

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \varepsilon_i \quad (1)$$

که در آن  $y$  متغیر مخارج مصرفی افراد (وابسته)،  $x_1$  متغیر توضیحی درآمد،  $x_2$  متغیر توضیحی ثروت انباشته افراد تحت بررسی،  $\beta$ ها پارامترهای مدل،  $\varepsilon_i$  جمله اخلاص و اندیس  $i$  نشان‌دهنده تعداد مشاهداتی است که از هر متغیر برای مقاطع و افراد جامعه در یک زمان خاص در دست داریم. مدل بالا را می‌توان با روش حداقل مربعات معمولی (Ordinary Least Squares) و با در نظر گرفتن فروض کلاسیک برآورد کرد.

### داده‌های سری زمانی (Time Series Data)

داده‌های سری زمانی شامل اطلاعات مربوط به مقادیر متغیرها طی یک دوره زمانی خاص است. گاهی نیز داده‌های سری زمانی در دو یا چند تناوب دوره زمانی ممکن است قابل مشاهده و جمع‌آوری باشد. برای مثال، این داده‌ها ممکن است، هم به صورت فصلی و هم سالانه وجود داشته باشد (نظیر آمار تولید ناخالص داخلی GDP در اغلب کشورها). همانند مثال مدل مقطعی، رابطه زیر به عنوان ساده‌ترین مدل سری زمانی در تحلیل‌های رگرسیونی به کار گرفته می‌شود:

$$y_t = \beta_0 + \beta_1 x_{1t} + \beta_2 x_{2t} + \varepsilon_t \quad (2)$$

که در آن  $y$  متغیر مخارج مصرفی در طول زمان (وابسته)،  $x_1$  متغیر توضیحی درآمد،  $x_2$  متغیر توضیحی ثروت انباشته افراد تحت بررسی،  $\beta$ ها پارامترهای مدل،  $\varepsilon_t$  جمله اخلاص و اندیس  $t$  نشان‌دهنده تعداد مشاهدات در طی دوره زمانی مورد بررسی است. در داده‌های سری زمانی چون مشاهدات متوالی است این مشاهدات ممکن است به یکدیگر همبسته باشند. در این صورت رگرسیون این داده‌ها ممکن است شامل مشکل خودهمبستگی و نامانایی باشند.

### داده‌های پانل و برآورد آن (Panel Data and It's Estimation)

## داده‌های پانل (داده‌های مقطعی - سری زمانی)

### Panel Data (Time Series- Cross Sectional Data)

داده‌های پانل ترکیبی از داده‌های مقطعی و زمانی است که در ادبیات اقتصادسنجی در مقایسه با استفاده از روش‌های مقطعی و سری زمانی برای انجام عمل رگرسیون دارای مزیت‌هایی است. با استفاده از داده‌های پانل می‌توان اثراتی را شناسایی و اندازه‌گیری کرد که در داده‌های مقطعی محض یا سری زمانی خالص قابل شناسایی نیست.

### داده‌های مقطعی (Cross- Sectional Data)

داده‌های مقطعی که به یک مقطع از جامعه مورد مطالعه مربوط است، با مشاهده تعداد زیادی متغیر (از جمله افراد جامعه، بنگاه‌ها، کشورها یا نواحی) در یک نقطه خاص از زمان یا قطع نظر از تفاوت در زمان گردآوری می‌شوند. به عبارت دیگر، بر اساس داده‌های مقطعی می‌توان اطلاعات متغیر در یک زمان خاص، برای مثال سال ۱۳۹۸ میلادی، را در یک جامعه آماری اندازه‌گیری کرد. برای مثال، اطلاعات مربوط به متغیر درآمد افراد گوناگون در سال ۱۳۹۸ به عنوان داده‌های مقطعی هستند.

جمع‌آوری اطلاعات به این شکل در واقع پیمایش یک مقطع از جامعه در یک زمان خاص است که در اصطلاح فنی‌تر آن را برش مقطعی (Cross- Section) می‌نامند. داده‌های مقطعی را می‌توان در رگرسیون مقطعی استفاده کرد که در واقع تحلیل رگرسیونی داده‌های مقطعی است. برای مثال، رگرسیون مخارج مصرفی افراد گوناگون در یک سال خاص را می‌توان بر روی درآمد و سطح ثروت انباشته افراد محاسبه کرد و چگونگی تفاوت رفتار مصرف‌کننده با تغییر در این متغیرهای توضیحی را تحلیل کرد.

رابطه زیر به مثابه ساده‌ترین مثال از مدل مقطعی در تحلیل‌های رگرسیونی به کار گرفته می‌شود:

در اینجا فرض می‌شود  $\eta_{it}$  جمله اخلاص تصادفی است و با  $x_{it}$ ها همبسته نیست. اما اثرات فردی را نشان می‌دهد که ممکن است با  $x_{it}$ ها همبسته باشند یا نباشند؛ به همین خاطر، روش‌های متفاوتی نیز برای برآورد آنها وجود دارد که به ترتیب به روش‌های اثرات ثابت (Effects Fixed) (Method Random Effects) و اثرات تصادفی (Method Random Effects) معروف هستند.

بنابراین، برای اینکه بتوان اثرات فردی را در مدل پانل دیتا برآورد کرد؛ یا حداقل آنها را در برآورد پارامترهای مدل در نظر گرفت؛ محقق باید از بین دو رویکرد متفاوت یکی را برگزیند. رویکرد اول اینکه؛ برای مثال در رابطه (۴) فرض شود فرد در تصمیم‌گیری خود درباره میزان مخارج مصرفی اش (که از بین بی‌نهایت تصمیم اتخاذ خواهد شد)، اثرات ناشی از عوامل حذف شده از مدل را به متغیرهای توضیحی موجود در مدل (در اینجا درآمد  $x_1$ ) و ثروت ( $x_2$ ) است) همبسته بدانند. به عبارت دیگر، در این رویکرد اثرات فردی ( $\mu_i$ ) با متغیرهای توضیحی مدل همبسته خواهد شد ( $cov(\mu_i, x) \neq 0$ ).

در رویکرد دوم نیز فرض می‌شود فرد در تصمیم‌گیری خود اثرات فردی را به‌طور مستقل از متغیرهای توضیحی ( $x$ )، در تصمیم خود دخالت می‌دهد. به عبارت دیگر، در این رویکرد اثرات فردی ( $\mu_i$ ) مستقل از متغیرهای توضیحی در نظر گرفته می‌شوند ( $cov(\mu_i, x) = 0$ ). برآورد مدل پانل دیتای (۴) با استفاده از رویکرد اول منجر به برآورد اثرات ثابت فردی و استفاده از رویکرد دوم نیز منجر به برآورد اثرات تصادفی فردی خواهد شد.

فراموش نکنیم که اثرات فردی در رابطه (۵) جزئی از جمله اخلاص است؛ حال اگر این جزء با متغیرهای توضیحی غیرتصادفی مدل ( $x$ ) همبسته باشند، در این صورت می‌توان آنها را مطابق با رابطه (۳) به شکل پارامتر ( $\beta_{i0}$ ) برای هر فرد یا مقطع تحت عنوان اثرات ثابت فردی برآورد

در داده‌های پانل متغیرها هم در میان مقاطع و گروه‌های جامعه آماری و هم در طول زمان اندازه‌گیری می‌شوند. در واقع داده‌های پانل ترکیبی از داده‌های مقطعی و سری زمانی است. برای مثال، میزان مخارج مصرفی گروه‌های جامعه (برای مثال، ساکنین به تفکیک ۳۱ استان کشور)، در طول یک فاصله زمانی مانند ۱۳۷۹ تا ۱۳۹۶ مجموعه داده‌های پانل درباره متغیر مخارج مصرفی را تشکیل می‌دهد. رابطه زیر به‌عنوان یک مدل پانل دیتا در تحلیل‌های رگرسیونی برای مثال بالا به‌کار گرفته می‌شود:

$$y_{it} = \beta_{i0} + \beta_1 x_{1it} + \beta_2 x_{2it} + \varepsilon_{it} \quad (3)$$

که در آن  $y_{it}$  متغیر مخارج مصرفی افراد در زمان‌های گوناگون (وابسته)،  $x_{1it}$  متغیر توضیحی درآمد،  $x_{2it}$  متغیر توضیحی ثروت انباشته افراد تحت بررسی در زمان‌های گوناگون،  $\beta$ ها پارامترهای مدل،  $\varepsilon_{it}$  جمله اخلاص، اندیس  $i$  نشان‌دهنده تعداد مشاهداتی است که از هر متغیر برای مقاطع و افراد جامعه در یک زمان خاص در دست داریم و اندیس  $t$  نشان‌دهنده تعداد مشاهدات در طی دوره زمانی مورد بررسی است.

پارامتر  $\beta_{i0}$  در رابطه بالا، نشان‌دهنده پارامتر اثرات فردی بر روی متغیر وابسته مدل است. یکی از مهم‌ترین و جذاب‌ترین نقاط قوت مدل پانل دیتا در برازش رگرسیون، توانایی برآورد اثرات فردی یا در نظر گرفتن این اثرات در تخمین پارامترهای مدل است؛ به‌طوری که این تخمین‌ها بدون تورش و سازگار خواهند بود. برای در نظر گرفتن اثرات فردی می‌توان رابطه (۳) را به‌صورت زیر نوشت:

$$y_{it} = \beta_0 + \beta_1 x_{1it} + \beta_2 x_{2it} + \varepsilon_{it} \quad (4)$$

که در آن ساختار زیر برای جمله اخلاص  $\varepsilon_{it}$  در نظر گرفته می‌شود:

$$\varepsilon_{it} = \mu_i + \eta_{it} \quad (5)$$

Wooldridge, J.M., (2010). *Econometric Analysis of Cross Section and Panel Data*, 2<sup>th</sup> edition, Cambridge: MIT Press, MA.

نادر مهرگان

هیئت علمی دانشگاه بوعلی سینا

یونس تیموری

سازمان برنامه و بودجه کشور

کرد. اما اگر برآورد مدل پانل دیتای (۴) با استفاده از رویکرد دوم بخواهد انجام گیرد؛ در این صورت چون اثرات فردی با  $\alpha$ ها رابطه همبستگی ندارد این اثرات جزء تصادفی جمله اختلال خواهد بود. بنابراین، باید یک توزیع آماری برای اثرات فردی  $\mu_i$  در جمله اختلال در نظر گرفته شود. برای این کار باید از جامعه آماری (جامعه افراد و مقاطع) یک نمونه برداری با حجم  $N$  انتخاب کنیم (یعنی به طور تصادفی  $N$  فرد را از جامعه انتخاب کنیم).

#### کتاب‌شناسی

اشرف‌زاده، حمیدرضا و مهرگان، نادر (۱۳۸۷). *اقتصادسنجی پانل دیتا*، چاپ اول، تهران: دانشگاه تهران، انتشارات مؤسسه تحقیقات تعاون. بالتاجی، بادی (۱۳۹۸). *تحلیل اقتصادسنجی پانل دیتا*، ترجمه محمد نصرافهانی و احسان رسولی‌نژاد، چاپ اول، همدان: انتشارات نور علم.

گجراتی، دامودار (۱۳۹۴). *اقتصادسنجی کاربردی*، ترجمه نادر مهرگان و لطفعلی عاقلی، چاپ سوم، همدان: انتشارات نورعلم.

گجراتی، دامودار (۱۳۹۷). *مبانی اقتصادسنجی*، ترجمه حمید ابریشمی، چاپ پانزدهم، تهران: انتشارات دانشگاه تهران.

Arellano M, Bonhomme S. (2011). *Nonlinear panel data analysis*, Annu., Rev., Econ. 3: 395-424

Arellano M, Bonhomme S. (2017). *Nonlinear panel data methods for dynamic heterogeneous agent models*, Annu., Rev., Econ. 9.

Boneva L, Linton O. (2017). "A discrete-choice model for large heterogeneous panels with interactive fixed effects with an application to the determinants of corporate bond issuance", *Journal of Applied Econometrics*, Vol. 32, Issue. 7, December 2017, pp. 1226-1243.

Davidson, R. & MacKinnon, J.G. (2004). *Econometric Theory and Methods*, Oxford University Press.

Greene, W.H. (2005). *Econometric Analysis*, 5<sup>th</sup> Edition, New Jersey, Prentice Hall International Inc.

Hsiao, C. (2002). *Analysis of Panel Data*, Second Edition, Cambridge University Press.

Matyas, L. (2017). *The Econometrics of Multi-dimensional Panels; Theory and Application*, Springer International Publishing, Vol 50, pp. 456.

Simonetta . L.A (???). *Practical Guide to Using Panel Data*.

Tsounis, N. & Vlachvei, A. (2018). *Advances in Panel Data Analysis in Applied Economic Research*, Springer International Publishing, pp. 705.