 1251–1271.

139. _____ (1983) Speci…cation and Estimation of Simultaneous Equation Models,” in Handbook of Econo-metrics, ed. by Zvi Griliches, and Michael Intriligator, vol. 1, pp. 391–448. North Holland, Amsterdam.

140. _____ (2001) Mismeasured Variables in Econometric Analysis: Problems from the Right and Problems from the Left,”Journal of Econometric Perspectives, 15, 57–67.

141. Hausman, Jerry, Whitney Newey, and Tiemen Wouterson (2006): “IV Estimation with Het-eroskedasticity and Many Instruments,” unpublished paper, Department of Economics, Massachusetts Institute of Technology.

142. Hay, Joel W., and Randall J. Olsen (1984): “Let Them Eat Cake: A Note on Comparing Alternative Models of the Demand for Medical Care,”Journal of Business & Economic Statistics, 2, 279–282.

143. Heckman, James J. (1978): “Dummy Endogenous Variables in a Simultaneous Equations System,”Econo-metrica, 46, 695–712.

144. Heckman, James J., Hidehiko Ichimura, and Petra E. Todd (1998): “Matching as as Econometric Evaluation Estimator,”Review of Economic Studies, 62, 261–94.

145. Heckman, James J., Jeffrey Smith, and Nancy Clements (1997): “Making the Most Out of Pro-gramme Evaluations and Social Experiments: Accounting for Heterogeneity in Programme Impacts,”The Review of Economic Studies, 64, 487–535.

146. Hirano, Keisuke, Guido W. Imbens, and Geert Ridder (2003): “E¢ cient Estimation of Average Treatment Effects Using the Estimated Propensity Score,”Econometrica, 71, 1161–89.

147. Hirano, Keisuke, Guido W. Imbens, Donald B. Rubin, and Xiao-Hua Zhou (2000): “Assessing the Effect of an In‡uenza Vaccine in an Encouragement

Design,”Biostatistics, 1, 69–88.

148. Hoaglin, David C., and Roy E. Welsch (1978): “The Hat Matrix in Regression and ANOVA,” The American Statistician, 32, 17–22.
149. Holland, Paul W. (1986): “Statistics and Causal Inference,”Journal of the American Statistical Associ-ation, 81, 945–70.
150. Holtz-Eakin, Douglas, Whitney Newey, and Harvey S. Rosen (1988): “Estimating Vector Autore-gressions with Panel Data,”Econometrica, 56, 1371–1395.
151. Horowitz, Joel L. (1997): “Bootstrap Methods in Econometrics: Theory and Numerical Performance,” in Advances in Economics and Econometrics: Theory and Applications, ed. by David M. Kreps, and Kenneth F. Wallis, vol. 3, pp. 188–222. Cambridge University Press, Cambridge.
152. _____ (2001) The Bootstrap,”in Handbook of Econometrics, ed. by James J. Heckman, and Edward E.
153. Leamer, vol. 5, pp. 3159–3228. Elsevier Science, Amsterdam.
154. Horvitz, Daniel G., and Donovan J. Thompson (1952): “A Generalization of Sampling Without Replacement From a Finite Population,”Journal of the American Statistical Association, 47, 663–85.
155. Hoxby, Caroline (2000): “The Effects of Class Size on Student Achievement: New Evidence from Popu-lation Variation,”The Quarterly Journal of Economics, 115, 1239–1285.
156. Hsia, Judith, Robert D. Langer, JoAnn E. Manson, Lewis Kuller, Karen C. Johnson, Susan L. Hendrix, Mary Pettinger, Susan R. Heckbert, Nancy Greep, Sybil Crawford, Charles B. Eaton, John B. Kostis, Pat Caralis, Ross Prentice, and for the Women’s Health Initia-tive Investigators (2006): “Conjugated Equine Estrogens and Coronary Heart Disease: The Women’s Health Initiative,”Archives of Internal Medicine, 166, 357–365.
157. Imbens, Guido (2000): “The Role of the Propensity Score in Estimating Dose-Response Functions,” Bio-metrika, 87, 706–10.
158. _____ (2004) Nonparametric Estimation of Average Treatment Effects Under Exogeneity: A Review,” The Review of Economics and Statistics, 86, 4–29.
159. Imbens, Guido, and Joshua Angrist (1994): “Identi…cation and Estimation of Local Average Treatment Effects,”Econometrica, 62, 467–476.
160. Imbens, Guido, and Thomas Lemieux (2008): “Regression Discontinuity Designs: A Guide to Practice,” Journal of Econometrics, 142, 615–635.
161. Inoue, Atsushi, and Gary Solon (2005): “Two-Sample Instrumental Variables Estimators,” National Bureau of Economic Research, Technical Working Paper No. 311.
162. Jappelli, Tullio, Jorn-Steffen Pischke, and Nicholas S. Souleles (1998): “Testing for Liquid-ity Constraints in Euler Equations with Complementary Data Sources,” The Review of Economics and Statistics, 80, 251–262.
163. Johnson, Norman L., and Samuel Kotz (1970): Distributions in Statistics:

Continuous Distributions, vol. 2. John Wiley, New York.

164. Kauermann, Goran, and Raymond J. Carroll (2001): “A Note on the E¢ ciency of Sandwich Covari-ance Estimation,”Journal of the American Statistical Association, 96, 1387–1396.
165. Kelejian, Harry H. (1971): “Two Stage Least Squares and Econometric Systems Linear in Parameters but Non-linear in the Endogenous Variables,”Journal of the American Statistical Association, 69, 373–4.
166. Kennan, John (1995): “The Elusive Effects of Minimum Wages,” Journal of Economic Literature, 33, 1950–1965.
167. Kézdi, Gábor (2004): “Robust Standard Error Estimation in Fixed-Effects Panel Models,” Hungarian Statistical Review, Special English Volume, 9, 95–116.
168. Kish, Leslie (1965): “Sampling Organizations and Groups of Unequal Sizes,”American Sociological Review, 30, 564–572.
169. Kloek, Teun (1981): “OLS Estimation in a Model Where a Microvariable is Explained by Aggregates and Contemporaneous Disturbances are Equicorrelated,”Econometrica, 49, 205–207.
170. Knight, Keith (2000): Mathematical Statistics. Chapman & Hall/CRC, Boca Raton, FL.
171. Koenker, Roger (2005): Quantile Regression. Cambridge University Press.
172. Koenker, Roger, and Gilbert Bassett (1978): “Regression Quantiles,”Econometrica, 46, 33–50.
173. Koenker, Roger, and Stephen Portnoy (1996): “Quantile Regression,”University of Illinois at Urbana-Champaign, College of Commerce and Business Administration, Working Paper No. 97-0100.
174. Krueger, Alan B. (1999): “Experimental Estimates of Education Production Functions,”Quarterly Jour-nal of Economics, 114, 497–532.
175. Kugler, Adriana, Juan F. Jimeno, and Virginia Hernanz (2005): “Employment Consequences of Restrictive Permanent Contracts: Evidence from Spanish Labor Market Reforms,”FEDEA Working Paper 2003-14.
176. LaLonde, Robert (1995): “The Promise of Public Sector-Sponsored Training Programs,” Journal of Economic Perspectives, 93, 149–68.
177. LaLonde, Robert J. (1986): “Evaluating the Econometric Evaluations of Training Programs Using Ex-perimental Data,”American Economic Review, 76, 602–20.
178. Lee, David S. (2008): “Randomized experiments from non-random selection in U.S. House elections,” Journal of Econometrics, 142, 675–697.
179. Lemieux, Thomas (2008): “The Changing Nature of Wage Inequality,”Journal of Population Economics, 21, 21–48.
180. Liang, Kung-Yee, and Scott L. Zeger (1986): “Longitudinal Data Analysis Using Generalized Linear Models,”Biometrika, 73, 13–22.
181. Machado, Jose, and Jose Mata (2005): “Counterfactual Decompositions of Changes in Wage Distribu-tions Using Quantile Regression,”Journal of

Applied Econometrics, 20, 445–65.

182. MacKinnon, James G., and Halbert White (1985): “Some Heteroskedasticity Consistent Covariance Matrix Estimators With Improved Finite Sample Properties,”Journal of Econometrics, 29, 305–325.
183. Maddala, Gangadharrao Soundalyarao (1983): “Methods of Estimation for Models of Markets with Bounded Price Variation,”International Economic Review, 24, 361–378.
184. Mammen, Enno (1993): “Bootstrap and Wild Bootstrap for High Dimensional Linear Models,” Annals of Statistics, 21, 255–285.
185. Manning, Willard G., Joseph P. Newhouse, Naihua Duan, Emmett B. Keeler, Arleen Lei-bowitz, and Susan M. Marquis (1987): “Health Insurance and the Demand for Medical Care: Evidence from a Randomized Experiment,”American Economic Review, 77, 251–77.
186. Manski, Charles F. (1991): “Regression,”Journal of Economic Literature, 29, 34–50.
187. Mariano, Roberto S. (2001): “Simultaneous Equation Model Estimators: Statistical Properties,” in A Companion to Theoretical Econometrics, ed. by B. Baltagi. Blackwell, Oxford.
188. McClellan, Mark B., Barbara J. McNeil, and Joseph P. Newhouse (1994): “Does More Intensive Treatment of Acute Myocardial Infarction Reduce Mortality? Analysis Using Instrumental Variables,” Journal of the American Medical Association, 272, 859–66.
189. McCrary, Justin (2008): “Manipulation of the Running Variable in the Regression Discontinuity Design:
190. A Density Test,”Journal of Econometrics, 142, 698–714.
191. McDonald, John F., and Robert A. Moffitt (1980): “The Uses of Tobit Analysis,” The Review of Economics and Statistics, 62, 318–321.
192. Melly, Blaise (2005): “Decomposition of Differences in Distribution Using Quantile Regression,” Labour Economics, 12, 577–590.
193. Meltzer, Allan H., and Scott F. Richard (1983): “Tests of a Rational Theory of the Size of Govern-ment,”Public Choice, 41, 403–418.
194. Messer, Karen, and Halbert White (1984): “A Note on Computing the Heteroskedasticity Consistent Covariance Matrix Using Instrumental Variables Techniques,”Oxford Bulletin of Economics and Statistics, 46, 181–184.
195. Meyer, Bruce, Kip Viscusi, and David Durbin (1995): “Workers’Compensation and Injury Duration: Evidence from a Natural Experiment,”American Economic Review, 85, 322–340.
196. Meyer, Bruce D., and Dan T. Rosenbaum (2001): “Welfare, the Earned Income Tax Credit, and the Labor Supply of Single Mothers,”The Quarterly Journal of Economics, 116, 1063–1114.
197. Milgram, Stanley (1963): “Behavioral Study of Obedience,”Journal of

Abnormal and Social Psychology, 67, 371–78.

198. Moffitt, Robert (1992): “Incentive Effects of the U.S. Welfare System: A Review,”Journal of Economic Literature, 30, 1–61.
199. Morgan, Mary S. (1990): The History of Econometric Ideas. Cambridge University Press, Cambridge, U.K.
200. Moulton, Brent (1986): “Random Group Effects and the Precision of Regression Estimates,”Journal of Econometrics, 32, 385–397.
201. Nelson, Charles R., and Richard Startz (1990a): “The Distribution of the Instrumental Variables Estimator and Its t-ratio when the Instrument is a Poor One,”Journal of Business, 63, 125–140.
202. Neumark, David, and William Wascher (1992): “Employment Effects of the Minimum and Submin-imum Wages: Panel Data on State Minimum Wage Laws,” Industrial and Labor Relations Review, 46, 55–81.
203. Newey, Whitney K. (1985): “Generalized Method of Moments Speci…cation Testing,”Journal of Econo-metrics, 29, 299–256.
204. _____ (1990) Semiparametric Efficiency Bounds,”Journal of Applied Econometrics, 5, 99–135. Newey, Whitney K., and Kenneth D. West (1987): “Hypothesis Testing with Method of Moments Estimation,”International Economic Review, 28, 777–787.
205. Nickell, Stephen (1981): “Biases in Dynamic Models with Fixed Effects,”Econometrica, 49, 1417–1426. Obenauer, Marie, and Berta von der Nienburg (1915): “Effect of Minimum Wage Determinations in Oregon,” Bulletin of the US Bureau of Labor Statistics, No. 176, US GPO, Washington, D.C.
206. Oreopoulos, Philip (2006): “Estimating Average and Local Average Treatment Effects of Education When Compulsory Schooling Laws Really Matter,”American Economic Review, 96, 152–175.
207. Orr, Larry L., Howard S. Bloom, Stephen H. Bell, Fred Doolittle, and Winston Lin (1996): Does Training for the Disadvantaged Work? Evidence from the National JTPA Study. Urban Institute Press, Washington, D.C.
208. Pfefferman, Daniel (1993): “The Role of Sampling Weights When Modeling Survey Data,”International Statistical Review, 61, 317–37.
209. Pischke, Jorn-Steffen (2007): “The Impact of Length of the School Year on Student Performance and
210. Earnings: Evidence from the German Short School Years,”Economic Journal, 117, 1216–1242.
211. Porter, Jack (2003): “Estimation in the Regression Discontinuity Model,”unpublished paper, Department of Economics, Harvard University.
212. Poterba, James, Steven Venti, and David Wise (1995): “Do 401K Contributions Crowd Out Other Personal Savings,”Journal of Public Economics, 58, 1–32.
213. Powell, James L. (1986): “Censored Regression Quantiles,”Journal of Econometrics, 32, 143–155.

214. “ : (1989) Semiparametric Estimation of Censored Selection Models,” unpublished paper, Depart-ment of Economics, University of Wisconsin-Madison.
215. Prais, Sig J., and John Aitchison (1954): “The Grouping of Observations in Regression Analysis,” Revue de l’Institut International de Statistique / Review of the International Statistical Institute, 22, 1–22.
216. Reiersol, Olav (1941): “Con‡uence Analysis by Means of Lag Moments and Other Methods of Con‡uence Analysis,”Econometrica, 9, 1–24.
217. Robins, James M., Steven D. Mark, and Whitney K. Newey (1992): “Estimating Exposure Effects by Modeling the Expectation of Exposure Conditional on Confounders,”Biometrics, 48, 479–95.
218. Rosenbaum, Paul R. (1984): “The Consequences of Adjustment for a Concomitant Variables that has Been Affected by the Treatment,”Journal of the Royal Statistical Society, Series A, 147, 656–66.
219. _____ (1995) Observational Studies. Springer-Verlag, New York.
220. Rosenbaum, Paul R., and Donald B. Rubin (1983): “The Central Role of the Propensity Score in Observational Studies for Causal Effects,”Biometrika, 70, 41–55.
221. _____ (1985) The Bias Due to Incomplete Matching,”Biometrics, 41, 106–116.
222. Rosenzweig, Mark R., and Kenneth I. Wolpin (1980): “Testing the Quantity-Quality Fertility Model:
223. The Use of Twins as a Natural Experiment,”Econometrica, 48, 227–240.
224. Rubin, Donald B. (1973): “Matching to Remove Bias in Observational Studies,”Biometrics, 29, 159–83.
225. _____ (1974) Estimating the Causal Effects of Treatments in Randomized and Non-Randomized Stud-ies,”Journal of Educational Psychology, 66, 688–701.
226. _____ (1977) Assignment to a Treatment Group on the Basis of a Covariate,” Journal of Educational Statistics, 2, 1–26.
227. _____ (1991) Practical Implications of Modes of Statistical Inference for Causal Effects and the Critical Role of the Assignment Mechanism,”Biometrics, 47, 1213–34.
228. Ruud, Paul A. (1986): “Consistent Estimation of Limited Dependent Variable Models Despite Misspeci…-cation of Distribution,”Journal of Econometrics, 32, 157–87.
229. Shadish, William R., Thomas D. Cook, and Donald T. Campbell (2002): Experimental and Quasi-Experimental Designs for Generalized Causal Inference. Houghton-Mi- in Company, Boston.
230. Sherman, Lawrence W., and Richard A. Berk (1984): “The Speci…c Deterrent Effects of Arrest for Domestic Assault,”American Sociological Review, 49, 261–272.
231. Shore-Sheppard, Lara (1996): “The Precision of Instrumental Variables

Estimates with Grouped Data," Princeton University, Industrial Relations Section, Working Paper No. 374.

232. Smith, Jeffrey A., and Petra E. Todd (2001): "Reconciling Con‡icting Evidence on the Performance of Propensity-Score Matching Methods,"American Economic Review, 91, 112–118.
233. " : (2005) Does Matching Overcome LaLonde's Critique of Nonexperimental Estimators?," Journal of Econometrics, 125, 305–53.
234. Stigler, Stephen M. (1986): The History of Statistics: The Measurement of Uncertainty Before 1900.
235. The Belknap Press of Harvard University Press, Cambridge, MA.
236. Stock, James H., and Francesco Trebbi (2003): "Who Invented Instrumental Variables Regression?," The Journal of Economic Perspectives, 17, 177–94.
237. Stock, James H., Jonathan H. Wright, and Motohiro Yogo (2002): "A Survey of Weak Instruments and Weak Identi…cation in Generalized Method of Moments,"Journal of Business & Economic Statistics, 20, 518–529.
238. Taubman, Paul (1976): "The Determinants of Earnings: Genetics, Family and Other Environments: A Study of White Male Twins,"American Economic Review, 66, 858–870.
239. Thistlewaite, Donald L., and Donald T. Campbell (1960): "Regression-Discontinuity Analysis: An Alternative to the Ex Post Facto Experiment,"Journal of Educational Psychology, 51, 309–317.
240. Trochim, William (1984): Research Designs for Program Evaluation: The Regression Discontinuity De-sign. Sage Publications, Beverly Hills, CA.
241. van der Klaauw, Wilbert (2002): "Estimating the Effect of Financial Aid Offers on College Enrollment: A Regression-Discontinuity Approach,"International Economic Review, 43.
242. Wald, Abraham (1940): "The Fitting of Straight Lines if Both Variables are Subject to Error,"Annals of Mathematical Statistics, 11, 284–200.
243. " : (1943) Tests of Statistical Hypotheses Concerning Several Parameters When the Number of Observations is Large,"Transactions of the American Mathematical Society, 54, 426–482.
244. White, Halbert (1980a): "A Heteroskedasticity-Consistent Covariance Matric Estimator and a Direct Test for Heteroskedasticity,"Econometrica, 48, 817–38.
245. _____ (1982) Instrumental Variables Regression with Independent Observations," Econometrica, 50, 483–499.
246. Wooldridge, Jeffrey (2003): "Cluster-sample Methods in Applied Econometrics," American Economic Review, 93, 133.
247. _____ (2005) Fixed-Effects and Related Estimators for Correlated Random-Coefficient and Treatment-Effect Panel Data Models,"The Review of Economics and Statistics, 87, 385–390.
248. ____ (2006) Introductory Econometrics: A Modern Approach. Thomson/South-Western, Mason, OH.

249. Wright, Phillip G. (1928): The Tariff on Animal and Vegetable Oils. Macmillan, New York.
250. Yang, Song, Li Hsu, and Lueping Zhao (2005): “Combining Asymptotically Normal Tests: Case Studies in Comparison of Two Groups,”Journal of Statistical Planning and Inference, 133, 139–158.
251. Yelowitz, Aaron (1995): “The Medicaid Notch, Labor Supply and Welfare Participation: Evidence from Eligibility Expansions,”The Quarterly Journal of Economics, 110, 909–939.
252. Yule, George Udny (1895): “On the Correlation of Pauperism with Proportion of Out-Relief,” The Economic Journal, 5, 603–611.
253. _____ (1897) On the Theory of Correlation,”Journal of the Royal Statistical Society, 60, 812–854.
254. _____ (1899) An Investigation into the Causes of Changes in Pauperism in England, Chiefly During the Last Two Intercensal Decades (Part I.),”Journal of the Royal Statistical Society, 62, 249–295.

مرکز پژوهش‌ها
مجلس شورای اسلامی

شماره مسلسل: ۱۶۴۵۷	شناسنامه گزارش

عنوان گزارش: راهنمای اقتصاددانان تجربه‌گرا، اقتصادسنجیِ تقریباً بی‌ضرر

نام دفتر: معاونت پژوهش‌های اقتصادی

ترجمه و تدوین: گروه مترجمان

واژه‌های کلیدی:

۱. اقتصادسنجی
۲. اقتصادسنجی کاربردی
۳. Mostly Harmless Econometrics

تاریخ انتشار: ۱۳۹۸/۳/۲۸