



TÜRKİYE CUMHURİYETİ CUMHURBAŞKANLIĞI
STRATEJİ VE BÜTÇE BAŞKANLIĞI

DEVLET YARDIMLARI ETKİ DEĞERLENDİRME REHBERİ



Devlet Yardımları Genel Müdürlüğü
Ankara, 2026

İÇİNDEKİLER

GRAFİK, TABLO, ŞEKİL VE KUTULAR.....	iii
KISALTMALAR.....	v
TANIMLAR.....	vi
GİRİŞ	1

1. ETKİ DEĞERLENDİRME	3
1.1. Etki Değerlendirme Kavramı	3
1.2. Etki Değerlendirmenin Zamanı	3
1.3. Etki Değerlendirmenin Faydaları	4
1.4. Etki Değerlendirmenin Gereklilikleri	4
1.5. Etki Değerlendirme Tasarımı	6
2. ETKİ DEĞERLENDİRME İÇİN VERİ YÖNETİMİ.....	8
2.1. Etki Değerlendirmede Verinin Önemi	8
2.2. Veri ve Veri Kaynakları	9
2.3. Etki Değerlendirmede Kullanılacak Verilerde Bulunması Gereken Özellikler	9
2.4. Etki Değerlendirmede Kullanılacak Verilerin Belirlenmesi	10
2.5. Örneklem ve Örneklem Büyüklüğü	11
2.6. Veri Toplama Stratejileri ve Eksik Veri Yönetimi	13
2.7. Verinin Korunması ve Gizlilik	14
3. NİCEL ETKİ DEĞERLENDİRME YÖNTEMLERİ	15
3.1. Nicel Etki Değerlendirmede İzlenecek Adımlar	15
3.2. Farkların Farkı Yöntemi	27
3.2.1. Farkların Farkı Yönteminde İzlenmesi Gereken Adımlar	31
3.2.2. Hedef Kitlenin Tamamının Yardımdan Yararlandığı Durumda Farkların Farkı Yönteminin Uygulanması	42
3.3. Sentetik Kontrol Yöntemi	50
3.4. Eğilim Skoru Eşleştirme Yöntemi	63
3.5. Araç Değişkenler Yöntemi	77
3.6. Regresyon Süreksizliği Tasarımı Yöntemi	79
3.7. Ters Olasılıklı Ağırlıklandırma Yöntemi	82
3.8. Rassal Kontrollü Deneyler	84
4. R PROGRAMI İLE NİCEL ETKİ DEĞERLENDİRME VE UYGULAMA ÖRNEKLERİ	86
4.1. R Programlama Dili (R Studio)	86
4.2. Etki Değerlendirme İçin Gerekli Adımlar	87
4.3. Uygulama Örnekleri	91
Örnek 1. Farkların Farkı	91
Örnek 2. Ters Olasılıklı Ağırlıklandırma Yöntemi	109
Örnek 3. Regresyon Süreksizliği Tasarımı	115

5. 102 SAYILI CUMHURBAŞKANLIĞI KARARNAMESİ KAPSAMINDA DEVLET YARDIMLARI HAKKINDA ETKİ DEĞERLENDİRMESİ	124
5.1. Ön Etki Değerlendirme.....	124
5.2. Uygulama Dönemi Etki Değerlendirme.....	132
5.3. Uygulama Sonrası Etki Değerlendirme.....	133
SONUÇ.....	135
KAYNAKÇA	136
EKLER.....	145
EK 1. Farkların Farkı Yöntemini Uygulayacak Araştırmacılar için Kontrol Listesi.....	145
EK 2. Nicel Etki Değerlendirme Karar Ağacı	148
EK 3. Değerlendirme Tasarımı için Karar Ağacı Süreci	149

GRAFİK, TABLO, ŞEKİL VE KUTULAR

Grafik 1. Yardımın Uygulanmasıyla Ortaya Çıkan Değişim	5
Grafik 2. Müdahale Sonrası Etki Durumu (Farkların Farkı Yöntemi için)	29
Grafik 3. Eşleştirme Sonrası Denge Kontrol Grafiği.....	39
Grafik 4. Yarı İşlenmemiş Grup Mevcudiyetinde AOE Doz -Yanıt Eğrisi.....	46
Grafik 5. YİG Mevcut Olmadığında AOEdmin Doz - Yanıt Eğrisi.....	47
Grafik 6. YİG ve Dmin Referans Gruplarında Gözlemlenen Etki Düzeyleri.....	48
Grafik 7. Sigara Satışlarındaki Eğilim: Kaliforniya ve Diğer Eyaletler	52
Grafik 8. Sigara Satışlarındaki Eğilim:Kaliforniya ve Sentetik Kaliforniya	53
Grafik 9. Sigara Satışlarındaki Açık (Kaliforniya ve 38 Kontrol Eyaleti).....	54
Grafik 10. Sigara Satışlarındaki Açık (Kaliforniya ve 34 Kontrol Eyaleti).....	55
Grafik 11. Öneri 99 MSPE Oranı: Kaliforniya ve 38 Kontrol Eyaleti	55
Grafik 12. Müdahale ve Kontrol Grubu Gelir Eğilimi	57
Grafik 13. Müdahale Sonrası Gelir Farkları	57
Grafik 14. Sentetik Kontrol Yöntemi vs. Sentetik Farkların Farkı Yöntemi Etkisi.....	61
Grafik 15. “Öneri 99” Üç Yöntem Sonrası Sonuçları.....	62
Grafik 16. ROC (Receiver Operating Characteristic) Eğrisi	70
Grafik 17. Ortak Destek Alanı (Common Support Area)	71
Grafik 18. Regresyon Süreksizliği.....	80
Grafik 19. Sıçrama Noktası	81
Grafik 20. Örnek 1 - Eksik Değerlerin Kontrolü	95
Grafik 21. Örnek 1 - Yıllara Göre Gelir Dağılımı	102
Grafik 22. Örnek 1 - Paralel Eğilim Grafiği	104
Grafik 23. Örnek 2 - Saçılım Grafiği.....	112
Grafik 24. Örnek 3 - Dağılım Nokta Grafiği	117
Grafik 25. Örnek 3 - Yoğunluk Grafiği.....	120
Grafik 26. Örnek 3 - Süreksizliğin Tespiti.....	121
Tablo 1. Evren Büyüklüğüne Bağlı Olarak Etki Değerlendirme İçin Seçilebilecek Örneklem Büyüklükleri	12
Tablo 2. Devlet Yardımı ve Etkileri Tablosu.....	15
Tablo 3. Anket Özellikleri.....	18
Tablo 4. OECD DAC Kriterleri	22
Tablo 5. Örnek Veri Formatı	33
Tablo 6. Heterojenliği Benimseme Tasarımlarında Etki Değerlendirme Yöntemleri.....	49
Tablo 7. Sigara Satışı Tahmin Edicilerinin Ortalaması.....	53
Tablo 8. Araştırma Çalışması - Grup Ortalamaları	57
Tablo 9. Elde Edilen Bulgular.....	62
Tablo 10. Mahalonobis Mesafesi İle Eşleştirme Yöntemi Özet Tablo.....	72
Tablo 11. Kaliperli En Yakın Komşu Eşleştirme Yöntemi Özet Tablo.....	72
Tablo 12. Yarıçap (Radius) Eşleştirme Yöntemi Özet Tablo	73
Tablo 13. Kernel Eşleştirme Yöntemi Özet Tablo	73
Tablo 14. Optimal Eşleştirme Yöntemi Özet Tablo	74
Tablo 15. Tam Eşleştirme Yöntemi Özet Tablo	74
Tablo 16. Sistematik Basitleştirilmiş Tam Eşleştirme Yöntemi Özet Tablo	75

Tablo 17. Örnek 3- “İş Sağlığı ve Güvenliği” Analiz Değişkenleri	116
Tablo 18. Örnek 3- Regresyon Süreksizliği Tasarımının Tahmin Sonuçları	123
Tablo 19. Devlet Yardımı ve Etkileri Tablosu.....	127
Tablo 20. Devlet Yardımının Performans Göstergeleri.....	128
Şekil 1. Farkların Farkı Yöntemi Kıyas Şeması	43
Şekil 2. Heterojenliği Benimseme Tasarımlarında Referans Grubu Tespit Şeması.....	44
Şekil 3. Heterojenliği Benimseme Tasarımlarında Etki Değerlendirme Yöntemi Karar Ağacı.....	49
Şekil 4. Sentetik Kontrol Grubu Oluşturma Adımları	59
Şekil 5. Birebir Eşleştirme ile Destek Alanında Olan ve Olmayan Gözlemler	64
Şekil 6. Müdahale, Karıştırıcı Değişken ve Çıktı Arasındaki İlişki.....	67
Şekil 7. Rassal Kontrollü Deneyler Yöntemi.....	85
Şekil 8. Değerlendirme Kriterleri	125
Şekil 9. Performans Göstergelerinde Olması Gereken Özellikler	128
Kutu 1. Bağımlı ve Bağımsız Değişken Türlerine Göre Örnekler.....	11
Kutu 2. Asgari Ücret Değişiminin İstihdam Üzerindeki Etkisinin Ölçümü	30
Kutu 3. Sentetik Kontrol Yöntemi Örneği.....	57
Kutu 4. Tarım ve Gıda Firmaları Üzerinde AB Katılım Öncesi Fonlarına İlişkin Etki Değerlendirme: Hırvatistan Örneği	65
Kutu 5. Kontrol ve Müdahale Grubunun Belirlenmesi.....	88
Kutu 6. Pipe Operatörü	94

KISALTMALAR

AB	Avrupa Birliđi
ABD	Amerika Birleşik Devletleri
AD	Araç Deđişken
Ar-Ge	Araştırma ve Geliştirme
ATE	Ortalama Müdahale Etkisi (Average Treatment Effect)
CPP	Toplumsal Planlama Ortaklığı (Community Planning Partnership)
DAC	Kalkınma Yardımı Komitesi (Development Assistance Committee)
EKK	En Küçük Kareler Yöntemi (Ordinary Least Squares)
FF	Farkların Farkı (Difference-in-Differences)
HBT	Heterojenliği Benimseme Tasarımları
ILO	Uluslararası Çalışma Örgütü (International Labour Organization)
IPW	Ters Olasılıklı Ağırlıklandırma (Inverse Probability Weighting)
LATE	Yerel Ortalama Müdahale Etkisi (Local Average Treatment Effect)
MSPE	Ortalama Kare Tahmin Hatası (Mean Squared Prediction Error)
QuIP	Nitel Etki Deđerlendirme Protokolü (Qualitative Impact Assessment Protocol)
OECD	Ekonomik Kalkınma ve İşbirliği Örgütü (Organisation for Economic Co-operation and Development)
RCT	Rassal Kontrollü Deneyler (Randomized Controlled Trial)
RST	Regresyon Süreksizliği Tasarımı (Regression Discontinuity Design)
STK	Sivil Toplum Kuruluşu
ToC	Deđişim Teorisi (Theory of Change)
TWFE	İki Yönlü Sabit Etkiler (Two Way Fixed Effects)
UNDP	Birleşmiş Milletler Kalkınma Programı (United Nations Development Programme)
UNEG	Birleşmiş Milletler Deđerlendirme Grubu (The United Nations Evaluation Group)
USAID	Amerika Birleşik Devletleri Uluslararası Kalkınma Ajansı (United States Agency for International Development)
YİG	Yarı İşlenmemiş Grup

TANIMLAR

Başkanlık Kararname	T.C. Cumhurbaşkanlığı Strateji ve Bütçe Başkanlığı 102 sayılı Devlet Yardımlarının Uygulanması, Koordinasyonu, İzlenmesi ve Değerlendirilmesi Hakkında Cumhurbaşkanlığı Kararnamesi
Uygulayıcı Kurum R	Devlet yardımı uygulamasını doğrudan veya dolaylı olarak yürütmekle yetkili ve sorumlu kamu kurum ve kuruluş Veri analizi ve araştırmalarda verileri temizlemek, analiz etmek, görselleştirmek ve anlamlı hale getirmek için yaygın olarak kullanılan açık kaynak bir programlama dili
Yardım	Kamu kurum ve kuruluşu tarafından uygulanan ve Kararname hükümlerine tabi devlet yardımları
Yönetmelik	7023 sayılı Devlet Yardımlarının Uygulanması, Koordinasyonu, İzlenmesi ve Değerlendirilmesine İlişkin Usul ve Esaslar Hakkında Yönetmelik

GİRİŞ

Günümüzde ülkeler, gelişmişlik düzeyine göre ekonomik ve sosyal alanlarda çeşitli amaçlar gütmektedir. Gelişmiş ülkeler istikrarı ve rekabet gücünü korumaya odaklanırken, az gelişmiş ve gelişmekte olan ülkeler ise kalkınmayı amaçlamaktadır. Kalkınma, yalnızca üretimin ve kişi başına düşen milli gelirin artırılmasından ibaret olmayıp gelir dağılımındaki eşitsizliklerin giderilmesi, bölgelerarası gelişmişlik farklarının azaltılması, kısaca ekonomik ve sosyo-kültürel yapının iyileştirilmesi anlamına gelmektedir (Ayanoğlu ve diğerleri, 1996). Bu kapsamda hükümetler tarafından kalkınma planları hazırlanmakta ve kalkınma planlarında yer alan hedeflere ulaşmak için piyasalara politika, plan, program ve/veya projeler ile müdahalelerde bulunmaktadır.

Bu müdahaleler rekabet gücünü artırmak, piyasalardaki birtakım aksaklıkları gidermek, sektörleri yönlendirmek veya özendirmek gibi çeşitli amaçlarla yapılmaktadır. Müdahale uygulanmadan önce potansiyel olarak ümit verici görünürken uygulamaya konulduktan sonra beklenen etki veya faydayı göstermeyebilir. Günümüze değin yaşanan tecrübeler, hükümetlerin bu amaçlarla yaptıkları müdahalelerin zaman zaman etkisiz olduğunu, hedeflere ulaşılması için yetersiz kaldığını, bazı durumlarda ise öngörülmeleyen başka sorunlara yol açabildiğini göstermiştir.

Bu tecrübeler, hükümetler tarafından piyasalara müdahalelerin daha iyi planlanması ve kurgulanması gereğini ortaya çıkarmıştır. Uygulaması planlanan müdahalelerin iyi bir şekilde tasarlanması ve hedeflenen etkileri doğurması için ihtiyaçların, bu ihtiyaçları doğuran nedenlerin, ilgili kesimlerin ve ihtiyacı karşılayacak müdahale kapsamının doğru bir şekilde tespit edilmesi gerekmektedir. Müdahale uygulanmaya başladıktan sonra planlandığı şekilde hayata geçirilip geçirilmediğinin izlenmesi ve etkilerinin ölçülmesi ise ihtiyaçların giderilmesi ve kamu kaynaklarının etkin kullanımı açısından zorunluluktur. Böylece başarılı olan müdahaleler, ihtiyaç karşılanana kadar devam edebilecek ya da başka ihtiyaç alanlarında uygunsa kullanılabilecek, başarısız olanlar ise gözden geçirilip iyileştirilecek veya durdurulacaktır.

Bu kapsamda ortaya çıkan etki değerlendirme kavramı, sadece uygulanan bir müdahalenin sonuçlarını ve etkilerini ölçmeyi değil, planlama aşamasından itibaren müdahaleye ihtiyaç duyan kesimlerin ve ihtiyaçların doğru bir şekilde tespit edilmesini, uygulanabilir bir müdahale geliştirilmesini ve müdahalenin ilgililik, tutarlılık, etkililik, etkinlik, etki ve sürdürülebilirlik kriterleri açısından bütüncül değerlendirilmesini sağlayan bir yaklaşımdır¹.

Devlet yardımları hükümetlerin uzun yıllardır kullandığı müdahalelerden biridir. Her geçen yıl daha çok ülke ve uluslararası kuruluş uyguladıkları yardımın etkisini değerlendirme, böylece kaynakları doğru yerlere yönlendirerek daha başarılı sonuçlar elde etme gayreti içerisine girmiştir (OECD, 2021).

¹ Bu rehberde kullanılan “Etki Değerlendirme” kavramı “Düzenleyici Etki Analizi” kavramından farklıdır. Düzenleyici Etki Analizi, kanun ve Cumhurbaşkanlığı kararnamesi taslaklarının bütçeye, mevzuata, sosyal, ekonomik ve ticari hayata, çevreye ve ilgili kesimlere etkilerini göstermek üzere yapılan ön değerlendirme olup ayrı mevzuatla düzenlenmiştir. Detaylar için bakınız <https://www.sbb.gov.tr/wp-content/uploads/2023/03/Duzenleyici-Etki-Analizi-Rehberi.pdf>

Ülkemizde 9/6/2022 tarihli ve 31861 sayılı Resmi Gazete’de yayımlanan Devlet Yardımlarının Uygulanması, Koordinasyonu, İzlenmesi ve Değerlendirilmesi Hakkında 102 sayılı Cumhurbaşkanlığı Kararnamesinde devlet yardımı şu şekilde tanımlanmıştır:

“Sürdürülebilir ve rekabetçi kalkınma amacıyla kamu kaynağı kullanılarak, belirli bir gerçek veya tüzel kişiye veyahut gerçek ve/veya tüzel kişi grubuna, normal piyasa şartlarında elde edemeyeceği bir ekonomik fayda sağlayan; hibe, kredi, taşınır/taşınmaz mal tahsisi, kamu gelir veya alacağından vazgeçme, sermaye katılımı, aynı destek gibi araçlarla gerçekleştirilen uygulamalardır.”

Anılan Kararnameye göre; kamu kurum ve kuruluşları tarafından ilk defa yürürlüğe konulması planlanan yardımlar için teklif formu ve beraberinde Ön Etki Değerlendirme Raporu, değerlendirilmek üzere Başkanlığa iletilmelidir.

Ayrıca, devam eden yardımlar için gerçekleştirildiği takvim yılına ait İzleme Raporu ile beraberinde Uygulama Dönemi Etki Değerlendirme Raporu ya da uygulaması tamamen sona eren yardımlar için Uygulama Sonrası Etki Değerlendirme Raporu bir sonraki yılın Mart ayı sonuna kadar Başkanlığa değerlendirilmek üzere iletilmesi gereken belgelerdendir.

Bu Rehber, Kararname’nin ilgili maddeleri doğrultusunda yukarıda adı geçen etki değerlendirme raporlarının hazırlanmasında kamu kurum ve kuruluşlarına yardımcı olmak amacıyla hazırlanmıştır. Rehberin ilk bölümünde etki değerlendirmenin tanımı, gerekliliği ve tasarımı gibi ana özellikleri açıklanmaktadır.

İkinci bölümde veri ve veri toplama araçlarına değinilmektedir.

Üçüncü bölümde, etki değerlendirme yöntemleri sınıflandırılmakta ve nicel etki değerlendirme yöntemlerinde izlenmesi gereken adımlar belirtilmekte ve her bir yöntem detaylı bir şekilde açıklanmaktadır.

Rehberin dördüncü bölümünde en çok kullanılan nicel etki değerlendirme yöntemlerinin daha iyi anlaşılabilmesi amacıyla R istatistik programında uygulama örneklerine yer verilmektedir.

Rehberin son bölümü olan beşinci bölümde Kararname’nin uygulanmasına dair yayımlanan Yönetmelik çerçevesinde devlet yardımlarının uygulanma aşamasına göre hazırlanması gereken etki değerlendirme raporlarına ilişkin bilgiler yer almaktadır.

Rehber hazırlanırken ulusal ve uluslararası yazından yararlanılmıştır. Bu kapsamda Dünya Bankası, Birleşmiş Milletler Kuruluşları, Avrupa Komisyonu gibi uluslararası kuruluşların hazırlamış oldukları rehber ve kılavuzlar, akademik ve görsel çalışmalar incelenmiştir.

1. ETKİ DEĞERLENDİRME

1.1. Etki Değerlendirme Kavramı

Etki değerlendirme konusunda yazında birçok tanım yer almaktadır. Esasında bu tanımların odağı aynı olmakla birlikte, ifade farklılıkları bulunmaktadır. Örneğin, etki değerlendirmeyi OECD (2009) politika müdahalelerinin yararlanıcıların refahını doğrudan veya dolaylı olarak nasıl etkilediği ve refahtaki değişimlerin ne ölçüde bu müdahalelere bağlanacağını araştırması; Khandker ve diğerleri (2010) refahtaki değişikliklerin gerçekten müdahalenin uygulanmasından kaynaklanıp kaynaklanmadığının araştırılması; Glewwe ve Todd (2022) ise bir projenin, programın veya politikanın hükümetlerin ve diğer ilgili tarafların ilgilendikleri bir sonuç üzerindeki etkisini ölçmeye çalışan bir araştırma olarak tanımlamıştır.

Yapılan tanımlardan anlaşılacağı üzere etki değerlendirme, yardımlar açısından ele alındığında, yardımın yararlanıcı üzerindeki olası etkisine odaklanmış olup bunu belirlemeye çalışmaktadır. Böylelikle yardımlar ve beklenen sonuçlar arasında neden - sonuç (nedensellik) ilişkisi incelenerek asıl etkiye dair yanıtlar bulunur.

Özetle, verimli kaynak kullanımının ve kamu mali sorumluluğunun öneminin gittikçe artması, nitel ve/veya nicel değerlendirme ile elde edilen sonuçlara dayanarak yardımların geliştirilmesi, değiştirilmesi ya da kaldırılması, diğer şehir ve/veya ülkelerdeki yardım programlarının mukayese edilebilmesi ve uygulanan politikalar için ayrılan kamu kaynaklarının etkin kullanılmasının takibi gibi gerekçelerle etki değerlendirme yapılmaktadır.

1.2. Etki Değerlendirmenin Zamanı

Değerlendirmenin zamanı ile amacı yakın ilişki içerisindedir. Bu sebeple yapılacak etki değerlendirme zamanı üç kategoriye ayrılmakta ve her birinin yapılaş amacı farklılık göstermektedir.

Uygulama öncesinde yapılan etki değerlendirme, yazında ön etki değerlendirme (ex-ante) olarak yer almaktadır. Ön etki değerlendirmenin amacı, tasarım aşamasında programın uygulanabilirliğini ve potansiyel etkisini tespit etmektir. Bu sebeple yardım uygulanmadan öncesinde yapılacak etki değerlendirme, ihtiyaçları, beklentileri ve karşılaşılması muhtemel sorunları görmek ve programı tasarlamak için önemli bir fırsat sunmaktadır.

Uygulama dönemi etki değerlendirme (interim) ise uygulayıcılara program başladıktan sonra programın doğru çalışıp çalışmadığına, hedeflenen performans göstergeleri üzerinden başarıya ulaşıp ulaşılamayacağına ilişkin önemli bilgiler sunmaktadır. Yardımın ara değerlendirmeler ile birlikte düzenli ve sürekli bir şekilde izlenmesi, uygulamada aksayan yönlerin tespit edilmesini sağlamakla birlikte yararlanıcıların durum hakkındaki farkındalıkları, gözlemleri ve davranışsal olarak değişiklikleri ile ortaya çıkmış olan hızlı etkileri görmek açısından önem taşımaktadır.

Yardım tamamlandıktan belli bir süre sonra yapılan etki değerlendirme ise, uygulama sonrası etki değerlendirme (ex-post) olarak ifade edilmekte olup uygulanmış olan yardımın kısa, orta veya uzun dönemde oluşturduğu etkiyi tespit etmeye imkân tanımaktadır. Yardım tamamlandıktan sonra yapılan değerlendirmeler, tasarım aşamasında belirlenen hedeflerin ne ölçüde gerçekleştirildiğini, yararlanıcılar ve/veya toplumun tamamı üzerinde yaratılan beklenen ve/veya beklenmeyen etkileri, uluslararası değerlendirme kıstasları çerçevesinde ölçmeye odaklanmaktadır.

1.3. Etki Değerlendirmenin Faydaları

Etki değerlendirme, bir müdahalenin başlatılması, geliştirilmesi, yönünün değiştirilmesi, devam ettirilmesi, sonlandırılması, büyütülmesi veya başka bir alanda uygulanması gibi önemli kararlar verilebilmesi için yapılabilmektedir. Etki değerlendirme, uygun bir şekilde yapıldığında birçok fayda sağlamaktadır. Bunlardan bazıları şunlardır (Glewwe ve Todd, 2022):

- Uygulanan yardım, büyük miktarda kamu kaynağı gerektirmekte ve belirli bir ihtiyacı karşılamayı amaçlamaktadır. Başarısız olması durumunda, sadece kamu kaynağının israf edilmesiyle kalmayacak, ihtiyacın karşılanmamasıyla sonuçlanacak ve aynı zamanda kamu kaynağının kullanılabileceği başka alandaki ihtiyaçların karşılanmasının ötelenmesine neden olabilecektir. Etki değerlendirme yapılarak politika, program veya projenin gerçekten ihtiyacı karşılayıp karşılamadığı anlaşılarak uygulamanın devam ettirilmesi, revize edilerek devam ettirilmesi veya sonlandırılması yönünde karar alınabilecektir.
- Etki değerlendirme, hükümetlerin vatandaşlara karşı şeffaf ve hesap verebilir olmasını sağlamaktadır. Bu durum aynı zamanda kamu kaynağının etkin kullanıldığının vatandaşlar tarafından da anlaşılmasını kolaylaştırmaktadır.
- Bunların dışında, son yıllarda gerek akademik camiada yapılan araştırmalar, gerek uluslararası kuruluşlar tarafından sağlanan fonlar ve uygulanan programlar etki değerlendirme yapılmasının önemine vurgu yapmaktadır. Günümüzde kanıta dayalı olarak politika, program ve projeler uygulanması gerektiği yönünde uluslararası ortak görüş oluşmuş olup etki değerlendirme bu kapsamda ön plana çıkan önemli mekanizmalardan biridir.

1.4. Etki Değerlendirmenin Gereklilikleri

Bir müdahalenin yalnızca o müdahaleye ilişkin sonuçlardaki değişikliğe ne ölçüde katkıda bulunduğunu ölçerek atfetme (attribution)² durumu irdelenmelidir. Başka bir ifadeyle, uygulanan müdahale ile ortaya çıkması beklenen değişikliğin söz konusu müdahaleye atfedilip atfedilemeyeceği, atfedilecekse ne ölçüde ve neden atfedilebileceği değerlendirmeye çalışılmaktadır.

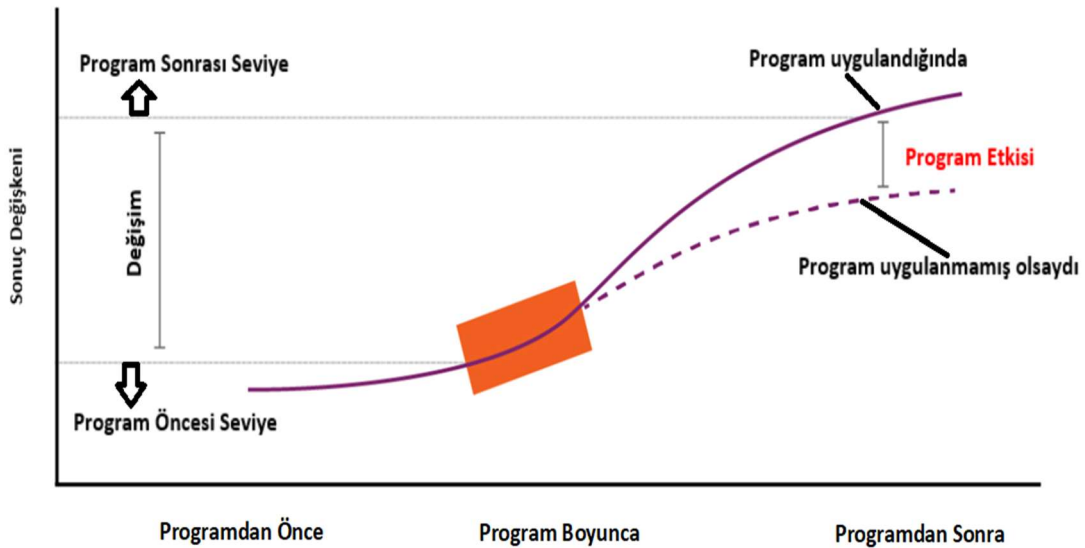
Örneğin, bir yardım kapsamında uygulanan iş danışmanlığı programı, “işverenler arasında memnuniyet düzeyini yükselterek işin devamlılığını sağlıyor mu?” ya da “start-

² Atfetme: Gözlenen (veya gözlemlenmesi beklenen) değişiklikler ile belirli bir yardım (müdahale) arasındaki nedensel bağlantının ilişkilendirilmesi.

up mentörlük programı iş alanı yaratmayı ve sürdürülebilirliği destekliyor mu?” gibi araştırma sorularının ilgilendiği işgücü piyasası çıktıları, sosyal ve ekonomik gelişme bağlamı, siyasi kişisel ve/veya bireysel koşullardaki değişiklikler gibi birçok karmaşık etken tarafından belirlenmektedir. Bu nedenle, bu tür çıktılardaki değişikliklerin ne dereceye kadar belirli bir müdahaleye atfedilebileceğini ortaya koymak zordur.

Bu sebeple yazında ağırlıklı olarak nicel etki değerlendirmenin dayandırıldığı sözde karşı-olgusalılık (counterfactual) çerçevesi üzerinden hareket edilerek atfetme durumu değerlendirilmektedir. Karşı-olgusalılık, müdahale uygulanmamış olsaydı ilgilenilen sonucun ne olacağını açıklamak için kullanılmaktadır. Ancak karşı-olgusalılık doğrudan gözlemlenemediğinden, kontrol grupları³ kullanılarak etki değerlendirme yapılmaktadır. Ayrıca yararlanıcılarda gözlemlenen sonuçların eğer yardım uygulanmamış olsaydı, nasıl olacağı konusunda sadece yararlanıcılara ait müdahale grubu⁴ verisini inceleyerek yanıt bulunamayacaktır. Bu sebeple eksik olan veriyi tamamlamak amacıyla yararlanıcılar için benzer veya makul bir kontrol grubu oluşturulur.

Grafik 1. Yardımın Uygulanmasıyla Ortaya Çıkan Değişim



İdeal olarak yararlanıcı yardımdan faydalanmadan önce ve faydalandıktan sonra aradaki farkın tamamen yardımdan kaynaklandığının gösterilmesi beklenir. Ancak aynı yararlanıcı hem müdahale grubunda hem de kontrol grubunda olamayacağı için bu yapılamaz. Bu sebeple müdahale ve kontrol gruplarının aşağıdaki en az üç karakteristik özelliği benzer olduğunda, iki grup arasındaki sonuçlarda görülen herhangi bir farklılığın uygulanan müdahaleye atfedilebileceği rahatlıkla söylenebilir. Bu özellikler:

- Hem gözlemlenebilir hem de gözlemlenemez özellikler açısından benzer olmalıdır: Gözlemlenebilir özellikler⁵ yaş, cinsiyet, eğitim düzeyi, sosyo-

³ Kontrol grubu, etki değerlendirmede karşı-olgusal tahmin etmek için kullanılan gruptur. Müdahale grubu üyelerinin aksine, kontrol grubunun üyeleri, değerlendirmek istediğimiz müdahaleye maruz kalmamıştır.

⁴ Bir müdahalede aktif olarak yer alan insan grubu, müdahale grubu, katılımcı grubu, yararlanıcı ya da hedeflenen kesim olarak bilinir. Diğer bir deyişle yardımdan yararlanan kişilerin oluşturduğu grubu temsil etmektedir.

⁵ Uygun veri toplama yöntemleriyle (Örn. anketler gibi) ölçülebilir. Genellikle yaş, cinsiyet, eğitim düzeyi, sosyo-ekonomik durum, aile özellikleri, çalışma durumu vb.

ekonomik durum, aile özellikleri, çalışma durumu ve benzerlerini içerebilir. Gözlemlenemeyen özellikler⁶, diğer etkenlerin yanı sıra motivasyon, ilgi, değerler ve ideolojileri içerebilir. Müdahale grubundaki her kişi, kontrol grubundaki her kişiyle birebir aynı olmak zorunda değildir, ancak her iki grup da benzer ortalama özelliklere sahip olmalıdır.

- Müdahale ve kontrol gruplarının müdahaleye benzer şekilde tepki vermesi beklenir: Örneğin, satış geliri gibi bir etkenin kontrol grubundakiler kadar müdahale grubu üyeleri için de artma olasılığı olmalıdır.
- Müdahale ve kontrol grupları, diğer müdahalelere benzer düzeyde maruz kalmalıdır: Örneğin, her iki grup da belediyeler, STK'lar vb. tarafından sağlanan diğer destek hizmetlerine aynı şekilde erişim imkânına sahip olmalıdır.

Müdahale grubunun kontrol grubundan karakteristiği bakımından farklı olduğu durumlarda seçim yanlılığı oluşur⁷. Başka bir deyişle, kontrol grubu müdahale grubundan gözlemlenebilir özellikleri bakımından önemli ölçüde farklıysa, müdahale grubu ve kontrol grubu arasındaki sonuçların karşılaştırılması yalnızca müdahalenin etkisini değil, aynı zamanda bu farklılıkların sonuçlarını da içerecektir. Söz konusu bu yanlılık, bir yardımın olduğundan daha başarılı görünmesine veya olması gerekenden daha az başarılı görünmesine neden olabilir. Uygun etki değerlendirme yöntemi kullanılıp, konuyla ilgili değişkenler dikkate alınarak değerlendirme yapılması bu sorunun çözümünü sağlamaktadır. Bu kapsamda kullanılabilecek yöntemler ve dikkat edilecek hususlar ilerleyen bölümlerde açıklanmaktadır.

Son olarak etki değerlendirmenin gerekliliklerinden biri de analiz edilecek uygun verinin toplanmasıdır. Kullanılacak veri, yardımı tasarlarırken belirlenmeli ve söz konusu verinin yararlanıcılar ve kontrol grubu için toplanabileceğinden emin olunmalıdır. Kullanılması planlanan verinin, yardım uygulanmaya başladıktan sonra toplanamaması durumunda etki değerlendirme yapılması çok zor hatta imkânsız bir hal alabilir.

1.5. Etki Değerlendirme Tasarımı

Başarılı bir etki değerlendirme yapılabilmesi için sürecin ve süreç içerisinde atılacak her adımın iyi planlanması ve uygulanması önem taşımaktadır. Etki değerlendirme sürecinde aşağıdaki adımların izlenmesinde yarar bulunmaktadır:

- Başlangıç ve ekibin oluşturulması
- Paydaşların belirlenmesi
- Paydaşlarla istişare
- Etki değerlendirme yönteminin belirlenmesi
- Veri toplama araçlarının belirlenmesi
- Araştırmacıların eğitimi
- Örneklem seçimi

⁶ Bir (etki) değerlendirmede ölçülemeyen veya ölçülmeyen değişkenlerdir ve motivasyon, ilgi, değerler ve ideolojiler vb. konuları içerebilir.

⁷ Seçim yanlılığı, bir araştırmada örnekleme veya analize dahil olma olasılığının, incelenen maruz kalma ya da sonuç değişkeniyle ilişkili olması nedeniyle, örneklemin hedef evreni temsil etmemesi ve bu durumun etki tahminlerini sistematik biçimde çarpıtmasıdır.

- Verinin toplanması
- Verinin analizi
- Ön rapor
- Bulgular hakkında paydaşlardan geri bildirim alınması
- Nihai rapor ve karar verme

Bu adımlardan veri toplama araçlarının belirlenmesi, araştırmacıların eğitimi, örneklem seçimi, verinin toplanması ve veri analizi başlıklı adımlar, verinin durumuna ve uygulanacak olan etki değerlendirme yöntemine göre farklılık göstermektedir. Etki değerlendirmede kullanılan yöntemler ve izlenen adımlar bir sonraki bölümde detaylı bir şekilde açıklanmaktadır.

2. ETKİ DEĞERLENDİRME İÇİN VERİ YÖNETİMİ

2.1. Etki Değerlendirmede Verinin Önemi

Etki değerlendirme, belirli bir politika, program veya müdahalenin hedeflenen sonuçlara ulaşp ulaşmadığını belirlemek amacıyla yürütülen sistematik bir analiz sürecidir. Kamu politikalarından sosyal programlara, devlet yardımlarından eğitim reformlarına kadar geniş bir yelpazede uygulanan etki değerlendirmesi, müdahalelerin etkilerini bilimsel yöntemlerle ölçerek politika yapıcılarının bilinçli ve veriye dayalı kararlar almasına imkân tanımaktadır.

Bu sürecin temel unsuru veridir. Doğru, güvenilir ve kapsamlı veriler olmaksızın gerçekleştirilen değerlendirmeler, bilimsel temelden uzaklaşmakta ve yanlış yönlendirmelere neden olabilmektedir. Verinin toplanması, temizlenmesi, analiz edilmesi ve yorumlanması, etki değerlendirme süreçlerinin sağlıklı ve etkin bir biçimde yürütülmesini sağlamaktadır.

Örneğin, bir devlet yardımı programının istihdam üzerindeki etkisinin ölçülmesi amacıyla yardımdan yararlanan ve yararlanmayan firmaların karşılaştırılması gerektiğinde, bu firmalara ilişkin veriler eksik veya hatalı ise, programın gerçek etkisi olduğundan farklı yorumlanabilir ve politika yapıcılar yanıltıcı sonuçlara ulaşabilirler. Benzer şekilde, eğitim alanında gerçekleştirilen reformların öğrenci başarısı üzerindeki etkisini analiz etmek amacıyla kullanılan veriler eksik, hatalı veya yanlış olduğunda değerlendirme sonuçları gerçeği yansıtmayacaktır.

Etki değerlendirmede veri, yalnızca mevcut politikaların değerlendirilmesinde değil, geleceğe yönelik politika ve programların tasarımında da kritik bir role sahiptir. Politika yapıcılar, geçmiş verilere dayalı analizler aracılığıyla benzer müdahalelerin hangi koşullar altında daha etkili olduğunu belirleyebilmekte ve bu doğrultuda daha verimli projeler planlayabilmektedir.

Etki değerlendirme sürecinin sağlıklı bir biçimde yürütülebilmesi, kullanılan verinin kalitesi ile doğrudan ilişkilidir. Bu noktada, var olan verinin düzenlenmesi aşamasına geçilmeden önce veriyi toplayan kaynağın güvenilirliğinin sağlanması önem arz etmektedir. Örneğin, veri kalitesinin idari kayıtlarla ilişkilendirilmesi ya da veri toplayıcıların belirli standartlara sahip olması, analizin temel dayanağını oluşturan veri setinin güvenilirliğini artıracaktır. Eksik, hatalı veya düzensiz veri setleriyle gerçekleştirilen analizler, hatalı sonuçlar üreterek politika yapıcılarını yanıltabilmekte ve kamu kaynaklarının etkin olmayan biçimde kullanılmasına yol açabilmektedir. Bu nedenle, veri setlerinin dönüşüm, temizleme ve bütünleştirme işlemlerinden geçirilerek analiz için uygun hale getirilmesi gerekmektedir.

Eksik verilerin uygun istatistiksel yöntemlerle tamamlanması, aykırı değerlerin tespit edilerek gerekli düzeltmelerin yapılması, zaman uyumsuzluklarının giderilmesi ve veri setlerinin standardize edilmesi, değerlendirme sürecinin doğruluk ve güvenilirliğini artırmaktadır. Ayrıca, verinin niteliğine ve kullanılan analiz yöntemine bağlı olarak veri setinin kesitsel veya panel veri formatında düzenlenmesi, nedensellik ilişkilerinin daha sağlıklı biçimde kurulmasına imkân tanımaktadır.

Sonuç olarak, etki değerlendirme süreçlerinde nitelikli ve doğru veri ile çalışılması; müdahale ve kontrol gruplarının etkin biçimde analiz edilmesini kolaylaştırmakta, nedensel ilişkilerin sağlam temeller üzerine kurulmasını sağlamakta ve güvenilir politika önerilerinin geliştirilmesine önemli katkı sunmaktadır.

2.2. Veri ve Veri Kaynakları

Devlet yardımlarının izlenmesi ve değerlendirilmesi süreçlerinde doğru, güncel ve yeterli verilerin kullanılması etki değerlendirme çalışmaları neticesinde tespit edilen etkinin güvenilir ve yansız olması açısından kritik öneme sahiptir. Etki değerlendirme çalışmaları tasarlanırken, araştırma sorularına yanıt verecek verinin nasıl toplanacağını önceden belirlenmesi ve veri toplama takviminin ve araçlarının netleştirilmesi gerekmektedir.

Veri; gözlemlerden, deneylerden veya ölçümlerden elde edilen ve işlenmeden önceki ham bilgilerdir. Veriler sayılar, gözlemler veya yanıtlardan oluşan değerler bütünüdür.

Veri kaynakları ikiye ayrılır:

- **Birincil Veri:** Doğrudan sahadan toplanan verilerdir.

Birincil veriler genellikle deneyler, gözlemler, anketler, soru formu, test, mülakatlar/ odak grup görüşmeleri, vb. veri toplama araçları kullanılarak toplanmaktadır.

- **İkincil Veri:** Mevcut veri tabanlarında tutulan izleme verileri, raporlar, istatistikler ve literatür taramaları gibi güvenilir kaynaklardan oluşur.

İkincil veriler idari kayıtlar, resmi istatistikler, devlet kayıtları, kitaplar, yayınlanmış kaynaklar, yapılan bilimsel çalışmalar, vb. kaynaklarda yer verilen verilerden türetilmektedir.

Devlet yardımlarının izlenmesi ve değerlendirilmesi için doğru, güncel ve yeterli verilerin kullanılması önemlidir. Ayrıca, etki değerlendirme tasarlanırken analizde kullanılacak araştırma sorularına yanıt vereceği düşünülen verinin nasıl toplanacağını önceden tespiti önemlidir.

2.3. Etki Değerlendirmede Kullanılacak Verilerde Bulunması Gereken Özellikler

Etki değerlendirme çalışmalarından elde edilen sonuçların geçerli ve güvenilir olabilmesi için analizde kullanılacak verilerin aşağıdaki koşulları sağlaması zorunludur:

- **Güncellik:** Verinin etki değerlendirme için güncel olması gerekmektedir.
- **Geçerlilik:** Verilerin tanımlanmış format, tür ve aralık kriterlerine uyması gerekmektedir.

- **Benzersizlik:** Gözlem birimlerinin (firma, şahıs) veri setinde yinelenen (mükerrer) kaydının bulunmaması gerekmektedir.
- **Doğruluk:** Verinin gerçek durumu doğru bir şekilde yansıtması gerekmektedir.
- **Eksiksizlik:** Etki değerlendirme için zorunlu olan tüm verilerin mevcut olmasıdır.
- **Tutarlılık:** Aynı verinin farklı zamanlar veya kayıtlar arasında çelişmemesi ve tek tip olması gerekmektedir.

2.4. Etki Değerlendirmede Kullanılacak Verilerin Belirlenmesi

Uygulanan bir devlet yardımının etki değerlendirmesinin yapılabilmesi için aşağıda hususlar dikkate alınarak gerekli verilerin belirlenmesi gerekmektedir.

- **Bağımlı Değişken:** Devlet yardımının amacını ve hedefini temsil etmeli ve performans göstergeleriyle uyumlu olmalıdır. Devlet yardımının amacını temsil etmeyen bir bağımlı değişken belirlenmemelidir.
- **Bağımsız Değişkenler:** Bağımlı değişkeni etkileyen, gruplar arasında doğru karşılaştırma yapmayı sağlayan değişkenler olmalıdır. Bağımlı değişken açısından kritik olmayan özellikler bağımsız değişken olarak kullanılmamalıdır.
- **Müdahale Grubu:** İdeal olarak etki değerlendirme yapılmak istenen yılda yardım alanlar arasından, daha önce aynı veya başka yardımdan yararlanmamış olanlardan oluşturulan bir grup olması gerekmektedir. Devlet yardımından yararlanan firmalar arasından ilk defa yararlanan yok ise, ilerleyen bölümlerde yer alan açıklamalar çerçevesinde müdahale grubunun oluşturularak veri toplanması gerekmektedir.
- **Kontrol Grubu:** İdeal olarak hiç yardım almamış ancak sektör, il, ölçek, yaş, cinsiyet, eğitim durumu gibi özellikler bakımından yardım alanlara çok benzeyen bir grup olması gerekmektedir. Devlet yardımından ilgili sektördeki/bölgedeki firma/kişilerin tamamının yararlanması nedeniyle kontrol grubu oluşturulamaması durumunda ilerleyen bölümlerde yer alan açıklamalar çerçevesinde kontrol grubunun oluşturularak veri toplanması gerekmektedir.
- **Etki Değerlendirme Yöntemi:** Devlet yardımının yapısına/tasarımına uygun olarak kullanılacak etki değerlendirme yöntemi esas alınarak hangi yıllara ilişkin veri toplanacağını belirlenmesi gerekmektedir. Örneğin Farkların Farkı yönteminde paralel eğilim varsayımının kontrolü gerektiğinden yardımın uygulandığı yıldan önceki en az iki yıla ilişkin veri toplanması gerekmektedir. Her durumda yardım öncesi ve sonrasına ilişkin veri olmalıdır.

Aşağıda devlet yardımının amacına bağlı olarak belirlenebilecek bağımlı ve bağımsız değişken türlerine örnekler yer almaktadır.

Kutu 1. Bağımlı ve Bağımsız Değişken Türlerine Göre Örnekler

Firmalar için Örnek Değişkenler	Kişiler için Örnek Değişkenler
• Firma Büyüklüğü • Firmanın Sektörü/Alt Sektörü • Çalışan Sayısı • Satış Hasılatı/Ciro • Teknoloji Seviyesi • Üretim Miktarı • Şirket Türü • Kaldıraç Oranı • Yatırım Harcaması • Ar-Ge Harcaması • Doktoralı Personel Sayısı • İstihdam Edilen Kadın Sayısı • Faaliyet Süresi • Faaliyette Bulunulan İl/Bölge • Faaliyette Bulunulan Sektör/Alt Sektör • İhracat Tutarı • İhraç Edilen Ürün/Hizmet Türü • Eser Sayısı • İzleyici Sayısı/Ziyaretçi Sayısı	• Yaş • Cinsiyet • Eğitim Durumu/Süresi • Medeni Durum • İkamet Edilen İl/Bölge • Saatlik Çalışma Süresi • Aylık Kazanç • İstihdam Edilen Meslek/Meslek Grubu • Daha Önce İstihdam Edilme Süresi/Tecrübe Süresi • İşsizlik Süresi • Sözleşme Türü • Aynı/Farklı İşyerinde Geçirilen Süre

2.5. Örneklem ve Örneklem Büyüklüğü

Etki değerlendirme çalışmalarında amaç, bir politika, program veya desteğin hedef kitle üzerindeki gerçek etkisini ortaya koymaktır. Ancak uygulamada çoğu zaman evreni oluşturan tüm birey, işletme veya birimlere ulaşmak; zaman, maliyet ve operasyonel kısıtlar nedeniyle mümkün olmamaktadır. Bu nedenle evreni temsil edebilecek nitelikte bir örneklem seçilerek analizler yürütülür.

Örnekleme süreci, evrenin temel özelliklerini doğru biçimde yansıtacak sayıda ve yapıda bir gözlem grubunun belirlenmesini ifade eder. Etki değerlendirme çalışmalarında seçilen örneklem yetersiz büyüklükte olması, yapılan istatistiksel analizlerin güvenilirliğini doğrudan zayıflatır. Özellikle minimum gerekli örneklem büyüklüğünün altında kalan çalışmaların sonuçları istatistiksel olarak anlamlı görünse dahi yanlış, hatalı ve yanıltıcı (sahte) sonuçlar üretebilir. Bu durum, politika kararlarının yanlış yönlendirilmesine yol açabilecek ciddi bir risktir.

Bu nedenle etki değerlendirme çalışmalarında kullanılacak örneklem, belirli istatistiksel varsayımlar altında asgari olarak kaç gözlemden oluşması gerektiği önceden hesaplanmalı ve analizler bu eşiklerin altına düşülmeden yapılmalıdır. Aşağıda yer alan tabloda evren büyüklüğüne bağlı olarak etki değerlendirme için seçilebilecek örneklem büyüklükleri verilmektedir.

Tablo 1. Evren Büyüklüğüne Bağlı Olarak Etki Değerlendirme İçin Seçilebilecek Örneklem Büyüklükleri

Evren Büyüklüğü	±0.03 örneklem hatası (d)			±0.05 örneklem hatası (d)			±0.10 örneklem hatası (d)		
	p=0.5 q=0.5	p=0.8 q=0.2	p=0.3 q=0.7	p=0.5 q=0.5	p=0.8 q=0.2	p=0.3 q=0.7	p=0.5 q=0.5	p=0.8 q=0.2	p=0.3 q=0.7
100	92	87	90	80	71	77	49	38	45
500	341	289	321	217	165	196	81	55	70
750	441	358	409	254	185	226	85	57	73
1000	516	406	473	278	198	244	88	58	75
2500	748	537	660	333	224	286	93	60	78
5000	880	601	760	357	234	303	94	61	79
10000	964	639	823	370	240	313	95	61	80
25000	1023	665	865	378	244	319	96	61	80
50000	1045	674	881	381	245	321	96	61	81
100000	1056	678	888	383	245	322	96	61	81
1000000	1066	682	896	384	246	323	96	61	81
100 milyon	1067	683	896	384	245	323	96	61	81

Kaynak: Yazıcıoğlu ve Erdoğan, 2004.

Tablo 1, evren büyüklüğüne, kabul edilen örneklem hatasına (hata payına) ve incelenen özelliğin evrendeki tahmini oranına bağlı olarak, etki değerlendirme çalışmalarında istatistiksel olarak güvenilir sonuçlar elde edebilmek için gerekli minimum örneklem büyüklüklerini göstermektedir.

Başka bir ifadeyle tablo şu soruya yanıt verir: “Belirli büyüklükte bir evren için, belirli bir hata payını kabul ederek, güvenilir bir etki değerlendirmesi yapabilmek için en az kaç gözlem gereklidir?”. Tabloda yer alan unsurların ne anlama geldiği ise aşağıda belirtildiği üzeredir:

Evren Büyüklüğü: Çalışmanın kapsadığı toplam birim sayısını ifade eder. Bu birimler birey, hane, işletme, proje veya firma olabilir. Örneğin 500, 5.000 ya da 100.000 kişilik bir hedef kitle bu sütunda yer alan değerlere karşılık gelir.

Örneklem Hatası (d): Tabloda üç farklı hata payı düzeyi bulunmaktadır:

±0,03 (%3 hata payı): Daha hassas ve ayrıntılı analizler için

±0,05 (%5 hata payı): Sosyal bilimlerde ve kamu politikası değerlendirmelerinde (devlet yardımları dahil) en yaygın kullanılan düzey

±0,10 (%10 hata payı): Daha genel, kaba tahminlerin yeterli olduğu çalışmalar için

Hata payı küçüldükçe, sonuçların hassasiyeti artmakta; buna karşılık gerekli örneklem büyüklüğü de yükselmektedir.

p^8 ve q Değerleri: Tabloda her hata payı için üç farklı p - q kombinasyonu yer almaktadır. p , incelenen özelliğin evrendeki tahmini oranını ifade eder.

$q = 1 - p$ 'dir. Örneğin:

$$p = 0,5 \rightarrow q = 0,5 \quad p = 0,8 \rightarrow q = 0,2 \quad p = 0,3 \rightarrow q = 0,7$$

⁸ Yapılan istatistiksel bir test sonucunda elde edilen bulgunun tesadüfen ortaya çıkmış olma olasılığını gösteren bir ölçüdür. P-değerinin küçük olması, elde edilen sonucun rastlantı ile açıklanma olasılığının düşük olduğunu ve analizde tespit edilen etkinin istatistiksel olarak anlamlı olabileceğini gösterir.

p değeri hakkında önceden bir bilgi yoksa $p = 0,5$ varsayımı kullanılır. Bu durum, evrendeki belirsizliğin en yüksek olduğu senaryoyu temsil eder ve en büyük örneklem ihtiyacını doğurur. Bu nedenle $p = 0,5$ uygulamada “temkinli ve güvenli” bir varsayım olarak kabul edilir.

Tabloyu doğru okumak doğru örneklem büyüklüğünü belirlemek için önem arz etmektedir. Dolayısıyla öncelikle çalışmanın kapsadığı evren büyüklüğü belirlenir. Analizde kabul edilecek örnekleme hatası ($\pm\%3$, $\pm\%5$ veya $\pm\%10$) seçilir. İncelenen değişkene ilişkin p değeri biliniyorsa buna uygun sütun, bilinmiyorsa $p = 0,5$ sütunu kullanılır. İlgili satır ve sütunun kesişiminde yer alan değer, asgari örneklem büyüklüğünü gösterir. Örneğin:

Evren büyüklüğü 1.000, kabul edilen hata payı $\pm\%5$, p değeri bilinmiyor ($p = 0,5$ varsayılıyor) ise tabloda yer alan değer 278’dir. Bu, güvenilir bir etki değerlendirmesi yapabilmek için en az 278 gözlem gerekliliği anlamına gelir.

Burada dikkat çekici bir husus, evren büyüklüğü arttıkça örneklem büyüklüğünün belli bir noktadan sonra çok sınırlı artış göstermesidir. Örneğin 50.000, 100.000 veya 1.000.000 büyüklüğündeki evrenler için gerekli örneklem sayıları birbirine oldukça yakındır.

Bu durum, evren büyüklüğü belirli bir eşiğin üzerine çıktığında örneklem büyüklüğünün artık evrenden ziyade hata payı ve değişim oranı tarafından belirlendiği ilkesine dayanmaktadır. Bu nedenle çok büyük evrenlerde dahi birkaç bin gözlem, doğru seçildiği sürece güvenilir sonuçlar üretebilir. Dolayısıyla tabloda yer alan örnekleme büyüklüğüne ilişkin değerlerin altına düşülmesi durumunda test sonuçları güvenilirliğini yitirir, p-değerleri yanıltıcı olabilir ve bulgular istatistiksel olarak anlamlı görünse bile gerçekte hatalı çıkarımlar üretilebilir.

Bu doğrultuda burada yer alan değerler “ideal” örneklem büyüklüğünü değil, analizin güvenilirliğini kaybetmemesi için altına inilmemesi gereken sınırı ifade eder. Özellikle alt gruplar, karşılaştırmalı analizler veya etki büyüklüğünün küçük olduğu durumlarda, bu minimum değerlerin üzerine çıkılması çoğu zaman gereklidir.

2.6. Veri Toplama Stratejileri ve Eksik Veri Yönetimi

Etki değerlendirme çalışmalarında kullanılacak olan gerekli verilerin toplanması planlama ve zaman yönetimi gerektiren bir süreçtir.

Yapılacak etki değerlendirme çalışmasında öncelikle analizi gerçekleştiren Kurumun idari kayıtlarında yer almakta olan veriler kontrol edilerek analiz için kullanılabilecek olan verilerin neler olduğu tespit edilmelidir.

Bu şekilde bir tespit çalışmasının hem verilerin hızlı ve etkin bir şekilde toplanması hem de ihtiyaç duyulan ancak mevcut durumda tutulmayan verilerin belirlenerek gelecek dönemlerde ihtiyaç duyulan verilerin Kurum tarafından toplanabilmesi için gerekli çalışmaların yapılmasında yol gösterici bir rol oynayacaktır.

Etki deęerlendirmede kullanılacak verilerin tamamının veya bir kısmının kurumun idari kayıtlarında bulunmaması durumunda eksik veya ilgisiz verilerle etki deęerlendirme yapılmamalı, gerekli verilerin tamamı ařaęıdaki yöntemler tek başına veya birlikte kullanılarak mutlaka toplanmalıdır:

- Potansiyel müdahale grubu ve kontrol grubu olarak belirlenen firmalardan anket yoluyla (yüz yüze, telefon, e-posta ve portal üzerinden),
- Halihazırda yapılmıř olan ve potansiyel müdahale grubu ile kontrol grubunu da içeren firma bazında veri bulunan (mikro veri) dięer anketlerden yararlanarak,
- Eksik olan verileri düzenli olarak kayıt altına alan dięer kurumlara iliřkin veri tabanları (idari kayıtlar).

Özellikle ilgili kurum veya kuruluřlardan veri talep edilmeden önce kapsamı ve içerięi belirlenen verilerin talep edilen kurum tarafından tutulduęundan emin olunmalıdır.

Verilerin ilgili kurum tarafından tutulmakta olduęundan emin olunmasının ardından talep edilen verilerin tanım, içerik, kapsam, dönem, asgari gözlem sayısı, vb. özellikleri somut olarak belirtilerek verinin ne zamana kadar gönderilmesi gerektięi hususlarında gerekli müzakereler ve planlamalar da yapılmalıdır.

Bu hususların yanı sıra ilgili kurum veya kuruluřtan talep edilen verilerin sadece etki deęerlendirmede kullanılacaęı ve veri gizlilięinin saęlanması için gerekli tedbirlerin alınacaęı hususu da ayrıca belirtilmelidir.

2.7. Verinin Korunması ve Gizlilik

Etki deęerlendirme çalıřmalarında gerek kurumun kendi kayıtlarından gerekse ilgili kurum veya kuruluřlardan temin edilen verilerin güvenlięinin saęlanması ve gizlilięinin korunması son derece önemli olmasının yanı sıra Kiřisel Verilerin Korunması Kanunu gereęince de kiřisel cezai sorumluluęa da yol açabilecek bir durumdur.

Bu çerçevede özellikle firmalara/kiřilere iliřkin kiřisel veriler kesinlikle gizlilik ilkelerine riayet edilerek korunmalı ve mümkün mertebe kiřisel verileri maskeleyecek tanımlama ve gizleme yöntemleri kullanılarak bu bilgiler gizlenmelidir.

Bu hususun yanı sıra ayrıca elde edilen veriler etki deęerlendirme çalıřmaları ile ilgili olmayan kurum personeli veya dięer kiřilerle paylařılmamalıdır. Özellikle veri analizi ve etki deęerlendirme sürecinde son dönemde kullanılmaya bařlanan yapay zeka platformlarında verilerin paylařılmaması konusunda gerekli dikkat ve özen gösterilmelidir. Aksi taktirde 6015 sayılı Devlet Desteklerinin İzlenmesi ve Denetlenmesi Hakkında Kanun gereęince kiřisel verilerin korunması ile ilgili hükümleri ihlal edenler, muhtelif kanunlar gereęince gizli kalması gereken bilgileri açıklayanlar ve elde edilen verileri devlet yardımlarının izlenmesi ve etkilerinin deęerlendirilmesi amacının dışında kullananlar hakkında 5237 sayılı Türk Ceza Kanunu ve ilgili düzenlemeler kapsamında gerekli cezalandırma işlemleri yapılabilecektir.

3. NİCEL ETKİ DEĞERLENDİRME YÖNTEMLERİ

Nicel yöntemlerle yardımın yararlanıcılar üzerindeki etkisi karşılaştırmalı analiz yaklaşımlarını içeren *Karşı Olgusal Etki Değerlendirme (Counter-Factual Impact Evaluation)* yöntemleri ile ölçülür. Nicel etki değerlendirme yapılırken izlenmesi gereken adımlara Rehber’de değinilmiştir. Aynı zamanda Ek 2’de değerlendirme çalışması öncesinde nicel yöntemlerin hangisinin seçilmesi gerektiğine dair bir karar ağacı şeması yer almaktadır.

3.1. Nicel Etki Değerlendirmede İzlenecek Adımlar

Adım 1: Başlangıç ve Ekibin Oluşturulması

Tüm etki değerlendirmelerde başlangıç noktasını, mevcut veya planlanan bir yardımın olası etkisini hesaplamak oluşturmaktadır. Başlangıç aşamasında odaklanılacak olan yardım ve sınırları konusunda ilgili kurumda mutabakat olması gerekmektedir. Bunun için yardımın değerlendirme çerçevesi önemli bir kolaylık sağlayacaktır. Değerlendirme çerçevesi içinde yer alması gereken unsurlar Tablo 2’de gösterilmiştir.

Kurumun kullanacağı girdi ile aşamalı olarak nasıl etkiye ulaşılacağı neden-sonuç mantığı (değişim teorisi) ile açıklanmaktadır. Bilgilerin bu tabloya yansıtılması, yeni tasarlanan bir yardım için planlamanın sağlıklı bir şekilde yapılmasına hizmet etmekte, uygulanmakta olan yardımlar için ise uygulama sonuçları ile hedeflerin karşılaştırılmasına yardımcı olmaktadır.

Tablo 2. Devlet Yardımı ve Etkileri Tablosu

YARDIM	YARDIM KAPSAMINDA ÖNGÖRÜLEN FAALİYETLER	YARDIMDAN YARARLANANLAR	YARDIMIN YARARLANICILARA ETKİSİ	YARDIMIN EKONOMİYE ETKİSİ
Devlet yardımının katkı türü ve toplam miktarı nedir? <i>Örneğin: Yıllık 1 Milyar TL tutarında nakdi yardım. Bu kapsamda ayrıca varsa diğer katkılar da belirtilecektir (Örn. uzman personel temini)</i>	Devlet yardımı kime, hangi koşullarda ve hangi amaçla sağlanacaktır? Yardımla ilgili yapılacak ilave faaliyetler (tanıtım, eğitim vb.) nelerdir? <i>Örneğin: Yardımla ilgili eğitim ve tanıtım yapılması, başvuru duyurusunun yapılması, başvuruların alınması, başvuruların değerlendirilerek uygun görülen firmalara yardımın sağlanması</i>	Devlet yardımı sağlanması öngörülen firma sayısı ve firma başına öngörülen yardım tutarı nedir? <i>Örneğin: 4.000 firmaya yıllık ... başvuruyu üzerinden ... TL tutarında yardım sağlanması</i>	Devlet yardımından yararlanan firmanın bu yardımdan elde edeceği fayda nedir? <i>Örneğin: Yardım sağlanan firmaların ... TL tutarında üretim veya ihracat yapması, dekar başına ... üretimini %... oranında artırması gibi</i>	Devlet yardımının ulusal ekonomiye sağlayacağı fayda nedir? <i>Örneğin: Türkiye’nin toplam üretiminin/ihracatının ... TL/Dolar ya da %..oranında artması</i>
Varsayımlar	Devlet yardımı teklifinin, hangi koşulların varlığı halinde yukarıda belirtilen etkiyi doğuracağı belirtilir. Söz konusu varsayımların belirlenmesinde idari ve hukuki altyapının mevcut durumu ile olası değişiklikler, ulusal ve küresel beklentiler vb. hususlar dikkate alınır.			
Riskler	Varsayımlardan hangilerinin gerçekleşmemesi halinde yardımın beklenen etkiyi doğurmayacağı ve nedenleri belirtilir.			

Sonuçlar zinciri modeli belirlenirken, değişim teorisine dayalı varsayımlar ve risklerin de belirtilmesinde yarar bulunmaktadır. Bu varsayım ve riskler kurumun ve yararlanıcıların dışında gelişen, ancak yardımı etkileyen dış etkenler gözetilerek

belirlenmelidir. Örneğin, varsayım olarak mevcut ekonomik koşulların devam etmesi, risk olarak ise doğal afet, savaş, ekonomik kriz vb. yazılabilecektir.

Bu ve izleyen adımların planlı bir şekilde yürütülebilmesi için bir ekip oluşturulması uygun olacaktır. Ekip uygulayıcı kurum tarafından Yönetmelikte bahsedilen ve Başkanlığın verdiği sertifikayı almış kişiler dâhil olacak şekilde ilgili kişilerin görevlendirilmesiyle oluşturulmalı ve değerlendirme bu ekip tarafından sonuçlandırılmalıdır.

Ekip içinde etki değerlendirmenin başından sonuna sorumlu olacak bir yönetici olmalıdır. Görev unvanı olarak bir yönetici olması şart değildir. Yönetici, değerlendirilecek yardımın konusuna hâkim olmasının yanı sıra bu yardımın kapsamına ilişkin veri toplama yöntemini, ekip içindeki baş araştırmacının belirlenmesini, soru setlerinin hazırlanmasını ve raporlanmasını sürecin başından sonuna kadar koordine etmelidir.

Baş araştırmacının veri toplanma sürecinin izlenmesi, veri toplayanları bilgilendirmesi gerektiğinde eğitim düzenlemesi, toplanan verinin derlemesi, araştırma tasarımına katkı sağlaması gibi görevleri vardır.

Son olarak veriyi analiz edecek analiste ihtiyaç olacaktır. Sonuçlar sadece analist tarafından değil, hem oluşturulan ekip hem de konuyla ilgili diğer kişiler tarafından da rahatlıkla yorumlanmalıdır.

Yukarıda belirtilen rolleri üstlenenlerin, gerekli bilgi ve tecrübeye sahip olması, etki değerlendirmenin kalitesini artıracaktır.

Adım 2: Paydaşların Belirlenmesi

Devlet yardımı ile ilgili etki değerlendirmeleri için belirlenen ilgili paydaşlarla istişarelerin yapılması önem taşımaktadır. Müdahalenin gerekli görüldüğü alan veya sektörde faaliyet gösteren kamu kurum ve kuruluşları, meslek veya ticaret odaları, akademisyenler ve/veya STK'ların belirlenmesi yararlı olacaktır.

Adım 3: Paydaşlarla İstişare

Planlanan devlet yardımıyla ilgili olarak yardıma ihtiyacın olup olmadığı, ihtiyaç varsa ihtiyacı doğuran sorunlar, hedef kitle, yardım miktarı, süresi ve türü, hâlihazırda bu alanda başka bir yardımın uygulanıp uygulanmadığı konuları paydaşlarla değerlendirilecektir. Uygulanmakta olan yardımlarla ilgili olarak ise yardımın sorunların çözümüne veya ihtiyaçlara hizmet edip etmediği, nedenleri, yapılabilir ilave çalışmalar gibi konuların istişare edilmesi büyük önem taşımaktadır. Paydaşlarla istişare ikili mülakatlar şeklinde yapılabileceği gibi toplantılar şeklinde de yapılabilir.

Adım 4: Veri Toplama Araçları

a. Nicel Anketler

Etki değerlendirme çalışmalarında veri toplama aracı olarak anketlerin kullanıldığı durumlarda, güvenilir sonuçlara ulaşılabilmesi için genellikle müdahale öncesi (baseline) ve müdahale sonrası (endline) olmak üzere iki zamanlı, dikkatlice tasarlanmış anketlerin uygulanması gerekmektedir. Bu yaklaşım, karşı olgusal tahminlerin sağlıklı biçimde yapılmasına olanak tanımaktadır.

İzleme değerlendirme sistemleri, programlar kapsamında belirlenen hedeflere ne ölçüde ulaşıldığına dair veriler sunabilse de bu tür veriler “program nasıl bir fark yarattı?”, “gözlenen değişiklikler ne ölçüde müdahaleye atfedilebilir?” gibi sorulara yanıt verememekte, bu sorulara yalnızca etki değerlendirme cevap verebilmektedir (White ve Raitzer, 2017).

Diğer yandan etki değerlendirme çalışmalarının izleme ve değerlendirme veri toplama süreçlerine entegre edilmesi gerçekleştirilen analize katkı sağlayacaktır. Bu iki sürecin birbiriyle entegrasyonu:

- Çakışan anketlerin önlenmesini,
- Proje izleme kalitesinin artmasını,
- Kısıtlı veri toplama kaynaklarının daha etkin kullanılmasını sağlar.

Etki değerlendirme yöntemleri, karşı-olgusal oluşturma ve seçim yanlılığını minimize etme prensipleri üzerine inşa edilmektedir. Özellikle etki değerlendirme bağlamında anket tasarımı, müdahale ve kontrol grupları arasında karşılaştırılabilirliği sağlamak, grupların başlangıç durumları ve davranışsal özelliklerine ilişkin gözlemlenebilir ve gözlemlenemeyen değişkenleri yakalamak, panel veya kesit veri yapısına uygun bilgi toplamak açısından kritik önem taşımaktadır. Örneğin, Eğilim Skoru Eşleştirme yönteminde, müdahale ve kontrol grupları arasındaki gözlemlenebilir farklılıkların minimize edilmesi gerekmekte bu farklılıkların belirlenmesi ve modellenebilmesi için gerekli değişkenler anket yoluyla elde edilebilmektedir.

Bu kapsamda anketler, etkiyi doğru şekilde ölçebilecek veri yapısını kurmak, araştırmacının veri üretme sürecini tasarladığı, kontrollü varyasyon yaratabildiği ve belirli değişkenleri doğrudan ölçebildiği bir sistem oluşturmak açısından önemlidir (Stantcheva, 2022).

Soru Tasarımı

Anketlerde soru tasarımı veri kalitesinin belirleyici unsurlarından biridir. Bu nedenle etki değerlendirme odaklı anketlerde soruların içeriği, sıralaması ve formatı ölçüm hatalarının azaltılması ve katılımcılardan güvenilir bilgi edinilmesi bakımından önem taşımaktadır. Soru tasarımına ilişkin önemli bazı hususlara aşağıda yer verilmiştir.

- Sorular, mümkün olduğunca mantıksal modelde belirlenen hipotezleri (varsayımları) kapsamalıdır. Ancak bazı hipotezlerin nicel anketler aracılığıyla ölçülmesi mümkün değildir. Bu durumda nitel araçlar ve ikincil veriler araştırmaya destek olacaktır.

- Net sorular olmalı ve her soru tek kavramı ölçmelidir. Belirsiz, çift anlamlı veya birden fazla kavramı aynı anda içeren sorular veri kalitesini düşürmektedir. Anketin pilot testi, soruların net ve anlamlı olmasını sağlamak için önemlidir. Anket içinde aynı konuya odaklanan farklı soru türleri kullanılarak iç tutarlılık sağlanabilmektedir.
- Kapalı uçlu sorular, genellikle açık uçlu sorulardan daha fazla tercih edilmektedir. Bunun nedeni kapalı uçlu soru kullanımının sonradan kodlama/sınıflandırma gerektiren açık uçlu sorulara göre daha kolay olmasıdır. Anket ne kadar büyükse ve cevapların çeşitliliği ne kadar fazlaysa, sonradan cevapları kodlamak o kadar zor olmaktadır. Bununla birlikte açık uçlu sorular, cevapların önceden makul bir şekilde tahmin edilemediği durumlarda veya derinlik arandığında kullanışlıdır (USAID, 2006). Anket soruları, etki değerlendirme hipotezini doğrudan test edebilecek veriyi sağlamalıdır. Bu nedenle karşı-olgusal oluşturma, potansiyel karıştırıcı değişkenleri (confounders) ölçme ve davranışsal değişiklikleri yakalama amaçlarına hizmet eden sorular ankete dâhil edilmelidir.
- Soru sıralaması nötral olmalı, yönlendirici olmamalıdır. Önce nötr sorular, ardından müdahale konusuna özgü sorular gelmelidir.
- Gelir veya sağlık durumu gibi hassas içerikli sorularda katılımcılara “*cevaplamak istemiyorum*” seçeneği sunulmalıdır.
- Teknik terimler sadeleştirilmeli ve hedef kitlenin anlayacağı şekilde formüle edilmelidir.

Anket katılımcılarından doğru bilgi edinmek ve anlaşılır kolay bir anket oluşturmak için aşağıda yer alan öneriler dikkate alınabilir.

Tablo 3. Anket Özellikleri

Öneriler	Açıklaması
Sorular anlaşılır olmalıdır.	Nispeten kısa sorular ve mümkün olduğunca basit, anlaşılır ifadeler kullanılmalıdır.
Hassas sorular görüşmenin sonuna saklanmalıdır.	Katılımcılara sorulan çoğu soru hassas değildir. Ancak yardım gibi mali konulara karşı tutumlar hassas olabilir. Faydalanıcılara yönelik seçim kriterleri gibi konuları içerdiğinden ve ayrıca genel kanı olarak yardımların her zaman gerekli olduğu düşüncesinden hareketle yanlış cevaplar verilmesi mümkündür. Bunlar göz önünde bulundurularak soruların içeriği ve sıralaması tasarlanmalıdır.
Uzun anket formlarından kaçınılmalıdır.	Araştırmacılar genellikle katılımcılardan büyük miktarda veri toplama eğilimindedir. Bunun nedeni birden fazla görüşme gerçekleştirmenin zaman maliyetini yükseltmesidir. Ancak uzun soru formları, katılımcılarda gönülsüzlük yaratabilir ve doğru cevap verme olasılığını düşürebilir. Bu nedenle verilecek yanıtın araştırmaya ne kadar faydası olacağı gözetilerek ankete sorular eklenmelidir.

Kaynak: Dünya Bankası, Uygulamada Etki Değerlendirme (Impact Evaluation in Practice) (2016) dokümanından derlenmiştir.

Etki değerlendirme sürecinde uygulanan tüm yöntemler için geçerli olmak üzere, anket tasarımında dikkat edilmesi gereken temel ilkeler aşağıda yer almaktadır.

- Zaman odaklı analizlerde panel veri kullanımı için aynı katılımcılardan farklı zaman noktalarında (müdahale öncesi-baseline, ara dönem-midline, müdahale sonrası-endline) veri toplamak önem arz etmektedir.

- Eşleştirme ve kovaryat ayarlamaları için yaş, eğitim, sektör, firma büyüklüğü, faaliyet yeri, ciro, üretim, ihracat, istihdam gibi demografik vesosyoekonomik göstergeler ile anılan göstergelere ilişkin geçmiş performansını yansıtan verilerin yanı sıra, mümkün olduğunda müdahale öncesi döneme ait sonuç değişkeni (pre-treatment outcome) verileri de toplanmalıdır; bu bir zorunluluk olmamakla birlikte paralel eğilim varsayımının değerlendirilmesini ve tahminlerin güvenilirliğini güçlendirmektedir.
- Anketin pilot uygulaması, anlaşılmayan veya hatalı işleyen soruların tespit edilmesi noktasında önemlidir.
- Anket sorularının, tutarlılığı sağlamak amacıyla her iki dönemde de aynı olması, korunması gerekmektedir.
- Hatırlama hatalarının önüne geçilebilmesi için zaman referanslarının açık ve net bir biçimde ifade edilmesi (örneğin son 12 ay içinde, 2025 yılında, devlet yardımı öncesinde, vb.) gerekmektedir.
- Ölçeklerin standartlaştırılması önemlidir; örneğin Likert tipi derecelendirmeler, memnuniyet/algı ölçümleri veya davranış derecelendirmeleri her iki dönemde de aynı biçimde ve dengeli (örneğin 5'li ya da 7'li, tamamen katılıyorum-katılmıyorum) seçeneklerle uygulanmalıdır.
- Tanımların ve kavramsal çerçevenin sabit tutulması gerekmektedir; kavramsal tutarlılığın sağlanması bakımından gelir, istihdam ve sağlık göstergeleri gibi kavramlar her iki dönemde de aynı şekilde tanımlanmalı ve ölçüm kriterleri dönemler arasında değiştirilmemelidir.
- Veri kalitesi anket sırasında kontrol edilmelidir. Eksik yanıtlar, uç değerler ve tutarsızlıklar anket sırasında tespit edilmeli ve doğrulama yapılmalıdır.
- Veri toplama koşullarındaki farklılıklar ciddi yanlılıklar yaratabilir. Bu nedenle anketin uygulandığı zaman dilimi ile uygulama yöntemlerinin, dönemler arasında mümkün olduğunca aynı tutulması (aynı ay/benzer zaman dilimi, aynı anket yöntemi, mümkünse aynı ekip veya aynısaha protokolü) gerekmektedir.

Anket Tasarımında Zaman Boyutu ve Veri Periyodu

Etki değerlendirme çalışmasında kullanılacak anketlerin zamanlaması, analizlerin gücünü doğrudan etkilemektedir. Bu nedenle, veri toplama süreci programın tasarım aşamasında dikkatle planlanmalıdır. Özellikle zaman boyutuna dayalı karşılaştırmalar yapan Farkların Farkı yöntemi gibi analizlerde, müdahale öncesi ve sonrası dönemlere ilişkin verilerin toplanması büyük önem taşımaktadır.

Analiz sonuçlarının güvenilirliği başlangıç verilerin varlığı ile önemli ölçüde artmaktadır. Bu da ancak erken planlama ile sağlanabilir. Orta vadeli (midterm) değerlendirme anketleri yardım programları gidişatını ve orta vadede ihtiyaç duyulan düzenlemeleri ortaya koymaktadır. Ancak, bitiş anketinin zamanlamasına bağlı olarak, orta vadeli anketler aynı zamanda programların geleceği ile ilgili kararların alınmasına da katkı sağlayabilmektedir. Yardım programlarının sonunda veya tamamlanmaya yakın bir zamanda bitiş (endline) anketi uygulanmakta olup bu, programın etkisinin değerlendirilebilmesini sağlamaktadır. Başlangıç (baseline) anketi, programın ilk bir veya iki yılı içerisinde gerçekleştirilmekte, bitiş anketi ise genellikle son yıl içinde uygulanmaktadır. Başarılı bir etki değerlendirme tasarımı için çoğu zaman iki tur veri toplama asgari gereklilik olarak görülse de, eğer bu iki tur anket arasında üç yıl veya daha fazla bir süre varsa ara dönem (midterm) anket yapılması da değerlendirilmelidir. (White ve Raitzer, 2017).

Başlangıç Anketi

Başlangıç anketini tasarlama aşamasında dikkate alınması gereken hususlar şunlardır (OECD, 2006):

- Başlangıç anketi mümkün olan en erken aşamada, devlet yardımı yürürlüğe konulmadan önceki dönemde gerçekleştirilmelidir.
- Anket tasarımı, değişim teorisini⁹ temel almalıdır. Sonuçlar zinciri modeli¹⁰ boyunca veriler toplanmalı, toplanacak veri girdileri, süreç göstergelerini, çıktı ve sonuçları içermelidir.
- Kontrol grubu örneği yeterli büyüklükte olmalıdır. Diğer taraftan örneklem büyüklüğü yalnızca gözlem sayısı ile sınırlı bir konu değildir. Müdahale ve kontrol gruplarının firma büyüklüğü, sektör, yaş, bölge, geçmiş performans, vb. bakımından karşılaştırılabilir (türdeş) olması ve ölçüm hatalarının düşük olması da yeterli örneklem koşullarının bir parçasıdır.
- Kontrol ve müdahale grupları aynı veya çok benzer soru formlarına tabi tutulmalıdır. Müdahale ile ilgili bazı sorular kontrol grubuna sorulamayacak olsa da, benzer nitelikte soru fazlalığı etkiyi ölçme noktasında daha başarılı olabilir.
- Anket tasarımına ihtiyaç duyulan sürenin ayrılması ve pilot çalışma yapılması önemlidir. Veriler toplandıktan sonra veri girişi ve temizleme aşamaları için zamana ihtiyaç duyulabileceği de göz önünde bulundurulmalıdır.
- Anket çalışmasının zaman içinde tekrarlanma ihtimaline karşın, adres/iletişim bilgileri, kimlik numaraları, firma bazlı tanımlayıcı değişkenler gibi durum takibi yapmaya imkân sağlayan verilerin temini önemlidir.
- Anketin tekrarlanması durumunda ideal olan, anket tasarımında büyük değişikliklerden kaçınılmasıdır. Araştırma kapsamında özellikle dikkate alınan ve incelenen değişkenlere anketlerde yer vermek önemlidir.
- Üçleme (triangulation), yani benzer sonuç elde edilebilirliğini doğrulamak için çeşitli veri toplama araçları kullanılması yoluyla anket sonuçlarının test edilmesi faydalı olacaktır. Zira tek veri kaynağı her zaman yeterli olmayabilir. Bu nedenle idari kayıtlar, ikincil veri setleri ve nitel görüşmeler yoluyla anket verilerinin tutarlılığı karşılaştırılmalıdır.
- Etkinin güvenilir biçimde ölçülebilmesi için kaç hane veya kaç firmanın değerlendirmeye dahil edilmesi gerektiği, rastlantısal tercihlere bırakılmadan istatistiksel olarak hesaplanmalı; bu amaçla güç analizi¹¹ (power analysis) yapılmalıdır.

⁹ Değişim Teorisi (ToC), karmaşık sosyal problemleri çözmeye çalışan programların geliştirilmesi ve sonrasında bu programların etkinliğinin değerlendirilmesi için kullanışlı bir araçtır (Blamey ve Mackenzie, 2007; Weiss, 1997). Uygun şekilde kullanıldığında, bir ToC yaklaşımı, ilgili problemleri, nedenlerini ve bu problemleri yaygınlaştıran bağlamsal faktörleri tanımlamak, keşfetmek ve çözümlemek için bize bir yol sunar, ardından kaynakları toplama ve müdahale başlatma aşamasına gelir. Dahası, bir programın kısa, orta ve uzun vadeli ilgi alanlarındaki sonuçlarını belirler (Dent, Mountford-Zimdars ve Burke, 2022).

¹⁰ Sonuçlar zinciri modeli (Results Chain Model) hedeflenen sonuçları başarmak için uygulanacak müdahalenin adımlarını ortaya koyarak müdahalenin sebep-sonuç (nedensellik) ilişkisini açıklayan mantıksal modeldir. İçerisinde ToC'un gömülü olduğu değerlendirme sürecidir. Bu yüzden ToC, hem program tasarımı hem de değerlendirme için önemlidir. Kararname kapsamında sonuçlar zinciri modeli, etki değerlendirme rapor formatlarında "Devlet Yardımları ve Etkileri Tablosu" adıyla yer almaktadır.

¹¹ Güç analizi (power analysis), bir değerlendirme çalışmasının gerçek bir etkiyi istatistiksel olarak tespit edebilmesi için gerekli örneklem büyüklüğünü hesaplayan yöntemdir. Beklenen etki büyüklüğü, anlamlılık düzeyi ve test gücü gibi parametreler belirlenerek uygun istatistiksel model üzerinden örneklem sayısı hesaplanır.

- Devlet yardımı uygulanmadan önce toplanan baseline verilerinin, program duyurusundan kaynaklanan davranış değişiklikleriyle bozulup bozulmadığı hususu (no treatment contamination) değerlendirilmelidir.

İzleme Anketleri

Program uygulamasından sonra orta ve uzun vadeli sonuçları ölçen izleme anketleri, genellikle programın etkileri hakkında en ikna edici kanıtları üretmektedir. Bu anket türü başlangıç (baseline) ve bitiş (endline) verilerine ek olarak zaman içindeki değişimin dinamiklerini ve belirli bir süre sonunda ihtiyaç duyulan verileri sağlayarak müdahalenin başarılı olup olmadığını analiz etmeye olanak tanımaktadır. Eğer hem kısa vadeli hem de uzun vadeli sonuçların ölçülmesi, etkinin kalıcılığının değerlendirilmesi, beklenmeyen olumlu veya olumsuz yan etkilerin tespit edilmesi ve müdahale ve kontrol grupları arasındaki ayrışmanın dinamik seyrinin incelenmesi hedefleniyorsa, birkaç izleme anketi yapılması faydalı olacaktır. Müdahale ve kontrol gruplarını birkaç yıl boyunca takip eden etki değerlendirme çalışmalarının nadir olması nedeniyle birden fazla izleme anketi yapmak değerlendirme sürecine önemli bir katkı sağlayacaktır (ILO, 2018).

Panel Anketler

Etki değerlendirme yöntemlerinde sıklıkla panel veriden yararlanılmakta olup panel veri tasarımında;

- En az iki zaman noktasına ilişkin veri toplanması (müdahale öncesi ve sonrası)
- Aynı birimlerde tekrar eden ölçümlerin elde edilmesi,
- Ölçüm aralıklarının eşitliği dikkat edilmesi gereken hususlardandır.

Müdahale öncesine ait yeterli veri bulunması paralel eğilim varsayımının test edilebilmesi açısından kritik önem taşımaktadır (Angrist ve Pischke, 2009). Litaratürde özellikle anketler aracılığıyla birden fazla yıl öncesine ait sonuç değişkeni (outcome) verisi toplanması önerilmektedir; böylece hem gruplar arasında eşleştirme yapılabilmekte hem de paralel eğilim varsayımı daha güvenilir şekilde sınanabilmektedir (Gertler vd., 2016). Klasik Eğilim Skoru Eşleştirme, grupları sabit başlangıç özelliklerine göre benzeştirse de, *pre-trend matching*¹² yöntemiyle sonuç değişkeninin geçmiş dönemlerdeki seyrine odaklanarak daha güçlü bir benzerlik sağlamak ve FF tahminlerinin güvenilirliğini artırmak mümkündür. Bu yaklaşımın uygulanabilmesi için mümkün olduğunca çok sayıda müdahale öncesi sonuç değişkeni verisine ihtiyaç duyulmaktadır (Egami ve Yamauchi, 2022).

OECD DAC Kriterleri ve Soru Tasarımı Sürecindeki Önemi

OECD DAC (Ekonomik Kalkınma ve İşbirliği Örgütü Kalkınma Yardımları Komitesi) kriterleri, kalkınma yardımlarının etkinliğini değerlendirmek için kullanılan bir dizi standarttır. Kriterler, kalkınma yardımı projeleri ve programları için bir değerlendirme çerçevesi sağlamakta olup beş ana başlık altında toplanmaktadır: Uygunluk (Relevance), Etkililik (Effectiveness), Etkinlik (Efficiency), Etki (Impact), Sürdürülebilirlik (Sustainability). Veri toplama sürecinde OECD DAC kriterlerine

¹²**Pre-trend matching** yöntemi, grupların sonuç değişkenlerinin müdahale öncesi dönemdeki eğilimlerini karşılaştırarak eşleştirme yapar. Bu sayede, sadece sabit özelliklere değil, aynı zamanda sonuç değişkeninin zaman içindeki seyrine göre de benzerlik sağlanır.

duyarlı sorular hazırlanması önemlidir. Söz konusu kriterlerin kapsamına Tablo 4’te yer verilmiştir.

Tablo 4. OECD DAC Kriterleri

KRİTER	TANIM	GÖZ ÖNÜNDE BULUNDURULMASI GEREKEN SORULAR
Uygunluk	Yardım faaliyetinin hedef grup, faydalanıcı ve yardım sağlayıcıların öncelikleri ve politikalarına ne derece uygun olduğunu ölçen kriterdir.	• Program hedeflerinin ne dereceye kadar geçerli olduğu, • Program faaliyetlerinin ve çıktılarının genel amaçla ve hedeflerle tutarlı olup olmadığı, • Program faaliyetlerinin ve çıktılarının amaçlanan etkiler ve sonuçlarla tutarlı olup olmadığı.
Etkililik	Bir programın veya projenin hedeflerine ne ölçüde ulaştığını ölçen kriterdir.	• Hedefler ne dereceye kadar başarılı / başarıya ulaşma ihtimali nedir? • Hedeflerin başarılması veya başarısızlığında etkili olan ana faktörler nelerdir?
Etkinlik	Etkinlik yardım sağlanması sürecinde istenen sonuçları elde etmek için mümkün olan en az maliyetli kaynakların kullandığını belirten ekonomik bir terimdir. Genellikle, aynı çıktıları elde etmek için alternatif yaklaşımları karşılaştırmayı, böylece en verimli sürecin benimsenip benimsenmediğini anlamak ister.	• Faaliyetler maliyet açısından verimli miydi? • Hedeflere zamanında ulaşıldı mı? • Program veya proje, alternatiflere kıyasla en verimli şekilde uygulandı mı?
Etki	Bir program ile doğrudan veya dolaylı olarak, hedeflenen veya hedeflenmediği halde gerçekleşen olumlu ve olumsuz değişikliklerle ilgilidir.	• Program veya proje sonucunda ne oldu? • Faaliyet, yararlanıcılar için gerçek anlamda ne gibi bir fark yarattı? • Kaç kişi etkilendi?
Sürdürülebilirlik	Sürdürülebilirlik, program tarafından sağlanan faydanın program sonrasında devam edip etmeyeceğinin ölçülmesiyle ilgilidir.	• Bir programın faydaları program için sağlanan kaynak sona erdikten sonra ne ölçüde devam etti? • Programın sürdürülebilirliğinin sağlanmasını veya varsa başarısızlığı etkileyen başlıca faktörler nelerdir?

Kaynak: “DAC Criteria For Evaluating Development Assistance”, <https://www.oecd.org/dac/evaluation/39119068.pdf> (Erişim: 12.12.2023)

b. Resmi Kayıtlar ve Veritabanları

Resmi kayıtlar ve veritabanları denildiğinde genellikle kamu kurum ve kuruluşları, özel sektör firmaları ve diğer ulusal veya uluslararası kuruluşlar tarafından düzenli olarak toplanan, saklanan ve yönetilen nicel veriler ifade edilmektedir. Bu idari kayıtlar ve veritabanları, çeşitli araştırma projelerinde kullanılmak üzere zengin ve değerli veri kaynakları olarak kullanılabilir. İdari kayıtlar genellikle tüm evreni (örneğin tüm firmalar, tüm çalışanlar, tüm ihracat işlemleri) kapsadığından, örneklem hatası ve seçim yanlılığı riskini azaltmaktadır. Ayrıca birçok kurum verileri yıllık, aylık veya işlem bazında topladığından söz konusu veriler panel veri oluşturulması, müdahale öncesi eğilimlerin incelenmesi ve nicel etki değerlendirme yöntemlerinin güvenilir şekilde uygulanmasına olanak sağlamaktadır.

Bu veriler kullanılarak yardımların ekonomiye etkisi analiz edilebilir, aynı zamanda geniş çaplı sektör, bölge, alan grupları üzerinde çıkarımlarda bulunulabilir. Yardımın etkisini değerlendirmek için yararlanıcılar ve kontrol grubu için çeşitli kurumlar tarafından işletilen veri tabanlarından firma performansı (ciro, istihdam, ihracat), finansal durum (bilançolar), sektör ve bölge bilgileri gibi kritik veriler eşleştirme yapılarak toplanabilir.

Resmi kayıtlar ve veritabanları, anketlere göre daha az zamanda ve daha az maliyetle veri toplanmasını sağlayabilmektedir. Birçok çalışmada idari kayıtların anketlerin tamamlayıcısı olarak kullanıldığı da görülmektedir.

Adım 5: Araştırma Sorularının Tasarlanması (Anket yapılması halinde)

Anketlerde soru tasarımı büyük önem taşımaktadır. Araştırmanın başında doldurulması beklenen Devlet Yardımları ve Etkileri Tablosu, araştırma sorularının da temelini oluşturmaktadır. Soru hazırlama ve katılımcı ile paylaşma sürecinde açık uçlu soruların hazırlığı önem taşımaktadır.

Açık uçlu soruların amacı, potansiyel katılımcıların ya da yararlanıcıların görüşlerini yönlendirme olmadan ifade etmelerini sağlamaktır. Açık uçlu sorulardan oluşan veri toplama araçlarında temel alınan yaklaşım katılımcıları araştırmanın merkezine almak ve deneyimlerini ifade etmelerine olanak sağlamaktır.

Hazırlanan soru formlarında dikkat edilmesi gereken bazı hususlar bulunmaktadır. Güncel akademik çalışmalarda, yardımla ilgili doğrudan soru sorulmaması gerektiği savunulmaktadır. Bunun sebebi görüşme yapılan kişilerin yardımı uygulayan merciinin duymak istediğini düşündüğü cevabı vermeye çalışacağıdır (Copestake ve Diğerleri, 2019). Diğer bir ifadeyle bu tür sorular, görüşmelerin yardımın olduğundan daha faydalı olduğu sonucuna neden olabilecektir. Bunun yerine katılımcılara belirli bir zaman diliminde yaşadıkları değişiklikler, bunların nedenleri ve kaynağı hakkındaki fikirleri sorulmalıdır (Copestake ve Diğerleri, 2019). Bu değişiklikler, yardımın uygulandığı alanla ilgili durumu ortaya koymaya çalıştığından soruların bu kapsamda belirlenmesi faydalı olacaktır.

Sorular hazırlanırken, uygulanmakta olan bir yardım için daha önce hazırlanmış olan Devlet Yardımları ve Etkileri Tablosu'ndan yararlanılmasında fayda bulunmaktadır. Bu tabloda yer alan hususlara ve hedeflenen sonuca odaklanacak şekilde sorular sorulmalıdır. Örneğin, tarım sektörüne ilişkin bir yardım için şu sorular sorulabilir: Bu yıl hangi ürün/ürünleri ektiniz? Kaç dönüm ektiniz? Üretiminiz ve veriminiz de bir artış/azalış yaşandı mı? Ne kadar artış/azalış yaşandı?

Soru formlarında yer alması gereken soru sayısı ile ilgili bir kural bulunmamakla beraber öneriler "Veri Toplama" başlıklı kısımda sunulmuştur. Sorular açık uçlu olduğundan soru sayısının çokluğu görüşmeyi zorlaştırabilir. Sorulara verilecek yanıtların analizi sonucunda etkiyle ilgili bir fikir elde edilmek istendiğinden kısa ve net soruların oluşturulmasında, ayrıca görüşme sırasında anketör tarafından açık uçlu sorulara açıklık getirilmesinde yarar bulunmaktadır.

Adım 6: Araştırmacıların Eğitimi (Anket yapılması halinde)

Araştırmacıların tercih ettikleri veri toplama yöntemleri ve kullandıkları veri toplama araçları hakkında ilgili yeterli bilgiye sahip olması ve veri toplarken etik ilkelere bağlı kalıp tarafsız olmaları, toplanacak verinin güvenilirliği açısından önem arz etmektedir. Araştırma sorularının doğru bir şekilde tasarlandığından emin olmak amacıyla uygulama öncesi paydaşlarla test edilmesi önemlidir. Böylelikle soruların katılımcılar üzerindeki etkisi ve hedeflenen bilgilere ulaşmak için soru tasarımında değişiklik ihtiyacı olup olmadığı anlaşılacaktır.

Araştırmacıların görüşme yanıtlarını yönlendirmeden çalışmalarını yürütmeleri önemlidir. Veri toplama sürecinde ihtiyaç duyulan bilgilerin sağlıklı bir şekilde temin edilebilmesi ve süreçte karşılaşılabilecek risklere ve senaryolara karşı hazırlıklı olunması için araştırma konusuna hâkim olunması beklenmektedir.

Adım 7: Örneklem Seçimi

Tamsayım imkânı varsa kitle üzerinden değerlendirme yapılabilir. Ancak derlenecek verinin büyüklüğü, bütçe, zaman, nitelikli personel gibi kısıtlar varsa örneklem seçimine gidilmektedir. Değerlendirme yapılacak konuya ilişkin istatistiksel olarak temsil kabiliyetine ve araştırma konusuyla yapısal uyuma sahip bir örneklem seçilmesi önemlidir. Örneklem hataları, örneklem hedef kitleyi yeterince temsil etmediğinde ortaya çıkmaktadır.

Örnekleme dâhil edilecek kişi sayısına yönelik tek bir evrensel formül bulunmamakla birlikte varyans¹³, beklenen etkinin büyüklüğü ve güven aralığı gibi parametreler dikkate alınarak durumdan duruma değişkenlik gösterecektir. Teknik olarak olabildiğince çok sayıda ve homojen; ayrıca araştırılan konuyla ilişkili değişkenleri yansıtan bir örneklem olması gereklidir. Öte yandan, örneklem hesaplaması için kullanılabilecek program sunan birçok internet sitesi bulunmaktadır.¹⁴

Nitel etki değerlendirme kısmında bahsedildiği şekilde örneklem seçiminde araştırmacı olasılıklı ya da olasılıklı olmayan örnekleme yöntemlerinden faydalanabilir.

Adım 8: Veri Toplama¹⁵

Araştırma sırasında gerçek/tüzel kişiler hakkında birincil (anketler, saha gözlemleri) ve ikincil verinin (idari kayıtlar, veri tabanları) toplanması, yardım etkilerinin istatistiksel olarak ayrıştırılabilecek, müdahale ve kontrol grupları arasındaki farkların doğru biçimde modellenmesine izin verecek bir veri yapısı oluşturulmasını sağlayacaktır. Veri mevcudiyeti ve kalitesi de yardım etkilerini değerlendirmenin ayrılmaz bir parçasıdır.

¹³ Bir veri grubundaki gözlemlerin ortalama değerden ne ölçüde sapma gösterdiğini ifade eden temel bir istatistiksel ölçüdür, verilerin birbirine ne kadar yakın ya da ne kadar dağınık olduğunu gösterir. Etki değerlendirme çalışmalarında varyans, incelenen değişkenin ne kadar heterojen bir yapıya sahip olduğunu anlamak açısından önem taşımakta ve örneklem büyüklüğünün belirlenmesinde doğrudan rol oynamaktadır. Varyans değeri büyüklüğü, veri dağılımının daha geniş aralıkta olduğunu göstermektedir.

¹⁴ Avustralya İstatistik Bürosu bunlardan birisidir. Bkz. <https://www.abs.gov.au/websitedbs/D3310114.nsf/Home/Sample+Size+Calculator+Examples?opendocument>

¹⁵ Etki değerlendirmede kullanılacak veriler ile veri toplamada dikkat edilmesi gereken hususlar hakkında detaylı bilgiler için “Etki Değerlendirme İçin Veri Yönetimi” başlıklı bölüme bakınız.

Bazı destek türlerinin etkisi ancak uzun vadede ortaya çıkmaktadır. Özellikle Ar-Ge ve inovasyon desteklerinde, yenilik faaliyetlerinin doğası gereği geleceğe ilişkin belirsizlikler bulunduğundan, yapılan analizlerin gerçek sonucu ancak yıllar içinde tam olarak yansıtabileceği değerlendirilmektedir. Bu nedenle, bu tür desteklerin başarısını ölçmek için, yardımın hedeflerine uzun vadede ulaşip ulaşmadığını gösterebilecek uygun ve sürdürülebilir performans göstergelerinin belirlenmesi (örneğin Ar-Ge harcaması, yatırım harcaması vb.) önem taşımaktadır.

Etki değerlendirmede amaç ekonomik ve ticari davranışların karmaşıklığını gerçeğe yakın nispeten basit bir model kullanarak tespit etmektir. Kurulan model ile yardımın amacını yansıtan bağımlı sonuç değişkeni üzerindeki birçok etkinin tamamını temsil etmek mümkün olmamakla beraber en etkili değişkenler seçilmeye çalışılır. Bu durum hangi verinin toplanabilir olacağına da bağlıdır.

Veri toplama sürecinde yalnızca müdahale değil kontrol grubuna ilişkin verilere de ihtiyaç duyulacağı göz önünde bulundurulmalıdır. Bu nedenle bir yardım programına ilişkin veri toplama aşamasında sadece yararlanıcılar bazında değil başvurup yararlanamayan veya kriterlere uyup başvurmayan firmalara ait verinin de depolanması fayda sağlayacaktır. Kontrol gruplarının oluşturulması için bu konuya en başından özen gösterilmesi gerekmektedir. Aksi takdirde, uygulanan yardımların sadece çıktı bazlı pozitif etkisi olduğunu savunmak, o yardımın etkili olduğunu göstermeyecektir. Her yardım uygulandığında alternatif maliyetlerin de olduğu unutulmamalı ve bu sebeple etkin bir izleme sistemi kurulmalıdır.

Adım 9: Veri Analizi

Yardımların nicel değerlendirmesi için deneysel ve deneysel olmayan yöntemler kullanılmaktadır. Deneysel yöntemlerde müdahale ve kontrol gruplarına yapılan rastgele seçimden ötürü yardım başlamadan önce iki grup ortalama olarak benzer olacaktır. Bu durumda araştırmacı iç geçerlilik sağlama konusunda daha şanslıdır. Ancak çeşitli nedenlerle, rastgeleleştirme her zaman mümkün olmaz veya kullanılmak istenmeyebilir.

Deneysel olmayan (yarı-deneysel, *quasi-experimental*) yöntemlerde¹⁶ ise yardımdan yararlanan bir müdahale grubu ve yardımdan yararlanmayan eşdeğer bir kontrol grubu şeklinde gözlemsel veri üzerinden karşı-olgusal etki değerlendirme yapılmaktadır.

Bu kapsamda aşağıdaki şekilde ifade edilen ve yazında en sık kullanılan deneysel ve deneysel olmayan yöntemler aşağıda yer almakta olup bu yöntemler bir sonraki başlıkta detaylı bir şekilde açıklanmaktadır.

- Farkların Farkı
- Sentetik Kontrol
- Eğilim Skoru Eşleştirme
- Araç Değişkenler

¹⁶ Deneysel yöntemlerin aksine deneysel olmayan yöntemlerin her biri, doğası gereği yardımın hedefi ve katılımı konusunda rastgele dağıtım olmadığından potansiyel seçim yanlılığına sahiptir ve yardım etkisini belirlemek için bu yanlılıkları göz önünde bulundurarak varsayım ve koşullara uygun olan yöntem ya da yöntemlerin kullanılması gerekecektir.

- Regresyon Süreksizliği Tasarı
- Rassal Kontrollü Deneyler

Adım 10: Ön Rapor

Analiz tamamlandıktan sonra varılan sonuçların raporlanması gerekir. Raporlanırken dikkat edilmesi gereken bazı hususlar bulunmaktadır.

Bunlardan ilki, kullanılan yöntemin neden tercih edildiğinin açıklanmasıdır.

İkinci olarak, etki değerlendirmenin amacının ve nasıl planlandığının açıklanması uygun olacaktır. Yukarıda sayılan dokuz adımın nasıl ve niçin o şekilde planlandığının, etki değerlendirme sonucunda ne elde edilmek istendiğinin değerlendirme sonuçlarına geçilmeden önce açıklanması uygun olacaktır.

Üçüncü olarak, planlanan faaliyetlerin nasıl gerçekleştirildiği açıklanmalıdır. Diğer bir ifadeyle dokuz adımın nasıl yapıldığının açıklanması uygun olacaktır.

Dördüncü olarak, yapılan etki değerlendirme sonucunda elde edilen bulgular, grafik veya tablolarla detaylı bir şekilde sunularak yardımın nasıl bir etki doğurduğu ve nedenleri açıklanmalıdır. Ayrıca kullanılan yöntemin gerekliliklerinin nasıl sağlandığının gösterilmesi gerekmektedir.

Beşinci olarak, etki değerlendirme uygulanacak, uygulanmakta olan veya uygulaması tamamlanmış bir yardımla ilgili olacağından, zamanlamaya uygun olarak çözüm önerileri sunulmalıdır. Örneğin ilk defa uygulanması planlanan bir yardımla ilgili olarak, yapılan çalışma bir kanıt veya gerekçe niteliği taşıyacağından, yardım teklifinin bu bulguları dikkate alarak tasarlanması gerekmektedir.

Etki değerlendirmeye konu olan uygulaması devam eden bir yardım ise, bulgular yardımın devamı, değiştirilmesi, geliştirilmesi veya sonlandırılması için bir gerekçe olabilecektir. Uygulaması tamamlanmış yardımlarda elde edilen bulgular ise, gelecekte uygulanabilecek veya başka alanda hâlihazırda uygulanmakta olan yardımlarla ilgili fikir vererek daha başarılı yardım veya yardım programları geliştirilmesine katkı sağlayabilecektir. Bu nedenle rapor sonunda gerekçeleriyle birlikte önerilere yer verilmesi yararlı olacaktır.

Raporda bu hususlara yer verilmesi değerlendirmenin güvenilirliğinin raporun sunulacağı kişi veya kurumlarca anlaşılması açısından önemlidir.

Adım 11: Ön Rapor Hakkında Paydaş Görüşlerinin Alınması

Etki değerlendirme, zaman içindeki sonuçlara bakarak, uygulanan yardımların insanların yaşamlarını nasıl değiştirdiğini takip edebilir ve istenen/ istenmeyen sonuçları ortaya çıkarabilir. Hazırlanan ön raporlar yardımların etkileri konusunda paydaşları bilgilendirmek ve görüşlerini almak için paylaşılmalıdır.

Ayrıca uygulayıcı kurumlar başarılı olduklarından emin olmak için de bu yolu kullanabilir. Böylelikle etkiyi nicelleştirmekten ziyade her bir katılımcının hikâyesinin de anlaşılır bir şekilde sunulması, değişimin nedensel süreçleri hakkında daha fazla kanıt

toplanmasını sağlayarak etki değerlendirme raporlarına zenginlik katacaktır. Bu sebeple nihai değerlendirme sonuçlarının güvenilir olarak değerlendirilmesini sağlamak için, ön raporun gözden geçirme sürecine paydaşların dâhil edilmesi iyi bir uygulama olacaktır.

Adım 12: Nihai Rapor

Ön raporda bahsedilen adımların ardından paydaşların görüşlerinin de alınmasıyla birlikte nihai raporlar oluşturulmalıdır.

Bulguların, değerlendirmelerin ve görüşlerin detaylı olarak yer aldığı nihai raporlar, sonuçların zaman içinde değerlendirilmesine izin verdiği için izleme ve değerlendirme alanında önemli bir araç olacaktır. Bu veri gelecekteki devlet yardımı tekliflerini değerlendirmek ve planlamak için de kullanılabilir.

3.2. Farkların Farkı Yöntemi

Farkların Farkı (FF) Yöntemi, bir müdahalenin nedensel etkisini, müdahaleyi alan bir grup (müdahale grubu) ile müdahaleyi almayan (kontrol grubu) arasında müdahale sonrası ve öncesi dönemdeki değişimleri karşılaştırarak tahmin eden yarı deneysel bir etki değerlendirme yöntemidir.¹⁷ Rassal atamanın başka bir deyişle randomizasyonun mümkün olmadığı durumlarda nedensel etkinin belirlenmesi için kullanışlı bir yöntem olduğu değerlendirilmektedir.

FF Yöntemi ile müdahale ve kontrol grupları için yardımın uygulanmasından öncesine ve sonrasına ilişkin verilerin mevcut olması durumunda yardımın etkisini yanlış olmayan (*unbiased*) şekilde tahmin etmek mümkündür. Bunun için, ilk olarak sonuç değişkeninin yardım uygulanmadan önce müdahale ve kontrol grupları için aldığı ortalama değerlerin arasındaki fark bulunur. İkinci olarak bu farkın yardım uygulandıktan sonra aldığı değer bulunur. Son olarak ise bu iki fark arasındaki fark hesaplanır ve bu fark yardımın etkisinin bir tahmini olarak ortaya konulmaktadır (Glewwe ve Todd, 2022).

FF, temelde müdahale ve kontrol gruplarının yardım uygulanmadan önce birbirlerinden farklı olmasına izin veren bir yöntemdir. Bu özelliği ile yardımın etkisinin belirlenmesinde basit bir katılımcı ve katılımcı olmayan kıyaslamasına nazaran daha isabetli sonuçlar elde edilmesini sağlar.

Söz konusu yöntem temel olarak *eğer müdahale gerçekleşmemiş olsaydı sonuçlardaki ortalama değişimin müdahale ve kontrol grupları için aynı olacağı anlamına gelen paralel eğilim varsayımına*, dayanmaktadır. Diğer bir deyişle, müdahale ve kontrol gruplarının paralel bir eğilim göstermekte olduğu varsayımı altında, müdahale ve kontrol grupları arasında yardım uygulanmadan önce var olan fark ile yardım uygulandıktan sonra var olan farkta ortaya çıkacak herhangi bir değişikliğin, yardımın etkisiyle gerçekleştiği sonucuna ulaşılmaktadır (Polat ve Aktakke, 2017).

Yukarıda bahsedilen açıklamalara ilişkin formül şu şekildedir:

Farkların Farkı Tahmini = Müdahale Sonrası Fark – Müdahale Öncesi Fark

¹⁷¹⁷ <https://erc.undp.org/methods-center/methods/evaluation-methods/differenceindifferences>

Bu formül, (1) no'lu eşitlikte gösterildiği şekilde yazılabilir.

$$\beta_{FF} = (E(Y_{\text{müdahale}}^{T=1} - Y_{\text{kontrol}}^{T=1})) - (E(Y_{\text{müdahale}}^{T=0} - Y_{\text{kontrol}}^{T=0})) \quad (1)$$

$\underbrace{\hspace{10em}}$
I. Fark

$\underbrace{\hspace{10em}}$
II. Fark

Y: Sonuç değişkeni (Örneğin ihracat miktarı, ciro ya da üretim miktarı gibi gözlemlenen bağımlı değişken)

$Y_{\text{müdahale}}^{T=t}$: Y'nin t zamanında müdahale grubundaki gözlemler için aldığı ortalama değer

$Y_{\text{kontrol}}^{T=t}$: Y'nin t zamanında kontrol grubundaki gözlemler için aldığı değer

I. Fark: Müdahale sonrası müdahale ve kontrol gruplarının Y değeri arasındaki fark

II. Fark: Müdahale öncesi müdahale ve kontrol gruplarının Y değeri arasındaki fark

β_{FF} (Farkların Farkı): Elde edilmek istenen yardım etkisinin tahmin edilen değeri işte bu iki fark arasındaki farktır. *I. Terim – II. Terim = Müdahale Etkisi.*

Yardımanın etkisinin doğru tanımlanabilmesi *paralel eğilim* varsayımına bağlıdır. Bazı durumlarda bu varsayım geçerli olmayabilmektedir. Kontrol ve müdahale gruplarının Y için aldığı ortalama değerlerin müdahaleden önce paralel eğilime sahip olmadığı durumlarda kontrol ve müdahale gruplarının çıktıların farklarındaki değişim sadece müdahaleye/devlet yardımına yüklenemez. Dolayısıyla bu varsayım geçerli olmadığında, tahmin edilen yardım etkisi değerinin yanlış olma ihtimali ortaya çıkacaktır (UNEG, 2013).

Özellikle sıralı lojistik regresyon modelleri başta olmak üzere Farkların Farkı yönteminin kullanılıp kullanılmayacağına karar verilebilmesi için bağımlı değişkenin kategorileri ile bağımsız değişkenler arasında paraleleğilim olup olmadığına bakılmalıdır (Hayat ve Özden, 2021). Paralel eğilim varsayımı, modeldeki katsayıların bağımlı değişkenin kategorilerinde sabit olup olmadığını ifade etmektedir. Bu varsayımın sağlanmaması, katsayıların yanlış tahmin edilip yanlış yorumlanması açısından önem arz etmektedir. Bu nedenle paralel eğilim varsayımının test edilmesi gerekmektedir. Bu varsayımda, bağımlı değişken ile bağımsız değişkenler arasındaki ilişki bağımlı değişkenin kategorilerine göre değişiklik göstermez.

Varsayımın kontrolüne ilişkin hipotez¹⁸:

H_0 : Modelin regresyon katsayıları, bağımlı değişkenin tüm kategorilerinde aynıdır.

$$(H_0 : \beta_1 = \beta_2 = \dots = \beta_{j-1} = \beta)$$

H_1 : Modelin regresyon katsayıları, bağımlı değişkenin en az bir kategorisinde farklıdır.

$$(H_1 : \text{En az bir } \beta_j \text{ farklıdır.})$$

¹⁸ İstatistiksel hipotez, kitleye ait parametrelere ilişkin kurulan önermelerdir. β katsayıları kitleyi temsil eden regresyon katsayılarıdır.

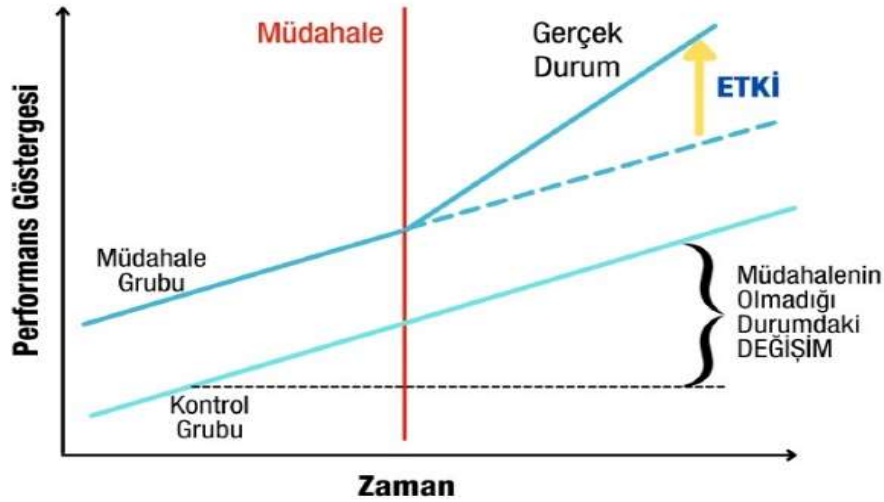
j: Bağımlı değişkene ait kategori sayısı¹⁹

olmak üzere uygun istatistiksel yöntemle test edilir. Test sonucu eğer katsayılar birbirine eşit ise paralel eğilim varsayımı sağlanmaktadır. Eşitsizlik söz konusu ise varsayım sağlanmamaktadır (Aydos ve Toker, 2021).

Paralel eğilim varsayımını test etmenin bir yolu, sonuç değişkeninin müdahale öncesi dönemde kontrol ve müdahale grupları için büyüme oranına bakmaktır. Bu oranlar iki grup için istatistiki olarak anlamlı bir şekilde birbirinden farklı değilse, paralel eğilim varsayımı geçerlidir. Aksi takdirde, paralel eğilim varsayımı geçersiz sayılır ve bu durumda *Farkların Farkı* yöntemi ile yapılan analiz yanlı tahminler verecektir. Paralel eğilim, kontrol ve müdahale grupları için müdahale öncesi sonuç değişkeninin zamana karşı grafiği çizilerek de test edilebilir. Ancak testin bu şekilde gerçekleştirilip sonuç değişkeninin hangi hızla büyüdüğünün hesaplanabilmesi için yardım uygulanmadan önceki döneme ait en az iki noktaya ait verinin olması gerekmektedir (The World Bank, 2010).

Belirli bir yardıma ilişkin etki değerlendirmeye yönelik Farkların Farkı yönteminin şematik gösterimine Grafik 2’de yer verilmiştir.

Grafik 2. Müdahale Sonrası Etki Durumu (Farkların Farkı Yöntemi için)



Kaynak: Çukurova Kalkınma Ajansı, 2010-2016 Mali Destek Programları Etki Değerlendirme Raporu, (2021)

Müdahale olmadığı durumda müdahale grubunun eğilimi ile kontrol grubununki aynı olmalı varsayımından hareketle uygulanan bu yöntemin güçlü yanı zaman içinde değişmeyen, gözlemlenebilen veya gözlemlenemeyen faktörlerin etkisini kontrol etmesi ve düşük bütçeyle uygulanabilmesidir.

¹⁹ Kategori sayısı, kategorik değişkenlerde olmaktadır. Kategorik değişken ise bir grup içinde nominal (örn. müdahale almış = 1 ve almamış = 0) ya da ordinal yani sıralı (örn. Eğitim=ilkokul, ortaokul, lise, üniversite) şekilde sınıflandırılabilen veridir. Sıralı lojistik regresyonda, adından anlaşılacağı üzere bağımlı değişken sıralı olarak gruplandırılmaktadır ve en az üç kategori sayısından oluşmaktadır.

Kutu 2. Asgari Ücret Değişiminin İstihdam Üzerindeki Etkisinin Ölçümü

Card ve Krueger tarafından gerçekleştirilen "*Minimum Wages ve Employment: A Case Study of the Fast-Food Industry in New Jersey ve Pennsylvania*" başlıklı çalışmada; 1 Nisan 1992'de yapılan yasal bir düzenleme ile New Jersey'de asgari ücretin saatlik 4.25 dolardan 5.05 dolara yükseltilmesinin etkisi Farkların Farkı yöntemi ile incelenmiştir. Bununla birlikte, adı geçen eyaletin komşularından olan Pensilvanya'da ise söz konusu dönemde saatlik asgari ücret değişmemiş ve 4,25 \$ olarak kalmıştır. Araştırmacılar tarafından hazır yemek sektöründe faaliyet gösteren toplam 410 firmadan veri toplanmıştır.

Çalışmada müdahale grubunu New Jersey eyaletindeki hazır yemek firmaları oluştururken kontrol grubunu Doğu Pensilvanya'daki hazır yemek firmaları oluşturmuştur. Farkların Farkı yönteminde zaman farkını oluşturmak için uygulama öncesi dönemi Şubat 1992, uygulama sonrası dönemi ise Kasım 1992 olarak belirlenmiştir. Çalışmada New Jersey'deki fast-food restoranlarında ortalama başlangıç ücretinin, asgari ücretteki artışı takiben yüzde 10 oranında arttığı belirtilmiş ve bu değişikliğe dair görsel amaçlı program öncesi ve sonrası iki eyaletteki başlangıç ücretlerinin dağılımlarını gösteren şekiller "Başlangıç Ücret Artış Oranlarının Dağılımı" başlıkla aşağıdaki gibi sunulmuştur.

Artış öncesinde, Pennsylvania'daki mağaza başına ortalama istihdam 23.3 tam zamanlı eşdeğer çalışan iken, New Jersey'de bu sayı ortalama 20.4 olarak verilmiştir. İki eyaletteki mağazalar arasında başlangıç ücretleri çok benzer görünmesine rağmen New Jersey'de bir "tam öğün"ün (orta boy soda, küçük boy patates kızartması ve bir ana yemek) ortalama fiyatının önemli ölçüde daha yüksek olduğu belirtilmektedir. İşletme saatleri ortalaması, tam zamanlı çalışanların oranı veya yeni çalışan çekmek için bonus programlarının yaygınlığı açısından ise bu iki eyalet arasında önemli farkların bulunmadığı ve benzer olduklarından hareketle, New Jersey eyaletinde hazır yemek sektöründe asgari ücret artışın "uygulama etkisi" tahmin edilmeye çalışılmıştır.

Farkların Farkı yöntemiyle asgari ücretteki değişim nedeniyle istihdamdaki değişimin tahmini için kullanılacak model şu şekildedir:

$$\hat{y} = (IST_{NJ\text{ SONRA}} - IST_{PA\text{ SONRA}}) - (IST_{NJ\text{ ÖNCE}} - IST_{PA\text{ ÖNCE}})$$

Değişken	Pensilvanya (PA)	New Jersey (NJ)	Fark NJ - PA
İstihdam, Önce	23,3312 (1,35)	20,4394 (0,51)	-2,8918 (1,44)
İstihdam, Sonra	21,1656 (0,94)	21,0274 (0,52)	-0,1382 (1,07)

Kaynak: Card ve Krueger, 1994, s. 780.

Model tahmininden elde edilen sonuçlar denkleme uygulandığında:

$$\hat{y} = (21,0274 - 21,1656) - (20,4394 - 23,3312)$$

$$\hat{y} = 2,7536$$

Not: İST: istihdam, PA: Pensilvanya, NJ: New Jersey, ÖNCE: Uygulama öncesi dönem, SONRA: Uygulama sonrası dönem, $\hat{y}$: farkların farkı tahmin edicisi. Elde edilen sonuç tahmin edici $\hat{y}$ için yazılan denkleme uygulanmıştır.

Yapılan çalışma asgari ücret artışı sonucunda istihdamın teoride beklenilenin aksine azalmadığı sonucunu ortaya koymuştur.

Son yıllarda Farkların Farkı (Difference-in-Differences) yöntemi üzerine çok sayıda çalışma gerçekleştirilmekte olup yöntemin farklı senaryolarda nasıl uygulanabileceğine ilişkin çeşitli yaklaşımlar geliştirilmektedir. Bu çalışmaların birinde, yöntemin uygulanmasına yönelik araştırmacılar için pratik bir kontrol listesi sunulmuş; söz konusu kontrol listesine bu çalışmanın sonunda EK-1'de yer verilmiştir.

3.2.1. Farkların Farkı Yönteminde İzlenmesi Gereken Adımlar

Bu bölümde Farkların Farkı yöntemiyle etki değerlendirme çalışması gerçekleştirilirken aşağıdaki adımların takip edilmesi ve hazırlanacak raporlarda bu adımlara ilişkin bilgi verilmesi büyük önem taşımaktadır.

- Veri setinin hazırlanması,
- Veri setinin kontrolü,
- Modele ilişkin varsayımların kontrolü,
- Tahmin modelinin kurulumu,
- Tahmin edilen modellerin karşılaştırılması

Söz konusu adımlara ilişkin detaylar ile her bir adıma ilişkin R programında yapılacak uygulamalar için örnek kodlar aşağıda sırayla paylaşılmaktadır.

Adım 1: Veri Setinin Hazırlanması

Farkların Farkı yönteminde, müdahale ve kontrol gruplarının müdahale sonrası ve müdahale öncesindeki dönemde farklarının farkı hesaplanarak uygulanan müdahalenin etkisi ölçülmektedir (Sant'Anna, 2025). Bu çerçevede, söz konusu yöntem kullanılarak etkinin ölçülmesi aşağıdaki denklemlerde gösterilmektedir (Wooldridge, 2010).

$$\text{Müdahale Grubunun Farkı} = \text{Müdahale Sonrası} - \text{Müdahale Öncesi}$$

$$\text{Kontrol Grubunun Farkı} = \text{Müdahale Sonrası} - \text{Müdahale Öncesi}$$

$$\text{Farkların Farkı} = \text{Müdahale Grubunun Farkı} - \text{Kontrol Grubunun Farkı}$$

Paylaşılan denklemlerden de anlaşılacağı üzere Farkların Farkı yönteminin kullanılabilmesi için veri setinde müdahale öncesi ve sonrası dönem ile müdahale ve kontrol gruplarının birbirlerinden ayırt edilebiliyor olması gerekmektedir. Bu kapsamda öncelikle devlet yardımıyla yararlanan gözlem birimlerinden (firmalardan, kişilerden vb.) oluşan **müdahale grubunun** belirlenmesi ve veri setinde yer alması gerekmektedir. Buna ek olarak müdahale grubu ile benzer özelliklere sahip olan ve devlet yardımıyla yararlanabilecek nitelikte olmakla birlikte devlet yardımıyla yararlanmayan gözlemlerden oluşan **kontrol grubunun** oluşturulması ve veri setinde yer alması gerekmektedir. Eğer etki değerlendirme çalışması için müdahale ve kontrol grupları oluşturulamıyorsa o zaman Farkların Farkı yönteminden vazgeçilmeli ve alternatif bir etki değerlendirme yönteminden yararlanılmalıdır.

Farkların Farkı yöntemiyle devlet yardımının etkisinin ölçülebilmesi için veri setinde **müdahale öncesi** ile **müdahale sonrası** zamanlara ait her bir gözlem birimine ait öncelikle bağımlı değişkene ilişkin verilerin yer alması gerekmektedir. Öncelikle bağımlı değişken denilmesinin temel nedeni, uygulamanın etkisinin ölçülmesi için bağımlı değişkene ihtiyaç duyulmasıdır. Bu çerçevede, bağımlı değişkene ek olarak bağımsız değişken ya da değişkenlere ait müdahale öncesi ve sonrası döneme ait verilere de veri setinde yer verilmesi mümkündür. Veri setinde müdahale ve kontrol grubunu birbirinden ayıran değişkene ek olarak, müdahale öncesi ve müdahale sonrası dönemi birbirinden

ayıran bir değişkene yer verilmesi gerekmektedir. Söz konusu tanımlamalar yapıldıktan sonra tahmin edilecek temel denklem aşağıda gösterilmektedir.

$$Y_{i,t} = \beta_0 + \beta_1 * Grup_t + \beta_2 * Zaman_t + \beta_3 * (Grup_t * Zaman_t)$$

Bahsedilen verilere ek olarak veri setinde gözlem birimlerine ait ilave değişkenlere yer verilmesi önem arz etmektedir. Zira, gözlem birimlerine ait ortak özellikler gözlemlenemeyen karıştırıcıların hesaba katılmasına yardımcı olabilmekte ve böylelikle tutarlılığı artırmaktadır (Farrell, 2015). İlerleyen aşamalarda da görüleceği üzere eğer paralel eğilim varsayımı doğrulanamazsa, yapılacak eşleştirme için gözlem birimlerinin gözlemlenebilir özelliklerini yansıtan değişkenlere ihtiyaç duyulacaktır. Bu çerçevede, eşleştirme yapılabilmesi için veri setinde söz konusu değişkenlerin yer alması önem arz etmektedir. Paralel eğilim varsayımının doğrulanamaması durumunda gözlemlenebilir özellikler bakımından eşdeğer gözlemler için yapılacak eşleştirme sonrasında paralel eğilim varsayımı yeniden kontrol edildikten ve varsayım doğrulandıktan sonra Farkların Farkı yöntemi kullanılabilir (Caetano, 2024). Temel denklemle yapılacak tahmin sonrasında denkleme gözlemlenebilir özellikleri yansıtan bağımsız değişkenlerin eklenmesinin, yapılacak tahminin tutarlılığını etkilemeden hassasiyetini artıracak değerlendirilmektedir (Imbens, 2009). Buna ek olarak, söz konusu değişkenlerin analize dahil edilmesinin sapmayı azaltacağı ve sonuçları sağlamlaştıracağı belirtilmektedir (Ghanem, 2023). Ayrıca gözlemlenebilir özellikleri yansıtan bağımsız değişkenler, uygulanan devlet yardımının müdahale grubunun farklı alt grupları üzerindeki etkisinin belirlenmesinde de kullanılabilir (Abadie, 2005). Buna ek olarak yapılacak sağlık testlerinde de alternatif modellerin kurulması mümkün olmaktadır.

Söz konusu gözlemlenebilir özellikleri yansıtan bağımsız değişkenler seçilirken dikkatli olunması gerekmektedir. Zira, seçilecek değişkenlerin bağımlı değişken arasında sistematik ve nedensel olmayan korelasyonlar oluşturma riski bulunmaktadır. Bu durum, tahmin edilen etkinin sapmalı olmasına yol açabilecektir. Bu nedenle söz konusu değişkenler, devlet yardımından yararlanma durumunu ve bağımlı değişkeni etkileyen özellikler arasından seçilmelidir. Devlet yardımından yararlanma için yapılacak başvurularda aranan kriterlerin bu kapsamda değişken olarak seçilebileceği değerlendirilmektedir. Ayrıca, değişkenler seçilirken araştırılan konuya ilişkin teori ve literatürden yararlanılması önerilmektedir (Chen G., 2024).

Adım 2: Veri Setinin Kontrolü

Etki değerlendirme aşamasına geçilmeden önce oluşturulan veri setinin kontrol edilmesi gerekmektedir. Bu kapsamda, öncelikle veri setinde yer alması gereken temel değişkenler üzerinde durulacaktır.

- Veri setinde yer alan her gözlem birimini (firma, kişi vd.) benzersiz şekilde tanımlayan bir değişken (firma_ID, kişi_ID vb.) bulunmalıdır.
- Her bir gözlemin ait olduğu dönemi (ay, çeyrek, yıl vb.) tanımlayan bir değişken yer almalıdır. Bu değişken paralel eğilim varsayımı bakımından önem arz etmektedir.
- Müdahalenin etkisinin görüleceği değişken, başka bir deyişle bağımlı değişken (yatırım harcaması, log_istihdam, vb.) veri setinde yer almalıdır.

- Gözlem biriminin müdahale grubunda olup olmadığını temsil eden kukla değişkenin (örneğin $müdahale = 1$, $kontrol = 0$) bulunması gerekmektedir.
- Müdahale öncesi ile müdahale sonrası dönemi birbirinden ayıran kukla değişkenin (örneğin sonrası = 1, öncesi = 0) veri setinde yer alması gerekmektedir.

Söz konusu değişkenler, "uzun format" (long format) adı verilen bir yapıda veri setinde yer alması tercih edilmektedir. Zira, uzun format, her satır tek bir gözlem biriminin tek bir zaman noktasındaki değerleri yer almaktadır. Uzun formata sahip örnek veri seti aşağıdaki tabloda gösterilmektedir.

Tablo 5. Örnek Veri Formatı

Firma_ID	Gözlem Birimi	Yıl	Grup	Zaman	Y (Bağımlı Değişken)	İmalat	Ölçek
1	A	2023	1	0	100	1	10-49
1	A	2024	1	0	110	1	10-49
1	A	2025	1	1	132	1	10-49
2	B	2023	0	0	90	0	1-9
2	B	2024	0	0	100	0	1-9
2	B	2025	0	1	110	0	1-9
3	K	2023	0	0	85	1	10-49
3	K	2024	0	0	98	1	10-49
3	K	2025	0	1	117	1	10-49
4	R	2023	1	0	110	0	1-9
4	R	2024	1	0	120	0	1-9
4	R	2025	1	1	135	0	1-9

Veri setinde yer alan gözlem birimlerinin dönem itibarıyla tekil olup olmadığı kontrol edilmelidir. Eğer tekrarlanan veri varsa bunların veri setinin dışında bırakılması gerekmektedir. Söz konusu tekillik kontrolü için aşağıda yer alan örnek R kodlarından yararlanılması mümkündür. Bu kapsamda değişken isimlerinin değiştirilmesi gerektiği unutulmamalıdır.

```
is_panel_unique <- data %>%
  count(firma_ID, yıl) %>%
  filter(n > 1)

if (nrow(is_panel_unique) > 0) {
  print("Tekrarlanan gözlemler var!")
  print(is_panel_unique)
} else {
  print("Veriler benzersiz.")
}
```

Veri setinde yer alan müdahale ve kontrol gruplarının doğru bir şekilde tanımlandığından emin olunmalıdır. Bunun için aşağıda yer alan R kodları, değişken isimlerine dikkat edilerek kullanılabilir. Aşağıda yer alan kod ile önce her bir ID'nin hangi grupta (müdahale grubu, kontrol grubu) yer aldığı ve sonra söz konusu gruplarda kaç farklı gözlem olduğu belirlenmektedir. Sonrasında her bir ID'nin dönemler itibarıyla

aldığı değerler kontrol edilerek grup değiştiren gözlem olup olmadığı araştırılmaktadır. Son olarak da eğer grubu değişken gözlem varsa bunun hakkında bilgi verilmektedir.

```
# Müdahale ve kontrol gruplarındaki gözlem sayısını kontrol etmek için
grup_sayilari <- data %>%
  distinct(Firma_ID, Grup) %>%
  count(Grup)

print("Gruplardaki benzersiz gözlem sayıları:")
print(grup_sayilari)

# Müdahale grubunun zamanla değişmediğinden emin olmak için
grup_degisimi <- data %>%
  group_by(Firma_ID) %>%
  summarise(grup_sayisi = n_distinct(Grup)) %>%
  filter(grup_sayisi > 1)

if (nrow(grup_degisimi) > 0) {
  print("Bazı birimler zamanla grup değiştirmiş!")
  print(grup_degisimi)
} else {
  print("Tüm birimler analiz dönemi boyunca aynı grupta kalıyor.")
}
```

Müdahale öncesi ve sonrasını birbirinden ayıran değişkenin doğru bir biçimde tanımlanıp tanımlanmadığı kontrol edilmelidir. Eğer söz konusu dönemleri ayıran bir değişken yoksa bunun veri setine eklenmesi, eğer veri setinde tanımlanmış bir değişken varsa bunun da kontrol edilmesi önem arz etmektedir. Aşağıda yer alan kodlar kullanılarak kontrol edilebilir.

```
# Müdahale yılının 2025 olduğu kabul edilmiş ve kontrol için yeni bir değişken tanımlanmıştır
data <- data %>%
  mutate(Zaman_kontrol = ifelse(yil >= 2025, 1, 0))

# Veri setinde yer alan "zaman" değişkeni ile kontrol için eklenen değişken arasında fark var mı?
if (!all(data$Zaman == data$Zaman_kontrol, na.rm = TRUE)) {
  print("Zaman değişkeninde tutarsızlık var!")
} else {
  print("Zaman değişkeni doğru kodlanmış.")
}

# Her grupta hem müdahale öncesi hem de sonrası gözlem olduğundan emin olmak için
donem_kontrolu <- data %>%
  group_by(Grup) %>%
  summarise(
    onemli_donem_sayisi = n_distinct(Zaman))

print(donem_kontrolu)
```

Eğer müdahale öncesi ve sonrasında yanlış bir tanımlama varsa düzeltildikten sonra çalışmaya devam edilmesi gerekmektedir.

Eksik veri, bir veri setindeki bazı bilgilerin ya da değişkenlerin bulunmaması olarak tanımlanabilmektedir (Zhou, 2024). Veri toplama ya da analiz sürecinde farklı nedenlerden dolayı eksik veriyle karşılaşılabilir. Bu çerçevede, veri setinde yer alan değişkenlerde, yıllarda, gruplarda eksik veri olup olmadığı kontrol edilmelidir.

```
# Değişkenlerdeki eksik veri sayısının kontrol edilmesi için
```

```
eksik_veri_sayisi <- sapply(data,  
                             function(x) sum(is.na(x)))  
  
print("Değişkenlerdeki eksik veri sayıları:")  
print(eksik_veri_sayisi)
```

Eksik veriler bir çalışmanın istatistiksel gücünü azaltabilmekte ve sapmalı tahminlere yol açarak geçersiz sonuçlara ulaşılmasına neden olabilmektedir (Kang, 2013). Ayrıca eksik verilere yol açan farklı mekanizmalar farklı çözümler gerektirdiğinden her duruma uygun standart bir çözüm bulunmamaktadır. Bu çerçevede eksik verisi fazla olan gözlem birimleri veri setinden atılması önerilmektedir. Zira imputasyon veri setinin yapısını bozacağından tercih edilmemelidir. Ayrıca, eksik veriler müdahale ve kontrol gruplarında sistematik şekilde farklılaşmamalıdır. Başka bir deyişle veri eksikliği belli gruplara veya belli dönemlere yoğunlaşmamalıdır. Eğer böyle bir durum söz konusuysa, veri setinin gözden geçirilmesi önerilmektedir.

Veri setinde kontrol edilecek bir diğer husus ise aykırı değerlerin olup olmadığıdır. Zira aykırı değerler tahmin sürecini etkileyerek değerlerin olduğundan daha yüksek ya da daha düşük tahmin edilmesine yol açabilmektedir (Kwak, 2017). Aykırı değer, bulunduğu örneklemin diğer birimlerinden belirgin şekilde sapma gösteren bir gözlem olarak tanımlanmaktadır (Grubbs, 1969). Bu çerçevede veri setinde aykırı değer olup olmadığı görselleştirme ya da istatistiksel testler ile belirlenebilir.

Görsel olarak aykırı değerlerin olup olmadığının belirlenmesi için kutu grafiğinden (box-plot) yararlanılabilir.

```
ggplot(data, aes(x = factor(Grup), y = Y)) +  
  geom_boxplot() +  
  facet_wrap(~Yil) +  
  labs(title = "Yıllara ve Gruplara Göre Sonuç Değişkeninin Boxplot'u",  
        x = "Grup Değişkeni", y = "Bağımlı Değişken")
```

Bunun yanında aykırı değerler Tukey tarafından geliştirilen interquartile range (IQR) yöntemiyle de tespit edilebilir.

Bu çerçevede, $Q1 = 1.$ çeyrek, $Q3 = 3.$ çeyrek, $IQR = Q3 - Q1$,

Alt sınır = $Q1 - 1,5 \cdot IQR$, Üst sınır = $Q3 + 1,5 \cdot IQR$ olarak tanımlanır (Tukey, 1977). Bu aralığın dışındaki gözlemler aykırı değer olarak tanımlanır. Daha önce belirtilen kutu grafiğin arkasındaki işleyiş de buradaki tanımlamayla aynıdır. 1,5 olarak kullanılan katsayı teorik bir zorunluluk değildir. Literatürde yaygın bir kullanıma sahip olduğu için seçilmiştir.

```
Q1 <- quantile(data$Y, 0.25, na.rm = TRUE)
Q3 <- quantile(data$Y, 0.75, na.rm = TRUE)
IQR <- Q3 - Q1
alt_sinir <- Q1 - 1.5 * IQR
ust_sinir <- Q3 + 1.5 * IQR

aykiri_degerler <- data %>%
  filter(Y < alt_sinir | Y > ust_sinir)

print("Aykırı değerler:")
print(aykiri_degerler)
```

Aykırı değerlerin belirlenmesinde z-skoru da kullanılabilir. Normal dağılım varsayımı altında ortalamadan 3 standart sapma uzak olan değerler, aykırı değer olarak belirlenmektedir.

```
z <- scale(data$Y)

aykiri_z <- data[abs(z) > 3, ]
aykiri_z
```

Adım 3: Varsayımların Kontrolü

Paralel eğilim varsayımı, eğer hiç müdahale uygulanmamış olsaydı, müdahale uygulanan grup ile uygulanmayan grup arasındaki farkın, müdahale öncesi dönemde olduğu gibi müdahale sonrası dönemde de aynı kalacağı düşüncesini içermektedir. Başka bir deyişle, eğer müdahale olmasaydı, sonuç değişkeninde zaman içindeki değişimlerin, müdahale ve kontrol gruplarının her ikisinde de tam olarak aynı olacağı varsayılmaktadır (Bertrand, 2004). Bu çerçevede, Farkların Farkı yönteminin kullanılabilmesi için paralel eğilim varsayımının kontrol edilmesi ve sağlanması gerekmektedir. Zira, eğer müdahale ve kontrol grupları için sonuç değişkeninin eğilimleri/trendleri farklıysa, Farkların Farkı yöntemiyle elde edilen müdahale etkisi tahmini geçersiz veya sapmalı olacaktır (Gertler, 2016). Bunun temel sebebi, kontrol grubuna ait eğilimin, karşıt olgusal eğilimin geçerli bir tahmini olmamasıdır. Bu noktada, kontrol grubunun seçimine dikkat edilmesi gerektiği hatırlanmalıdır. Zira, daha öncede ifade edildiği üzere kontrol grubu istatistiksel olarak müdahale grubunun eş değeri olmalıdır.

Paralel eğilim varsayımının kontrolü için müdahale öncesi dönemin esas alınması ve sonuç değişkeninin, başka bir deyişle bağımlı değişkenin kullanılması gerektiğine dikkat edilmelidir. Ayrıca, müdahale öncesindeki dönemde en az 3 yılın verisi kullanılmalıdır. Bu kapsamda, aşağıda paylaşılan R kodları ile paralel eğilim varsayımı görsel olarak test edilebilir.

```

veri_did %>%
  group_by(treat, year) %>%
  summarise(mean_log_investment = mean(log_investment, na.rm = TRUE)) %>%
  ggplot(aes(x = year, y = mean_log_investment, color = factor(treat))) +
  geom_line(size = 1.2) +
  geom_point(size = 2) +
  labs(color = "Grup",
        x = "Years",
        y = "Ortalama Log Investment") +
  theme_minimal()

```

Paralel eğilim varsayımına ek olarak Kararlı Birim Müdahale Değeri Varsayımı (Rubin, 1980) (the stable unit treatment value assumption, SUTVA) sağlanmalıdır. Söz konusu varsayım iki hususun sağlanmasına dikkat çekmektedir. İlk olarak, müdahale ve kontrol grupları arasında etkileşim olmamalıdır. Başka bir deyişle, bir firmaya verilen devlet yardımı, diğer firmaların davranışını/durumunu değiştirmemelidir. İkinci olarak, Müdahale tüm yararlanıcılar için aynı şekilde uygulanmalı, şartlar ve içerik değişkenlik göstermemelidir (VanderWeele, 2013).

Paralel eğilim varsayımı Farkların Farkı yöntemi için en kritik unsurdur. Bu varsayımın sağlanamaması durumunda, kontrol grubunun trendi müdahale grubunun karşı olgusal durumu olmaktan çıkmaktadır. Bu durumda elde edilen sonuçlar sapmalı hale gelmekte ve etki, olduğundan daha büyük ya da daha küçük tahmin edilebilmektedir. Bu nedenle, Farkların Farkı yöntemiyle yapılacak etki değerlendirme çalışmalarında paralel eğilim varsayımının doğrulanamaması durumunda koşullu paralel eğilim varsayımının sağlanması önerilmektedir. Başka bir deyişle, müdahale ve kontrol grupları gözlemlenebilir ortak özellikler bakımından benzer hale getirilerek paralel eğilim varsayımı sağlanmaktadır (Abadie, 2005). Başka bir deyişle, müdahale ve kontrol grubu gözlemlenebilir özellikler kullanılarak eşleştirilmekte ve sonrasında paralel eğilim varsayımının sağlanıp sağlanmadığı kontrol edilmektedir.

Eşleştirme, büyük bir potansiyel kontrol havuzundan örnekleme yaparak, ortak değişkenlerin dağılımının müdahale grubundaki dağılımla benzer olduğu, makul büyüklükte bir kontrol grubu oluşturma yöntemidir (Rosenbaum, The Central Role of the Propensity Score in Observational Studies for Causal Effects, 1983). Bu kapsamda, yapılacak eşleştirmede gözlemlenebilir ortak özellikler kullanılmalı ancak bağımlı değişkenin, başka bir deyişle sonuç değişkeninin kullanımından kaçınılmalıdır (Rubin, Using Propensity Scores to Help Design Observational Studies: Application to the Tobacco Litigation, 2001). Bu durum koşullu bağımsızlık varsayımı (conditional independence assumption, CIA) olarak da ifade edilmektedir (Rosenbaum, 1983).

$$(Y^0, Y^1) \perp\!\!\!\perp D \mid X$$

Varsayım, ortak değişkenlerin X kontrol edilmesi durumunda, müdahale ve kontrol gruplarının herhangi bir gözlenebilir ya da gizli yanlılıktan arındırılmış hâle geleceğini ifade etmektedir. Böylelikle, örneğin devlet yardımından yararlanma durumunun potansiyel sonuçlardan bağımsız olduğu varsayılmakta ve hangi firmanın veya kişinin devlet yardımından yararlanacağı rassal gibî hareket edilebilmektedir. Gözlemlenebilir ortak özelliklerin, başka bir deyişle devlet yardımından yararlanma durumunu ve bağımlı değişkeni etkileyen gözlem birimlerinin özelliklerinin eşleştirmede

kullanılmasının sapmalı sonuçları önlediği farklı çalışmalarda da ortaya konulmuştur (Ham, 2024).

Eşleştirme yapılırken müdahale öncesi döneme ait veriler kullanılmalıdır (Rosenbaum, Constructing a Control Group Using Multivariate Matched Sampling Methods That Incorporate the Propensity Score, 1985). Bu çerçevede, birden fazla müdahale öncesi döneme ait veri bulunması durumunda müdahaleye en yakın olan dönem, eşleştirme için tercih edilebilir.

Eşleştirme için aşağıda belirtilen yöntemlerden yararlanılabilir:

1. Eğilim Skoru Eşleştirme (Propensity score matching),
2. Mahalanobis mesafe eşleştirmesi (Mahalanobis distance matching),
3. Tam eşleştirme (Exact matching),
4. Kabalaştırılmış Tam Eşleştirme (Coarsened exact matching).

Eşleştirmede kullanılacak değişkenlerin seçiminde teori ve literatürden yararlanılmalıdır (Imbens G. W., 2004). Buna ek olarak, sadece bağımlı değişkeni etkileyen değişkenler ile müdahale sonrasında gözlemlenebilen değişkenlerin eşleştirmede kullanılmaması önerilmektedir. Ayrıca çok fazla sayıda değişkenin eşleştirmede kullanılmasının sorun oluşturacağı ve eşleşen gözlem sayısının azalacağı belirtilmektedir (Cunningham, 2021). Öte yandan, kullanılan değişken sayısının az olması gözlemlerin birbirine benzerliğini azaltacağından tahmin sonuçlarını olumsuz etkileyecektir.

Eşleştirme ile müdahale ve kontrol grupları arasında kovaryat/gözlemlenebilir özellik dengesini artırmak, söz konusu grupları karşılaştırılabilir hale getirmek amaçlandığından yapılan eşleştirme sonrasında grupların dengesi kontrol edilmelidir. Gruplar arasında değişkenler bakımından dengenin kontrol edilebilmesi için kullanılabilecek iki farklı değer bulunmaktadır.

Bunlardan ilki Standartlaştırılmış Ortalama Fark (Standardized Mean Difference, SMD) değeridir. İki grup arasındaki ortalama farkını, ilgili değişkenin standart sapması ile normalize etmektedir (Hedges, 2024). Değişken bazında SMD değerinin 0,1 değerinden küçük olması eşleştirme dengesinin iyi olduğu, 0,1 ile 0,25 arasında olması ise ideal olmadığı ancak kabul edilebilir olduğu şeklinde değerlendirilmektedir (Austin P. C., 2009). Ancak SMD değerinin 0,25 üzerinde olması durumunda eşleştirmenin yeniden yapılması önerilmektedir.

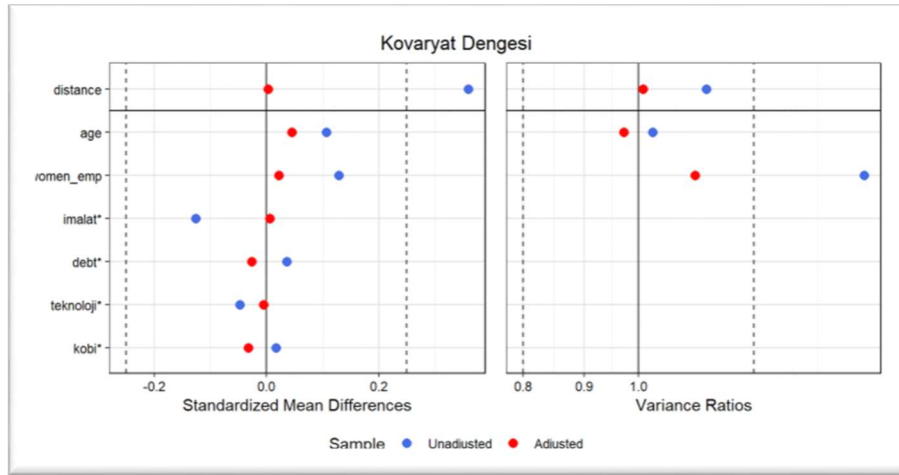
Eşleştirmenin dengesinin kontrolünde kullanılabilecek bir diğer gösterge, Varyans Oranıdır (Variance Ratio). Varyans oranı, bir değişkenin müdahale ve kontrol gruplarındaki varyanslarının birbirine ne kadar benzediğini ölçmektedir. Eğer grupların varyansları birbirlerinden oldukça farklıysa, eşleştirme zayıflamakta ve tahmin edilen etki yanıltıcı olmaktadır. Bu çerçevede, varyans oranının 1'e yakın olması beklenmektedir (Rubin, Using Propensity Scores to Help Design Observational Studies: Application to the Tobacco Litigation, 2001). Bununla birlikte, 0,8 ile 1,25 arasında olabileceği belirtilmektedir. Ayrıca, 0,5 ya da 2 gibi değerlerin oldukça uzak değerler olduğu belirtilmektedir.

SMD ve varyans oranı değerlerine ek olarak yapılan eşleştirmenin dengesi görsel olarak da kontrol edilebilir. Bunun için örnek R kodları aşağıda yer almaktadır.

```
love.plot(bal.tab(eslestirme),
  stat = c("m", "v"),
  grid = TRUE,
  thresholds = c(m = .25, v = 1.25),
  colors = c("royalblue", "red"),
  position = "bottom",
  title = "Kovaryat Dengesi",
  stars = "raw")
```

Paylaşılan kodların kullanılması sonucunda elde edilecek örnek grafik aşağıda gösterilmektedir.

Grafik 3. Eşleştirme Sonrası Denge Kontrol Grafiği



Eğer eşleştirme dengesi bazı değişkenler için sağlanırken bazıları için sağlanmıyorsa, sonuç tahminine geçilmemelidir. Bunun yerine eşleştirme tasarımının yenilenmesi önerilmektedir. Zira, eşleştirme tekrarlamalı bir süreç olarak değerlendirilmektedir. Öte yandan, eğer paralel eğilim varsayımı gibi eşleştirme sonrasında da müdahale ve kontrol grupları arasında denge sağlanamıyorsa o zaman etki değerlendirme için farklı bir yöntemin denenmesi önerilmektedir.

Eğer yapılan eşleştirme sonrasında gruplar arasında denge sağlandıysa o zaman Farkların Farkı yönteminin kullanılabilmesi için paralel eğilim varsayımının kontrol edilmesi gerekmektedir. Paralel eğilim varsayımının doğrulanması sonrasında etki tahminine geçilebilecektir.

Adım 4: Model Kurulumu

Paralel eğilim varsayımı doğrulandıktan sonra Farkların Farkı yöntemiyle etki tahminine geçilmektedir. Bu çerçevede, söz konusu yöntemle etki tahmini için oluşturulan standart regresyon denklemi aşağıda gösterilmektedir.

$$y_{it} = \beta_0 + \beta_1 * Grup + \beta_2 * Zaman + \beta_3 * (Grup * Zaman) + \varepsilon_{it}$$

Söz konusu denklemde yer alan “ y_{it} ” terimi bağımlı değişkeni temsil etmektedir. Grup ile ifade edilen terim, veri setinde müdahale ve kontrol grubunu birbirinden ayıran kukla değişkeni, Zaman terimi ise müdahale öncesi ve sonrasını birbirinden ayıran kukla değişkeni temsil etmektedir. Buna ek olarak Grup ve Zaman terimlerinin çarpımı ise etkileşim terimi olarak ifade edilmekte olup etkiyi temsil etmektedir. Bu çerçevede, Farkların Farkı yöntemi kapsamında müdahale ve kontrol gruplarını birbirinden ayıran değişken ile müdahale öncesi ve sonrası dönemi birbirinden ayıran değişkenin etkileşiminin tahmin edilen regresyon denkleminde yer alması gerektiğine dikkat edilmelidir. Uygulamaya ilişkin örnek R kodları aşağıda paylaşılmaktadır.

```
tahmin_modeli <- lm(Y ~ grup + zaman +  
                    (grup * zaman),  
                    data = veri_did)  
  
summary(tahmin_modeli)
```

Model tahmin edildikten sonra elde edilen tahmin sonucunun kullanılabilmesi için değişen varyans²⁰ ile çoklu doğrusal bağlantı²¹ olup olmadığı kontrol edilmeli ve raporlanmalıdır.

Değişen varyans sorunu olup olmadığı Breusch-Pagan (BP) testi ile kontrol edilebilir. Bu kapsamda yapılacak test sonucunda eğer p-değeri < 0,05 ise değişen varyans sorunu olduğu sonucuna ulaşılır. Eğer değişen varyans sorunu varsa bu durumda kümelenmiş standart hatalar²² kullanılmalıdır (Wooldridge J. M., 2010). Ayrıca, R programında “feols” fonksiyonu kendiliğinden kümelenmiş standart hataları verdiği için tercih edilebilir.

```
coeftest(tahmin_modeli,  
         vcov = vcovHC(tahmin_modeli, type = "HC1"))
```

Buna ek olarak çoklu doğrusal bağlantı sorunu olup olmadığı VIF (Variance Inflation Factor) değişkeni ile test edilebilmektedir. Eğer hesaplanan VIF değişkeni 1 ile 5 arasındaysa bu düşük seviyede çoklu doğrusal bağlantının varlığına işaret etmekte olup sonuçlar kullanılabilir olarak değerlendirilmektedir. VIF değeri 5’ten büyük olmakla birlikte 10’dan küçükse, ideal olmamakla birlikte sonuçların kullanılması mümkündür. Ancak VIF değerinin 10’dan daha büyük olması ciddi çoklu doğrusal bağlantı sorunu olduğuna işaret etmekte olup araştırılması gerekmektedir (O’Brien, 2007).

```
vif(tahmin_modeli, type = "predictor")
```

²⁰ Bir istatistiksel modelde hata terimlerinin varyansının tüm gözlemler için sabit olmaması durumunu ifade eder. Bu durumda, bazı gözlemler etrafında hatalar küçükken, bazı gözlemler etrafında daha büyük sapmalar görülebilir. Etki değerlendirme çalışmalarında değişen varyansın dikkate alınmaması, politika etkilerinin olduğundan daha güçlü veya daha zayıf görünmesine yol açabileceğinden, bu durumun tespit edilmesi ve uygun yöntemlerle düzeltilmesi önemlidir.

²¹ Bir regresyon modelinde yer alan bağımsız değişkenlerin birbirleriyle yüksek düzeyde ilişkili olması durumunu ifade eder. Etki değerlendirme çalışmalarında, program etkisini açıklamak için kullanılan değişkenlerin birbirine aşırı derecede benzer bilgi taşıması, politika etkilerinin doğru biçimde ölçülmesini engelleyebilir.

²² Bir örneklemden hesaplanan istatistiksel bir tahminin (örneğin ortalama veya etki katsayısı) ne kadar belirsizlik içerdiğini gösteren bir ölçüdür. Standart hata küçüldükçe, yapılan tahminin daha hassas ve güvenilir olduğu kabul edilir.

Eğer veri seti panel yapıdaysa etki tahmini için “lm” fonksiyonu yerine “plm” fonksiyonu kullanılmalıdır. Aksi takdirde, veri yapısı ihmal edilmiş olacaktır. Bu durum tahmin sonuçlarını olumsuz etkileyecektir. Bu çerçevede, panel veri kullanılması durumunda kullanılabilecek örnek R kodu aşağıda gösterilmektedir. Bununla birlikte panel verinin yapısına göre “model” ve “effect” argümanlarının değişebileceği unutulmamalıdır.

```
model_plm <- plm(Y ~ grup * zaman,  
  data = panel_data_did,  
  model = "within",  
  effect = "twoways")  
summary(model_plm)
```

Adım 5: Modellerin Karşılaştırılması

Uygulanan devlet yardımının etkisinin ölçülmesi için tahmin edilen modeller karşılaştırılırken dikkat edilmesi gereken hususlar aşağıda belirtilmektedir.

- Tahmin edilen etki katsayısının mantıksal geçerliliği olup olmadığı kontrol edilmelidir. Eğer uygulanan devlet yardımı ile örneğin firmaların yatırım harcamalarının artırılması hedefleniyorsa, tahmin sonucunda ilgili katsayı pozitif olmalıdır. Eğer söz konusu katsayı negatif çıkıyorsa tahmin gözden geçirilmelidir.
- Tahmin edilen katsayının ve modelin anlamlı olup olmadığı kontrol edilmelidir. Bu çerçevede, p-değerinin seçilen anlamlılık düzeyinden (yüzde 1, yüzde 5, yüzde 10 vb) küçük olması gerektiği hatırlanmalıdır.
- En az 2 model tahmin edileceğinden tahmin edilen etkinin işareti ve büyüklüğü modeller arasında tutarlı olmalıdır. Eğer etkileşim katsayısının işaretleri modeller arasında farklılaşıyorsa bu bir sorun olduğuna işaret etmektedir. Bunun yanında, eğer tahmin edilen etkinin büyüklüğü ciddi oranda değişim gösteriyorsa bu durumda da yapılan çalışmalar gözden geçirilmelidir.
- Daha fazla gözlem sayısına sahip olan ve bu bakımdan daha fazla temsil sağlayan modelin etki sonucu, raporlama sürecinde tercih edilmelidir.
- RMSE (hataların ortalama karekökü) değişkeni, modelinin tahminlerinin gerçek değerlerden ortalama olarak ne kadar saptığını göstermektedir. Bir başka deyişle, modelin tahmin performansını gösterir. Aynı veri ve aynı değişken için RMSE değeri ne kadar küçükse, model o kadar iyi tahmin ediyor denilebilmektedir. Modellerin uyumunu karşılaştırmak için kullanılmakla birlikte modelin doğru ya da hatalı olduğu anlamına gelmediğine dikkat edilmelidir.

```
#RMSE değeri elle aşağıdaki formül kullanılarak hesaplanabilir  
RMSE <- sqrt(mean((veri_did$Y - predict(tahmin_modeli))^2))
```

```
#Ayrıca “Metrics” paketi kullanılarak da hesaplanabilir  
rmse(veri_did$Y, predict(tahmin_modeli))
```

- Artık Standart Hatası (RSE), modele ait tahmin hatalarının (artıkların) ortalama olarak ne kadar yayıldığını, başka bir deyişle ortalama büyüklüğünü gösteren bir değerdir. RSE ne kadar küçükse, modelin bağımlı değişkendeki varyasyonu o kadar iyi açıklamaktadır. Bu çerçevede RSE değeri küçük olan model tercih edilebilir. RSE, modeldeki bağımsız değişken sayısını dikkate aldığı için bağımsız değişken sayısı farklı olan modellerin karşılaştırmasında RSE değeri kullanılmamalıdır. RSE değeri R programında standart olarak verilmektedir. Bununla birlikte hesaplama için aşağıdaki kodlar kullanılabilir.

```
artiklar <- residuals(tahmin_modeli)
# SSR
SSR <- sum(artiklar^2)
# n ve k
n <- nobs(tahmin_modeli)
k <- length(coef(tahmin_modeli))
# RSE hesaplama
RSE <- sqrt(SSR / (n - k))
RSE
```

R^2 , modelde yer alan bağımsız değişkenlerin, bağımlı değişkendeki değişimi ne ölçüde tahmin edebildiğini veya açıklayabildiğini göstermektedir. Bu çerçevede, R^2 nedensellik ile ilgili bir bilgi vermemektedir. Bu nedenle, bilgi kriteri olarak kullanılabilir ama seçim kriteri olmamalıdır.

3.2.2. Hedef Kitlenin Tamamının Yardımdan Yararlandığı Durumda Farkların Farkı Yönteminin Uygulanması

Heterojenliği Benimseme Tasarımları (HBT) bir müdahalenin farklı gözlem birimlerinde (firma, ilçe, bölge, kişi vb.), farklı zamanlarda (Callaway, B. ve Sant’Anna, P. H. C., 2021) uygulandığı ve/veya müdahale etkisinin yararlanıcılar ve zaman boyunca heterojen olduğu (de Chaisemartin, C. ve D’Haultfœuille, X., 2020) durumları ifade etmektedir (Roth, J., Sant’Anna, P. C. H., Bilinski, A. ve Poe, J., 2023).

Müdahalenin kademeli olarak uygulandığı, müdahale etkisinin dinamik olduğu ve zamanla değiştiği, müdahale etkisinin yararlanıcılar arasında farklılaştığı durumlarda HBT tercih edilmektedir. Zira, iki yönlü sabit etkiler ile uygulanan Farkların Farkı yönteminde, müdahale etkisinin dönemler ve gruplar arasında homojen olduğu varsayılmaktadır. Bu çerçevede söz konusu homojenliğin sağlanmadığı durumlarda HBT

HBT tasarımlarına asgari ücret ayarlamaları, vergi ve sigorta primi düzenlemeleri, tarifeler vb. uygulamalar Örnek olarak gösterilebilmektedir.

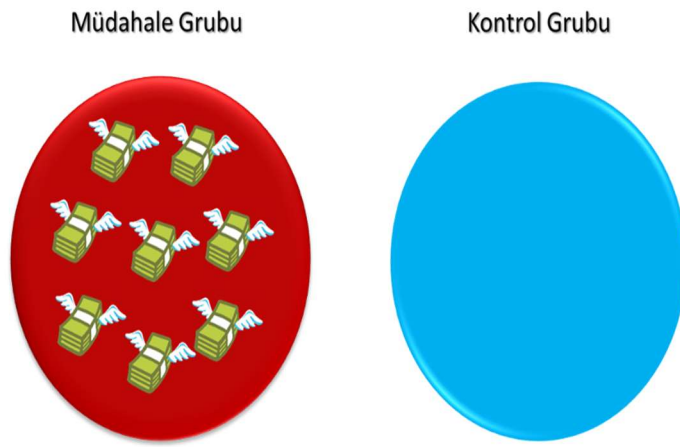
Bu çerçevede, HBT devlet yardımından yararlanmayan birimlerin bulunmadığı ve yardımın yoğunluğunun ve etkilerinin birimler arası farklılaştığı durumlar için de kullanılabilir.

1- Heterojenliği Benimseme Tasarımlarında Karşı Olgusallığı Sağlama

Karşı olgusallık, bir bireyin ya da grubun müdahaleye maruz kalmamış olması durumunda ortaya çıkacak olan, ancak doğası gereği gözlemlenemeyen sonuçtur (Rubin, 1974).

FF yaklaşımında karşı olgusallık, kontrol grubunun zaman içi eğiliminin müdahale grubunun müdahaleye maruz kalmaması durumunda izleyeceği eğilimi temsil ettiği paralel eğilim varsayımıyla oluşturulur (Angrist ve Pischke, 2009). Dolayısıyla FF yönteminde karşı-olgusallık referans olarak alınan kontrol grubunun varlığı ile sağlanmaktadır.

Şekil 1. Farkların Farkı Yöntemi Kıyas Şeması



HBT’de ise kontrol grubunun bulunmaması başka bir ifade ile tüm birimlerin müdahaleye maruz kalması sebebiyle karşı olgusallığı sağlayacak referans grubunun tespiti **müdahalenin dozu** ve **çıktıda meydana gelen potansiyel değişimlere** göre şekillenmektedir (de Chaisemartin, Ciccia, d’Haultfoeuille ve Knau, 2025).

Yarı-İşlenmemiş Grup (YİG) : Müdahale dozunun 0’a çok yakın olduğu ($\lim_{d \rightarrow 0}$) ve çıktı düzeyi müdahale sonucunda tıpkı kontrol grubuymuş gibi değişim gösteren birimlerdir.

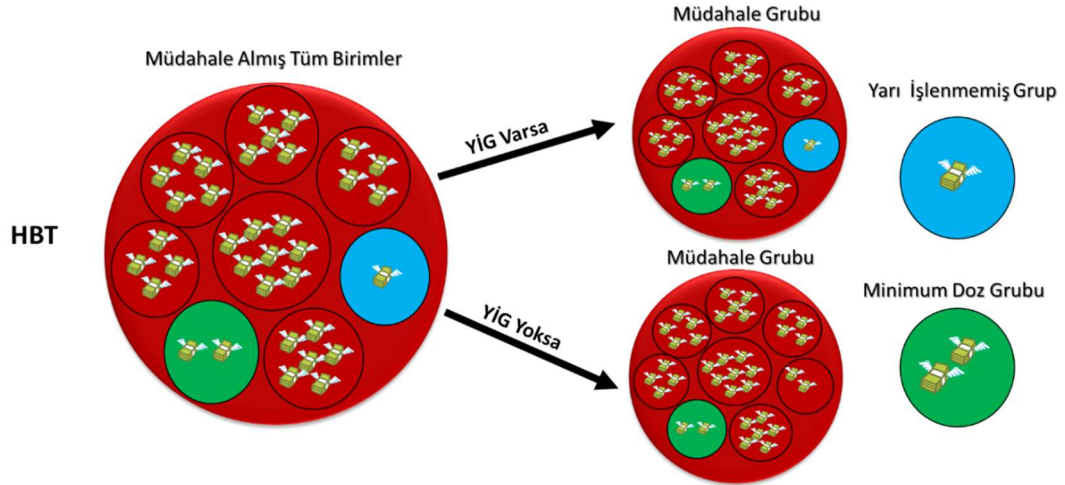
YİG olmadığı durumda ise birimlerin tamamı pozitif müdahale dozları ($d_{min} > 0$) almaktadır.

HBT tasarımlarında karşı olgusallık:

- ✓ **YİG varlığında** herkes müdahale almasaydı ne olurdu?
- ✓ **YİG olmadığı** durumda ise herkes minimum düzeyde müdahale alsaydı ne olurdu? Sorularıyla sağlanmaktadır.

2- Analiz Yöntemleri

Şekil 2. Heterojenliği Benimseme Tasarımlarında Referans Grubu Tespit Şeması



- Klasik FF yöntemi, sonucu müdahale değişkeni binary (müdahale aldı/müdahale almadı) iken, Heterojenliği benimseme tasarımlarında ise müdahalenin sürekli (0,1 – 0,5 – 1) ve birimler için farklı dozlarda uygulanıyor olması FF’nin söz konusu tasarımlara uygun olmadığını göstermektedir. (de Chaisemartin vd, 2020)
- HBT ise; analizde ortaya konulacak yöntem **YİG** varlığına göre değişim göstermektedir.
 - **YİG varsa;** yani müdahaleye maruziyet dozu 0’a çok yakın birimler mevcut ise ve bu birimler müdahale sonrası dönemde kontrol grubunun gösterdiği gibi bir tepki gösteriyor ise bağımlı değişkende gözlemlenen eğilim “müdahalesiz eğilim” ile aynıdır. Başka bir deyişle YİG bize doğal bir kontrol grubu mekanizmasını sunmaktadır.
 - ✓ YİG varlığında herkes müdahale almasaydı $\Delta Y(0)$ ne olurdu? Sorusuna cevap aranır.
 - ✓ Bu durumda karşı olgusallık tıpkı bir kontrol grubu varmış gibi doğrudan gözlemlenebilmektedir. Böylece tahmini gerçekleştiren varsayımlara dayanmadan doğrudan gerçek veriyle hareket edecektir.
 - ✓ Modelde müdahalenin sürekli ve heterojen olması, varsayımlara dayanmaması tahmin yönteminin doğrudan nonparametrik olmasını sağlayacaktır.

a) Ağırlıklı Ortalamalar Eğimi (AOE) Yöntemi

AOE yarı işlenmemiş grupların varlığında müdahale ve bağımlı değişken arasındaki ilişkiyi gösteren doz–yanıt eğrisinin eğimlerinin ağırlıklı ortalaması alınarak oluşturulur.

Müdahale dozu bir birim arttığında, sonuç değişkeni ortalama olarak ne kadar değişir? Sorusunun cevabını vermektedir. AOE müdahalenin bağımlı değişken üzerindeki marjinal etkisini göstermektedir.

AOE tıpkı RDD mantığı ile çalışmaktadır. RDD’de cutoff (eşik değer) değerin hemen altında kalanlar hemen üzerinde kalan birimler için müdahale almasalardı ne olurdu durumunu temsil etmektedir. AOE’de 0 yada 0’a yakın müdahale değerine maruz kalmış birimlerde cutoff’un altında kalan gözlemlerin görevini görmektedir. (de Chaisemartin vd, 2025)

Genel Tanım:

$$AOE = E \left[\frac{\Delta Y}{\Delta D} \right]$$

ΔD : Müdahale dozundaki değişimi ifade eder.

ΔY : Müdahale dozundaki değişime görece çıktıda meydana gelen değişimi ifade eder.

YİG ile Tahmin;

$$AOE = E \left[\frac{\Delta Y_i - \Delta Y^{YIG}}{D_i - 0} \right]^{23}$$

ΔY_i : İ biriminine (Müdahale grubu) ait bağımlı değişkeninde meydana gelen ortalama değişim.

ΔY^{YIG} : YİG’E ait bağımlı değişkende meydana gelen ortalama değişim.

D_i : İ (müdahale grubunun) maruz kaldığı ortalama müdahale dozu – şiddeti

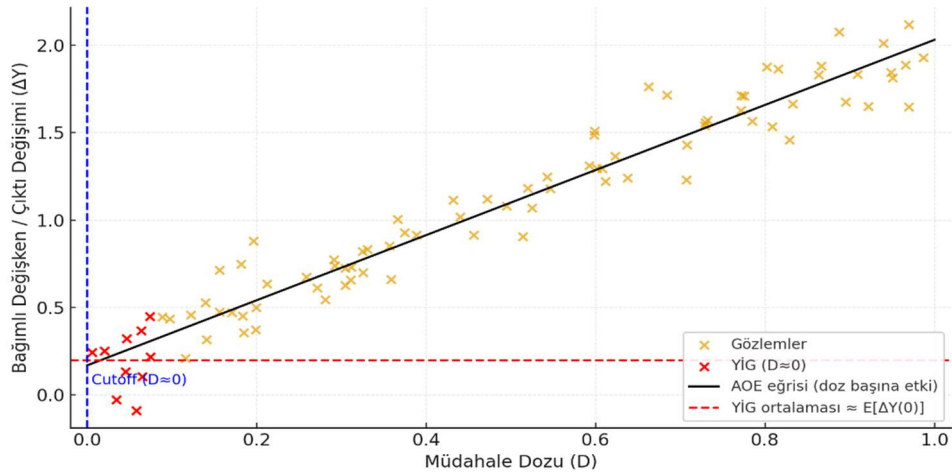
Kısaca AOE yöntemi:

Yarı-işlenmemiş grupların varlığında müdahale dozu bir birim arttığında, çıktının ortalama olarak ne kadar değiştiğini hesaplayarak müdahalenin etkisini ortaya koymaktadır. Yani AOE 0’a (YİG müdahale düzeyi) kıyasla doz başına düşen etkiyi tahmin etmektedir.

$$AOE = \left[\frac{(Toplam Cıktıdaki Ortalama Değişim - YIG Ortalama Cıktı Değişimi)}{(Ortalama Mudahale Duzeyi - 0)} \right]$$

²³ Paydada yer alan 0 düzeyi Yarı işlenmemiş grubun tanımından gelen 0’a çok yakın ($\lim_{d \rightarrow 0}$) değerde müdahale alması sebebiyle belirtilmiştir.

Grafik 4. Yarı İşlenmemiş Grup Mevcudiyetinde AOE Doz -Yanıt Eğrisi



Grafik 4'ten de anlaşılacağı üzere 0'a yakın müdahaleye maruz kalmış birimler cutoff (eşik değeri) etrafında toplanmakta ve referans grubunu oluşturmaktadır. AOE eğrisinin (doz-yanıt eğrisi) eğimi ise cutoff müdahale düzeyine ek her 1 br doz artışına karşılık bağımlı değişkenden meydana gelen değişimi göstermektedir.

Heterojenliği benimseme tasarımlarında yarı işlenmemiş bir grubun olmaması halinde ise etki değerlendirmesi,

- **YİG yoksa;** Yani bütün birimlerin farklı düzeylerde ancak pozitif dozda müdahaleye maruz kaldığı sonucunda etkinin heterojen olarak görüldüğü durumda nasıl bir yol izlenmeli?
- ✓ Yarı-işlenmiş grubun mevcudiyeti bize doğal bir kontrol grubu mekanizması sunarken, YİG olmaması durumunda $\Delta Y(0)$ sağlanamayacağından karşı-olguşallığı yakınsamak için gerekli referans grubu oluştururken baz alınacak yöntem en az doza maruz kalmış birimlerden oluşturulmasıdır.
- ✓ YİG olmaması halinde herkes en düşük dozu alsaydı ne olurdu? Sorusuyla karşı-olguşallık aranır.

b) Minimum Dozla Ağırlıklı Ortalamalar Eğimi (AOE_{dmin}) Yöntemi

AOE'dan farklı olarak AOE_{dmin} tüm birimlerin müdahale aldığı ancak müdahale düzeylerinin sıfıra yakınsamadığı durumda ($d_{min} > 0$) izlenecek etki değerlendirme yöntemidir.

$$AOE_{dmin} = \frac{E[\Delta Y] - E[\Delta Y(d_{min})]}{E[D_2 - d_{min}]}$$

ΔY : Gözlemlenen ortalama çıktı değişimi.

$\Delta Y(d_{min})$: Tüm birimler en az dozu alsaydı gözlemlenecek ortalama çıktı değişimi

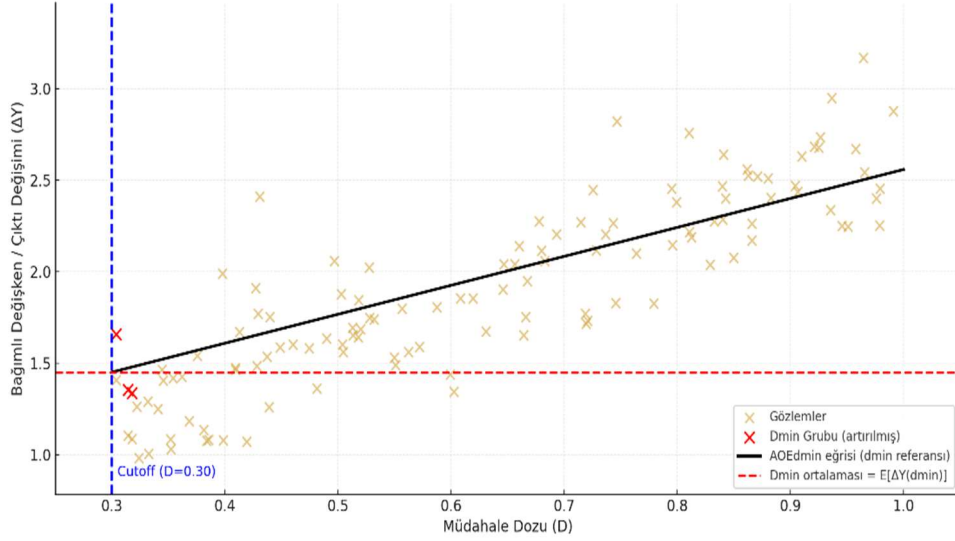
$D_2 - d_{min}$: Ortalama ek doz miktarı, birimin aldığı doz, minimum doza kıyasla ne kadar fazla olduğunun tespiti.

Kısaca AOE_{dmin} yöntemi:

Müdahale düzeyi minimum doza ilave olarak bir birim artırıldığında, çıktının ortalama olarak ne kadar değişim gösterdiğini hesaplamaktadır. Başka bir ifade ile minimum doz müdahaleye kıyasla doz başına düşen etkiyi tahmin etmektedir.

$$AOE_{dmin} = \frac{\text{Toplam Cıktıdaki Ort. Değişim} - \text{Minimum Doz Grubu Ort. Cıktı Değişimi}}{\text{Ortalama Müdahale Dozu} - \text{Minimum Müdahale Dozu}}$$

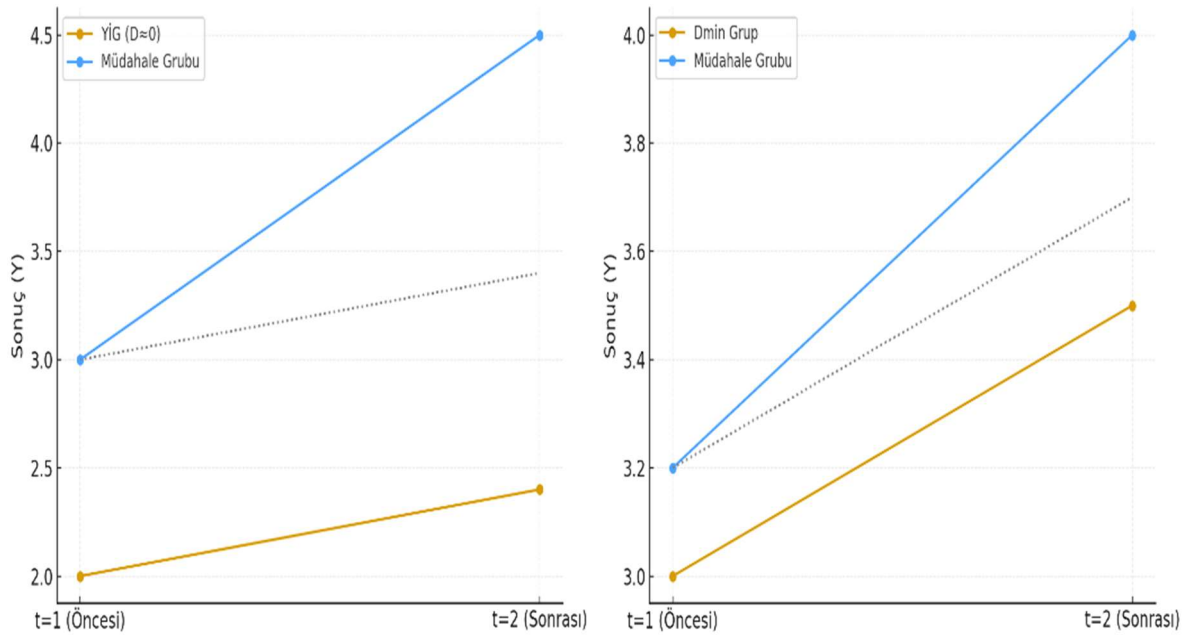
Grafik 5. YİG Mevcut Olmadığında AOEdmin Doz - Yanıt Eğrisi



Grafik 5'te cutoff değerinin (minimum müdahale dozunun) 0'a yakınsamadığı dolayısıyla referans grubunu oluşturacak olan birimlerin 0,3 müdahale dozu düzeyinde yoğunlaştığı görülmektedir. AOE_{dmin} eğrisinin (doz-yanıt eğrisi) eğimi ise 0,3'e (Dmin) ek her 1 br doz artışına karşılık bağımlı değişkende meydana gelen değişimi göstermektedir.

AOE_{dmin} yaklaşımında referans noktası müdahaleyi en az alan birimler olduğu için, karşı olgusal eğri daha yüksek bir başlangıç seviyesinden başlar. Bu durum, müdahale dozundaki artışın etkisini daha temkinli ve daha düşük bir ortalama eğimle yansıtır. Bu sebeple AOE_{dmin} ile hesaplanan etki doğal olarak YİG varlığında hesaplanan AOE karşısında daha küçük görünür, çünkü karşılaştırma daha yüksek bir referans seviyesine göre yapılmaktadır (Bkz. Grafik 6).

Grafik 6. YİG ve Dmin Referans Gruplarında Gözlemlenen Etki Düzeyleri



c) Two-Way Fixed Effects – İki Yönlü Sabit Etkiler (TWFE) Yöntemi

TWFE hem **birim sabit etkilerini** (α_i) hem de **zaman sabit etkilerini** (λ_t) kontrol eder. Böylece gözlemlenemeyen ama birime ait sabit kalan heterojenlikleri denkleme dahil ederek yanlılığı azalmaktadır (Wooldridge, 2010).

TWFE'nin FF ile farkı ise; FF müdahale değişkeni binary yapıda olup sürekliliği kabul etmezken TWFE ise müdahalenin sürekli olması katsayısının hesaplanmasına engel olmayacaktır.

$$Y_{it} = \alpha_i + \lambda_t + \beta D_{it} + \varepsilon_{it}$$

Y_{it} : i birimini t dönemindeki çıktısı

α_i : i birimine özgü karakteristik (değişmeyen) sabit etkiler.

λ_t : zamana özgü sabit etkiler

D_{it} : müdahale değişkeni

β : tahmin edilmek istenen katsayı (Müdahale 1 birim arttığında çıktının ne kadar değişeceğini gösteren katsayı)

ε_{it} : hata terimi

İki yönlü sabit etkiler yöntemi HBT tasarımlarında (müdahale sürekli – kontrol grubu yok)

- YİG varlığı söz konusu olmadığında,
- AOE_{dmin} yöntemi için veri zayıf kalıyor ise durumlarında kullanılabilmektedir.

Tüm birimlerin müdahaleye maruz kaldığı başka bir ifade ile kontrol grubunun bulunmadığı durumlarda etki değerlendirme yöntemlerini bir tabloda özetlemek gerekirse:

Tablo 6. Heterojenliği Benimseme Tasarımlarında Etki Değerlendirme Yöntemleri

<i>Durum</i>	<i>Yöntem</i>	<i>Neden?</i>	<i>Avantajları</i>	<i>Dezavantajları</i>
<i>YİG Varsa</i>	AOE	YİG (sıfıra yakın yardım alanlar) müdahalesiz karşı-olgusalı temsil eder. Bu sayede 0'a göre marjinal etkiyi nonparametrik olarak tahmin edebiliriz.	Nonparametrik: doğrusal varsayım gerekmez. Heterojen etkileri doğru yansıtır. Daha güvenilir ve teorik olarak sağlam.	Güç, YİG grubunun büyüklüğüne bağlıdır. Eğer YİG küçükse tahminler gürültülü olabilir.
<i>YİG Yoksa</i>	AOEdmin	0'a yakın müdahale alan grup yoksa 0'a göre etki tanımlanamaz. Bunun yerine en az doz (dmin>0) referans alınır.	Nonparametrik: etki doğrusal varsayılmaz. Heterojenliği tolere eder.	Yorum "0'a göre" değil, en az doz göre yapılır. En az doz grubunda az gözlem varsa güven aralıkları genişler.
<i>YİG Yoksa</i>	TWFE	AOEdmin belirsizse, tek yol varsayımlarla model kurmaktır. TWFE, doğrusal ve homojen etki varsayımı altında β tahmin eder.	Yöntem parametrik olması sebebiyle güçlü istatistiksel tahmin ve dar güven aralıkları vardır	Doğrusallık ve homojenlik varsayımına bağlıdır. Heterojenlik veya doğrusal olmayan ilişki varsa yanlış olabilir.

Şekil 3. Heterojenliği Benimseme Tasarımlarında Etki Değerlendirme Yöntemi Karar Ağacı



HBT - Örnek Uygulama

Bir bölgedeki bulunan **bütün tarım işletmeleri (Kontrol grubu yok)**, ürün satışlarından elde ettikleri yıllık gelire göre gelir vergisi ödemektedir. Devlet, bu işletmeleri desteklemek ve üretim kapasitesini artırmak amacıyla **kademeli bir gelir vergisi indirim programı (Heterojenlik)** başlatır. Program basittir: Vergi indirimi oranı yükseldikçe çiftçilerin elinde daha fazla kaynak kalacak, bu kaynaklar da tarımsal üretime geri dönecektir.

Politika tasarım ekibinin merak ettiği kritik soru şudur;

“Vergi indirimi oranı ne kadar artırılmalı ki tarımsal gelirden anlamlı bir artış sağlansın?”

Vergi İndirim Oranı	Birim Sayısı	Teşvik Sonrası Ortalama Tarımsal Gelir Artışı (Bin TL)
0.10 = dmin	100	2,5
0.30	300	6,5
0.50	350	14,5
0.70	250	26,5
Toplam	1.000	—

Birimler Arası Heterojen Müdahale Dozu

Birimler Arası Heterojen Etki

$$\text{Tarımsal Gelirdeki Ortalama Değişim} = \frac{100 \cdot 2,5 + 300 \cdot 6,5 + 350 \cdot 14,5 + 250 \cdot 26,5}{1000} = 13,9 \text{ bin TL}$$

$$\text{Minimum Doz Grubu Ortalama Gelir Değişimi} = 2,5 \text{ bin TL}$$

$$\text{Ort. Tarımsal Gelir Değişimi} - \text{Minimum Ort. Gelir Değişimi} = 13,9 - 2,5 = 11,4 \text{ bin TL}$$

$$\text{Ortalama Vergi İndirim Oranı} = \frac{100 \cdot 0,10 + 300 \cdot 0,30 + 350 \cdot 0,50 + 250 \cdot 0,70}{1000} = 0,45$$

$$\text{Ortalama Vergi İndirim Oranı} - \text{Minimum Vergi İndirim Oranı} = 0,45 - 0,10 = 0,35$$

$$\left[\frac{\text{Ort. Tarımsal Gelir Değişimi} - \text{Minimum Ort. Gelir Değişimi}}{\text{Ortalama Vergi İndirim Oranı} - \text{Minimum Vergi İndirim Oranı}} \right] = \frac{11,4}{0,35} = 32,6$$

Sonuç: Vergi indirimi oranı 1 birim arttığında (%100), tarımsal gelir ortalama 32,6 bin TL artmaktadır.

Vergi indirimi oranı 0,10 (%10) artırılması $\approx$ 3,26 bin TL ek gelir artışı olarak yorumlanabilmektedir.

3.3. Sentetik Kontrol Yöntemi

Etki değerlendirme, devlet yardımıyla yararlananlar (müdahale grubu) ile yardımdan yararlanmayanlar (kontrol grubu) arasında yardıma atfedilebilecek fark oluşup oluşmadığını araştırmaktadır. Yardımdan yararlanan çok küçük bir kesim hatta tek yararlanıcı olan durumlarda ve kontrol grubu için büyük miktarda gözlem bulunamadığında sentetik kontrol yöntemi kullanılabilir. Sentetik kontrol yöntemi, ilk olarak çok az sayıda müdahale grubu olan durumlar için önerilmiştir (Abadie, 2021). Örneğin, müdahale grubunda toplam 10 gözlem olduğunda ve ilgili tüm değişkenler itibarıyla hiçbir kontrol grubu hatta kontrol gruplarının ortalaması dahi müdahale grubuna benzerlik göstermediğinde, sentetik kontrol yöntemi olmadan, en azından daha çok veri toplanıncaya kadar bu veriler heba olabilir (Thompson, 2017).

Sentetik kontrol yöntemi, az sayıda birime müdahalede bulunulduğunda, kontrol için müdahaleden etkilenmeyen tek bir birim yerine birden çok etkilenmeyen birimin kombinasyonunun daha uygun olacağı fikrine dayanmaktadır (Abadie, 2021). Özellikle alan/bölge bazlı yardımlar için etki değerlendirme yapılırken, yardım uygulanan bölge ile belirlenen değişkenler itibarıyla benzer özellik gösteren farklı bölgelerin kombinasyonundan kontrol grubu oluşturularak yapılacak kıyaslama ile yardıma atfedilecek etkiyi tahmin etmek mümkündür.

Sentetik kontrol yöntemi sadece devlet yardımlarının etkisini ölçmek için değil, birçok alanda ortaya çıkan değişikliklerin etkisini ölçmek için de kullanılabilir. Nitekim sentetik kontrol yöntemi ilk defa Abadie ve Gardeazabal (2003) tarafından İspanya’da Bask bölgesinin terörden kaynaklı zararının tespit edilmesi için önerilmiştir. Terör olayları başlamadan önce birçok parametre açısından iyi bir performans gösteren Bask bölgesiyle İspanya’nın Bask bölgesine benzer özellik gösteren şehirleri karşılaştıran çalışmada, söz konusu diğer şehirlerden ağırlıklı ortalama uygulanarak “Sentetik Bask” oluşturulmuştur. Ayrıca Abadie ve diğerleri (2010), ABD’nin Kaliforniya eyaletinde 1989 yılında uygulamaya konulan Kaliforniya Sigara Kontrol Programı Öneri 99 (Proposition 99)’un etkisini ölçmek için de bu yöntemi kullanmıştır.²⁴

Bu iki çalışma sentetik kontrol yönteminin çıkış noktası olarak kabul edilmekte olup özellikle yöntemin daha iyi anlaşılması için Abadie ve diğerleri (2010)’nin çalışmasını detaylı bir şekilde incelemekte fayda bulunmaktadır.

Kaliforniya’da kabul edilen “Öneri 99” isimli düzenleme ile sigara tüketim vergisi paket başına 25 cent artırılmıştır. Bu düzenlemenin etkisinin araştırıldığı çalışmada 1970-2000 dönemine ait eyalet verileri kullanılmıştır. İlk defa 1970 yılında sigara satış kayıtlarının söz konusu karşılaştırma eyaletleri için tutulmaya başlanması ve sigara azaltıcı benzer mevzuat düzenlemelerinin 2000 yılından sonra bu eyaletlerde uygulamaya konulmasından ötürü kontrol grubu bulunmasının zor olacağı, bu sebeple 1970-2000 döneminin analiz edildiği belirtilmiştir. Kaliforniya’da böyle bir yasa çıkmasaydı ne olurdu? sorusuna cevap olarak “sentetik Kaliforniya” grubu oluşturulmuştur. Sentetik Kaliforniya, potansiyel karşılaştırma eyaletlerinin ağırlıklandırılmış ortalaması alınarak oluşturulmuştur. Öncelikle karşılaştırma eyaletleri arasında 1989-2000 yılları arasında “Öneri 99”a benzer yasa çıkarmış veya sigara paketlerine ilave vergi getirmiş olan 11 eyalete yer verilmemiş ve 38 eyalet kalmıştır. Dolayısıyla söz konusu dönemde sigara tüketimini azaltabilecek bir gelişmesi olan herhangi bir eyalet, kontrol grubuna dahil edilememiştir. Bağımlı değişken olarak kişi başına düşen yıllık sigara satış miktarı (paket) belirlenmiş ve sigara paketinin ortalama perakende fiyatı, logaritması alınmış kişi başına düşen milli gelir, 15-24 yaş aralığı nüfus oranı ve kişi başına bira tüketimi açıklayıcı değişkenler olarak kullanılmıştır. Öncelikle “Öneri 99” öncesi Kaliforniya’nın durumunu yansıtan sentetik Kaliforniya oluşturulmuş ve “Öneri 99” sonrası dönem için etki Kaliforniya ile sentetik versiyonları arasındaki fark ölçülmüştür. Ardından bir dizi plasebo testleri²⁵ yapılarak Kaliforniya için tahmin edilen etkinin diğerlerine göre yüksek olduğunu teyit etmişlerdir.

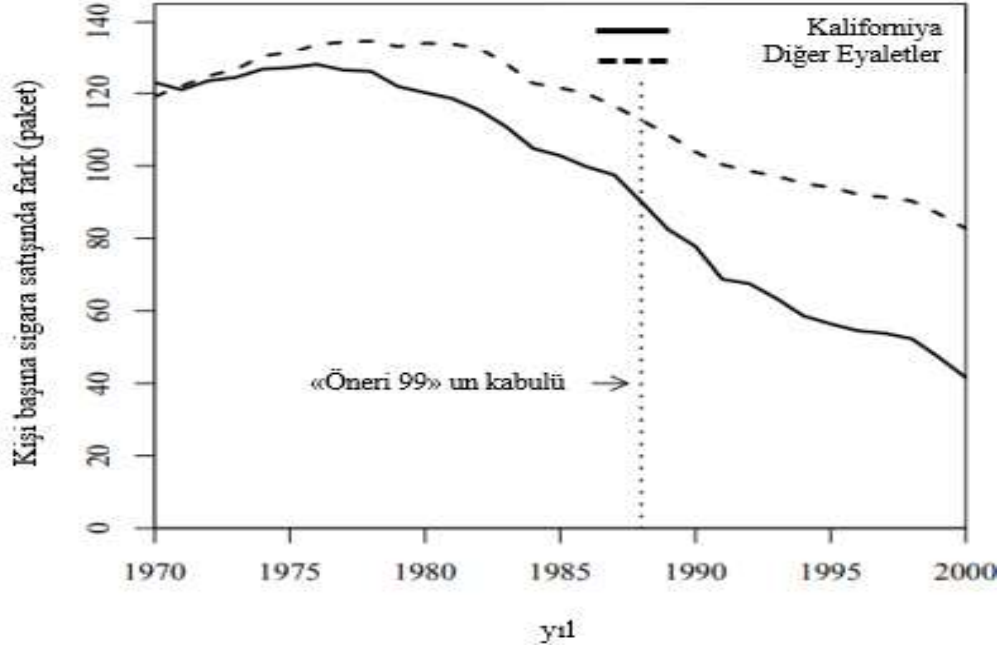
²⁴ Söz konusu çalışma aynı zamanda Sentetik Kontrol yönteminin dayandığı istatistiki modeli genel hatlarıyla açıklamaktadır. Yöntemin matematiksel açıklaması için bakınız Abadie ve Diğerleri (2010).

²⁵ Plasebo testleri, bir istatistiksel analizin sonuçlarının tesadüfen meydana gelip gelmediğini kontrol etmek için kullanılan bir yöntemdir. Sentetik Kontrol yöntemi kullanılarak yapılan çalışmalarda, plasebo testleri genellikle Sentetik Kontrol biriminin gerçek kontrol birimiyle ne kadar iyi eşleştiğini ve elde edilen etki tahminlerinin ne kadar güvenilir olduğunu değerlendirmek için kullanılmaktadır.

Bu çalışmada akla ilk gelen soru, her ne kadar 1989-2000 arasındaki dönemde benzer düzenlemeler yapan eyaletler kontrol grubundan çıkartılsa da, kalan eyaletler ne derece Kaliforniya ile karşılaştırılabilir?

Abadie ve diğerleri (2010) Kaliforniya ile diğer eyaletlerin normal karşılaştırmasını aşağıdaki grafikte açıklamaktadır. Grafik 7’de görüldüğü gibi Kaliforniya ve diğer eyaletler arasında süreç içerisinde farklar bulunmaktadır. Bu nedenle diğer eyaletler iyi bir kontrol grubu teşkil etmemektedir.

Grafik 7. Sigara Satışlarındaki Eğilim: Kaliforniya ve Diğer Eyaletler



Kaynak: https://www.nber.org/system/files/working_papers/w12831/w12831.pdf

Bu kapsamda, sentetik Kaliforniya grubu oluşturmak için “Öneri 99” öncesi dönemde Kaliforniya’ya en yakın olan eyaletlerin dışbükey kombinasyonu alınmış ve 38 eyaletin nüfusa göre ağırlıklandırılmış ortalamasıyla karşılaştırılmıştır. Tablodan da görüleceği gibi sentetik Kaliforniya, 1989-2000 döneminde “Öneri 99” olmasaydı ne olurdu? sorusuna yanıt için dikkat çekici bir yakınlık göstermektedir.²⁶

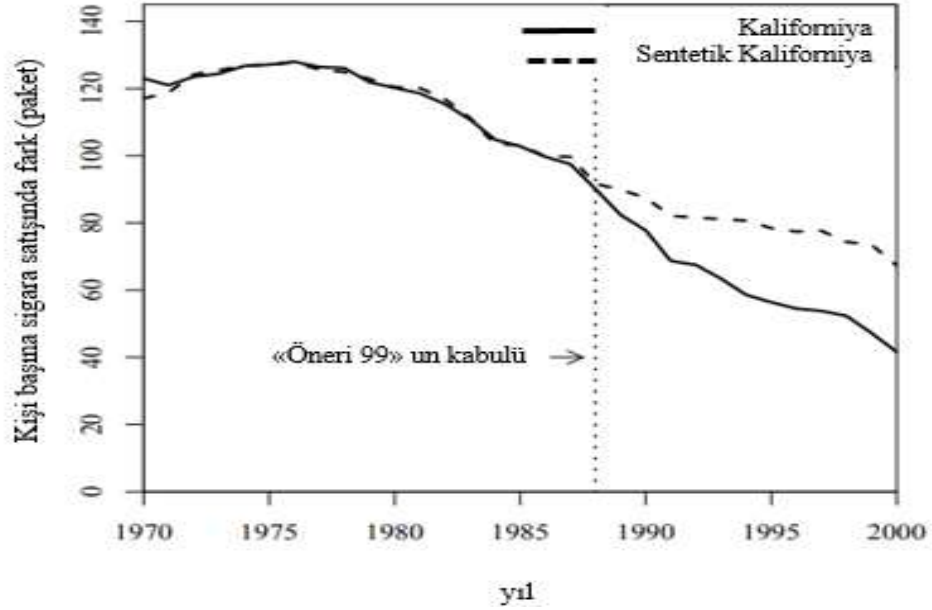
²⁶ Yapılan ağırlıklandırma çalışması sonrasında sentetik Kaliforniya’yı Colorado, Connecticut, Montana, Nevada ve Utah eyaletlerinin en iyi temsil ettiği anlaşılmıştır.

Tablo 7. Sigara Satışı Tahmin Edicilerinin Ortalaması

Değişken	Kaliforniya		38 Kontrol Eyaletinin Ortalaması
	Gerçek	Sentetik	
Kişi Başına GSYH (ln)	10,08	9,86	9,86
15-24 yaş ortalaması	17,40	17,40	17,29
Perakende sigara fiyatı	89,42	89,41	87,27
Kişi başına düşen bira tüketimi	24,28	24,20	23,75
Kişi başına düşen sigara satışı 1988	90,10	91,62	114,20
Kişi başına düşen sigara satışı 1980	120,20	120,43	136,58
Kişi başına düşen sigara satışı 1975	127,10	126,99	132,81

Not: Tüm değişkenler 1980-1988 dönem için (bira tüketimi 1984-1988) ortalamadır. Sigara satışları paket olarak ölçülmüştür.

Kaliforniya ve sentetik Kaliforniya'nın karşılaştırması Grafik 8'de gösterilmiştir. Grafikten de görüleceği gibi yasa öncesi "tek çizgi" gibi ilerlerken yasanın çıkmasından sonra fark oluşmaya başlamıştır. Dolayısıyla yasanın etkisi iki çizgi arasındaki fark olarak tahmin edilmektedir. Sonuçların 1989-2000 yılları arasında satılan sigara paketi sayısında ortalama 20 (yaklaşık %25 azalma) olduğu tahmin edilmiştir. Grafikteki düz çizgi Kaliforniya'daki gerçekleştirmeyi, kesikli çizgi ise "Öneri 99" öncesi benzerlik gösterdiği için sentetik Kaliforniya'yı temsil eden 5 eyaletteki gelişmeyi göstermektedir. Diğer bir ifadeyle kesik çizgi, düzenleme olmasaydı Kaliforniya'daki gelişme nasıl olacaktı sorusunu cevaplamaktadır.

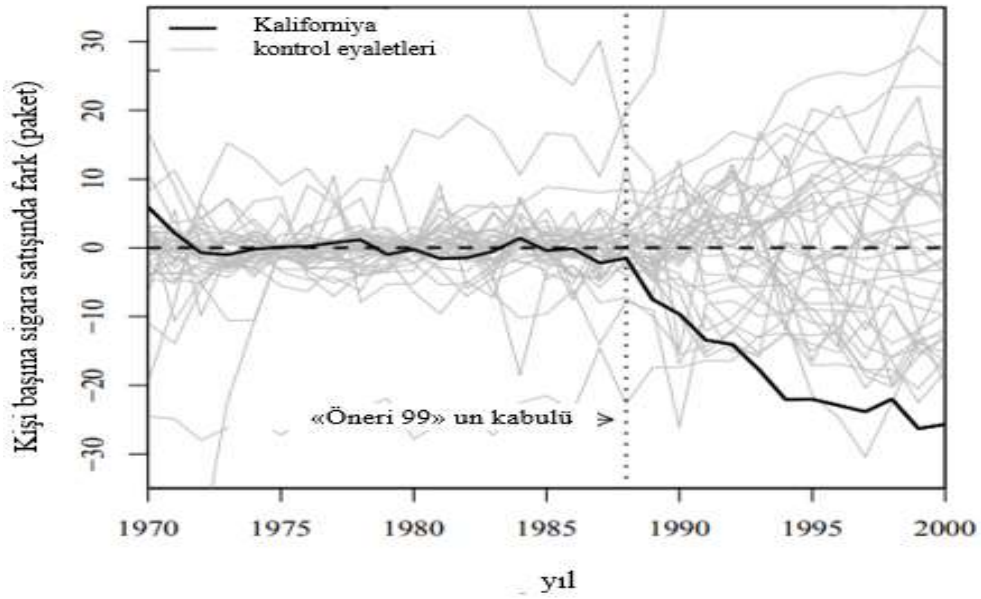
Grafik 8. Sigara Satışlarındaki Eğilim: Kaliforniya ve Sentetik Kaliforniya

Kaynak: https://www.nber.org/system/files/working_papers/w12831/w12831.pdf

Sonuçların sağlamlığını test etmek için ilave tahmin ediciler eklenmiştir. Bunlar arasında eyalet düzeyinde işsizliği azaltmaya yönelik alınan tedbirler, gelir eşitsizliği, yoksulluk, sosyal yardımlar, suç oranı, uyuşturucuyla bağlantılı tutuklanma oranları, sigara vergileri, nüfus yoğunluğu ve demografik yapıya dair bazı değişkenler yer almaktadır. Bu değişkenler durumu görsel olarak etkilememiştir.

Ayrıca, sonuçların tesadüfi olup olmadığını anlamak için plasebo testleri yapılmıştır. Plasebo testleri, söz konusu dönemde etkin bir sigara kontrol programı uygulamamış olan eyaletlere sentetik kontrol yöntemi uygulanması suretiyle gerçekleştirilmiştir. Bu testlerde söz konusu eyaletler için de Kaliforniya için uygulanan çalışmaya benzer boşluklar yakalanması halinde, çalışmanın geçersiz olacağı değerlendirilmiştir. Bu kapsamda 38 eyaletin her biri için sentetik kontrol yöntemi uygulanmış ve bu çalışmalarda Kaliforniya donör havuzunda²⁷ tutulmuştur. Plasebo test sonuçları aşağıdaki grafikte gösterilmektedir. Grafik 9'dan da anlaşılabacağı üzere Kaliforniya'daki fark diğer eyaletlere göre oldukça yüksektir.

Grafik 9. Sigara Satışlarındaki Açık (Kaliforniya ve 38 Kontrol Eyaleti)

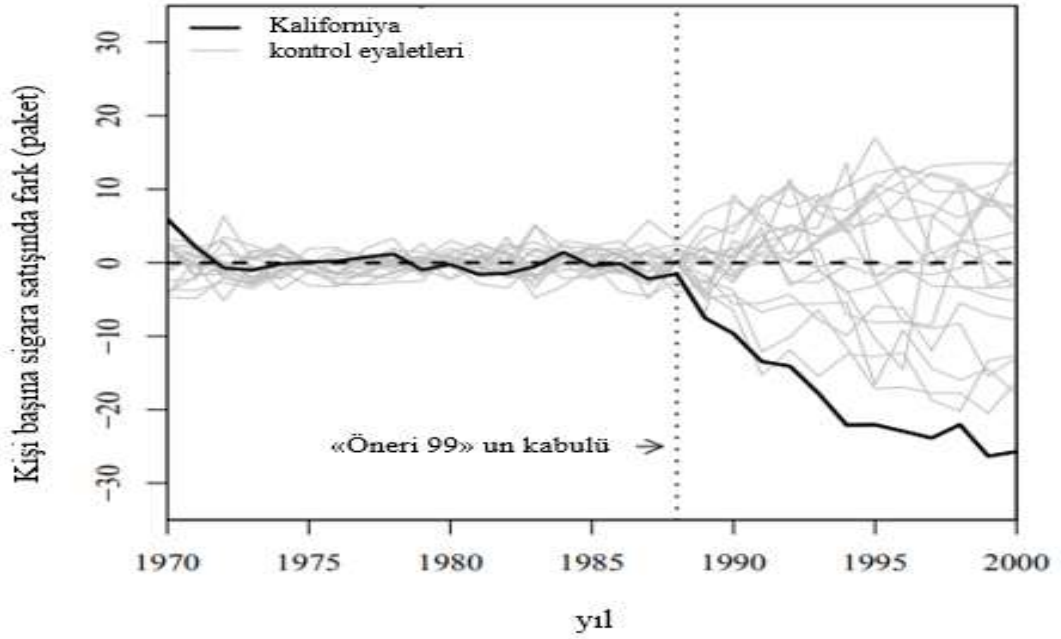


Kaynak: https://www.nber.org/system/files/working_papers/w12831/w12831.pdf

Yazarlar, müdahale öncesi ortalama kare tahmin hatası (MSPE- Mean Squared Prediction Error) yani 1970-1988 döneminde Kaliforniya ile sentetiğinde kişi başına düşen sigara satışları arasındaki farkların karelerinin ortalaması 3 iken, diğer 38 ülke için 6'dır. Bu, "Öneri 99" öncesi dönem için sentetik kontrol yönteminin sigara tüketimine iyi bir uyum sağladığına işaret etmektedir. Ancak, 38 eyalet içerisinde sigara tüketiminin çok olduğu eyaletlerde MSPE'nin yüksek olduğunu (Örneğin New Hampshire için 3437) belirtmiştir. Plasebo çalışması Kaliforniya'nın MSPE'sinin 2 katından fazla olanları dışarıda tutarak yapıldığında Grafik 10'daki gibi sonuç çıkmaktadır.

²⁷ Sentetik Kontrol yöntemlerinde donör havuzu potansiyel kontrol grubunu temsil etmektedir.

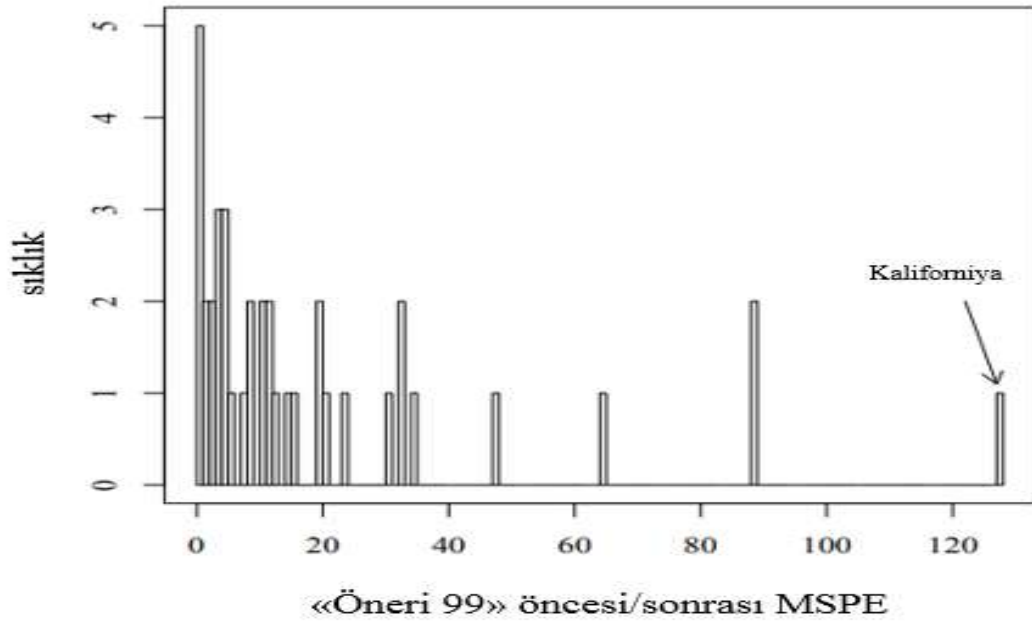
Grafik 10. Sigara Satışlarındaki Açık (Kaliforniya ve 34 Kontrol Eyaleti)



Kaynak: https://www.nber.org/system/files/working_papers/w12831/w12831.pdf

Çalışmada son olarak, Kaliforniya için bulunan fark ile plasebo testleri sonucu diğer eyaletler için bulunan farklarının “Öneri 99” öncesi ve sonrası MSPE oranlarının dağılımı incelenmiştir. Oranlara bakmanın temel avantajının, uygun olmayan plasebo çalışmalarını dışarıda bırakmak için bir kesme noktası seçmeyi bertaraf etmesi olduğunu belirten yazarlar, Kaliforniya için “Öneri 99” sonrası ile öncesi MSPE değerleri arasındaki farkın yaklaşık 130 kat olduğunu bulmuştur (bkz Grafik 11).

Grafik 11. Öneri 99 MSPE Oranı: Kaliforniya ve 38 Kontrol Eyaleti



Kaynak: https://www.nber.org/system/files/working_papers/w12831/w12831.pdf

Abadie ve Gardeazabal (2003) ile Abadie ve Diğerleri (2010) sentetik kontrol yönteminin alan/bölge bazlı politika ve değişikliklerde kullanılabileceğiyle ilgili bu bulgular dikkat çekmiş ve çeşitli alanlarda uygulanmıştır. Örneğin, İskoçya'nın 32 Toplumsal Planlama Ortaklığından (CPP) birinde uygulanan yangın ve kurtarma hizmetlerine yönelik yeni bir yaklaşımın yangına bağlı ölümler üzerindeki etkisini değerlendirmek istediğimizi varsayıldığında etkilenmeyen 31 CPP bir “kontrol havuzu” oluşturmaktadır. Müdahale ile aynı anda yangın ve kurtarma hizmetlerinde başka değişikliklerin de gerçekleştiği tüm CPP'leri kontrol havuzunun dışında tutulmakta ve sentetik kontrol için yangına bağlı ölümlerde müdahale öncesi seviyelere ilişkin bilgiler ve geri kalanlarda bu seviyelerin tahminçileri kullanılmaktadır. Kontrol grubu, söz konusu havuzda, müdahale alanındaki müdahale öncesi eğilime benzeyen tüm alanların ağırlıklı ortalamasıdır. Daha sonra müdahalenin etkisi, müdahale ve sentetik kontrol gruplarının müdahale sonrası eğilimleri arasındaki fark olarak tahmin edilmektedir (Craig, 2015). Bir başka çarpıcı örnek Kutu 3’de sunulmuştur.

Kutu 3. Sentetik Kontrol Yöntemi Örneği

2011 yılında Şikago Cubs takımının genel yöneticisi olarak görevlendirilen Theo Epstein, 103 yıllık hasretten sonra 2016 yılında dünya serisini kazandı ve Cubs taraftarları tarafından doğaüstü güçleri olduğu savunuldu. Peki Epstein bu zaferle Illinois'deki gelirleri artırdı mı?

Bu sorunun yanıtını araştırmak için Amerika Topluluk Araştırması Biriminin (American Community Survey) 2006-2015 dönemine ait 500 gözlem verisinden faydalanılmıştır. Veriler, eyalet düzeyi gelirler (bağımlı değişken), yaş, eğitim, medeni durum ve gazilik durumu kullanılmıştır. Illinois müdahale birimi iken kontrol grubu için toplam 10 kontrol birimi kullanılmıştır. Müdahale olarak ifade edilen ise Theo Epstein'in kendisi olarak ifade edilmiştir.

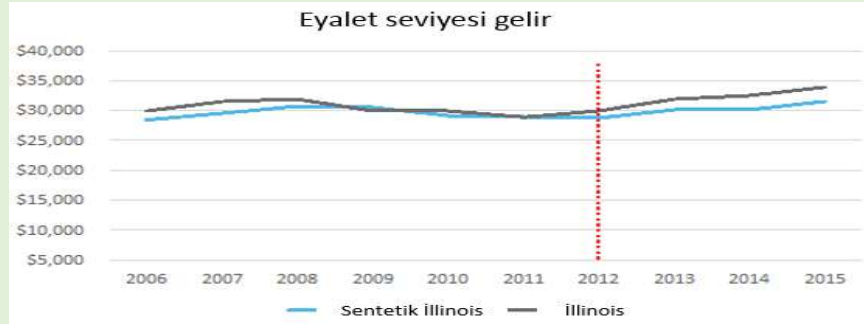
Öncelikle müdahale öncesinde Illinois için oluşturulan sentetik kontrol grubunun Illinois ile ne ölçüde benzeştiğine bakılmaktadır.

Tablo 8. Araştırma Çalışması - Grup Ortalamaları

Değişken	Illinois	Sentetik Illinois	ABD Ortalama
Yaş	52	52	52,2
Üniversite mezunu	%28,4	%28,4	%26,5
Evli	%84,5	%84,4	%86,1
Gazi	%11,3	%11,3	%12,9

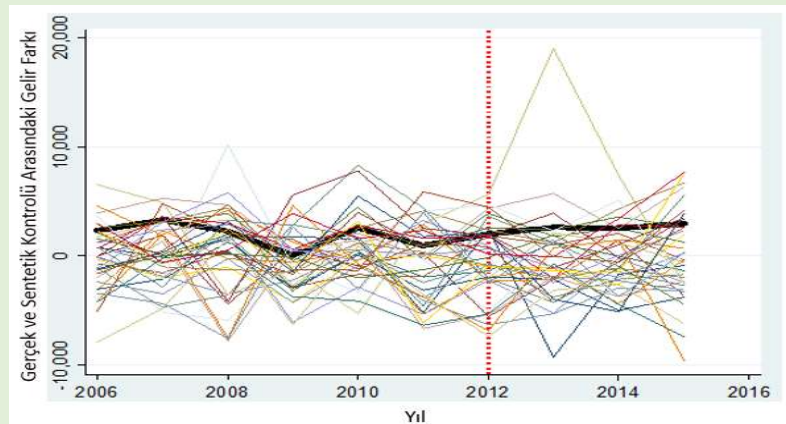
Tablodan da görüleceği üzere gözlemlenebilir değişkenler itibarıyla sentetik Illinois, gerçek Illinois ile neredeyse mükemmel eşleşmektedir. Grafik 12'de Theo Epstein öncesi iyi bir örtüşme görülmüyor.

Grafik 12. Müdahale ve Kontrol Grubu Gelir Eğilimi



Bir sonraki adımda tüm kontrol eyaletleri için plasebo sentetik kontrolleri uygulanmıştır. Ardından, müdahale sonrasında Illinois ile sentetik Illinois arasında gelir farkı araştırılıp sonuç kontrol eyaletlerinin kendi plasebo sentetik kontrolleriyle karşılaştırılmıştır. Eğer Illinois'deki fark diğerlerinden anormal bir şekilde farklı çıkarsa bu fark Epstein'e atfedilebilecekken, sonuç beklendiği gibi çıkmamıştır. Aşağıdaki grafikteki kalın siyah çizgi Illinois'e ait. Grafik 13'te anlaşılabileceği üzere Illinois'deki gelir farkı, diğer eyaletlerden çok da farklı değildir. Bu sebeple Theo Epstein'in eyalet gelirine bir etkisi olmadığı şeklinde çıkarımda bulunulmuştur.

Grafik 13. Müdahale Sonrası Gelir Farkları



Not: Grafik tüm eyaletlerde mevcut olan gelir farklarını sunar. Illinois siyah renkle kalın olarak vurgulanmıştır.
Kaynak: Thompson, 2017.

Sadece coğrafi olarak değil, belirli sınırlar ihtiva eden veya başka alanlarda uygulanan değişikliklerin etkisinin sentetik kontrol yöntemiyle değerlendirilebileceğiyle ilgili literatürde örnekler bulunmaktadır.²⁸

Brezilya’da tropikal ormanların tükenmesini önlemeye yönelik uygulanan bir müdahalenin etkilerini araştıran Sills ve diğerleri (2015) ise, söz konusu araştırmada uygulanan sentetik kontrol yönteminden hareketle yöntemde takip edilen adımları şu şekilde açıklamıştır:

1. Müdahalenin uygulandığı alan ile müdahale öncesi ve sonrasında gözlemlenebilen ilgili sonuç değişkeni tanımlanır.
2. Sonucu doğuran süreçler bakımından yapısal olarak benzerlik gösterdiği değerlendirilen kontrol birimlerinden oluşan muhtemel kontrol grubu seçilir. Bu havuz bölgedeki tüm birimler olabileceği gibi regresyon ve eşleştirmede kullanılan ortak değişkenler bakımından benzer birimler de olabilir. Bunlar sosyoekonomik özellikler (kişi başına düşen gelir, ulaştırma altyapısı, eğitim durumu vb.), biyofiziksel durum (yağış miktarı, arazi kullanımı için kıyıdaş yüzdesi vb.) veya siyasi rejim (aynı ülke veya eyalet) içerebilir. Özelliklerin doğrusal olmayan etkilere sahip olması bekleniyorsa, doğrusallığa yakın bir değer elde etmek için havuzu, müdahale edilen birimin özelliklerine makul bir ölçüde yakın bulunan birimlerle sınırlamak oldukça önemlidir.
3. Sentetik kontrol yöntemi söz konusu havuzdan kontrol sonuçlarının doğrusal bir kombinasyonunu tanımlamak için bu potansiyel kontrol birimleri üzerindeki ağırlıkları (W) seçer. Bu ağırlıklar etki tahminini belirler. Bu, herhangi bir müdahale sonrası dönem için ağırlıklı ortalama veya sentetik sonuç ile müdahale edilen alandaki sonuç arasındaki farktır.
4. Sonuç değişkeninin (Y), gözlemlenen (Z) ve gözlemlenmeyen (U) faktörlerden etkilendiği varsayılırsa $Y = \beta * Z + U$ ’dur. Sentetik kontrol yöntemi, benzer müdahale öncesi Y ve benzer Z ile sonuçlanan W ağırlıklarını arar ki bu da benzer U anlamına gelir. Bazı özelliklerin kombinasyonu ile sonuç farklılıkları hariç tüm farklar aynı anda minimuma indirgenemeyeceğinden, Z değişkenlerine ve müdahale öncesi Y için her yıla ağırlıklar atamak için başka bir V vektörü gerekir. Böylece, kullanılan yazılıma programlanmış bir minimizasyon uygulanarak sentetik kontrol oluşturulur.
5. Sentetik kontrolün kalitesi, ağırlıklı sentetik sonuçların müdahaleden önceki yıllarda, müdahale edilen birimin sonuçlarıyla ne kadar yakından eşleştiğiyle ölçülür. Bunun bir ölçüsü ortalama kare tahmin hatasıdır (MSPE). Ancak en iyi uyum her zaman bir yanlışlık meselesidir ve kısmen sentetik bir grubun, müdahale öncesi yıllarda belirlenen sonuçların dönüm noktaları açısından müdahalede bulunulan alanı yansıtip yansıtmadığına bağlı olabilir. Eğer dönüm noktaları benzer ise, müdahalede bulunulan alan ile sentetik kontrol birimleri büyük olasılıkla ortak faktörlere karşı benzer duyarlılığa sahiptir. Eğer müdahale öncesi dönemde sentetik sonuç ile gerçek sonuç arasında sistematik bir hata varsa, bu durum müdahale sonrasına da yansiyarak etki

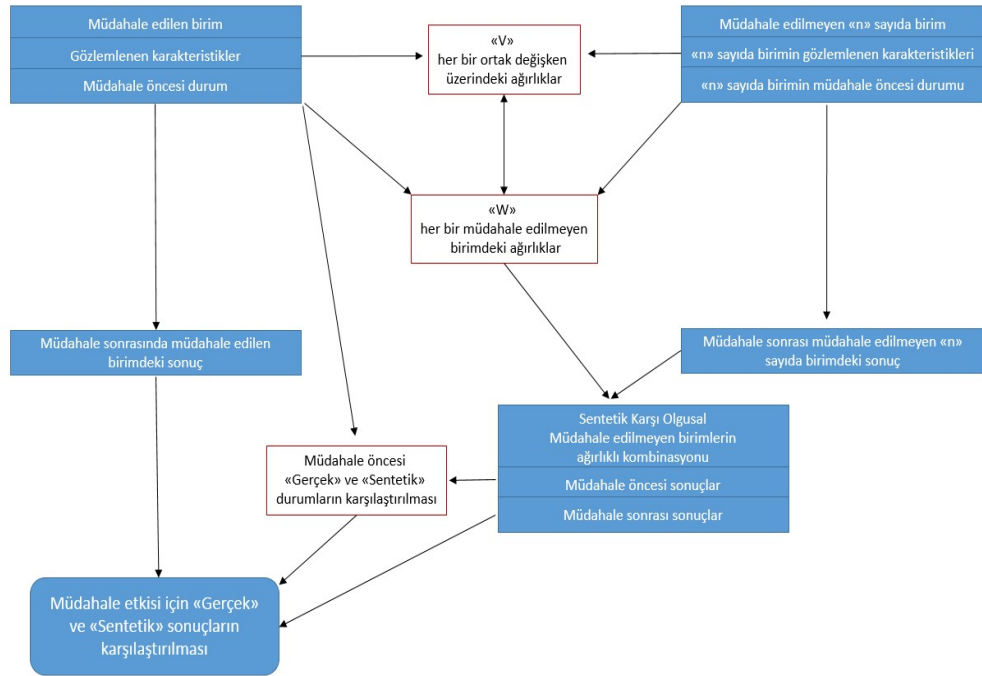
²⁸ Bakınız <https://www.latentview.com/blog/a-b-testing-in-the-absence-of-a-control-group/> (Erişim tarihi 27.02.2024)

tahmininde yanlılığa yol açabilir. Eğer sentetik kontrol kalite kontrolünü geçerse, etki gerçek ve sentetik sonuçlar arasındaki fark olarak ölçülür.

6. Fark elde edildikten sonra farkın sıfırdan önemli ölçüde farklı olup olmadığı değerlendirilmelidir. Müdahale edilen tek bir gözlem olan durumlarda ortak değişken 'dengesini' gibi standart istatistiksel benzerlik testleri mümkün değildir. Bu nedenle, tahmindeki 'gürültüyü (hata)' karakterize etmek ve dolayısıyla tahmini etkilerin bu gürültüden ayırt edilip edilemeyeceğini değerlendirmek için plasebo testleri gibi diğer yöntemlerin kullanılması gerekir. Bir yaklaşım, ağırlıklı, karşıolguasal sonuç için ampirik bir güven aralığı oluşturmak amacıyla donör havuzunun alt örneklerini çizerek sentetik kontrolü başlatmaktır.

Sills ve Diğerleri (2015)'nin önerdiği adımların yol haritası Şekil 4'de sunulmuştur.

Şekil 4. Sentetik Kontrol Grubu Oluşturma Adımları



Kaynak: Sills ve diğerleri (2015)

Sentetik kontrol yöntemi ile Farkların Farkı yöntemi arasında bazı benzerlikler vardır. Aradaki en temel fark, Farkların Farkı yönteminde paralel eğilim varsayımının ön koşul olmasıdır. Sentetik kontrol yönteminde bu durum esnetilmiştir. Farkların Farkı daha çok örneklemin olduğu durumlarda daha etkilidir. Farkların Farkı yönteminde alan veya bölgeye özel ortaya çıkabilecek gelişmeler, yeteri kadar hesaba katılmayabilmektedir. Sentetik kontrol yöntemi bu sorunu, yardım uygulanan alandaki eğilimi sentetik olarak oluşturulan bileşik alandaki eğilimi karşılaştırarak aşmaya çalışmaktadır (Craig, 2015). Kreif ve diğerleri (2015) de Farkların Farkı yönteminin gözlemlenemeyen karıştırıcıların etkilerinin zaman içinde sabit olduğunu varsaydığını ifade ederek, sentetik kontrol yönteminin, kontrol grubunu müdahale grubuna benzer olması için müdahale öncesi özelliklere göre yeniden ağırlıklandırarak bu etkilerin zaman içinde değişmesine izin verdiğini belirtmektedir. Öte yandan, Abadie (2021) sentetik kontrol yönteminde

müdahale edilen birimin birden fazla olmasının ilave kavramsal zorluklara yol açmayacağını, her birim için müdahale etkisinin ayrı ayrı tahmin edilebileceğini ve gerekiyorsa ikinci adımda bir araya getirilebileceğini; ancak birden fazla müdahale edilen birimin varlığının tahmin konusunda bazı uygulama sorunlarına yol açabileceği gibi ilişki konusunda avantaj ve dezavantaj sunabileceğini savunmuştur.

Bazı çalışmalar, etki değerlendirmeyi sentetik kontrol yöntemi ile Farkların Farkı yöntemini ayrı ayrı kullanmak suretiyle gerçekleştirmektedir. Bu çalışmalarda zaman zaman etkinin farklı çıkmaktadır. Örneğin, Verheuvél ve diğerleri (2023), Hollanda Devlet Hazine Ajansı'nın fonlama stratejisindeki değişikliğin piyasada yol açtığı etkiyi araştırmıştır. Yazarlar, iki yöntemle farklı sonuç bulmuş ve bunun nedenini şu şekilde açıklamıştır.

Kontrol grubu ülkelerinin seçimi farklı olduğundan yöntemler arasındaki sonuçlar farklılık göstermektedir. Sentetik kontrol yöntemi ve kısıtlı regresyon, kontrol grubu olarak Finlandiya, Fransa, Almanya ve Avusturya'ya en yüksek ağırlığı verirken, diğer ülkeler küçük bir ağırlık atamaktadır. Bu ülkelerdeki getiri nispeten Hollanda'daki getiriye yakındır. Bununla birlikte, Farkların Farkı yöntemi aynı zamanda Hollanda'dan daha yüksek getiriye sahip olan Belçika, İtalya, Portekiz ve İspanya'yı da kontrol birimleri olarak eşit ağırlık atamaktadır. Bu ülkelerin getirisi Hollanda'nın getirisinden daha fazla düşmüştür; çünkü müdahale öncesi dönemde önemli ölçüde daha yüksek seviyelerde kalmışlardır. Bu durum, sonuçlardaki farkı açıklamaktadır. Örneklem büyüklüğü sınırlı olduğunda her kontrol biriminin tahmin edilen katsayı üzerinde büyük etkisi vardır. Bu nedenle müdahale birimine benzer kontrol birimleri seçilmelidir. Sentetik kontrol yönteminin ve kısıtlı regresyonun ana özelliği budur.

Kreif ve diğerleri (2015) ise daha önce Farkların Farkı yöntemiyle etkisi değerlendirilen ve birden fazla müdahale birimi bulunan sağlıkta performansa bağlı ödeme sisteminin (P4P) riske göre ayarlanmış hastane ölüm oranına etkisini sentetik kontrol yöntemiyle değerlendirmiştir. Farkların farkı analizinin aksine, sentetik kontrol yöntemi, teşvik edilen koşullar için P4P sisteminin ölüm oranını önemli ölçüde azaltmadığını tespit etmiştir. Aynı zamanda teşvik edilmeyen koşullar için ise ölüm oranında istatistiksel olarak anlamlı bir artış ortaya çıkmıştır. Bu bulgulardan hareketle, yardım uygulanan alan, bölge veya birimin tek veya çok az olması durumunda sentetik kontrol yönteminin kullanılabileceği, yardımdan yararlananların sayısının fazla ve kontrol grubu için yeterli gözlem olan durumlarda ise Farkların Farkı yönteminin tercih edilebileceği değerlendirilmektedir.²⁹

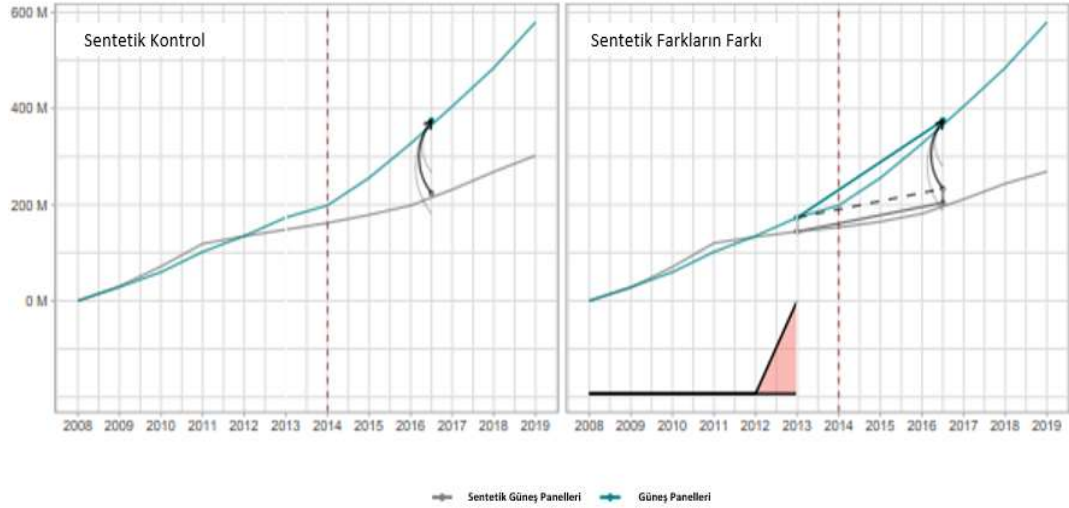
Arkhangelsky ve diğerleri (2019), Farkların Farkı ve Sentetik Kontrol yöntemlerinin ampirik olarak uygulandığı durumların farklı olmasına karşılık³⁰ her iki yöntemin dayandığı temel varsayımların birbiriyle ilişkili olduğunu belirterek “Sentetik Farkların Farkı (SFF)” yöntemini ortaya atmıştır. SFF yöntemi, gözlemsel çalışmalarda özellikle paralel eğilim varsayımının ihlal edildiği durumlarda nedenselliği iyileştirmeyi amaçlamaktadır (Villabona, 2022).

²⁹ Sentetik Kontrol yönteminin R programında uygulanması hakkında bilgi için bakınız <https://www.codingthepast.com/2023/07/21/Synthetic-Control.html>.

³⁰ Çalışmanın bahsettiği farklılık, Farkların Farkının daha çok müdahale birimi olması durumunda uygulanması, Sentetik Kontrol yönteminin ise tek veya çok az birimin müdahaleye tabi tutulması durumudur.

Bu yöntem, Sentetik Kontrol yönteminde olduğu gibi müdahale öncesi eğilimleri ağırlıklandırarak ve eşleştirme yaparak paralel eğilim varsayımını esnetirken, Farkların Farkı yönteminde olduğu gibi birim düzeyindeki toplam değişimlere göre geçerli geniş panel çıkarımını mümkün kılmaktadır (Arkhangelsky ve diğerleri, 2019). Bu sebeple, her iki yöntemin de güçlü yanlarını kullanmaktadır. Grafik 14’te etki Sentetik Kontrol ve SFF yöntemiyle gösterilmektedir.

Grafik 14. Sentetik Kontrol Yöntemi vs. Sentetik Farkların Farkı Yöntemi Etkisi



Kaynak: Villabona (2022)

SFF yönteminde kullanılan bazı özellikleri ve gerekçelerini Arkhangelsky ve diğerleri (2019) şu şekilde açıklamıştır:

- Farkların Farkı yöntemi, müdahalenin etkisini zaman ve birim ağırlıklandırması yapmadan iki yönlü etkiler regresyon problemini çözerek tahmin etmektedir. SFF’de ise zaman ve birim ağırlıklandırmaları da dikkate alınmaktadır. Birimlerin ve zamanın müdahale birimleriyle benzer olduğunu ortaya koyarak tahmini daha güçlü kılmaktadır. Ancak, zaman ve birimlerin sonuçlarında küçük bir sistematik farklılık olursa, birim ve zamanın eşit ağırlıklandırılmaması Farkların Farkı yönteminden daha zayıf sonuç vermektedir.
- Birim ağırlıklarının kullanılmasının nedeni müdahale edilen birimler ile kontrol birimlerinin ağırlıklı ortalamasıyla neredeyse paralel olmasını sağlamaktır. Zaman ağırlıklarının kullanılmasının nedeni ise, müdahale sonrası ortalama sonucun her bir kontrol birimi için aynı kontrol birimleriyle bir “sabit” değerde farklılaşması içindir. Her iki ağırlık birlikte Farkların Farkı stratejisinin daha makul olmasını sağlamaktadır.
- Sentetik Kontrol yöntemi ise SFF’nin aksine birim sabit etkisi ile zaman ağırlıklarını regresyon fonsiyonundan çıkarmaktadır. SFF’nin zaman ağırlıklandırmasını kullanmasının nedeni yanlılığı ortadan kaldırabilir ve müdahale sonrası dönemden çok farklı zaman aralıklarının rolünü ortadan kaldırarak sonucun doğruluğunu artırabilir. Birim sabit etkisinin kullanılmasının da faydası bulunmaktadır. Bunlardan birincisi modelin daha esnek olması ile sağlamlık özellikleri güçlendirilir. İkincisi, sonuçlardaki farklılıkları açıklar ve sonuçların doğruluğunu artırır.

Bu yöntemin daha iyi anlaşılması için yukarıda örnek olarak verilen Kaliforniya “Öneri 99” yasa çalışması, Arkhangelsky ve diğerleri (2019) tarafından bu yöntem dahil başka yöntemlerle de tekrarlanmıştır. Söz konusu çalışmada hatırlanacağı gibi 1970-2000 yıllarına ait Kaliforniya ve diğer 39 eyalet verisi kullanılmıştır ve aşağıdaki tanımlamalar yapılmıştır:

$T_{\text{öncesi}} = 19$ yıl müdahale öncesi dönem,
 $T_{\text{sonrası}} = T - T_{\text{öncesi}} = 12$ yıl müdahale sonrası dönem,
 $N_{\text{kontrol}} = 38$ müdahale edilmeyen eyalet ve
 $N_{\text{müdahale}} = 1$ müdahale edilen eyalet (Kaliforniya).

Yapılan çalışma sonrasında Tablo 9’daki sonuçlara ulaşılmıştır.³¹ Sonuçlardan görüleceği üzere sonuç tüm yöntemlerde farklı çıkmaktadır. Çalışma, zaman ağırlıklarının ve birim sabit etkilerinin kullanılmasının sonucu değiştirdiğini, Sentetik Kontrol ile sentetik Farkların Farkı yoluyla bulunan sonuçların farklı çıkması durumunda sentetik Farkların Farkı yöntemiyle bulunan sonucun daha güvenilir olduğunu savunmuştur.³²

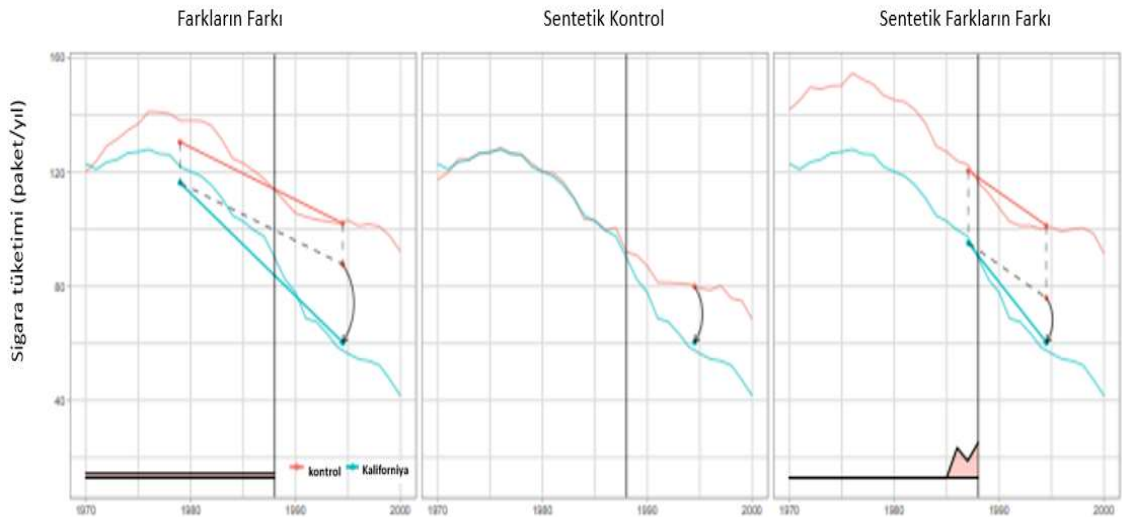
Tablo 9. Elde Edilen Bulgular

	Sentetik Farkın Farkı	Sentetik Kontrol	Farkların Farkı	Matriks Tamamlama	Kesikli Sentetik Kontrol
Tahmin	-15,6	-19,6	-27,3	-20,2	-11,1
Standart hata	(8,4)	(9,9)	(17,7)	(11,5)	(9,5)

Kaynak: Arkhangelsky ve Diğerleri (2019)

Grafik 15’te söz konusu “Öneri 99” un üç yöntemle uygulanması grafikleri gösterilmektedir.

Grafik 15. “Öneri 99” Üç Yöntem Sonrası Sonuçları



Kaynak: Arkhangelsky ve Diğerleri (2019)

³¹ Çalışmada kullanılan algoritma için bakınız Arkhangelsky ve Diğerleri (2019).

³² Sentetik Farkların Farkı yönteminin R programında uygulanması hakkında bilgi için bakınız <https://synth-inference.github.io/synthdid/>.

3.4. Eğilim Skoru Eşleştirme Yöntemi

Temelde müdahale öncesi gruplardaki gözlem birimlerinin özellikleri kullanılarak müdahale ve kontrol grubundaki gözlemlerin eşleştirilmesi fikrine dayanan *Eğilim Skoru Eşleştirme Yöntemi*, karşılaştırılabilir müdahale ve kontrol grupları elde edilmesini ve böylelikle müdahalenin nedensel etkisinin tahmin edilmesini sağlayan bir yöntemdir.

Eğilim Skoru Eşleştirme Yöntemi aşağıda yer alan üç aşamalı bir yaklaşım üzerine kurulmaktadır (Çukurova Kalkınma Ajansı, 2021).

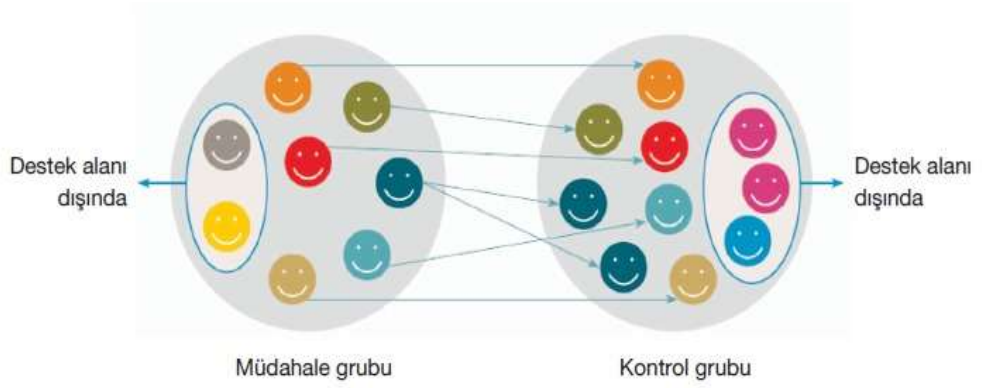
1. Probit³³ veya lojistik regresyon analizi yapılarak yardımdan yararlanmayı ölçen bağımlı değişkenin müdahale öncesi özellikler üzerine regresyonu alınır. Böylece her bir özellik için katılımla ilişkiyi gösteren bir katsayı tahmin edilmiş olur. Örneğin, bağımlı değişken yardımdan yararlanma (0: yararlanmamış, 1: yararlanmış), bağımsız değişkenler ise firmanın yardıma başvurmasını etkileyen değişkenler (son üç yıl içinde ihracat yardımlarına ilişkin bilgilendirme programına katılmış olma, son üç yıl içinde ihracat yapmış olma, satışlar içindeki ihracatın oranı, firma türü, sektörü, üretim bölgesi, verimliliği) olsun. Yapılan regresyon analizi sonucunda bağımsız değişkenlerin firmaların yardımdan yararlanma olasılıkları tespit edilir.
2. Müdahale ve kontrol grubundaki her bir gözlem için yararlanıcı olma olasılığı tahmin edilir. Başka bir ifadeyle yardımdan yararlanan ve yararlanmayan firmalar için eğilim skoru hesaplanır. Böylelikle gözlemlenen özelliklere yani seçilen bağımsız değişkenlere dayalı olarak bir firmanın yardımdan yararlanma olasılığı hesaplanmış olur.
3. Müdahale ve kontrol grubundaki gözlemler bu tahmin edilmiş olasılık değerleri kullanılarak eşleştirilir. Sonuç olarak, müdahale ve kontrol gruplarının birbirine benzeyen birimlerden oluşması ve böylece bu grupların birbirleri için iyi bir karşı-olgusallık oluşturmaları söz konusu olacaktır.

Üçüncü adımda bahsedilen eşleştirme, çeşitli yöntemlerle yapılabilmektedir. Yazında, bire bir eşleştirme, en yakın komşu eşleştirme, tam eşleştirme, kabalaştırılmış tam eşleştirme, Kernel eşleştirme, yarıçap eşleştirme, lokal lineer regresyon eşleştirme, Spline eşleştirme ve Mahalanobis eşleştirme gibi farklı eşleştirme yöntemlerinin kullanıldığı görülmektedir. Ancak sonuçlar kullanılan eşleştirme yöntemine göre değişiklik gösterebileceğinden bunların sağlamlığından emin olmak için birden fazla eşleştirme yöntemi ile sonuçları test etmek gerekmektedir (Glewwe ve Todd, 2022).

Bu yöntemin en büyük avantajı, hesaplamanın net tanımlanan aşamalar çerçevesinde gerçekleştirilmesidir. Yeterli sayıda ilgili veri toplandığı durumda müdahale ve kontrol gruplarının birbirinin karşı-olgusallığını iyi şekilde yansıtmaları olasıdır.

³³ Lojistik regresyona alternatif olarak kullanılan bir model olup kategorik veri analizlerinden biridir. Bağımlı değişken ikili ya da çoklu, bağımsız değişkenler ise kategorik veya sürekli olabilmektedir.

Şekil 5. Birebir Eşleştirme ile Destek Alanında Olan ve Olmayan Gözlemler



Kaynak: Kalkınma Programlarının Değerlendirilmesinde Etki Değerlendirmesi Yöntemleri ve Türkiye’de Uygulamalar (2017)

Yöntemin zayıf olduğu iki nokta bulunmaktadır. Birincisi, bu yöntem müdahale ve kontrol grupları arasındaki farkların yalnızca gözlemlenebilir özelliklere dayalı olduğu varsayımına dayanmaktadır. Bu çok güçlü bir varsayımdır. Müdahale ve kontrol grubu arasında gözlemlenemeyen değişkenler olabilmekte ve bu değişkenler yardımdan yararlanma durumu ya da sonuçlarla ilişkili olabilmektedir. Bu nedenle, gözlemlenemeyen özelliklerden dolayı ortaya çıkacak seçim yanlılığı sorunu Eğilim Skoru Eşleştirme ile çözümlenememektedir. Yöntemin bir diğer dezavantajı ise büyük çaplı veri gerektirmesidir. Müdahale ve kontrol gruplarının en azından gözlemlenebilen özellikler bakımından doğru bir şekilde eşleştiğinden emin olunması için mümkün olduğunca çok sayıda değişkenle ilgili veri toplanması gerekmekte olup bu durum her zaman kolay olmayabilmektedir (Polat ve Aktakke, 2017).

Kutu 4. Tarım ve Gıda Firmaları Üzerinde AB Katılım Öncesi Fonlarına İlişkin Etki Değerlendirme: Hırvatistan Örneği

Avrupa Birliği (AB) aday ülkelere sağlanan AB katılım öncesi fonları, bu ülkelerin üyeliğe uyum sağlamalarında önemli bir rol oynamaktadır. Hırvatistan aday ülke statüsünde iken Tarım ve Kırsal Kalkınma için AB Katılım Öncesi Programı hedeflerinden biri olan tarım ve gıda üretimi sektörlerindeki firmaların rekabet gücünü artırmak için iki önemli fon kullanmıştır. SAPARD (Tarımsal ve Kırsal Kalkınma İçin Özel Eylem Programı) ve IPARD (Katılım Öncesi Mali Yardım-Kırsal Kalkınma Aracı) olarak adlandırılan bu fonların kullanım süresi, iki AB programlama dönemini ve Hırvatistan’da 2009’dan 2014’e kadar süren bir durgunluk dönemini kapsamıştır.

Katılım Öncesi AB fonlarının Hırvatistan’da tarım ve gıda sektöründeki yararlanıcılar üzerindeki etkisini değerlendirmek için yarı-deneysel tasarım kullanılmıştır. Bu tasarım aynı zamanda katılım öncesi programların yararlanıcılar üzerindeki net etkisinin algılanan değerlendirme eksikliğine karşı-olgusal bir yaklaşım kullanarak yanıt vermektedir. Çalışmadaki etki değerlendirmenin temeli, fonlardan yararlanan firmalar (müdahale uygulanan firmalar) ile yararlanmayan benzer firmalar (kontrol grubu firmaları) arasında belirlenen sürdürülebilirlik, üretkenlik ve iş performansı göstergelerinin karşılaştırılmasıdır. Bu çalışmada, Eğilim Skoru Eşleştirme ve Farkların Farkı (FF) yöntemleri bir arada kullanılmıştır. Fonların alındığı yıl “t” ile ifade edilirken, analiz dönemi “t – 1” ile “t + 5” yıllarını kapsamaktadır.

Yapılan hesaplamalar sonucunda, SAPARD ve IPARD fonlarının, “dikkate alınan döneme”, “performans göstergesinin türüne” ve “firma ölçeğine” göre farklı etkilere sahip olduğu görülmüştür. Ayrıca fonlardan yararlanan firmaların beş yıl boyunca istihdam, banka kredisi temini ve ciro, işgücü verimliliği ve sermaye girdisinde daha yüksek bir büyüme yaşadığı; ancak Altman’ın Z-skor modeli ile ölçülen ücret artışının ve finansal istikrarın etkilenmediği gözlenmiştir. Elde edilen sonuçlar, üretim kapasitelerinin modernizasyonu ve yükseltilmesi olan fonların amacını büyük ölçüde doğrulamaktadır.

Firma ölçeği bazında sonuçlara bakıldığında ise “t+1” yılında piyasada kalabilme etkisinin, iç pazara odaklanan ve Orta Hırvatistan’da yerleşik olan mikro ve küçük firmalarda yoğunlaştığı görülmüştür. Fonları kullanan mikro ölçekli firmalar, satışlarda ve katma değerde büyük firmalara göre on kat daha fazla büyüme kaydetmiştir. Yine, mikro ölçekli firmalarda istihdam ve TFV (toplam faktör verimliliği)’de daha fazla artışın yaşandığı ve borçluluğun azaldığı gözlenmiştir.

Kaynak: Marin Kukoć, Bruno Škrinjarić ve Josip Juraćak, 2020

3.4.1. Eğilim Skoru Modelinin Kurulması ve Kovaryat Seçimi

Eğilim skoru modelinde, müdahale etkisine yönelik sonuç değişkeni yerine müdahaleye maruz kalma durumu bağımlı değişken olarak kullanılmaktadır. Modele dahil edilebilecek olası değişken setleri, başlangıçta ölçülen kovaryatlar, müdahaleyle ilişkili kovaryatlar, sonuçla ilişkili kovaryatlar ve hem müdahale hem de sonuçla ilişkili kovaryatlardır(Totik Doğan, Püren, Karakaya ve Alparslan, 2023).

Kovaryatlar genellikle istatistiksel modellemede, özellikle neden-sonuç ilişkilerinin araştırıldığı çalışmalarda, analiz sonuçlarının doğruluğunu artırmak amacıyla kontrol edilen değişkenlerdir (Fan, 2010). Bağımsız değişkene benzer şekilde bağımlı değişkeni etkileyen, kontrol edilmesi gereken açıklayıcı değişkenlerdir. Bir değişken, eğer bağımlı değişkenle istatistiksel olarak ilişkilirse, kovaryat olarak kabul edilir. Bu tanıma göre, ölçülebilir olan ve bağımlı değişkenle anlamlı bir ilişkisi olduğu düşünülen,

aynı zamanda müdahaleye maruz kalma olasılığını etkileyen gözlemlenebilir başlangıç özelliklerine ilişkin herhangi bir değişken, potansiyel bir kovaryat niteliği taşır.

Özellikle rassallığın mümkün olmadığı gözlemsel araştırmalarda kovaryatlar büyük önem taşımaktadır. Rassal Kontrollü Deneylerde müdahale ataması rastgele yapıldığından, gruplar arasında başlangıçta kovaryatlar açısından bir denge oluşmakta ve bu da nedensel çıkarımların güvenilirliğini artırmaktadır. Ancak rassallığın olmadığı durumlarda, müdahale ve kontrol grupları başlangıçta benzer özellikler göstermeyebilir. Bu durum, kovaryatlar açısından gruplar arasında dengesizlik yaratarak müdahale etkisinin yanlı (biased) tahmin edilmesine yol açabilmektedir.

Bu nedenle, gözlemsel çalışmalarda kovaryatların dikkatle belirlenmesi gerektiği değerlendirilmektedir. Bu bağlamda kullanılan en yaygın yöntemlerden biri, eğilim skoru (propensity score matching) eşleştirmedir. Eğilim skoru, bireylerin başlangıçtaki kovaryatlarına göre müdahale grubuna atanma olasılığını ifade etmektedir. Her ne kadar bu olasılık doğrudan gözlemlenemese de eğilim skorları kullanılarak müdahale grubundaki gözlemler kontrol grubundakilerle eşleştirilebilmekte ya da benzer skorlara sahip gözlemlerden alt gruplar oluşturulabilmektedir. Böylelikle, kovaryatlar açısından dengelenmiş gruplar elde edilerek, Rassal Kontrollü Deneylere benzer bir yapı oluşturulabilmektedir.

Rosenbaum ve Rubin'e (1983) göre belirli varsayımlar altında, eğilim skoruna göre eşleştirme yapmak, gözlenen yardımcı değişkenlere ilişkin özelliklerden kaynaklanan karıştırıcı etkiden (confounding) doğan yanlılığı ortadan kaldırmak için yeterlidir.

Sonuç olarak, eğilim skoru eşleştirme yönteminin güvenilir sonuçlar üretebilmesi için hangi değişkenlerin modele dahil edileceğine ilişkin teorik temelli ve dikkatli bir seçim süreci büyük önem taşımaktadır (Pattanayak, Rubin ve Zell, 2011).

Kovaryat Seçimi Nasıl Olmalıdır?

Karıştırıcı etki, sonuç³⁴ üzerinde etkisi olan, diğer bir deyişle müdahale ve sonuç değişkeni arasında *nedensel olmayan* bir ilişki varmış gibi görünmesine yol açan başlangıçtaki kovaryatlar nedeniyle ortaya çıkmakta olup deneysel olmayan çalışmalarda müdahale etkisi tahminlerinin geçerliliği bakımından önemli bir tehdit oluşturmaktadır (Yang, ve diğerleri, 2019). Müdahale ve kontrol grupları başlangıçta kovaryatlar açısından dengeli değilse, gözlenen etkinin tamamı veya bir kısmı müdahaleden ziyade, başlangıçtaki farklılıklarından kaynaklanıyor olabilmektedir. Bu durum, müdahale etkisi tahminlerinin yanlı (biased) ve güvenilmez olmasına yol açmaktadır.

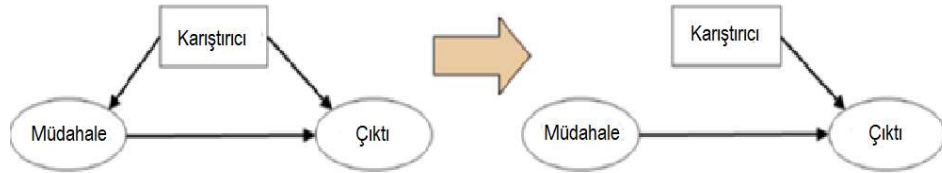
Yanlılığı önlemek için bu karıştırıcı etkileri ortadan kaldırmak ya da en aza indirmek gerekmektedir. Ancak gözlemsel çalışmalarda rassallık olmadığı için karıştırıcı etkilerin ortadan kaldırılması doğrudan mümkün olmamaktadır. Bu nedenle karıştırıcı etkilerin istatistiksel yöntemlerle kontrol edilmesi gerekmektedir.

³⁴ Müdahaleye maruz kalma olasılığı

Eğer eğilim skorları, karıştırıcı etkilere sebep olan yardımcı değişkenlere dayanarak doğru şekilde hesaplanırsa ve ardından müdahaleye maruz kalan ve kalmayan gözlemler eğilim skorlarına göre eşleştirilirse, gruplar arasında başlangıçta yardımcı değişkenlerden kaynaklanan dengesizlikler önemli ölçüde azaltılmakta ya da ortadan kaldırılmaktadır.

Bu yöntemle, **RCT'ye** yaklaşan, yani karıştırıcı etkilerden arınmış ve daha güvenilir müdahale etkisi tahminlerinin elde edilmesi mümkün olmaktadır.

Şekil 6. Müdahale, Karıştırıcı Değişken ve Çıktı Arasındaki İlişki



Yukarıda yer alan şekil, müdahale, karıştırıcı değişken ve sonuç arasındaki ilişkiyi göstermektedir. Burada karıştırıcı, hem müdahaleyi hem de sonucu etkileyen herhangi bir kovaryanttır (soldaki şekil). Eğer karıştırıcı değişken, müdahale grupları arasında başlangıçta dengelenmemişse (örneğin müdahale grubunda sistematik olarak daha yüksek veya daha düşük değerlere sahipse), elde edilen müdahale etkisi tahmini yanlış olmaktadır. Zira bu durumda karıştırıcının etkisi, müdahalenin etkisiyle karışmaktadır.

Rassallık ve eğilim skoru eşleştirme, karıştırıcı değişkenler ile müdahaleye atama arasındaki ilişkiyi ortadan kaldırarak karıştırıcı etkileri kontrol altına almayı hedeflemektedir (sağdaki şekil). Böylelikle, müdahale grupları arasındaki ölçülmüş başlangıç özelliklerindeki dengesizlikler giderilmekte ve müdahale ve kontrol grupları, karşılaştırılabilir hâle gelmektedir.

Eğilim skoru, her bireyin ölçülen başlangıçtaki gözlemlenebilir özelliklerin (kovaryatların) kendine özgü kombinasyonuna dayanarak, müdahaleye maruz kalma olasılığının tek bir skorla özetlenmesidir. Bu nedenle, eğilim skoru, seçilen kovaryatların spesifik kombinasyonuna bağlı olarak değişebilmekte ve seçilen kovaryat seti, tahmin edilen müdahale etkisinin yanlılığı ve kesinliği üzerinde anlamlı bir etkiye sahip olabilmektedir.

Literatürde eğilim skoru ile eşleştirmeye yönelik olarak kovaryat seçimi süreci için iki husus vurgulanmaktadır:

- Sonucu etkilediği düşünülen tüm kovaryatlar, eğilim skorunun hesaplandığı modele dahil edilmelidir. Bu tür kovaryatlar, karıştırıcı etkileri ortadan kaldırmaya yardımcı olmaktadır. Ayrıca sonuç üzerindeki açıklanamayan varyansı azalttığı için tahminin kesinliğine (precision) katkı sağlamaktadır. Sonuçla ilişkili olmayan değişkenler, dengeleme açısından anlamlı katkı yapmamakta ve modelin karmaşıklığını artırmaktadır.
- Sadece müdahale ile ilişkili olan ancak sonuç değişkeni ile ilişkisi olmayan kovaryatlar, eğilim skoru modeline dahil edilmemelidir. Bu tür değişkenler, müdahale alma olasılığını güçlü şekilde tahmin etmektedir. Bu durum eğilim

skorlarını uç değerlere itmekte (0 veya 1'e yaklařır), ortak destek problemlerine yol açmakta, eşleřtirmeyi zorlařtırmakta ve tahminin varyansını řiřirmektedir., Ayrıca müdahale etkisinin doęruluęuna katkı saęlamamaktadır.

Sonucu etkileyen kovaryatların (hem karıřtırıcılar hem de sonuç için baęımsız etkisi olan bařlangıç özellikleri) dahil edilmesi, yanlılıęı ortadan kaldırmaktadır. Bununla birlikte sonuç deęiřkeni ile ilgili olmayan kovaryatların dahil edilmesi ise standart hataların büyümesine neden olmakta, bu doęrultuda müdahale etkisi tahmininin kesinlięini azaltmaktadır. Bu nedenle, sonuç deęiřkeni ile iliřkili olmayan kovaryantlar modele dahil edilmemelidir (Yang, ve dięerleri, 2019).

Bir deęiřkenin yalnızca devlet yardımından yararlanma durumunu (müdahale) etkileyip, sonuç deęiřkeni (örneğin ciro) üzerinde doğrudan veya dolaylı bir etkisinin bulunmaması durumunda, bu deęiřkene göre grupların eşleřtirilmesi gerekmemektedir. Çünkü bu tür deęiřkenler, karıřtırıcı etki içermediklerinden sonuç deęiřkeni açısından sistematik bir yanlılıęa neden olmamakta ve dolayısıyla dengeleme yapılması gerekmemektedir. Eğilim skoru modelleri, esas olarak bireylerin müdahaleye maruz kalma olasılıklarını tahmin etmeye yönelik olup bu nedenle müdahale atamasını etkileyen kovaryatların modele dahil edilmesi yerinde bir yaklařımdır. Ancak yalnızca müdahale atamasını etkileyen kovaryatların modele eklenmesi, modelin parsimonisini³⁵ azaltabilmekte ve tahmin performansını olumsuz etkileyebilmektedir. Bu nedenle, hem müdahale deęiřkeni hem de sonuç deęiřkeni ile iliřkili olan kovaryatların tanımlanarak modele dahil edilmesi, potansiyel karıřtırıcıların kontrolü ve nedensel etkinin sapmasız tahmini açısından önem arz etmektedir. Yalnızca sonucu etkileyen (ancak müdahale ile iliřkisi olmayan) kovaryatlar eğilim skoru modeline dahil edilmemelidir. Öte yandan, bu tür kovaryatlar etki deęerlendirmede baęımsız deęiřken olarak kullanılabilmektedir.

Zayıf prognostik³⁶ potansiyele sahip kovaryatların dahil edilmesi teorik olarak önerilmekle birlikte bu yaklařım, müdahale grubunun az sayıda gözlemden oluřtuęu durumlarda model kararlılıęını olumsuz etkileyebilmektedir.

Eğilim skoru tahmin etmek için lojistik regresyonun kullanıldıęı durumlarda, genel bir kural olarak modele dahil edilen her kovaryat için yaklařık 6-10 müdahale gören birim olması önerilmektedir (Yang, ve dięerleri, 2019).

3.4.2. Eğilim Skoru Model Anlamlılıęın Teřhisi

Eğilim skoru modelinin anlamlılıęı tanısı, Eğilim Skoru Eşleřtirme yönteminin doęru çalışıp çalışmadıęını ve elde edilen skorlarla müdahale ve kontrol grupları arasında karřılařtırılabilirlik (denge) saęlanıp saęlanmadıęını kontrol etmek için yapılan bir dizi test ve görsel analizdir.

Eğilim Skoru Eşleřtirme modelinin gücü ve ayırt edicilięi Pseudo R², ROC eęrisi ya da histogramlar ve daęılım göstergeleri ile test edilebilmektedir.

³⁵ Parsimoni, istatistiksel modelleme ve ekonometrik analizde, bir modelin veri yapısını en az sayıda parametre veya açıklayıcı deęiřkenle en iyi řekilde temsil etme derecesini ifade eder.

³⁶ Sonuç üzerinde etkisi az ya da zayıf olan anlamına gelmektedir.

Pseudo R²

Pseudo R² kavramı, özellikle lojistik regresyon gibi doğrusal olmayan modellerde, klasik R² (açıklama katsayısı – açıklanan varyans oranı) teriminin doğrudan uygulanamadığı durumlarda modelin açıklayıcılığını değerlendirmek için geliştirilmiş alternatif ölçülerden biridir.

Bilindiği üzere, Eğilim Skoru Eşleştirme yönteminde, bağımlı değişkenin kukla (örneğin müdahale alma durumu: evet/hayır) değişken olması nedeniyle, klasik doğrusal regresyon yerine ikili sonuç değişkenleri için uygun lojistik regresyon modeli kullanılmaktadır. Lojistik regresyon ile veri analiz edilirken, R²'ye eşdeğer bir istatistik bulunmamaktadır. Lojistik regresyon model tahminleri, yinelemeli süreçle elde edilen maksimum olabilirlik (maximum likelihood) tahminleridir³⁷. Bu tahminler varyansı en aza indirecek şekilde hesaplanmadıkları için, klasik EKK yaklaşımındaki uygunluk ölçütleri burada geçerli değildir.

Bununla birlikte, lojistik modellerin uygunluğunu değerlendirmek için çeşitli pseudo R² hesaplamaları geliştirilmiş olup genellikle 0 ile 1 arasında değer almaktadırlar. Ancak, bu katsayı EKK yaklaşımındaki R² gibi yorumlanamamaktadır. Farklı pseudo R² türleri aynı veri üzerinde çok farklı değerler üretebilmektedir (Statistical Methods and Data Analytics, 2025). Bu modellerin performansını değerlendirmede önemli ölçütlerden biri McFadden'in R² değeridir. Söz konusu değer, literatürde sık kullanılan bir ölçüttür ve 0.2 - 0.4 arası değerler, iyi bir uyumu göstermektedir. R² değerinin 0.40 üzeri olması çok güçlü uyum olarak değerlendirilmektedir (Lee, 2025).

Receiver Operating Characteristic (ROC) Eğrisi ve C istatistiği

Model anlamlılığı ve uyumunun değerlendirilmesinde, özellikle sınıflandırma performansının değerlendirilmesinde kullanılan önemli araçlardan biri ROC eğrisi ve C istatistiğidir. ROC eğrisi, modelin farklı eşik değerlerinde doğru pozitif oranı ile yanlış pozitif oranını göstererek, modelin pozitif (devlet yardımı alanlar) ve negatif (devlet yardımı almayanlar) sınıfları ayırt etme başarısını görselleştirmektedir. C istatistiği veya AUC (Eğri Altında Kalan Alan, Area Under the Curve) ise ROC eğrisinin altındaki alanı temsil etmekte ve modelin genel ayırt edicilik gücünü sayısal olarak özetlemektedir. Bu testler, modelin tahminleri ile gerçek durumlar arasındaki uyumun daha ayrıntılı ve performansa dayalı bir değerlendirmesini sağlamaktadır. Böylece modelin ne kadar güvenilir ve etkili olduğunu anlamamıza yardımcı olmaktadır.

Bir ROC eğrisi, dikey ekseninde gerçek pozitif oranı (TPR³⁸) ile yatay ekseninde yanlış pozitif oranı (FPR³⁹) karşılaştırılarak çizilmektedir. Gerçek pozitif oranı, tüm pozitif gözlemler içinde pozitif olduğu doğru tahmin edilen gözlemlerin oranıdır (TP/(TP + FN)). Benzer şekilde, yanlış pozitif oranı, tüm negatif gözlemler içinde pozitif olduğu yanlış tahmin edilen gözlemlerin oranıdır (FP/(TN + FP)). Örneğin, tıbbi testlerde gerçek pozitif oranı, kişilerin söz konusu hastalık için doğru şekilde pozitif test sonucu alma

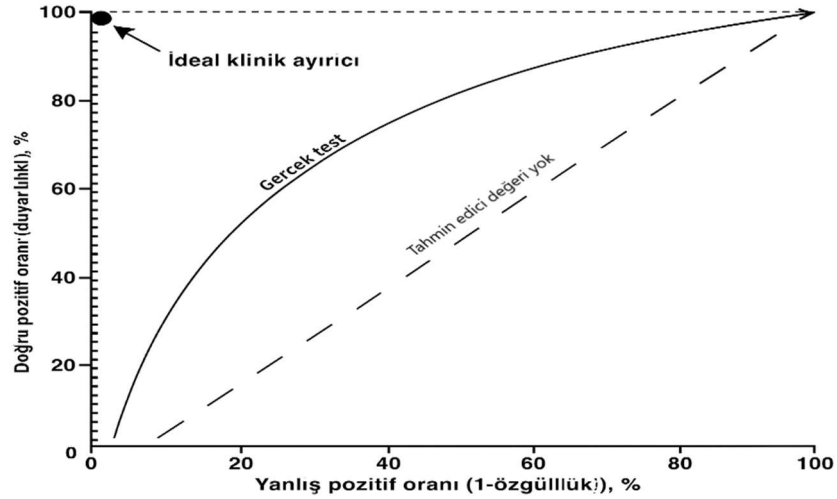
³⁷ Lojistik regresyonun beta katsayıları (model parametreleri) doğrudan formülle hesaplanamamaktadır. Bunun yerine, başlangıçta bazı tahmini değerler atanır. Ardından, bu değerler adım adım güncellenir. Her güncellemede modelin veri üzerindeki uyumu biraz daha iyileştirilir. Bu işlem, yeterince iyi sonuç (yani katsayılar sabitlenene kadar) elde edilene kadar tekrarlanır.

³⁸ True Positive Rate

³⁹ False Positive Rate

oranıdır. Sol üst köşeye daha yakın eğriler veren sınıflandırıcılar daha iyi bir performansa işaret etmektedir. Zira bu nokta TPR'nin 1'e, FPR'nin ise 0'a yakın olduğu ideal bir sınıflandırıcıyı ifade etmektedir. Temel değer olarak, rastgele bir sınıflandırıcının diyagonal boyunca uzanan noktalar vermesi beklenmektedir ($FPR = TPR$). Eğri, ROC uzayının 45 derecelik diyagonaline ne kadar yaklaşırsa, testin doğruluğu o kadar azalmaktadır (Chan). Modelin ayırt edicilik gücünün değerlendirilmesinde C istatistiği (AUC) 0.7–0.8 aralığında 'kabul edilebilir', 0.8–0.9 aralığında 'iyi', 0.9 üzerinde ise 'mükemmel' olarak sınıflandırılmaktadır.

Grafik 16. ROC (Receiver Operating Characteristic) Eğrisi



Farklı modelleri karşılaştırmak için, her modelin performansını tek bir ölçüte özetlemek faydalı olabilmektedir. Yaygın bir yaklaşım, ROC eğrisinin altında kalan alanı hesaplamaktır. Bu, rastgele seçilen pozitif bir örneğin, rastgele seçilen negatif bir örnekten daha yüksek sırada yer alma olasılığına eşdeğerdir. Yüksek AUC'ye sahip bir sınıflandırıcı, zaman zaman belirli bir bölgede daha düşük AUC'ye sahip başka bir sınıflandırıcıdan daha düşük puan alabilir. Ancak pratikte, AUC genel bir tahmin doğruluğu ölçüsü olarak iyi performans göstermektedir (Chan).⁴⁰

3.4.3. Ortak Destek Alanı Kontrolü (Common Support)

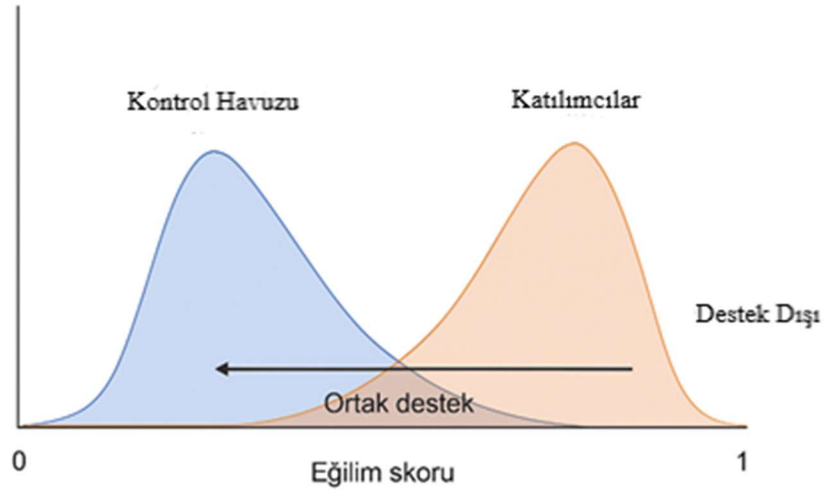
Ortak destek alanı, müdahale ve kontrol grupları arasındaki eğilim skoru dağılımındaki örtüşmeyi ifade etmektedir. Başka bir deyişle, hem müdahale hem de kontrol gruplarında gözlemlerin sahip olduğu eğilim skoru aralığındaki kesişimi göstermektedir.

Ortak destek alanı, müdahale ve kontrol grupları arasında birbirine benzer gözlem bulmanın mümkün olduğu eğilim skoru aralığını temsil etmektedir. Eğilim skoru dağılımında örtüşme bulunmaması durumunda, bu gözlemler için karşılaştırılabilir bir karşı-olgusal üretilmesi mümkün olmadığından analiz geçerli olmamaktadır (Common Support Assumption in Propensity Score Matching, 2023). Zira müdahale grubuna benzeyen, ancak müdahaleyi almayan bir gözlem yok ise nedeni sonuçtan ayırmak mümkün olmayacaktır. Eğilim Skoru Eşleştirme yönteminin etkinliği özellikle, geniş bir

⁴⁰ <https://www.displayr.com/what-is-a-roc-curve-how-to-interpret-it/> Erişim Tarihi: 12.12.2025

ortak destek bölgesinin bulunabilmesi, müdahale grubu ile kontrol grubunun büyüklüklerinin yeterli olması, iki grubun başlangıç özelliklerinin belirli ölçüde benzerlik göstermesi ve eğilim skorunun iyi kalibre edilmiş bir modelden elde edilmesine bağlıdır (Aditya ve Subash, 2019).

Grafik 17. Ortak Destek Alanı (Common Support Area)



Kaynak: <https://www.gov.uk/government/publications/employment-data-lab-information-and-guidance/employment-data-lab-methodology-report>

Ortak destek varsayımının ihlali durumunda eğilim skorlarının yalnızca her iki grubun da gözlemlendiği ortak aralıkta yer alan bireylerin analize dahil edilmesi yaygın bir çözümdür. Eşleştirme yapılamayan gözlemleri dışlamak üzere çalışmanın uygunluk kriterleri yeniden tanımlanabilmektedir. Bu işlem, genellikle aşağıdaki biçimde uygulanır:

- Müdahale grubundaki en düşük eğilim skoru $\geq$ kontrol grubundaki en düşük eğilim skoru
- Müdahale grubundaki en yüksek eğilim skoru $\leq$ kontrol grubundaki en yüksek eğilim skoru.

Bu durum, elde edilen bulguların dışlanan alt gruplara genellenemeyeceği gerçeğini açıkça yansıtmak adına çalışmada net bir şekilde belirtilmelidir.

Hiçbir kontrol birimiyle eşleşmeyen gözlemlerin eşleştirme sürecinden çıkartılması da mümkündür. Ancak bu yöntem dikkatle uygulanmalıdır. Zira bu tür bir uygulama, müdahale ve kontrol grupları arasında yanlılığa ve yapısal farklara yol açabilmektedir (Common Support Assumption in Propensity Score Matching, 2023).

3.4.4. Eşleştirmede Kullanılacak Yöntemin Seçimi ve Denge Testi

Etki değerlendirme sürecinde müdahale etkisinin daha net şekilde gözlemlenebilmesi amacıyla kontrol ve müdahale grupları arasında görece en yakın (en benzer) birimlerin birbirleri ile eşleştirilerek müdahale etkisi dışında diğer etkilerin minimize edildiği varsayılmaktadır. Optimal eşleştirmenin yakalanması için uygulanacak yöntem, verinin türüne (panel, kesit) kovaryatların türü ve boyutuna, örtüşmenin durumu (zayıf, güçlü) gibi birçok koşula bağlıdır. Bu bölümde etki değerlendirmede sık kullanılan

alternatif eşleştirme yöntemlerinin hangileri olduğuna, nasıl çalıştıklarına, ne zaman kullanılması gerektiğine ve avantajları ile dezavantajları ortaya konulacaktır.

Mahalanobis Mesafesi ile Eşleştirme Yöntemi

Mahalanobis mesafesi ile eşleştirme, iki gözlem arasındaki farkı ölçerken değişkenlerin kovaryans yapısını dikkate alan çok değişkenli bir mesafe ölçümüne dayanmaktadır. Bu yaklaşımda mesafe ölçümü, değişkenlerin ölçeklerinden bağımsız olarak yapılmakta ve aralarındaki korelasyon dikkate alınmaktadır. Temel istatistiksel mantık, her müdahale biriminin kovaryat uzayında en yakın kontrol birimi ile eşleştirilmesidir. Hedef, gözlemlenebilir kovaryat farklılıklarını minimize ederek müdahale etkisinin yanlılıktan arındırılmış biçimde tahmin edilmesidir (Abadie ve Imbens, 2006).

Tablo 10. Mahalanobis Mesafesi İle Eşleştirme Yöntemi Özet Tablo

Metodoloji	Avantajlar	Dezavantajlar	Ne zaman kullanılmalı?
- Her gözlemin kovaryat bilgisi kullanılarak Mahalanobis mesafesi hesaplanır. - Mesafe, kovaryans matrisinin tersi ile ağırlıklandırılır. - En küçük mesafeye sahip kontrol birimi ile eşleştirme yapılır.	- Kovaryatlar arası korelasyonu hesaba katar. - Ölçek bağımsızdır. Değişkenlerin farklı birimlerde olması (TL, kişi sayısı, yıl vb.) eşleştirmeyi etkilemez. - Küçük veri setlerinde yüksek doğruluk sağlar.	- Kovaryat sayısı arttığında performans düşer. - Kategorik değişkenlerde doğrudan uygulanması zordur.	- Kovaryatlar az (8'den az) ve sürekli olduğunda. - Kovaryatlar arasında yüksek korelasyon olduğunda.

Kaliperli En Yakın Komşu Eşleştirme Yöntemi

Kaliperli en yakın komşu eşleştirmesi, her müdahale birimini eğilim skoru bakımından en yakın kontrol birimi ile eşleştirirken, yalnızca belirli bir skor farkının (kaliper) altında olan/ belirli bir tolerans aralığında olan (kaliper) eşleşmeleri kabul eden bir yöntemdir. Temel mantık, uzak eşleşmelerin dâhil edilmesiyle oluşabilecek yanlılığı azaltmaktır. İstatistiksel olarak hedef, eşleştirmeyi yalnızca anlamlı derecede benzer birimler arasında gerçekleştirmektir (Austin, 2011).

Tablo 11. Kaliperli En Yakın Komşu Eşleştirme Yöntemi Özet Tablo

Metodoloji	Avantajlar	Dezavantajlar	Ne zaman kullanılmalı?
- Eğilim skorları tahmin edilir. - Bir kaliper değeri belirlenir (ör. $0.2 \times$ eğilim skoru standart sapması). - Kaliperin dışındaki eşleşmeler hariç tutulur. - En yakın komşu seçilir.	- Yanlılığı önemli ölçüde azaltır. - Uzak eşleşmelerin etkisini ortadan kaldırır.	- Katı kaliper koşulu örneklem kaybına yol açabilir. - Çok dar kaliper, analizi temsil gücünden yoksun bırakabilir.	- Ortak destek alanı zayıf olduğunda. - Uzak eşleşmelerin dâhil edilmesinin ciddi yanlılık yaratabileceği durumlarda. - Kontrol grubunun büyük olduğu durumlarda.

Yarıçap (Radius) Eşleştirme Yöntemi

Yarıçap eşleştirme, kaliperli eşleştirmeye benzer şekilde bir eşik değer üzerinden belirlenen yarıçap içinde kalan tüm kontrol grubuna ait birimleri eşleştirmeye dâhil ederek gerçekleştirilmektedir. Böylelikle, en yakın tek bir kontrol birimi yerine birden fazla kontrol gözlemi kullanılabilir. Temel istatistiksel mantık, birden fazla kontrol birimini kullanarak varyansı azaltmaktır (Dehejia ve Wahba, 2002).

Tablo 12. Yarıçap (Radius) Eşleştirme Yöntemi Özet Tablo

Metodoloji	Avantajlar	Dezavantajlar	Ne zaman kullanılmalı?
- Eğilim skorları tahmin edilir. - Bir yarıçap değeri belirlenir. - Belirlenen yarıçap içinde yer alan tüm kontrol grubu birimleri eşleştirilir. - Her kontrol biriminin katkısı (weight) eşit veya ters ağırlıklandırma yaklaşımıyla modele taşınır.	- Tüm uygun kontrol grubu gözlemlerini kullanarak varyansı azaltır. - Kovaryant dengesini iyileştirir. - Örneklem kaybını azaltır.	- Yarıçap geniş tutulursa, uzak eşleşmeler yanlışlık ekleyebilir. - Müdahale grubundan daha fazla kontrol grubu gözleminin kullanılması bağımlı eşleşme durumunda standart hataları artırabilir. - Aşırı eşleşme durumunda ağırlıklandırma gereklidir.	- Ortak destek alanı iyi örtüşen veri setlerinde. - Varyans azaltılmak istendiğinde.

Kernel Eşleştirme Yöntemi

Kernel eşleştirme yöntemi, her müdahale birimi için tüm kontrol birimlerine eğilim skoru uzaklığına göre ağırlık vererek yapılan bir yöntemdir. Temel istatistiksel mantık, yalnızca en yakın bir veya birkaç komşuyu kullanmak yerine, uygun ağırlıklarla tüm kontrol birimlerinden bilgi çekerek varyansı azaltmaktır (Heckman, Ichimura ve Todd, 1997).

Tablo 13. Kernel Eşleştirme Yöntemi Özet Tablo

Metodoloji	Avantajlar	Dezavantajlar	Ne zaman kullanılmalı?
- Eğilim skorları tahmin edilir. - Bir çekirdek fonksiyonu (Gaussian, Epanechnikov vb.) seçilir.⁴¹ - Bir bandwidth (h) değeri belirlenir (örneğin Silverman kuralı veya çapraz doğrulama ile). - Yakın skor değerlerine sahip kontrol grubu birimleri yüksek, uzak olanlar düşük ağırlık değerleri alır.	- Tüm kontrol gözlemleri kullanılır. - Düşük varyans sağlanır. - Örneklem kaybı olmaz.	- Ağırlık fonksiyonu ve bant genişliği seçimi sonuçları etkilemektedir. - Uzak gözlemler düşük ağırlıklandırmalar a sahip olsa dâhi eşleştirmeye dahil edilmektedir.	- Örneklemin büyük olduğu ve nearest-neighbor eşleştirmenin varyansı gereksiz yere artırdığı durumlarda. - Kovaryant dengesinin iyileştirmesi gerektiğinde. - Eğilim skoru dağılımının geniş olduğu ve nearest-neighbor eşleştirmenin “kötü yakın komşular” getirdiği durumlarda.

⁴¹ Çekirdek fonksiyon; belirli bir noktadaki yoğunluğu veya fonksiyon değerini komşuluk bilgisinden türetmeye yarayan ağırlıklandırma fonksiyonudur.

Optimal Eşleştirme Yöntemi

Optimal eşleştirme, tüm eşleşmeleri bir arada değerlendirerek toplam eşleşme mesafesini minimize eden algoritmalara dayanmaktadır. Mantığı, tek tek eşleştirmeler yerine tüm örneklem için en iyi toplam eşleşme kombinasyonunu bulmaktır. Her müdahale birimi için “lokal olarak en yakın” kontrol birimi seçmek yerine, tüm eşleşmelerin toplam mesafesini minimize eden global optimum bulunmaktadır (Rosenbaum, 1989).

Tablo 14. Optimal Eşleştirme Yöntemi Özet Tablo

Metodoloji	Avantajlar	Dezavantajlar	Ne zaman kullanılmalı?
• Mesafe matrisi oluşturulur. Tüm olası eşleşmeler değerlendirilir. • Toplam mesafeyi minimize eden kombinasyon seçilir.	• Genel eşleşme kalitesini artırır. • Yanlılığı azaltır.	• Büyük veri setlerinde hesaplama maliyeti yüksektir.	• Küçük/orta büyüklükteki veri setlerinde. • Eşleşme kalitesi kritik olduğunda.

Tam Eşleştirme Yöntemi

Tam eşleştirme, tüm gözlemleri kapsayan, ortaya çıkan her alt grupta en az bir müdahale ve kontrol grubu birimi bulunan ve müdahale ve kontrol birimlerini, seçilen tüm kovaryantlar açısından tamamen aynı değerlere sahip olacak şekilde eşleştiren bir yöntemidir. Hedef, örneklem kaybını en aza indirerek dengeyi artırmaktır (Hansen, 2004). İlgili yöntemde müdahale grubunda yer alan her gözlem, kovaryant vektörü tamamen özdeş olan bir (veya daha fazla) kontrol grubu gözlemiyle eşleştirilmektedir. Eşleşemeyen gözlemler ise veri seti dışında bırakılmaktadır. Böylece eşleştirmede kullanılan kovaryantlar bakımından müdahale ve kontrol grupları arasında tam denge sağlanmaktadır.

Tablo 15. Tam Eşleştirme Yöntemi Özet Tablo

Metodoloji	Avantajlar	Dezavantajlar	Ne zaman kullanılmalı?
• Optimal eşleştirme algoritması kullanılarak, tüm gözlemleri içeren subclass'lar oluşturulur. • Her gruba dengeli şekilde müdahale ve kontrol grubu birimi atanır.	• Örneklem kaybı minimumdur. • Tam denge sağlar.	• Yüksek miktarda veri gerektirmektedir.	• Az sayıda kovaryantın olduğu durumda, Yeterli sayıda gözlem olduğunda

Kabalaştırılmış Tam Eşleştirme Yöntemi (CEM – Coarsened Exact Matching)

Bu yöntem, kovaryatların önceden belirlenen aralıklara bölünerek kabalaştırılması ve sonrasında oluşturulan yeni kategoriler üzerinden tam eşleştirme yapılmasına dayanmaktadır. İstatistiksel olarak model bağımlılığını azaltır ve eşleştirme sürecini şeffaflaştırır (Iacus, King ve Porro, 2012).

Tablo 16. Sistematik Basitleştirilmiş Tam Eşleştirme Yöntemi Özet Tablo

Metodoloji	Avantajlar	Dezavantajlar	Ne zaman kullanılmalı?
- Kovaryatlar kritik kategorilere ayrılır. - Aynı kategorideki müdahale ve kontrol birimleri eşleştirilir.	- Yüksek kovaryat dengesi sağlar. - Model bağımlılığını azaltır.	Kategorilerin aşırı dar seçimi örneklem kaybı yaratabilir, çok geniş seçilmesi ise dengeyi zayıflatabilir.	- Şeffaf ve kontrol edilebilir eşleştirme gerektiğinde. - Büyük veri setleri ile çalışılırken hızlı bir eşleştirme gerekiyorsa. - Kovaryatların kategorik olarak doğal sınıfları varsa (ör. sektör, il, destek türü, büyüklük sınıfı)

3.4.5. Kovaryat Denge Kontrolü

Denge kontrolü, nedensel etki değerlendirmede müdahale ve kontrol gruplarının, müdahale öncesi döneme ait kovaryat bakımından benzer olup olmadığını sistematik olarak değerlendirmeyi amaçlamaktadır. Amaç, iki grup arasındaki farklılıkların yalnızca müdahalenin etkisinden kaynaklanmasını sağlamak ve karıştırıcı değişkenlerden (confounder) gelen yanlılığı azaltmaktır. Eğer denge sağlanmazsa, sonuçlar güvenilir olmamaktadır (Austin, 2009).

Denge kontrolünde genellikle Standartlaştırılmış ortalama farklar kullanılmaktadır. SMD kontrol ve müdahale grupları için eşleştirme gerçekleştirildikten sonra kovaryatların ortalamalarının karşılaştırılmasının yapılması ile denge kontrolü sağlanmaktadır. SMD, müdahale ve kontrol grupları için bir kovaryata ait ortalamalar arasındaki farkın ilgili kovaryatın varyansının kareköküne başka bir deyişle kovaryatın standart sapmasına bölünerek hesaplanmaktadır. (Stuart, 2010)

$$SMD_x = \frac{\overline{X_T} - \overline{X_C}}{\sqrt{\frac{S_T^2 + S_C^2}{2}}}$$

$\overline{X_T}$: Müdahale grubunun X kovaryatı için ortalama değeri

$\overline{X_C}$: Kontrol grubunun X kovaryatı için ortalama değeri

S_T^2 : Müdahale grubunun X kovaryatı için varyansı

S_C^2 : Kontrol grubunun X kovaryatı için varyansı

SMD değeri;

- 0,1 ve altında bir değer alırsa ilgili kovaryat için dengede kabul edilir.
- 0,1 ve 0,25 arasında ise kabul edilebilir bir denge mevcut olduğu varsayılır.
- 0,25 üstünde ise eşleştirme ya da ağırlıklandırma yeniden gözden geçirilmelidir.

SMD Değeri Yetersiz Olduğunda İzlenecek Adımlar

SMD değerinin 0,25 ve üzerinde olması durumunda, kovaryatlar arasındaki dengenin sağlanamadığı kabul edilir ve aşağıdaki iyileştirme adımlarının değerlendirilmesi gerekmektedir:

Adım 1: Eğilim Skoru Modelinin Yeniden Belirlenmesi

Dengenin sağlanamadığı durumlarda ilk olarak eğilim skoru modelinin spesifikasyonu gözden geçirilmelidir. Bu kapsamda:

- Müdahale kararını etkileyebilecek ancak modele dahil edilmemiş kovaryatların eklenmesi, (Caliendo ve Kopeinig, 2008)
- Lineer olmayan ilişkileri yakalayabilmek amacıyla değişken dönüşümlerinin (logaritma, kare terimleri vb.) dahil edilmesi ile dağılımlardaki dengesizlik azaltılabilir.

Adım 2: Eşleştirme Yönteminin Revize Edilmesi

Eşleştirme algoritmasının değiştirilmesi veya daha kısıtlayıcı eşleştirme kurallarının uygulanması dengeyi iyileştirebilir. Bu doğrultuda:

- Caliper (eşik) mesafesi kullanılarak aşırı uzak eşleşmelerin sınırlandırılması,
- 1:1 yerine 1:2 veya 1:3 gibi çoklu eşleştirme yöntemlerinin tercih edilmesi,
- Yenilemeli eşleştirme (replacement) uygulanarak kontrol birimlerinin birden fazla müdahale birimiyle eşleşmesine imkân tanınması gibi yaklaşımlar dengeyi artırmaya yardımcı olabilir (Staffa ve Zurakowski, 2018).

Adım 3: Ortak Destek Koşulunun Değerlendirilmesi

Dengenin sağlanamamasının bir nedeni, eğilim skoru dağılımlarının müdahale ve kontrol gruplarında yeterince örtüşmemesi olabilir. Bu durumda:

- Eğilim skoru dağılımının uç noktalarında yer alan gözlemlerin örneklemden çıkarılması (trimming) dengeyi önemli ölçüde iyileştirebilir (Caliendo ve diğerleri, 2018).

Adım 4: Hibrit ve Gelişmiş Yöntemlerin Kullanılması

Yukarıdaki yöntemlerle denge sağlanamaması durumunda daha gelişmiş veya hibrit yöntemlere başvurulabilir. Bu kapsamda:

- Eşleştirme ve ağırlıklandırmanın birlikte uygulandığı hibrit modeller (örneğin PSM + IPTW) dengeyi güçlendirmeye yönelik etkili alternatiflerdir (Lee ve Little, 2017).

3.5. Araç Değişken Yöntemi

Araç Değişken yöntemi (AD), yansız çıkarım için gerekli tüm varsayımları yerine getiremeyen deneysel veriye, doğal deneylerden⁴² elde edilen veriye ve çeşitli koşullar altında gözlemsel veriye uygulanabilen istatistiksel bir yöntemdir. Temel varsayımı, seçilen araç değişken ile yardım programından yararlanma arasında yüksek bir korelasyon⁴³ olması; ancak yardıma ilişkin sonuç değişkeni üzerinde doğrudan bir etki bulunmamasıdır.

Gözlem birimleri müdahale grubuna bir içsel değişken⁴⁴ referans alınarak atanırsa, uygulanan müdahalenin doğrudan çıktı değişkeniyle veya çıktıyla ilişkili değişkenlerle arasında korelasyon olduğu anlamına gelecektir. Bu durumda müdahalenin etkilerini modelde var olan diğer değişkenlerden ayırmak mümkün olmayacak ve hesaplanan etki yanıltıcı olacaktır. AD yöntemi, bu sorunu araç değişken adında üçüncü bir değişkenden yararlanarak çözmeye çalışmaktadır. Böylelikle bu üçüncü değişken yani dışsal değişken⁴⁵ referans alınarak müdahale grubuna atama yapılacaktır ve içsel değişkenlerden etkilenmeyen bu değişken üzerinden etki tespit edilecektir. Bu durumda yanlılık etkisi ortadan kaldırılmış olacağından asıl müdahale etkisi tespit edilmiş olacaktır.

AD yönteminin yanıltıcı olmayan etki tahmini yapabilmesi için AD değişkeninin (i) programa katılıma etki eden katılımcı özellikleri ile yüksek korelasyona sahip olması ve (ii) program çıktıları ile tek ilişkisinin program katılımını ölçen değişken üzerinden olması gerekir. Araç değişken bu iki özelliği sağlayabildiği takdirde geçerliliğini kanıtlamış sayılır. İlk şartın sağlanıp sağlanmadığını sınamak için araç değişken ile programa katılım değişkeni arasındaki korelasyona bakılır ve bulunan korelasyon değeri yeterince yüksekse bu şart sağlanmış olur. İkinci şartın sağlanıp sağlanmadığını doğrudan söyleyebilecek bir istatistik yoktur. Değerlendirmeyi yapan araştırmacının belli bir mantık çerçevesinde tartışarak araç değişkenin program çıktısı ile yalnızca programa katılım değişkeni üzerinden ilişkili olduğunu kanıtlaması gerekir (Polat ve Aktakke, 2017).

Eşitlik (1)'de Y müdahale sonucunu gösteren bağımlı değişken, i ise programa katılan ve katılmayan gözlemleri gösterir. T, eşitlikte kukla (gösterge) değişken olup programa katılanlar için 1 değerini, katılmayanlar için 0 değerini alacaktır. X değişkeni gözlemlenen özellikleri temsil eden bağımsız değişken setidir. ϵ ise hem gözlenmemiş değişkenlerin yükünü taşıyan hem de Y sonuç değişkenini etkileyen hata terimidir.

$$Y_i = \alpha X_i + \beta T_i + \epsilon_i \quad (1)$$

⁴² Doğal deney, sosyal araştırmalarda kullanılan bir deney türüdür. Doğal deneyde araştırmacı, doğal bir değişimden ya da olgudan faydalanarak onu bağımsız değişken olarak kullanır. Etkisi incelenen durum yardıma ilişkin bir program, yeni bir kanun, seçim, afet vb. olabilir.

⁴³ Aynı öge grubundaki iki veya daha fazla özelliğin veya ölçümün birlikte değişme eğilimi gösterme derecesidir. Örneğin, ihracat ve satış karlılığı değişkenleri hem temsil ettiği parametreler açısından ortak bir yapıda hem de araştırmada mesela "A" adlı tek bir firmadan alınan ölçümler oldukları için ilişkilendirilebilir.

⁴⁴ İstatistiksel bir modelde içsel (endojen) değişken, model içindeki diğer değişkenlerle olan ilişkisi tarafından değeri değişen veya belirlenen bir değişkendir. Başka bir deyişle, bağımlı bir değişkenle eş anlamlıdır, araştırılan durumda diğer faktörlerle ilişkili olduğu anlamına gelir. Bu nedenle, değerleri diğer değişkenler tarafından belirlenebilir. Örneğin, ihracat ile ilgili uygulanacak Devlet yardımı içsel değişken olarak modellenebilir.

⁴⁵ Endojen (içsel) değişkenlerin aksine, eksojen (dışsal) değişkenler bağımsız kabul edilir. Bu sebeple model içindeki bir değişken, diğer değişkenlerle doğrudan ilişkili değildir. Örneğin, yardım yapılacak firma sahiplerinin eğitim durumu, firmanın kuruluş yılı, ülkedeki doğalgaz fiyatları, enflasyon, döviz kuru vs. dışsal faktörler olarak kabul edilecektir.

(1) no'lu eşitlik T programının, sonuç değişkeni (Y) üzerindeki doğrudan etkisini ölçmek için etki değerlendirmede yaygın olarak kullanılan bir yaklaşımı yansıtır. T'nin rastgele olduğu yani müdahale grubuna atamalarda rastgelelik olduğu durumlarda seçim yanlılığı sorun teşkil etmeyecektir ve deneysel yöntemlerle çözüme ulaşılabilecektir.

Ancak iki durumda deneysel düzen bozulabilir:

- Programa katılımda içsellik mevcut olabilir; yani gözlemlenebilen veya gözlemlenemeyen belirli özelliklere (kazanç fırsatları veya sosyal normlar gibi) sahip olan ve sonuç değişkeni ile arasında ilişki bulunan kişilerin seçilerek müdahale grubunun oluşturulması,
- Yararlanıcıların programa kendi seçimlerinden kaynaklı gözlemlenmemiş heterojenlik ortaya çıkarması.

Bu durumda hata terimi ε ile T kukla değişkeni arasında korelasyon oluşacaktır. Yani, $\text{cov}(T, \varepsilon) \neq 0$ olduğu için En Küçük Kareler Yöntemi-EKK (OLS - Ordinary Least Squares)⁴⁶ için gerekli olan yansız tahmin elde etme varsayımı ihlal edilmiş olacaktır. Bu korelasyon, programa katılımın Y üzerindeki etkisini gösteren β katsayısının da sapmasına yol açacaktır.

Bu durumun çözümü için hata terimi ε ile korelasyonu sıfır olan Z adlı bir araç değişken modele ilave edilirse etkin bir tahmin yapılabilir.

Geçerli bir araç değişken bulunduktan sonra, AD'nin iki aşamalı regresyon analizi uygulanır. Birinci aşamada (2) no'lu eşitlikte gösterildiği gibi T'nin, Z üzerine regresyonu alınır.

$$T_i = \gamma Z_i + \phi X_i + v_i \quad (2)$$

Buradan bulunan katsayılar kullanılarak T'nin tahmin edilen değeri hesaplanır. Tahmin edilen T değerlerinde içsellik sorunu olmayacaktır. Sadece Z'deki bir birimlik değişim, T'de γ katsayısı kadar değişime yol açacaktır, böylelikle diğer karıştırıcı faktörlerden olan değişim ise arındırılmış olacaktır.

Regresyonun ikinci aşamasında Y'nin gerçek T yerine tahmin edilmiş T değeri üzerine regresyonu alınır. Böylelikle (3) no'lu eşitliğe, (2) no'lu eşitlikteki katsayıların tahmin değerleri konularak oluşturulacaktır.

$$Y_i = \alpha X_i + \beta (\gamma Z_i + \phi X_i + v_i) + \varepsilon_i \quad (3)$$

İkinci aşamanın sonunda hesaplanan yardımın etkisi seçim yanlılığından arındırılmış olacaktır.

Bu yöntemde kullanılan dışsal araç değişkenler ne kadar iyiye AD tahmin edicileri de bir o kadar iyi sayılmaktadır. Bu tür değişkenleri bulmak çoğu zaman kolay olmayabilir. Ancak tüm istatistiksel-ekonometrik analizlerde beklenen rassal deneylerin kullanılması olduğundan, müdahalenin olduğu sistemlerde göreceli olarak EKK yerine bu tür yöntemlerle etki değerlendirme yapılabilir.

⁴⁶ En Küçük Kareler Yöntemi bağımlı ve bağımsız değişken arasındaki ilişkiyi matematiksel denklem olarak ortaya koymaya yarayan standart bir regresyon yöntemidir.

AD yöntemi ile elde edilen tahmin değerlerinin EKK ile elde edilecek tahminlerden farklı olduğu ya da AD güven aralığının EKK güven aralığından farklı olduğu test edilirse tahminler iyileştirilmiş olur. Diğer türlü EKK varsayımları üzerinden basit bir model kurgulayarak da ortaya çıkarılmak istenen etki değerlendirilebilir.

Örneğin, yardımın firmaların istihdamı üzerindeki etkisini, yardımdan yararlanma değişkeni ve diğer gözlemlenebilir değişkenler üzerinden regresyon ile belirlenebilecektir. Yardımın amacının, zorlayıcı iç pazar koşullarına maruz kalma olasılığı yüksek olan firmaları hedeflemek olduğu varsayıldığında piyasa koşulları değerlendirici tarafından gözlemlenebilir olmadığından doğrudan kontrol edilemeyecektir. Ancak, bu değişken açıklanmadan bırakıldığında, yardımdan yararlanmanın içselliklerinden dolayı yardımın etkisi değerlendirici tarafından düşük görülebilir. Firmanın zorlayıcı piyasa koşullarıyla karşı karşıya olup olmaması, hem yardımdan yararlanma hem de istihdam üzerinde bir etkiye sahip olacaktır, yani yardımdan yararlanma, sonuç değişkeni olan istihdama ait hata terimi ile ilişkilidir. Piyasa koşullarının yardımdan yararlananlar üzerindeki etkisinin olması, yardımdan yararlanma ile istihdam arasındaki tüm ilişkiyi nedensellik etkisine atfetmenin imkânsız olduğu anlamına gelecektir.

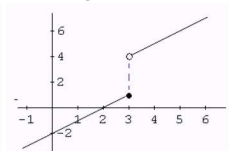
Ayrıca, yardımdan yararlanmayı açıklayıp istihdamı açıklayamayan başka etkenler de olacaktır. Örneğin, coğrafi konum bölgedeki program için mevcut toplam para miktarını belirleyebilir (Criscuolo ve Diğerleri, 2012). Belirli bir bölge için yardım bütçesi dış nedenlerle zaman içinde değiştiyse (örneğin kişi başına düşen milli gelirde azalış), bu durumun yardımdan yararlanma üzerinde etkisi olacaktır; ancak firmanın iç pazar koşulları üzerinde etkisi olmayacaktır. Bu sebeple yardımdan yararlanma değişkeninin bu kısma odaklanarak, iç pazar koşullarının müdahalesi olmaksızın, yardımın firmanın istihdamı üzerindeki gerçek etkisini AD yöntemi ile ayırtırmak mümkün olabilecektir.

Yöntemin avantajı doğru araç değişkeninin tespiti halinde birimlerin programa katılmasını sağlayan gözlemlenemeyen etkilerin kontrol edilebilir olmasıdır.

3.6. Regresyon Süreksizliği Tasarımı Yöntemi

Regresyon Süreksizliği Tasarımında (RST), bir politikadan etkilenme olasılığı üzerinde süreksiz⁴⁷ bir etkiye sahip olan bir değişkenin varlığından yararlanılır. Bu yöntem ilk olarak Thistlethwaite ve Campbell (1960) tarafından liyakat ödülleri öğrencilerin gelecekteki akademik sonuçlarına (kariyer hedefleri, lisansüstü programlara kayıt vb.) etkisi çalışmalarında tanıtılmıştır. Müdahale ve kontrol grubu birimlerinin eşik değer etrafında toplanan gözlemleri için yerel ortalama müdahale etkisi (Local Average Treatment Effect – LATE) tahmin edilmiştir.

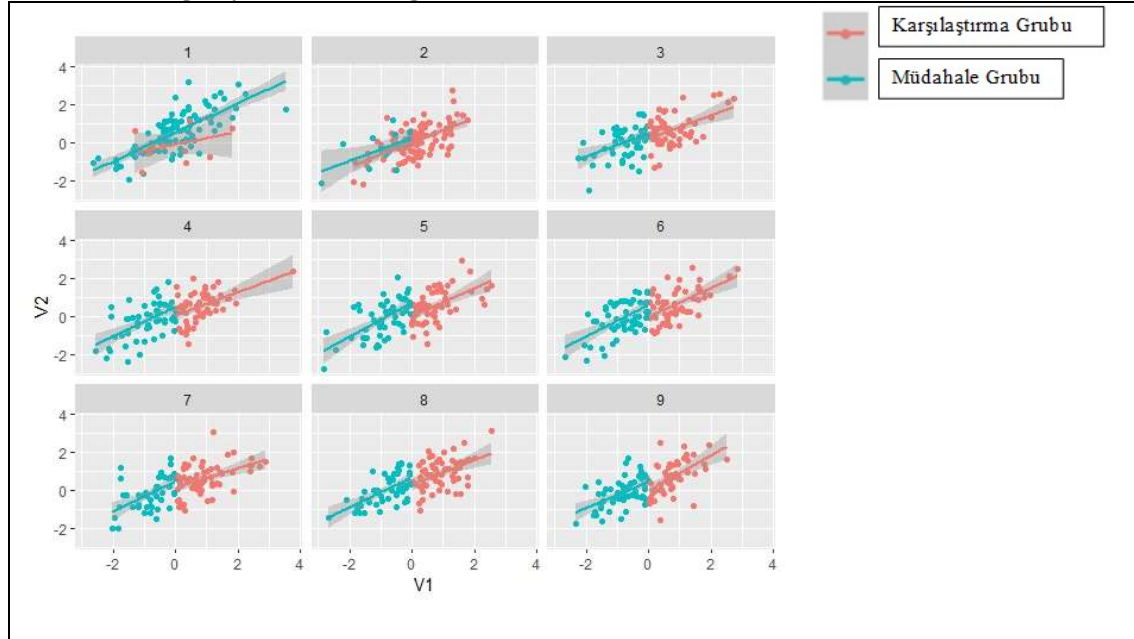
⁴⁷ Grafiğin birbiriyle bağlantılı olmayan fonksiyonu süreksiz fonksiyon olarak bilinir. Aşağıdaki grafikte, fonksiyonun soldaki ve sağdaki limitleri eşit değildir ve dolayısıyla $x = 3$ 'teki limit artık yoktur. Böyle bir fonksiyona, fonksiyonun süreksizliği denir.



Yardımdan yararlanmayı belirleyen belli kurallar olduğu varsayıldığında, katılımcılar önceden belirlenmiş bir eşik değeri kullanılarak müdahale ve kontrol grubu olarak ikiye ayrılabilir. Bu yöntem müdahale ve kontrol gruplarına atanan katılımcıların mukayese edilebilmesini sağlamaktadır. Örneğin, yardım için oluşturulan bir programda müdahale ve kontrol grubunu ayırmak için belli bir ciro tutarı eşik değeri seçilebilir. Alternatif olarak, eşik değeri farklı gerekçelerle de seçilebilir. Örneğin, önerilecek programın kapasitesi 30 adet firma kontenjanını içeriyorsa ve 100 firma başvurmuşsa, yardıma en çok ihtiyacı olduğu düşünülen ilk 30 firmayı ayıran bir eşik değeri seçilebilir. Yöntemin dezavantajı seçilen eşik değerden ötürü yanlışlığa sebep olabileceğinden programın etkisinin olduğundan daha yüksek çıkabilmesidir.

Buradan hareketle müdahale ve kontrol grubuna atama rastgele değil, yaş, gelir ve puan gibi gözlenen bir değişkenin belirlenen eşik değerine dayalı olacaktır. Eğer katılımcıya ait değişkenin değeri eşik değerden büyük ise o kişi, müdahale grubuna; küçük ise kontrol grubuna atanacaktır. Ardından müdahale söz konusu olduğunda katılımcıların tekrar atanması sırasında ortaya çıkan farklılık ise “süreksizlik” olarak adlandırılmaktadır. Ortaya çıkan bu süreksizlik ise etki tahminlerinde kullanılmaktadır. Grafik 18’de V1 başlangıç skoru, V2 bitiş skoru iken müdahale öncesi ve sonrası oluşan süreksizliğe ilişkin görsel grafikler sunulmuştur

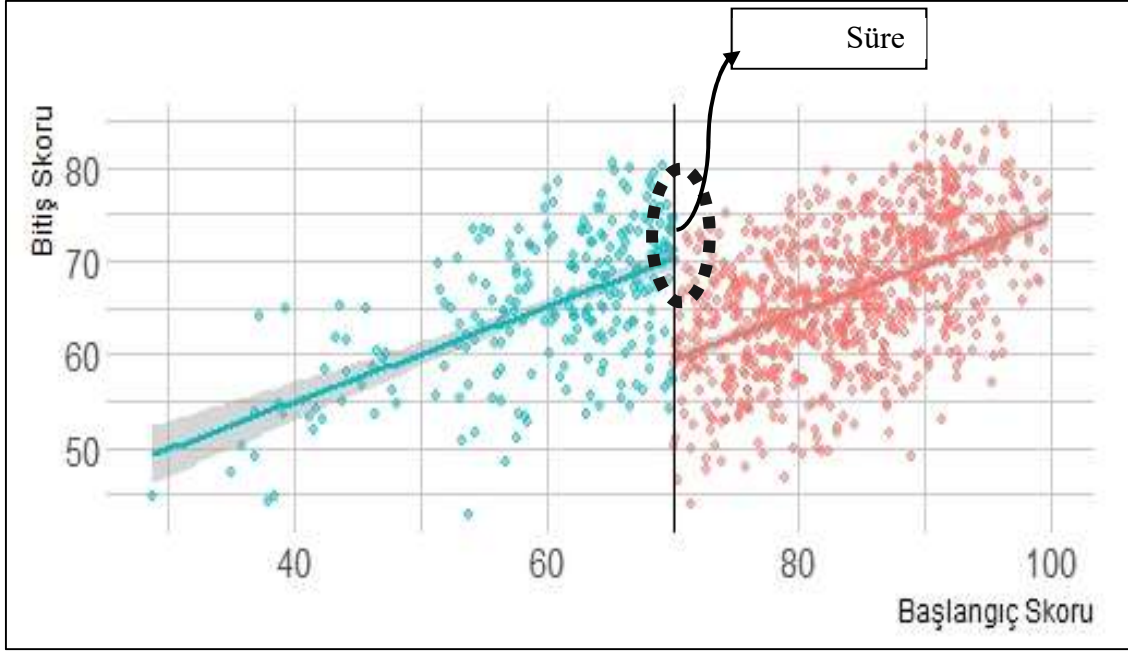
Grafik 18. Regresyon Süreksizliği



Kaynak: Başkanlık tarafından örnek amaçlı oluşturulmuştur.

Ayrıca Grafik 19’da yöntemin adını nasıl aldığı görülebilmektedir. Kesme noktasındaki regresyon çizgilerinde bir “sıçrama” veya “süreksizlik” var ve bu bize yardımın etkisinin olduğunu söylemektedir. Ancak istatistiki olarak bunun önemli bir etki mi olup olmadığı yapılacak hipotez testleri ile ortaya çıkacaktır.

Grafik 19. Sıçrama Noktası



Kaynak: Başkanlık tarafından örnek amaçlı oluşturulmuştur.

Eşik değerin üstünde kalan katılımcılar (müdahale grubu) ile eşik değerin altında kalan katılımcılar (kontrol grubu) sonuç değişkeni bakımından birbirine benzer olmasına rağmen, seçilen eşik değerden ötürü bir kısmı yardım programına dâhil olmuş diğer kısmı olmamıştır. Bu sebeple seçilen eşik değeri bu yöntemde bir araç değişken görevi görmektedir.

Eşik değere uzak kalan katılımcılar örnekleme dâhil edilmediğinden küçük hacimli bir örneklem üzerinden yöntem uygulanır. Eşik değerin olduğu noktada sonuç değişkeninde görülen ani yükselme (ya da alçalma), yardım etkisinin yansız tahminini vermektedir.

Yöntem yardımın bir eşik değerde⁴⁸ kesilmesine veya genişletilmesinin gerekip gerekmediğine dair yararlı kanıtlar sağlayacaktır. Her katılımcı için bağımlı değişkenin tahmini değeri ve ardından bu tahminlerin ortalama değeri hesaplanacaktır.

Eşik değere yakın katılımcılar, gözlemlenen ve gözlemlenmeyen değişkenler açısından benzer olmalıdır. RST çerçevesinde bu, gözlemlenen ve gözlemlenmeyen değişkenlerin dağılımının eşik civarında *sürekli* olması gerektiği anlamına gelir. Gözlemlenen ortak değişkenler arasındaki benzerlik kontrol edilebilirken, gözlemlenmeyen özellikler arasındaki benzerlik varsayımsal olarak kabul edilmelidir. Bu, özellikle eşik değere çok yakın bireyler için, yani nispeten dar bir örneklem için makul bir varsayım olarak kabul edilir.

RST belirsiz (fuzzy) ve keskin (sharp) olmak üzere iki farklı şekilde uygulanabilir. Keskin RST kullanıldığında, potansiyel yararlanıcıların belirlenmesinde kural sıkı bir şekilde uygulanır. Belirsiz RST’de ise bazı birimler belirlenen ölçüte göre yardım programına katılmaya hak kazanmış olsalar dahi programdan yararlanmayabilir ya da programa katılmaya hak kazanmış olmayanlar programdan yararlanabilir (Polat ve Aktakke, 2017).

⁴⁸ Burada bahsedilen değerin altındayken ekonomik bir faaliyetin devam edemediği sınırdır.

Yazında RST'nin iç geçerlilik açısından güçlü olduğu, bu yönüyle rastgele deney tasarımı yapılamayan yerlerde kullanımının yaygınlaştığı belirtilmektedir. Ancak aynı düzeyde istatistiki doğruluk elde etmek için, RST'de rastgele bir deneye göre 2,75 kat daha fazla katılımcı gerekmektedir. Örneğin, bir rastgele deneyin belirli bir güç düzeyine ulaşması için 100 katılımcıya ihtiyacı varsa, RST'de 275 kadar katılımcıya ihtiyacı olabilir. Ayrıca katılımcıların %20'si rastgele seçildiğinde 2,48 kat daha fazla, %40'ı için 1,96 kat %60'ı için 1,46 ve %80 rastgele seçimin olduğu yerde de 1,14 kat daha fazla katılımcı gereklidir (CappelleriveTrochim, 2015). Sadece eşik değere çok yakın gözlemler dikkate alındığından analiz için yeterli sayıda örneklem hacminin olması gerekmektedir.

Regresyon süreksizliği haricindeki coğrafi süreksizlik tasarımı (geographical discontinuity design) da bir başka süreksizlik tasarımı yöntemidir. Bu yöntem adından anlaşılacağı üzere uygun olan alan/bölgelerdeki uygulamalardır. Bu tasarım yararlanıcıların bulundukları konuma bağlı olarak il, bölge ya da ülke bazında uygulanan müdahalelerde kullanılabilir. Bu sayede karakteristik özellikleri benzer gruplar seçileceğinden seçim yanlılığı azaltılmış olacaktır. Yöntemin dezavantajı bölgesel karakteristik özelliklere göre ön planda olduğu için çıkan sonuçların genelleştirilmemesidir. Diğer bir deyişle, yöntemin dış geçerliliği zayıftır.

3.7. Ters Olasılıklı Ağırlıklandırma Yöntemi

Etki değerlendirme çalışmalarında rastgele deneylerin en önemli faydalarından biri müdahale ve kontrol gruplarının rastgele dağıtımını sağlamasıdır. Böylelikle müdahale ve kontrol grupları gözlemlenen ve gözlemlenemeyen değişkenler bakımından sadece rassal bir şekilde birbirlerinden farklılaşmaktadırlar. Bu çerçevede, gözlemsel (rassal olmayan) veriler kullanılarak gerçekleştirilen etki değerlendirme çalışmalarında, ortak değişkenlerin benzer dağılımlara sahip olduğu müdahale ve kontrol grupları elde edilerek rassal bir deneye mümkün olduğunca yakın olacak şekilde analizin yapılması hedeflenmektedir. Bu hedefe genellikle müdahale ve kontrol gruplarından iyi eşleşen örneklerin seçilmesi, başka bir deyişle eşleştirmenin gerçekleştirilmesi ve böylelikle ortak değişkenlerden kaynaklanan yanlılığın ve etkinsizliğin azaltılmasıyla ulaşılabileceği değerlendirilmektedir.⁴⁹ Bu çerçevede, eşleştirme, gözlemsel bir veri seti içerisinde gizli olan deneysel verileri belirlemeye yönelik bir yöntem olarak da tanımlanabilmektedir. Bunun yanında, parametrik varsayımlarda bulunmaya gerek kalmaksızın uygulanan müdahalenin etkilerinin tahmin edilebilmesi için eşleştirme yönteminden yararlanılmaktadır.⁵⁰

⁴⁹ Ho, D. E., Imai, K., King, G. and Stuart, E. A. (2007). Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. *Political Analysis* 15 199–236.

⁵⁰ Hill, Jennifer. 2008. "Discussion of research using propensity-score matching: Comments on A critical appraisal of propensity-score matching in the medical literature between 1996 and 2003 by Peter Austin, *Statistics in Medicine*." *Statistics in medicine* 27(12):2055–2061.

Her ne kadar eşleştirme alanındaki çalışmalar zaman içerisinde artmış olsa da araştırmacıların kullanabileceği bir reçete bulunmamaktadır. Aksine, eşleştirme ile ilgili yayınlar ve kaynaklar istatistik⁵¹, iktisat⁵², sosyoloji⁵³ ve siyaset bilimi gibi farklı bilim alanlarına yayılmıştır.

Eşleştirme yöntemleri içerisinde Eğilim Skoru Eşleştirme araştırmacılar tarafından sıklıkla kullanılan yöntemlerden biridir. Eğilim Skoru Eşleştirme ile birlikte her ne kadar veri seti daha dengeli hale gelse de bazı gözlemler veri setinin dışında kalmaktadır. Bu çerçevede ortaya bir çelişki çıktığı değerlendirilmektedir.⁵⁴ Zira, eğer veri setinde dengesizlik varsa bu durumda güçlü modelleme varsayımları olmaksızın geçerli bir nedensel çıkarım yapmak mümkün olmayacaktır. Öte yandan, Eğilim Skoru Eşleştirme dengesizlikleri gidermesine rağmen veri seti herhangi bir yöntemle nedensel etki tahmin edilmesi için kullanışlı olmayabilecek ve bu durumda da söz konusu çelişki ile karşı karşıya kalınmış olacaktır.

Mahalanobis ya da kaba kestirim gibi eşleştirme yöntemleriyle Rassal Kontrollü Deneylere (RCT) benzer bir yaklaşımla araştırmanın gerçekleştirilmesi durumunda müdahale ve kontrol grupları deneyin başlangıcında tam olarak gözlemlenen ortak değişken(ler) üzerinde eşleştirilir. Böylece ortak değişken(ler)e ilişkin tüm muhtemel yanlılıklar örneklemde tam olarak ortadan kaldırılır ve gözlemlenmeyen değişkenlerden kaynaklanan muhtemel yanlılık, örneklemde kontrol edilmez; ancak ortadan kaldırılmış olur. Bu durumda da eşleştirmeye bağlı olarak gözlemlerin veri setinden atılması, bilgiyi göz ardı edecek ve dengesizliği, model bağımlılığını, araştırmacının takdir yetkisini ve yanlılık riskini artıracaktır.

Özellikle örneklemde yeterli büyüklükte gözlem olmaması durumunda eşleştirme yönteminin kullanılması ile ortaya çıkacak gözlem kaybı⁵⁵ yaşamamak ve model tanımlanmasındaki muhtemel hataların önüne geçilmesi⁵⁶, gruplar arasındaki dengesizliğin artmaması ve yanlılık riskinin yükselmemesi için alternatif olarak “Ters Olasılıklı Ağırlıklandırma (Inverse Probability Weighting, IPW)” yönteminden yararlanılabileceği değerlendirilmektedir.

Ters Olasılıklı Ağırlıklandırma müdahale etkisinin ölçülmeye çalışıldığı ve müdahale grubunun rassal bir şekilde belirlenmediği gözlemsel veriye dayalı etki değerlendirme çalışmalarında kullanılan istatistiksel bir yöntemdir. Her gözleme, müdahaleye fiilen maruz kalmasının, öngörülen müdahaleye maruz kalma olasılığıyla ne kadar iyi eşleştiğine bağlı olarak bir ağırlık atanmakta ve söz konusu ağırlıklar müdahalenin nedensel etkisinin ölçülmesine ilişkin gerçekleştirilen tahminde kullanılmaktadır.

⁵¹ Rosenbaum, P. R. (2002). *Observational Studies*, 2nd ed. Springer, New York.

⁵² Imbens, G. W. (2004). Nonparametric estimation of average treatment effects under exogeneity: A review. *Review of Economics and Statistics* 86 4–29.

⁵³ Morgan, S. L. and Harding, D. J. (2006). Matching estimators of causal effects: Prospects and pitfalls in theory and practice. *Sociological Methods ve Research* 35 3–60.

⁵⁴ King, G. and Nielsen, R. (2019), Why Propensity Scores Should Not Be Used for Matching, *Political Analysis*, 27, 4, Pp. 435-454.

⁵⁵ Stuart, E. A. (2010). Matching methods for causal inference: A review and a look forward. *Statistical Science*, 25(1), 1–21.

⁵⁶ Austin, P. C. (2011). An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies. *Multivariate Behavioral Research*, 46(3), 399–424.

3.8. Rassal Kontrollü Deneyler (RCT)

Rassal Kontrollü Deneyler yöntemi, rastgeleleştirme olarak da bilinen deneysel yöntemdir. Potansiyel yararlanıcılardan, yardım alacak olanların rastgele bir şekilde seçilmesi ile oluşturulan müdahale grubu ile yine bu potansiyel yararlanıcılardan müdahale grubundakilerle istatistiksel olarak benzer olanlardan oluşturulan kontrol grubunun mukayese edilerek değerlendirilmesidir. Gözlem birimlerinden kontrol grubu ve müdahale grubu seçimi üç aşamada belirlenir.

İlk aşamada yardımdan yararlanabilecek uygun bir gözlem grubu seçilir.

İkinci aşamada bu gruptan değerlendirmeyi yapmak için uygun büyüklükte bir örneklem seçilir. İlk aşamada belirlenen ve yardımın uygulanabileceği gözlem grubunun çok büyük olmadığı ve yardım bütçesinin bu grup içindeki her bir gözlem biriminden detaylı bilgi toplamaya yeterli olduğu durumlarda örneklem seçme aşaması atlanıp söz konusu gözlem grubu ile devam edilebilir.

Değerlendirmenin yapılacağı örneklemin seçilmesi önemlidir. Bu seçim son aşamada çalışmanın dış geçerliliğini (*external validity*)⁵⁷ belirleyecektir. Örneklemdaki gözlem sayısı çok küçük seçilirse ya da kitleyi temsil etme özelliği zayıf olursa, değerlendirmede ulaşılan sonuçların değerlendirmenin yapıldığı örneklem dışında geçerliliği olmayabilir. Öte yandan, örneklemden yer alan gözlem sayısı çok büyükse yardım bütçesi verinin toplanması için yetmeyebilir. Dolayısıyla hem dış geçerliliği sağlayacak hem de bütçeyi aşmayacak ideal örneklem hacminin bulunması önemlidir.

Üçüncü aşamada ise örneklemdaki tüm birimler kontrol ya da müdahale grubuna atanır. Bu aşamada örneklemdaki her bir birimin yardımdan yararlanma şansı eşittir. Rastgeleliği sağlamak için genelde bir çekiliş yapılır. Son aşamadaki yardımdan yararlanma rastgeleliğinin ölçüsü ise yardımın iç geçerliliğini (*internal validity*)⁵⁸ belirlemektedir (Glewwe ve Todd, 2022).

Örneklemdaki gözlem sayısı yeterli olduğu sürece rastgele dağıtım, kontrol ve müdahale gruplarının birbirine benzer olmalarını sağlayacaktır. Bu yönteminin aşamaları genel bir çerçevede Şekil 7’de belirtilmiştir.

⁵⁷ Etki değerlendirmede dış geçerlilik, gözlemlenen nedensel etkinin tüm uygun bireylere genellenebileceği anlamına gelir. Bu nedenle, bir değerlendirmenin harici olarak geçerli olabilmesi için, değerlendirme örnekleminin tüm uygun bireyleri temsil eden bir örneklem olması gerekir.

⁵⁸ İç geçerlilik için, bir etki değerlendirmenin geçerli bir karşı-olgusal tahmini bulabileceği bir kontrol grubuna sahip olması gerekir. İç geçerliliği sağlanmış bir etki değerlendirme, müdahale ve kontrol grubu arasındaki olası tüm farklılıkları kontrol ederek, müdahalenin çıktılarındaki değişiklikleri açıkça ilişkilendirebilecektir. Bu, deneysel veya yarı deneysel teknikleri uygun şekilde uygulayarak elde edilebilir.

Şekil 7. Rassal Kontrollü Deneyler Yöntemi



Kaynak: Glewwe ve Todd (2022).

Bu yöntemin karşı-olguşallığın iyi bir tahminini sunması ve seçim probleminden dolayı yaşanabilecek yanlılığı gidermesi açısından “altın standart” olduğu değerlendirilmektedir. Rassal Kontrollü Deneyler yöntemi teorik olarak karşı-olguşallığı iyi tahmin edebilse dahi uygulama kısmında bunun doğru olup olmadığının test edilmesi gerekmektedir. Bunun için gözlem birimlerinin belirli özellikler ve sonuç değişkeni için aldığı değerlere ilişkin veri; müdahale ve kontrol grupları için yardım uygulanmadan önce toplanır. Buna *ön test verisi* denir. Ampirik açıdan altın standardın sağlanıp sağlanmadığını test etmek için bu ön test verisi kullanılarak *Dengeli Testler* (*Balancedness Tests*) uygulanır. Yöntem uygun müdahale ve kontrol gruplarını oluşturmakta başarılı olduysa, bu iki grubun gözlemlenebilen özellikleri arasında istatistiki olarak anlamlı bir fark olmaması gerekmektedir. Testler ile bu varsayımın sağlanmış olduğu görülmüşse, kontrol ve müdahale gruplarının çıktıları arasında yardımın uygulanması sonrası ölçülen fark, yardımın gerçek ve yansız etkisini gösterecektir (Cavallo, 2016).

Yöntemin güçlü yanı etkiyi mukayese edebilecek iyi bir kontrol grubu ile uygulanmasıdır. Zayıf yanı her yardıma uygulanamaz ve elde edilen sonuçlar, yardım geniş çaplı farklı hedef kitlelere uygulanmak istendiğinde dış geçerlilik sorununa sebep olabilir.

4. R PROGRAMI İLE NİCEL ETKİ DEĞERLENDİRME VE UYGULAMA ÖRNEKLERİ

Veri analizinde kullanılan çok sayıda yazılım aracı bulunmaktadır. Bazıları SPSS, SAS ve Excel gibi lisanslı program kategorisinde bazıları ise R Programlama dili, Python gibi açık kaynak kodlu kategorisindedir. Açık kaynak kodlu yazılımlar lisanslı programlara alternatif olarak daha esnek ve maliyet etkin çözüm sunmaktadır. Rehber'in bu bölümünde R programı kısaca tanıtılmış, uygulama örneklerinin veri analizi R'da yapılmış ve kaynak kodları ile birlikte açıklamalara yer verilmiştir.

4.1. R Programlama Dili (R Studio)⁵⁹

Geniş kütüphanesi ve geliştirilmesini sağlayan topluluk desteği ile hem basit veri analizleri hem de karmaşık ekonometrik modellemeler için kullanışlı bir araçtır. R'da paket adı altında çeşitli fonksiyonları barındıran nesneler yer almaktadır. Paketler vasıtasıyla çeşitli analizleri pratik şekilde gerçekleştirmek mümkündür. Örneğin, FF yöntemi için geliştirilmiş çeşitli paketler mevcuttur. Bu paketlerin detayları ve nasıl kullanılması gerektiği de resmi web sayfasında yer alan kılavuzlarda açıklanmaktadır.

R programını kullanmak için öncelikle bilgisayarınıza indirilmesi ve kurulumunun yapılması gerekmektedir. Bunun için R resmi internet sitesi⁶⁰ ziyaret edilerek sırasıyla aşağıdaki adımlar takip edilmelidir:

- **CRAN Mirrors:** Sayfanın sol üst köşesinde yer alan “CRAN” bağlantısına tıklanır. CRAN (**C**omprehensive **R** Archive **N**etwork), R yazılımı ve ilgili belgelerin bulunduğu ana platformdur.
- **İşletim Sistemi Seçimi:** İşletim sistemine göre sürüm seçilir (Windows, macOS, Linux, Ubuntu vb.).
- **İndirme Sayfası:** Seçilen uygun sürüm “Download R” bağlantısı tıklanarak bulunur.
- **İndirme İşlemi:** Seçilen sürümün indirme işlemi başlatılır.⁶¹
- **Kurulum:** İndirme işleminin tamamlanmasından sonra, kurulum sihirbazı ile R kurulumu gerçekleştirilir.

Kurulum tamamlandıktan sonra R'da bir proje oluşturulup çalışma başlatılabilir. Oluşturulan proje; kod, veri dosyaları, analiz sonuçları ve dokümantasyon gibi ilişkili tüm dosyaları bir arada tutmaktadır. Bu sayede analizi yapılan her yardım için oluşturulacak proje dosyaları, analizin daha düzenli ve yönetilebilir olmasını sağlamaktadır. Bu sayede bir projede yapılan değişikliklerin diğer projeleri etkilemesi önlenmektedir.

⁵⁹ Detaylı bilgi edinmek için: CRAN - The Comprehensive R Archive Network (R paketleri, dokümantasyon ve kullanım kılavuzları gibi kaynaklar) ve "R for Data Science" - Hadley Wickham ve Garrett Grolemund (R ile veri bilimi uygulamalarını kapsayan temel konular)

⁶⁰ <https://www.r-project.org/>

⁶¹ Bazı sürümlerde önceki sürümde çalışan paketlerin çalışmama durumu söz konusu olabilir. Bu durumda eski sürüm paketin varsa yeni sürümü ya da alternatif paketler kullanılarak analiz yapılabilmesi mümkündür.

Menü Aracılığıyla proje dosyası oluşturmak için:

- **R Studio** açılır.
- Üst menü çubuğunda **“File”** (Dosya) sekmesine tıklanır.
- Ardından **“New Project”** (Yeni Proje) seçeneği bulunur ve tıklanır.
- Açılan menüden dosyanın ismi ve konumu belirlenir.
- Ardından **“Create Project”** (Proje Oluştur) düğmesi tıklanarak proje dosyası oluşturulur.

Kısayol tuşu ile proje dosyası oluşturmak için:

Klavyeden **“Ctrl + Shift + N”** (Windows/Linux) veya **“Command + Shift + N”** (Mac) tuş kombinasyonu kullanılarak yeni proje oluşturma penceresi açılabilir.

Proje dosyanın açılması ile birlikte yapılması gereken bir diğer adım kodların yazılacağı sayfanın (R Script)⁶² açılmasıdır.

Menü aracılığıyla R script sayfasının açılması:

1. Üst menü çubuğundan **“File”** (Dosya) sekmesi seçilir.
2. Ardından **“New File”** (Yeni Dosya) seçeneğini bulunur ve tıklanır.
3. Açılan alt menüden **“R Script”** seçeneğini seçilir.

Kısayol tuşu ile klavyeden **“Ctrl + Shift + N”** (Windows/Linux) veya **“Command + Shift + N”** (Mac) kısayolu kullanılarak yeni bir R Script sayfası açılabilmesi de mümkündür.

Açılan script sayfasında kod yazılabilir ya da yazılan kodları düzenlenip yeniden çalıştırılabilir. Yapılan çalışmaların kaybolmaması için ara ara kaydedilmesi gerekmektedir. Bunun için **“File”** (Dosya) sekmesinden **“Save”** (Kaydet) veya **“Save As”** (Farklı Kaydet) seçenekleri kullanılmaktadır.

Yeni script **“.R”** uzantısı ile proje dosyasına kaydedilir. Bir R script içerisindeki kodlar bütün olarak ya da satır satır çalıştırılabilir. Bu özelliği sayesinde özellikle tekrar eden veya karmaşık görünen analizlerde zaman ve emek tasarrufu sağlamaktadır.

Ayrıca R’da paketler mevcuttur. Bu paketler, belirli fonksiyonları, veri setlerini, işlevleri ya da analiz yöntemlerini içeren modüller olarak tanımlanmaktadır. Örneğin, veri manipülasyonu için ‘dplyr’, veri görselleştirme için ‘ggplot2’, zaman serisi analizi için ‘forecast’ gibi paketler kullanılmaktadır. Bu örnek verilen paketlere alternatif benzer paketler de mevcuttur. Paketler sayesinde, karmaşık görünen işlemler basitleştirilmiş şekilde yapılabilir.

4.2. Etki Değerlendirme İçin Gerekli Adımlar

Bu başlık altında, R programlama dilini kullanarak bağımlı değişken üzerinde devlet yardımının etkisini değerlendirirken adım adım izlenebilecek yollar sıralanmaya çalışılmıştır. Bu adımlar, genel bir çerçeve sunabilir; ancak, her bir değerlendirmede farklı ihtiyaçlar olabileceğini ve bu yüzden sürecin yardımın ve değerlendirmenin özelliklerine göre uyarlanması gerektiği unutulmamalıdır.

⁶² R script, R programlama dilinde yazılan komutların bulunduğu bir dosyadır.

Adım 1: Veri Setinin Oluşturulması

Etki değerlendirme yapabilmek için veri seti iki gruba ait şekilde oluşturulmalıdır. İki gruptan biri müdahale grubu olarak adlandırılan ve yardımdan yararlanmış grup iken, diğer grup yardımın etkisinin araştırılmasını sağlayacak olan kontrol grubu yani yardımdan yararlanmamış grup olacaktır. Bu sebeple ilk aşamada bu iki grubun tanımlanması gerekmektedir.

Kutu 5. Kontrol ve Müdahale Grubunun Belirlenmesi

Önemli Notlar:

Bir müdahaleden geçerli bir etki ölçümü elde etmek için, diğer faktörlerin karıştırıcı olmadığı bir kıyaslama tekniğine ihtiyaç vardır. Temel bir yaklaşım, kontrol (kontrol) grubu kullanmaktır. Bu grup, müdahale alan grubun özelliklerine yakından benzeyen; ancak müdahale almayan bir grup birey/hane/firmalardan oluşmaktadır. Kontrol grubu, müdahalenin yapıldığı aynı geniş coğrafi alandan çekilen bireyler, haneler, firmalar veya başka bir birim örneğidir. Kontrol grubunun, başlangıçta, müdahale grubu ile ortalama olarak aynı özelliklere sahip olması gerekir; buna denge denilmektedir. Bir kontrol grubunu tanımlamak için, araştırmacının seçilen yardım alanlarını ve yardıma başvuru mekanizmasını anlaması gerekmektedir.

Bir kontrol grubunu tanımlamak için, hem deneysel hem de deneysel olmayan tasarımlar mevcuttur:

Deneysel tasarım, müdahale alan ve almayan grupların, müdahaleyle ilişkili olmayan dışsal bir faktör tarafından rastgele atandığı durumlar için kullanılan terimdir.

Rassal Kontrollü Deneyler (RCT'ler): Bazı kaynaklarda rastgele deneyler veya rastgele değerlendirmeler de denilmektedir. Bu yaklaşım, müdahalenin kime veya hangi alana verileceğinin rastgele atanmasını içermektedir. Bu, müdahale grubundaki ve kontrol grubundaki bireylerin, program için uygun olan hedef kitleden rastgele seçilmesi anlamına gelmektedir. Müdahale ve müdahale dışı gruptakilerin rastgele örneklemelerden seçilmesi ile aynı değildir. Belirlenmiş bir hedef kitle içerisinde rastgele seçim ile kontrol ve müdahale grubu oluşturulması ifade edilmektedir.

Deneysel olmayan tasarımlar, maliyetler veya rastgele atamayı engelleyen diğer düşünceler nedeniyle deneysel bir tasarımın mümkün olmadığı durumlarda kullanılmaktadır. Müdahale gerçekleştirildikten sonra bir etki değerlendirme yapma kararı alındığında genellikle deneysel olmayan deneyler düşünülmektedir.

Doğal deneyler: Bir doğal koşul, coğrafi konumlar veya bir hükümet politikası, doğası gereği homojen olan bir nüfusu, yalnızca bir müdahaleye maruz kalma ile farklılaşan iki veya daha fazla gruba ayırıyorsa, bu bir doğal deney olarak kabul edilebilir. Örneğin, yeni inşa edilmiş bir otoyol bir köyü ikiye böldüğünde bir sağlık kliniği bir yarıda yer alır ve bu, köylülerin o tarafta sağlık tesisine diğer tarafa göre daha kolay erişim sağlamasını sağlamaktadır. Bu durum doğrudan bir tarafın müdahale, diğer tarafın kontrol grubu olmasına sebep olmaktadır. Doğal deneyler, etki değerlendirme için benzersiz şekilde uygun olsa da, pratikte yarı deneysel tasarımlar veya regresyona dayalı yaklaşımlar aracılığıyla ele alınmaktadır.

Yarı deneysel tasarımlar: Bu tasarımlarda, müdahale grubu hariç, müdahale grubu ile aynı özelliklere sahip bir kontrol grubu oluşturmak için istatistiksel yöntemler kullanılmaktadır. Temel yarı deneysel yaklaşımlar Farkların Farkı, Eğilim Skoru Eşleştirme veya regresyon süresizliğidir. Sayılanlar yazında en yaygın olanlar yaklaşımlardır. Örneğin, bir eşik uygunluk kuralı olduğunda regresyon süresizliği tasarımı kullanılabilir.

Regresyona dayalı yaklaşımlar: Müdahalenin genellikle bir kukla değişken aracılığıyla tespit edildiği regresyon modelleri de tahmin edilmektedir. Çok yaygın olmamakla birlikte müdahaleyi temsil eden değişken, kredi büyüklüğü veya eğitim süresi gibi sürekli bir değişkendir. Bu yaklaşımlar arasında içsel (endojen) müdahale modelleri, araç değişkenler, robust regresyon bulunmaktadır. Kaynak: WhiteveRaitzer, 2017

Ardından araştırma sorusu yardımın uygulanması ile ortaya çıkacak etkinin tespiti ile ilişkili olduğundan, bu kapsamda analizde kullanılacak bağımlı ve bağımsız değişkenlerin belirlenmesi gerekmektedir. İlgili değişkenlerin doğru şekilde belirlenmesi, yardım uygulanmamış olsaydı ne olurdu sorusunun cevabını alabilmek için önemlidir. Örneğin, “Tekstil sektöründe yer alan KOBİ’lere sağlanan ihracat yardımlarının, bu KOBİ’lerin toplam ihracat hacmi üzerinde etkisi nedir” şeklinde bir araştırma sorusu için etki değerlendirme yapılacağı varsayıldığında bağımlı değişken doğrudan araştırma sorusunun içerisinde yer almaktadır. Anlaşılacağı üzere “Toplam İhracat Hacmi” bağımlı değişken olacaktır ve bu kapsamda veri toplanması gerekmektedir. Aynı zamanda bağımsız değişken de yine araştırma sorusunun içerisinde yer almakta olup “İhracat Yardımından Yararlanma Durumu” bağımsız değişken olarak çalışmada olacaktır. Bir bağımlı bir bağımsız değişken ile kurulan model basit modeldir. Hem müdahale grubu hem de kontrol grubu birbirine benzerliği konusunda belli başlı karakteristik özelliklere (demografik özellikler, firma bilgileri, sektör bilgisi, çalışan profili, pazarın çeşitliliği, faaliyet süresi, firmanın ölçeği vb.) göre seçildiyse, bu durumda tek bir bağımsız değişken üzerinden de basit bir model kurularak analiz yapılabilmektedir. Fakat yine bulguların yorumlanması konusunda hem katsayıların hem modelin anlamlılığının irdelenmesi gerekmektedir.

Buradan hareketle araştırmalarda tek bir bağımsız değişken ile çalışılması mümkün olmayabilir. Hatta çok gözlem olan çalışmalarda yeterli sayıda bağımsız değişken olması bağımlı değişkenin açıklanma durumunu artıracaktır. Örneğin, verilen örnekte ikinci bir bağımsız değişken olarak “Yardımın Tutarı” çalışmaya eklenebilir. Bu durumda basit modelden çoklu modele geçiş yapılmış olacaktır. Analizde hem değişkenlerin katsayıları hem de model birlikte yorumlandığından, sonuçların anlamlılığına göre bağımsız değişken sayısı azaltılabilir ya da artırılabilir. Bağımsız değişkenler ile ilgili ihracat konusunda yapılan ampirik (deneysel) çalışmalar ve aynı zamanda teori temelli konuların araştırılması da araştırmacıya analiz öncesinde fikir vermesi açısından gereklidir. Unutulmaması gereken önemli husus, araştırmanın konusunun ve kurgusunun araştırmacı tarafından tasarlandığıdır. Bu sebeple iyi bir yazın araştırması yapmak hem yardımın etkisini hem de yardım uygulanmamış olsaydı ne olurdu sorusunu araştırabilmek açısından faydası olacaktır.

Ardından bir yardıma ilişkin müdahale ve kontrol grupları için örneklem seçilmektedir. Seçilen örneklem ilgilenilen kitlenin tamamı da olabilir. Bu seçim araştırmacının zaman, maliyet, işgücü ve araştırmanın gerekliliklerine göre karar vermesi gereken bir durumdur. Bir diğer önemli konu ise müdahale grubunun referans grup olması ve seçilecek kontrol grubunun gözlemlenen karakteristik özellikleri bakımından birbirine benzer olmasıdır. Çalışmanın genelleştirilebilmesi ve seçim yanlılıklarının azaltılması açısından benzerlik önemlidir. Aynı zaman da iki grubun kıyaslanması konusunda daha doğru sonuçlar ortaya çıkaracaktır. Benzerliğin yeterli olmadığının görüldüğü çalışmalarda eşleştirme yöntemleri ile gereken homojenlik sağlanabilmektedir. Örneklemenin seçilmesi ile çalışmanın en başında belirlenen bağımlı ve bağımsız değişkenler için veri toplanması gerekmektedir.

Adım 2: Verinin İncelenmesi ve Derlenmesi

İlk adım her zaman veriyi anlamak ve doğru bir şekilde derlemektir. Basit tanımlayıcı istatistikler kullanılarak keşifsel veri analizi yapılması mümkündür. Bu sayede, veri setinin yapısı, içerdiği değişkenlerin türleri ve potansiyel problemleri (eksik veriler, aykırı değerler, tutarsız hatalı girişler vb.) tanımlanabilmektedir. Bu yapıların asıl amacı veri kalitesini artırmak ve analiz sonuçlarının doğruluğunu ve güvenilirliğini artırmaktır. Bu nedenle, veriyi anlama ve derleme sürecine yeterli zaman ve özenin gösterilmesi gerekmektedir. Aşağıda bir veri seti için neler yapılabileceği birkaç adımda özetlenmiştir.

Veri Setinin İncelenmesi: Veri seti R ortamına `'read.csv()'`, `'readRDS()'`, `'load()'`, vs. gibi fonksiyonlar aracılığıyla yüklendikten sonra veri setinin yapısı `'str()'`, `'summary()'` veya ilk birkaç satırı `'head()'` fonksiyonu kullanılarak incelenebilir ve veri hakkında genel bir bilgi edinilebilir.

Veri Temizliğinin Yapılması: Eksik değerler (NA⁶³) `'na.omit()'`, `'is.na()'`, gibi fonksiyonlarla temizlenebilmektedir. Aynı zamanda aykırı değerler `'boxplot()'` fonksiyonu ve tutarsız veri girişleri de `'dplyr'` paketinin `'filter()'` gibi fonksiyonları kullanılarak tespit edilebilmektedir. Büyük miktarda eksik verisi olan değişkenler, analizin doğruluğunu etkileyebilir. Eksik veri oranı yüksek olan değişkenlerin analize dahil edilip edilmeyeceği ya da eksik verilerin nasıl ele alınacağına bakılmalıdır.

Değişken Seçimi ve Dönüşümünün Yapılması: Bir veri seti içerisinde yer alan tüm değişkenlerin analizde kullanılması mümkün olmayabilir. Örneğin çok değişkenin yer aldığı bir veri setinde gözlem sayısının az olması veride tahmin edilen değerlerin olması gerekenden daha büyük çıkmasına sebep olabilecektir. Bu sebeple öncelikle araştırma sorusu ve hipotez üzerinden ilgili değişkenlerin tespiti, yeterli gözlem sayısının olup olmadığına bakılmalıdır. Ayrıca bazı değişkenlerin dönüştürülmesi gerekebilmektedir. Örneğin, yardımdan yararlanma diye bir değişken oluşturulacaksa bu değişkenin 0-1 değerlerini mantıksal olarak alan kategorik değişken olması beklenir. Bu durumda bu değişkenin R'da faktör olarak kodlanması gerekmektedir. Ayrıca verinin dağılımı `'hist()'`, `'boxplot()'` fonksiyonları ile incelenebilir ya da değişkenler arasındaki ilişkiler için `'plot()'`, `'cor()'` fonksiyonları ya da `'ggplot2'` paketi kullanılabilir. Verinin dağılımının incelenmesinin çeşitli sebepleri vardır. Bunlardan ilki veri homojen mi heterojen mi dağılıyor bunu görmek içindir. Etki değerlendirme analizlerinde karşı olgusalılık için kontrol grubu oluşturulması söz konusu olduğundan bu iki grubun birbirine göre homojen olması beklenmektedir. Bu sebeple verilerin yayılımı⁶⁴ incelenmelidir. Ayrıca ekonomik veriler ile çalışılması gerektiğinden verilerin yayılımı genellikle geniş olmaktadır. Bu tür verilerin yorumlanmasının daha kolay hale gelmesi için de büyük ekonomik verilere uygulanan logaritmik dönüşümün⁶⁵ yapılması gerekmektedir.

⁶³ R programlama dilinde, "NA" ifadesi "Not Available" (Mevcut Değil) ifadesinin kısaltmasıdır. Eksik veya mevcut olmayan verileri temsil etmek için kullanılır.

⁶⁴ Veri setinin standart sapma, varyans, aralık ve çeyrekler arası aralık gibi ölçülerine bakılmaktadır. Bu, verilerin ne kadar homojen veya heterojen olduğunu göstermektedir.

⁶⁵ Veri setinin dağılımı, bazı analizler için gerekli olan veri dönüşümlerinin belirlenmesine yardımcı olur. Örneğin, logaritmik dönüşüm, veri dağılımını daha normal hale getirebilmektedir.

Adım 3: Modelin Kurulması

Araştırma sorusuna göre uygun istatistiksel testler seçilmektedir. Devlet yardımının etkisini incelemek için Doğrusal Regresyon Modelleri, Eşleştirme gibi rehberde bahsedilen yöntemler kullanılabilmektedir. Ardından R'da `'lm()'`, `'glm()'`, `'plm()'`, `'MatchIt'` gibi fonksiyon ve paketler ile seçilen model kurulmaktadır. Burada en önemli konu seçilen yöntemin varsayımlarının sağlanıp sağlanmadığının kontrol edilmesidir.

Adım 4: Bulguların Elde Edilmesi ve Yorumlanması

Modelden elde edilen katsayılar, p-değerleri ve güven aralıkları incelenmelidir. Asıl amaç etkinin büyüklüğünün hesaplanmasıdır. Bilindiği üzere devlet yardımının bağımlı değişken üzerindeki etkisinin büyüklüğü ve önemi çalışmanın yorumlanması gereken nihai kısmı olacaktır. Ayrıca bulguların da görselleştirilmesi mümkündür. Örneğin, Farkların Farkı yönteminde elde edilen bulguların kontrol ve müdahale grupları için nasıl bir değişim yarattığı `'ggplot2'` ya da `'plotly'` gibi paketlerle görselleştirilebilmektedir.

Adım 5: Sonuçların Raporlanması

Analiz sürecinde⁶⁶ çeşitli değişkenler üzerinden çeşitli modellerin kurulması söz konusu olacaktır. Bunun sebebi gözlemsel kısıtlar olabileceği gibi varsayımlarla alakalı doğrulamalar da olabilmektedir. Ayrıca sonuçların anlamsız çıkması da bir sebeptir. Bu sebeple farklı model ve yöntem denemeleri de gerçekleştirilebilmektedir. Ancak, raporlama safhasında araştırma sorusuyla en ilişkili ve sonuçları anlamlı olan modelin yorumlanması yerinde olacaktır.

Aşağıda burada anlatılanlar ve ilave fonsiyonlar örnekler aracılığıyla açıklanmaktadır.

4.3. Uygulama Örnekleri⁶⁷

Örnek 1. Farkların Farkı Yöntemi

“Tarımda Üretkenliği Artırma” Devlet Yardımı Programı

Kırsal kalkınmayı ve tarım sektörünü desteklemek amacıyla "Tarımda Üretkenliği Artırma" adında bir yardım programı başlatılmıştır. Program, çiftçilere, modern tarım tekniklerinin uygulanması, kaliteli tohum ve gübre edinimi için mali destek sağlamaktadır. Programın temel amacı, çiftçilerin gelirini artırarak tarımsal üretkenliği ve sürdürülebilirliği teşvik etmektir. Yardım programından yararlanan çiftçiler, belirlenen başvuru kriterlerine göre seçilmiş ve programın sağladığı kaynaklardan faydalanmaya başlamışlardır. Yardım programının amacı gözetilerek araştırma sorusu,

⁶⁶ Analiz sürecinde karşılaşılan problemler için R dökümantasyonu ve çeşitli R topluluk forumları (örneğin, Stack Overflow, RStudio Community) kullanılabilir.

⁶⁷ Uygulama örnekleri, kurgusal geliştirilmiş ve sentetik veriler kullanılarak sonuçlar türetilmiştir. Sentetik veri setlerine Başkanlığın <https://www.sbb.gov.tr/devlet-yardimlariyla-ilgili-belge-ve-yayinlar/> adresinden erişilebilmektedir.

"Tarımda Üretkenliği Artırma" yardım programının, programdan yararlanan çiftçilerin yıllık gelirleri üzerindeki etkisi nedir? şeklindedir. Bu araştırma sorusuna yönelik oluşturulacak yokluk hipotezi:

H₀: Programın çiftçilerin yıllık gelirleri üzerine etkisi yoktur.

Bu kapsamda etki değerlendirme yapılarak programdan yararlanan ve yararlanmayan çiftçiler arasında karşılaştırmalı bir analiz ile programın hedeflerine ulaşılıp ulaşılmadığı ve tarımsal hasılatı nasıl etkilediği araştırılacaktır.

R'da gerçekleştirilen analize ilişkin adımlar aşağıda sırasıyla sunulmuştur.

Adım 1: Kullanılacak Paketlerin Yüklenmesi ve Kütüphaneden Çağırılması

Bölümün başında R'da belirli bir paketin kullanabilmesi için öncelikle o paketin sisteme yüklenmesi gerektiğinden bahsedilmişti. Bu işlem, `install.packages("paket_adi")` fonksiyonu aracılığıyla gerçekleştirilmektedir. Burada "paket_adi" yerine, yüklenecek paketin adını yazılmalıdır. Örneğin, veri görselleştirme için **ggplot2** yüklenmek istenirse, `install.packages("ggplot2")` komutu yazılır ve çalıştırılır. R'da yazılan kodlar klavyede Ctrl + Enter ya da Alt + Enter tuşlarının birlikte kullanımı ile pratik çalıştırılabilmektedir.

Bu komut, ggplot2 paketini R kütüphanesine kaydetmektedir. Bu yükleme işlemi sadece bir kez yapılmakta olup paket yüklendikten sonra, her R oturumunda çağrılmak üzere hazır hale gelmiş olacaktır. Çağırma işlemi ise `library()` komutu ile gerçekleştirilmektedir.

```
install.packages("tidyverse") #Veri işleme, görselleştirme ve modelleme için kullanılır.
install.packages("readxl") #Excel dosyalarını R Studio'ya aktarmak için kullanılır.      install.packages("broom")
#Modelleri veri çerçevelerine dönüştürmek için kullanılır.
install.packages("scales") #Grafiklerde ölçekleme ve etiketleme işlemlerini kolaylaştırır.
install.packages("modelsummary") #Regresyon tablolarının özelleştirilmesini sağlar.
install.packages("skimr") #Özet istatistikler, dağılım analizi ve eksik değerlerin kontrolü
install.packages("dlookr") #R'da DAG68 grafiği oluşturmak için kullanılır.

library(tidyverse) # Veri manipülasyonu ve görselleştirme.
library(readxl) # Excel dosyalarını okuma aracı.
library(broom) # Modelleri veri çerçevelerine dönüştürmek için kullanılır.
library(scales) # Grafiklerde ölçekleme ve etiketleme işlemlerini kolaylaştırır.
library(modelsummary) # Regresyon tablolarını özelleştirmenizi sağlar.
library(skimr) # Özet istatistikler, dağılım analizi ve eksik değerlerin kontrolü
library(dlookr)
```

Belirli bir işlevi gerçekleştirmek için hangi R modülünün kullanılacağından emin olunmadığında "[CRAN Task Views](#)" sayfası ziyaret edilerek, görev kategorilerinin (konuların) uygun bir listesi ve her birinin kısa bir açıklamasına erişmek mümkündür. Gerekli konuyu seçerek, o kategoriyle ilgili tüm R paketlerinin yanı sıra her birinin kapsamlı kullanım açıklamasına erişilebilmektedir.

⁶⁸ Directed acyclic graph (DAG) : Değişkenlerin, birbirlerini etkileme durumlarını bir bütün olarak görmeyi sağlayan grafiklerdir.

Adım 2: Verinin Yüklenmesi

Aşağıdaki kod satırı bir Excel dosyasını okumak için kullanılmaktadır. `read_excel` fonksiyonu, `readxl` paketi tarafından sağlanmaktadır ve excel dosyalarını (.xlsx veya .xls formatında) doğrudan R ortamına veri çerçevesi (dataframe) olarak yüklemek için kullanılmaktadır.

```
tarim_veri <- read_excel("./tarim_data.xlsx")69
```

Kodun açıklaması:

- `tarim_veri` adında bir değişken ile veri çerçevesi tanımlanmıştır.
- `<-` operatörü kullanılarak, `read_excel` fonksiyonunun çıktısı bu değişkene atanmaktadır. R'da `<-` operatörü, sağdaki ifadenin sol tarafta belirtilen nesneye atanmasını sağlar ve atama operatörü olarak işlev görmektedir.
- `read_excel("./tarim_veri.xlsx")` ifadesi, çalışma dizinindeki `tarim_veri.xlsx` adlı Excel dosyasını okur. Dosya yolu, dosyanın konumunu gösterir. Burada `./` ile başlayan yol, dosyanın mevcut çalışma dizininde olduğunu belirtmektedir.

Adım 3: Verinin İncelenmesi

Script sayfasına `tarim_veri` yazıldığında yükleme işlemi yapılmış olan veri setinin bir bölümü gözden geçirilebilmektedir. `tarim_veri` adlı veri çerçevesini çağırma işlemi yapıldığında görüldüğü üzere ham veri setinin ilk 10 satırı konsol ekranında görüntülenmektedir. Aynı zamanda tabloya ilişkin veri setinin boyutunu (7834 satır ve 12 sütun) ve verinin türünü⁷⁰, setinin bir özetini gösteren özet bilgiler de gelmektedir.

```
tarim_veri
## # A tibble: 7,834 × 12
##   id yardım gelir_2018 gelir_2019 gelir_2020 gelir_2021 gelir_2022 alan_2018
##   <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
## 1 1 1 NA NA 431 54411 NA
## 2 2 0 25831 NA NA NA NA 2
## 3 3 0 491173 NA NA NA NA 8
## 4 4 0 46980 NA 36407 NA NA 3
## 5 5 0 35635 22866 6258 NA NA 2
## 6 6 0 43110 16672 3214 21483 NA 2
## 7 7 0 471719 293056 277229 NA NA 7
## 8 8 0 2876334 5650624 3121731 2938289 2360823 33
## 9 9 0 25521 5289 14935 11356 5291 2
## 10 10 1 112479 28542 29907 19817 261401 2
## # 7,824 more rows
## # 4 more variables: alan_2019 <dbl>, alan_2020 <dbl>, alan_2021 <dbl>,
```

id: Her bir gözlemin, yani veri setindeki her bir çiftçinin tekil olarak tanımlanmasını sağlayan numardır. Bu numara, veri setindeki her bir kaydın diğerlerinden ayrılmasını ve referans alınmasını kolaylaştırmaktadır.

yardım: Bu değişken, çiftçinin devlet tarafından sağlanan yardım programından yararlanıp yararlanmadığını gösterir. Eğer kişi devlet yardımı almışsa bu alanın değeri **1**, almamışsa **0** olarak belirlenir. Bu bağımsız kategorik değişken, devlet yardımının gelir üzerindeki etkisini analiz etmek için kullanılmaktadır.

⁶⁹ Dosya adresinin uygulama yapılan bilgisayardaki konumuna göre güncellenmesi gerekmektedir.

⁷⁰ Tabloda <dbl> olarak görünen veri çerçevesi içindeki sütunlarda reel sayı olduğunu ifade etmektedir. R'da tam sayı (integer), mantıksal (logical), metin (character) vb. türler de mevcuttur.

gelir_2018 - gelir_2022: 2018-2022 yıllarına ait tarım faaliyetlerden elde edilen geliri temsil etmektedir. Veri setinde her yıl için ayrı sütun yer almaktadır.

alan_2018 - alan_2022: 2018-2022 yıllarında ekilen alanının büyüklüğünü hektar cinsinden ifade etmektedir.

Adım 4: Yardım Durumuna Göre Gözlem Sayısı

Devlet yardımı alan (**yardim = 1**) ve almayan (**yardim = 0**) çiftçilere ait gözlem sayılarına göz atılır.

```
yardim_sayisi <- tarim_veri %>%  
  group_by(yardim) %>%  
  summarize(count = n())
```

```
yardim_sayisi  
## # A tibble: 2 × 2  
##   yardim count  
##   <dbl> <int>  
## 1     0 7228  
## 2     1  606
```

Kodun açıklaması:

- **yardim_sayisi <-**: Sonucu saklamak için **yardim_sayisi** adında bir değişken oluşturulmuştur.
- **tarim_veri %>%** : ‘%>%’ , **tarim_veri** veri çerçevesini sonraki işlemlere aktaracak olan operatördür⁷¹ . **dplyr** paketinin bir parçasıdır ve kodu okunabilir ve mantıksal bir sırayla zincirlemek için kullanılmaktadır.
- **group_by(yardim) %>%** : **group_by** fonksiyonu, **yardim** sütunundaki değerlere göre veri çerçevesini gruplara ayırır. Yani, **yardim** sütunundaki her benzersiz değer için bir grup oluşturur. Bu durumda, devlet yardımından yararlananlar ve yararlanmalar şeklinde iki grup üzerinden frekansları hesaplamaktadır.
- **summarize(count = n())**: **summarize** fonksiyonu **yardim** değişkenine göre gözlem sayısını hesaplamaktadır.

Sonuç olarak 7.824 çiftçiden 606’sının yardım aldığı (yani “1”), 7.228’inin ise yardım almadığı (yani “0”) görülmektedir.

Kutu 6. Pipe Operatörü

“%>%”: R programlama dilinde “**pipe operatörü**” olarak bilinen ve **tidyverse** paketlerinde sıkça kullanılan bir operatördür. R Script’inde klavye kısayolu olarak, Alt ve **shift** tuşlarına basılı tutup ardından M tuşuna basmak %>% sembolünü yazmak için bir yöntemdir. Bu operatör, bir işlevin çıktısını başka bir işleve veya ifadeye iletmek için kullanılır. Kodun okunabilirliğini artırır ve işlem zincirlemesini kolaylaştırır. Özellikle **dplyr**, **tidyr**, **ggplot2** gibi **tidyverse** paketlerinde kullanımı yaygındır.

⁷¹ Bkz. Kutu 6.

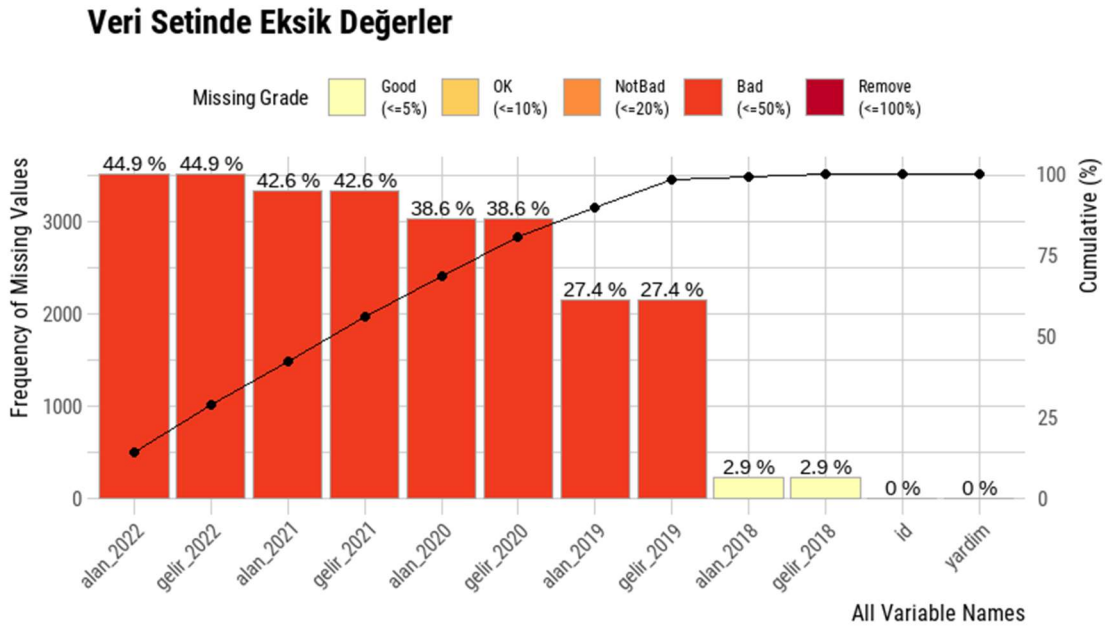
Adım 5: Veri Dönüşümü

➤ Eksik Değerlerin (NA) Kontrolü ve Veri Setinden Çıkarılması

Eksik veriler, analiz sonuçlarını yanıltıcı hale getirebileceğinden ve eşleştirme gibi yöntemler için “NA” veriler analiz sonuçlarını etkileyebileceğinden bu verilerin kontrolü hususunda dikkatli olunması gerekmektedir. Bu kapsamda değerlerin grafiksel gösterim ile kontrolü aşağıda yer verilen kod ile gerçekleştirilmiştir. *dlookr* paketinde yer alan *plot_na_pareto()* fonksiyonu, *tarim_veri* adlı veri çerçevesindeki eksik verileri pareto grafiği olarak görselleştirmektedir.

```
plot_na_pareto(tarim_veri, main = "Veri Setinde Eksik Değerler")
```

Grafik 20. Örnek 1 - Eksik Değerlerin Kontrolü



Eksik verilerin kümülatif yüzdesi, Grafik 20’de siyah çizgi ile gösterilmektedir. Görüldüğü üzere eksik verilerin %80’den fazlası 2020-2022 yılları arasında yer almaktadır. Eksik verilerin, analizin doğruluğunu etkileme potansiyelinden ötürü *na.omit* fonksiyonu ile veri setinden çıkartılıp çalışmaya yeni oluşturulan *tarim_veri_2* veri seti ile devam edilmektedir.

```
tarim_veri_2 <- na.omit(tarim_veri)
tarim_veri_2
## # A tibble: 3,419 × 12
##   id yardım gelir_2018 gelir_2019 gelir_2020 gelir_2021 gelir_2022 alan_2018
##   <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
## 1 8 0 2876334 5650624 3121731 2938289 2360823 33
## 2 9 0 25521 5289 14935 11356 5291 2
## 3 10 1 112479 28542 29907 19817 261401 2
## 4 11 0 41665 248927 152295 281711 125092 3
## 5 13 0 1045571 636919 450274 398085 99178 4
## 6 15 0 60576 95059 49077 41644 156298 3
## 7 16 0 2217403 1323270 462911 734596 822196 20
## 8 19 0 10674 179022 219110 117799 171238 2
## 9 20 0 94408 65936 42969 64882 40344 4
## 10 27 0 419399 389786 153745 65118 194983 3
## # 3,409 more rows
```

```
## # i 4 more variables: alan_2019 <dbl>, alan_2020 <dbl>, alan_2021 <dbl>,  
## # alan_2022 <dbl>
```

na.omit() fonksiyonu, ***tarim_veri*** veri çerçevesindeki ***NA*** değerlerini içeren tüm satırları kaldırır.

Daha temiz bir veri çerçevesi elde edilir ve bu yeni veri çerçevesi ***tarim_veri_2*** olarak kaydedilir.

Bu yeni veri seti, eksik veri içermeyen tam gözlemleri içerir ve daha sonra yapılabilecek analizler için kullanılabilir bir temel sağlar.

```
yardim_sayisi_2 <- tarim_veri_2 %>%  
  group_by(yardim) %>%  
  summarize(sayisi = n())  
  
yardim sayisi 2  
## # A tibble: 2 × 2  
##   yardım sayisi  
##   <dbl> <int>  
## 1     0  3082  
## 2     1   337
```

Sonuç olarak, ham ***tarim_veri*** veri seti başlangıçta **7834** satır içerirken, ***na.omit(tarim_veri)*** fonksiyonu uygulandığında, **3419** satırlı yeni ***tarim_veri_2*** veri çerçevesi oluşturulmuştur. Yardımdan yararlanan sayısı 337, yararlanmayan kişi sayısı 3082 olacak şekilde çalışmaya devam edilmiştir. Aynı zamanda tanımlayıcı istatistikleri de iki grubun kıyaslanması açısından önemlidir. Ona ilişkin bilgiler de hem müdahale grubu hem de kontrol grubu bazında aşağıda sunulmuştur.

```
Yardımdan_Yararlanmayanların_Tanımlayıcı_İst_Almayan <- tarim_veri_2 %>%  
  filter(yardim == 0)  
  
Yardımdan_Yararlanmayanların_Tanımlayıcı_İst  
## # A tibble: 3,082 × 12  
##   id yardım gelir_2018 gelir_2019 gelir_2020 gelir_2021 gelir_2022 alan_2018  
##   <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>  
## 1  8  0  2876334  5650624  3121731  2938289  2360823  33  
## 2  9  0  25521  5289  14935  11356  5291  2  
## 3 11  0  41665  248927  152295  281711  125092  3  
## 4 13  0 1045571  636919  450274  398085  99178  4  
## 5 15  0  60576  95059  49077  41644  156298  3  
## 6 16  0 2217403 1323270  462911  734596  822196  20  
## 7 19  0  10674 179022  219110  117799  171238  2  
## 8 20  0  94408  65936  42969  64882  40344  4  
## 9 27  0 419399 389786  153745  65118  194983  3  
## 10 28  0 248981 335536  26577  14255  67803  2  
## # i 3,072 more rows  
## # i 4 more variables: alan_2019 <dbl>, alan_2020 <dbl>, alan_2021 <dbl>,  
## # alan_2022 <dbl>  
summary(Yardımdan_Yararlanmayanların_Tanımlayıcı_İst)  
##   id      yardım gelir_2018 gelir_2019  
## Min. : 8 Min. :0 Min. : 228 Min. : 291  
## 1st Qu.:1970 1st Qu.:0 1st Qu.: 68794 1st Qu.: 74012  
## Median :3897 Median :0 Median : 204024 Median : 235286  
## Mean :3896 Mean :0 Mean : 976171 Mean : 985163  
## 3rd Qu.:5831 3rd Qu.:0 3rd Qu.: 608975 3rd Qu.: 678452  
## Max. :7834 Max. :0 Max. :255991639 Max. :88452023  
  
## gelir_2020 gelir_2021 gelir_2022 alan_2018  
## Min. : 115 Min. : 158 Min. : 21 Min. : 2.000  
## 1st Qu.: 62186 1st Qu.: 78472 1st Qu.: 68328 1st Qu.: 2.000  
## Median : 208114 Median : 249418 Median : 246374 Median : 4.000
```

```
## Mean : 872449 Mean : 1067910 Mean : 1051424 Mean : 6.282
## 3rd Qu.: 615101 3rd Qu.: 772362 3rd Qu.: 773908 3rd Qu.: 7.000
## Max. :93559277 Max. :129971954 Max. :172971449 Max. :154.000
```

```
## alan 2019 alan 2020 alan 2021 alan 2022
## Min. : 2.000 Min. : 2.000 Min. : 2.000 Min. : 2.000
## 1st Qu.: 2.000 1st Qu.: 2.000 1st Qu.: 2.000 1st Qu.: 2.000
## Median : 4.000 Median : 4.000 Median : 4.000 Median : 4.000
## Mean : 6.488 Mean : 6.333 Mean : 6.948 Mean : 6.972
## 3rd Qu.: 7.000 3rd Qu.: 7.000 3rd Qu.: 8.000 3rd Qu.: 8.000
## Max. :159.000 Max. :106.000 Max. :102.000 Max. :101.000
```

Müdahale grubu tanımlayıcı istatistikleri:

```
Yardımdan_Yararlananların_Tanımlayıcı_İst <- tarim_veri_2 %>%
  filter(yardim == 1)

Yardımdan_Yararlanmayanların_Tanımlayıcı_İst
## # A tibble: 337 × 12
##   id yardım gelir_2018 gelir_2019 gelir_2020 gelir_2021 gelir_2022 alan_2018
##   <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
## 1 10 1 112479 28542 29907 19817 261401 2
## 2 84 1 8441730 1156871 98704 36711 1038604 17
## 3 90 1 433860 400669 460418 266540 286090 24
## 4 111 1 9067300 11706981 4027061 13942188 12083063 3
## 5 128 1 772465 918261 664630 1514170 1403004 10
## 6 143 1 166435 154479 158571 323209 581612 6
## 7 176 1 62056 814574 626511 1024316 1133155 2
## 8 187 1 8108695 10566414 9262265 9575628 5999815 66
## 9 223 1 1114415 1271087 369271 1600207 470898 7
## 10 272 1 18336271 14667557 12626835 17145104 21304260 42
## #> 327 more rows
## #> 4 more variables: alan_2019 <dbl>, alan_2020 <dbl>, alan_2021 <dbl>,
## # alan_2022 <dbl>
summary(Yardımdan_Yararlanmayanların_Tanımlayıcı_İst)
## id yardım gelir_2018 gelir_2019
## Min. : 10 Min. :1 Min. : 706 Min. : 1105
## 1st Qu.:2111 1st Qu.:1 1st Qu.: 353090 1st Qu.: 336038
## Median :4006 Median :1 Median : 1330943 Median : 1271087
## Mean :4108 Mean :1 Mean : 5314661 Mean : 4908583
## 3rd Qu.:6331 3rd Qu.:1 3rd Qu.: 3497260 3rd Qu.: 3684424
## Max. :7813 Max. :1 Max. :237502224 Max. :184826549
## gelir_2020 gelir_2021 gelir_2022 alan_2018
## Min. : 3013 Min. : 11123 Min. : 1590 Min. : 2.00
## 1st Qu.: 346787 1st Qu.: 379694 1st Qu.: 490324 1st Qu.: 6.00
## Median : 1034580 Median : 1520565 Median : 1441324 Median :13.00
## Mean : 4840878 Mean : 7119972 Mean : 7909828 Mean :18.51
## 3rd Qu.: 3713450 3rd Qu.: 5173223 3rd Qu.: 4500126 3rd Qu.:25.00
## Max. :212504420 Max. :509110091 Max. :632497177 Max. :83.00
## alan_2019 alan_2020 alan_2021 alan_2022
## Min. : 2.00 Min. : 2.00 Min. : 2.00 Min. : 2.00
## 1st Qu.: 6.00 1st Qu.: 6.00 1st Qu.: 7.00 1st Qu.: 8.00
## Median :13.00 Median :13.00 Median :15.00 Median :15.00
## Mean :18.72 Mean :17.75 Mean :19.19 Mean :20.14
## 3rd Qu.:26.00 3rd Qu.:25.00 3rd Qu.:25.00 3rd Qu.:26.00
## Max. :90.00 Max. :79.00 Max. :84.00 Max. :84.00
```

➤ Logaritmik Dönüşüm

Logaritmik dönüşüm, genellikle ekonomik verilere uygulanmaktadır. Bu çalışmada veri setindeki büyük değerler arasındaki farkları küçültmek ve böylece aykırı değerlerin etkisini azaltarak, verileri daha dar bir aralığa toplamak amaçlı yapılmaktadır. Aşağıdaki R kod bloğu, *tarim_veri_2* veri çerçevesinde yıl bazında ayrı ayrı yer alan gelir değişkenine logaritmik dönüşüm uygulayarak veriyi yeniden düzenleme işlemi gerçekleştirmektedir.

```
tarim_veri_3 <- tarim_veri_2 %>%
  mutate(across(starts_with("gelir"), log, .names = "log_{.col}"),
         id = as_factor(id),
         yordim = as_factor(yordim)) %>%
  select(id, yordim, starts_with("log"), everything(), -starts_with("gelir"))

tarim_veri_3
## # A tibble: 3,419 × 12
##   id yordim log_gelir_2018 log_gelir_2019 log_gelir_2020 log_gelir_2021
##   <fct> <fct>          <dbl>          <dbl>          <dbl>          <dbl>
## 1 8 0 14.9 15.5 15.0 14.9
## 2 9 0 10.1 8.57 9.61 9.34
## 3 10 1 11.6 10.3 10.3 9.89
## 4 11 0 10.6 12.4 11.9 12.5
## 5 13 0 13.9 13.4 13.0 12.9
## 6 15 0 11.0 11.5 10.8 10.6
## 7 16 0 14.6 14.1 13.0 13.5
## 8 19 0 9.28 12.1 12.3 11.7
## 9 20 0 11.5 11.1 10.7 11.1
## 10 27 0 12.9 12.9 11.9 11.1
## # 3,409 more rows
## # 6 more variables: log_gelir_2022 <dbl>, alan_2018 <dbl>, alan_2019 <dbl>,
## # alan_2020 <dbl>, alan_2021 <dbl>, alan_2022 <dbl>
```

Kodun açıklaması:

- *tarim_veri_3 <-*: İşlem sonucunu *tarim_veri_3* adında yeni bir değişkene atamaktadır.
- *tarim_veri_2 %>%*: *tarim_veri_2* veri çerçevesini *%>%* operatörü ile sonraki işlemlere aktarmaktadır.
- *mutate(across(starts_with("gelir"), log, .names = "log_{.col}"))*: *mutate* ve *across* fonksiyonları kullanılarak, “gelir” ile başlayan tüm sütunlara logaritmik dönüşüm uygulanır ve yeni sütun adları “log_” ile başlayacak şekilde belirlenir. Örneğin, “gelir_2018” sütunu “log_gelir_2018” olarak dönüştürülmüştür.
- *id = as_factor(id)*: *id* sütununu faktör veri tipine dönüştürür, bu işlem kategorik veriler için yapılmaktadır.
- *yordim = as_factor(yordim)*: *yordim* sütunu da faktör veri tipine dönüştürülmüştür.
- *select(id, yordim, starts_with("log"), everything(), -starts_with("gelir"))*: “log” adı ile başlayan yeni sütunları, *id* ve *yordim* sütunları ile birlikte seçip, ham veride “gelir” adı ile başlayan sütunları seçimden çıkarmaktadır. Bu işlem, veri çerçevesini yeniden düzenlemektedir. Yani, *tarim_veri_3* yeni veri çerçevesi sadece istenen sütunları içerecek şekilde filtrelenerek oluşturulmuştur.

➤ Aykırı Değerlerin Veri Setinden Çıkarılması

Aykırı değerler, genellikle veri setinin geri kalanından önemli ölçüde farklı değerler olarak tanımlanmaktadır. Veri setinden, genel dağılımın dışına çıkan ve analiz sonuçlarını yanıltabilecek değerler temizlenmektedir. Bu tür değerler analiz aşamasında sonuçları etkileyebileceğinden analiz öncesinde incelenmelidir. Bu çerçevede, görsel inceleme başta olmak üzere diğer araçlardan (IQR vb.) yararlanılarak aykırı değerlerin hatalı veri girişi, ölçüm hatası ya da veri yapısının doğasından kaynaklanıp kaynaklanmadığı değerlendirilmelidir.

Aykırı değerler hatalı veri girişi, ölçüm hataları nedeniyle de ortaya çıkabilmektedir. Analize başlamadan önce veriyi tanıma aşamasında bu ihtimaller göz önünde bulundurularak veriye göz atılmalıdır. Aykırı değerlere ilişkin yapılan işlemler, analizin yeniden türetilmesini sağlayacak şekilde raporlanmalıdır.

Aşağıdaki kod, *tarim_veri_3* veri setimizde yer alan *log_gelir_2021* değişkeni için aykırı değerleri belirleyip, bu değerleri içeren gözlemleri veri setinden çıkarmaktadır. Bu işlem, yalnızca 2021 yılı için uygulanmıştır. Sebebi 2021 yılı ile 2022 yılının etki değerlendirmesi de kullanılacak olmasıdır. 2018, 2019 ve 2020 yılları ise iki grup arasında eğilimlerin kontrolü amacıyla kullanılmıştır ve bu yıllar için aykırı değerler veri setinden çıkarılmamıştır.

```
# quantile fonksiyonu ile Q1 ve Q3 değerlerin hesaplanması
Q1 <- quantile(tarim_veri_3$log_gelir_2021, 0.25, na.rm = TRUE)
Q3 <- quantile(tarim_veri_3$log_gelir_2021, 0.75, na.rm = TRUE)

# IQR (Interquartile Range) hesaplanması
IQR <- Q3 - Q1

# Alt ve üst sınırların hesaplanması
alt_sinir <- Q1 - 1.5 * IQR
ust_sinir <- Q3 + 1.5 * IQR

# Aykırı olmayan değerlerin filtresinin oluşturulması
normal_deger_filtresi <- (tarim_veri_3$log_gelir_2021 >= alt_sinir) & (tarim_veri_3$log_gelir_2021 <= ust_sinir)

# Aykırı olmayan değerleri içeren veri setinin güncellenmesi
tarim_veri_3_2021 <- tarim_veri_3[normal_deger_filtresi, ]

# Sonuç
print(tarim_veri_3_2021)
## # A tibble: 3,374 × 12
##   id   yarim log_gelir_2018 log_gelir_2019 log_gelir_2020 log_gelir_2021
##   <fct> <fct>      <dbl>      <dbl>      <dbl>      <dbl>
## 1 8     0         14.9       15.5       15.0       14.9
## 2 9     0         10.1       8.57      9.61       9.34
## 3 10    1         11.6       10.3      10.3       9.89
## 4 11    0         10.6       12.4      11.9       12.5
## 5 13    0         13.9       13.4      13.0       12.9
## 6 15    0         11.0       11.5      10.8       10.6
## 7 16    0         14.6       14.1      13.0       13.5
## 8 19    0          9.28      12.1      12.3       11.7
## 9 20    0         11.5       11.1      10.7       11.1
## 10 27   0         12.9       12.9      11.9       11.1
## # 3,364 more rows
## # 6 more variables: log_gelir_2022 <dbl>, alan_2018 <dbl>, alan_2019 <dbl>,
## # alan_2020 <dbl>, alan_2021 <dbl>, alan_2022 <dbl>
```

Kodun açıklaması:

- **Quantile ile Hesaplama:** İlk olarak, *log_gelir_2021* değişkeninin birinci (%25) ve üçüncü (%75) çeyrek değerleri (Q1 ve Q3) *quantile* fonksiyonu kullanılarak hesaplanmaktadır. *na.rm = TRUE* parametresi, eksik (NA) değerlerin hesaplamalarda dikkate alınmamasını sağlamaktadır. Veri setimizde bir önceki adımda NA değerler atıldığı için, bu aşamada *na.rm* argümanının yazılmaması da mümkündür.
- **IQR (Interquartile Range) Hesaplama:** IQR, Q3 ve Q1 arasındaki fark olarak hesaplanır. IQR, veri setinin orta %50'sinin yayılımını ölçmektedir. Aykırı değerlerin tespitinde kullanılmaktadır.
- **Alt ve Üst Sınırların Hesaplanması:** Aykırı değerleri belirlemek için alt ve üst sınırlar, sırasıyla $Q1 - 1.5 * IQR$ ve $Q3 + 1.5 * IQR$ formülleri kullanılarak hesaplanmaktadır. Bu sınırlar dışında kalan değerler aykırı olarak kabul edilmektedir.
- **Aykırı Olmayan Değerlerin Filtresinin Oluşturulması:** *normal_deger_filtresi*, *log_gelir_2021* değişkeni alt ve üst sınırlar arasında kalan (yani aykırı olmayan) değerler için bir mantıksal filtre oluşturmaktadır.
- **Aykırı Olmayan Değerleri İçeren Veri Setinin Güncellenmesi:** *tarim_veri_3* veri çerçevesi, oluşturulan filtre kullanılarak güncellenmektedir. Bu işlem, sadece aykırı değerlerin çıkartıldığı yeni bir veri çerçevesi oluşturmaktadır (*tarim_veri_3_2021*).
- **Sonuç:** Güncellenmiş veri seti *tarim_veri_3_2021* adını almıştır ve analizler için daha temiz ve daha tutarlı bir veri seti haline getirilmiştir.

➤ Verinin Uzun (Long) Formata Çevrilmesi

Veriyi temizledikten sonra Farkların Farkı yöntemini uygulayabilmek ve öncesinde görselleştirme ile varsayımlarını kontrol edebilmek için veri seti, mevcut geniş formattan uzun formata dönüştürülmektedir. Uzun format, örneğin ayrı yıl sütunları olarak verilen “Gelir” değişkenini tek bir sütun haline dönüştürmektedir. Ancak bu birleşim ayrıdan bir de “Yıl” değişkeni tanımlanmasını gerektirmektedir. Benzer verilerin bu şekilde tek bir sütunda toplanması ile R’da gruplar arası karşılaştırma yapmak basitleşmektedir. Aşağıdaki kod bloğu, *tarim_veri_3_2021* veri çerçevesini geniş formattan uzun formata dönüştürmektedir.

```
# Veri çerçevesini geniş formattan uzun formata çevirme
long_tarim_veri <- tarim_veri_3_2021 %>%
  pivot_longer(cols = starts_with("log_gelir_") | starts_with("alan_"),
    names_to = c(".value", "yil"),
    names_pattern = "(.*)_(\\d+)"
  ) %>%
  mutate(yil = as_factor(yil),
    yil_num = as_factor(ifelse(yil == "2022", 1, 0)),
    yardim = as_factor(yardim)) %>%
  select(id, yil, yil_num, everything())

# Uzun formattaki veriyi gösterme
long_tarim_veri
## # A tibble: 16,870 × 6
##   id   yil   yil_num yordim log_gelir alan
##   <fct> <fct> <fct>   <fct>   <dbl> <dbl>
## 1 8    2018 0       0      14.9    33
## 2 8    2019 0       0      15.5    29
```

```
## 3 8 2020 0 0 15.0 19
## 4 8 2021 0 0 14.9 20
## 5 8 2022 1 0 14.7 21
## 6 9 2018 0 0 10.1 2
## 7 9 2019 0 0 8.57 2
## 8 9 2020 0 0 9.61 2
## 9 9 2021 0 0 9.34 2
## 10 9 2022 1 0 8.57 2
## # i 16,860 more rows
```

Kodun açıklaması:

- **Uzun Formata Çevirme:** *pivot_longer* fonksiyonu kullanılarak, *log_gelir_* veya *alan_* ile başlayan tüm sütunlar uzun formata çevrilmektedir. Bu işlem, birden fazla sütunu tek bir sütuna indirgerken, her bir gözlem için birden fazla satır oluşmuş olacaktır. Bu şekilde, veri seti daha esnek bir yapıya kavuşur ve çeşitli analiz türleri için daha uygun hale gelmektedir.
- **names_to ve names_pattern:** *names_to* parametresi ile sütun isimleri iki bölüme ayrılmaktadır: *.value* ve *yil*. *.value* bir yer tutucudur ve *names_pattern* ile eşleşen grupları korumaktadır. Örneğin, *log_gelir* veya *alan*. *names_pattern* parametresi ise sütun isimlerindeki deseni tanımlamaktadır ve *(.*)\\d+* regex'i ile sütun isimleri iki parçaya (*gelir/alan* ve *yil*) ayrılmaktadır.
- **Yıl Sütununun Faktöre Çevrilmesi:** Elde edilen *yil* sütununu *as.factor* fonksiyonu ile faktör tipine dönüştürerek, bu sütun üzerinde kategorik analizlerin yapılabilmesini sağlamaktadır.
- **yil_num Sütununun Oluşturulması:** *yil* sütununa dayanarak *yil_num* adında yeni bir sütun oluşturulmuştur. Eğer *yil* "2022" ise *yil_num*'a "1" değeri atanmakta, değilse "0" değeri atanmaktadır. 2022 yılının 1 değerini alması etki değerlendirme de 2022 yılında bir değişim olup olmadığının araştırıldığını göstermektedir.
- **yardim Sütununun Faktöre Çevrilmesi:** *yardim* sütunu da faktör tipine çevrilmiştir. Bu, *yardim* sütununun kategorik bir değişken olarak analiz edilmesine olanak tanımaktadır.
- **select ile Sütunların Seçilmesi ve Yeniden Düzenlenmesi:** *select* fonksiyonu, sonuç veri çerçevesindeki sütunların sıralamasını ayarlamaktadır. Görüldüğü üzere *id*, *yil*, *yil_num* ve diğer tüm sütunlar belirlenen sırayla seçilmiştir.

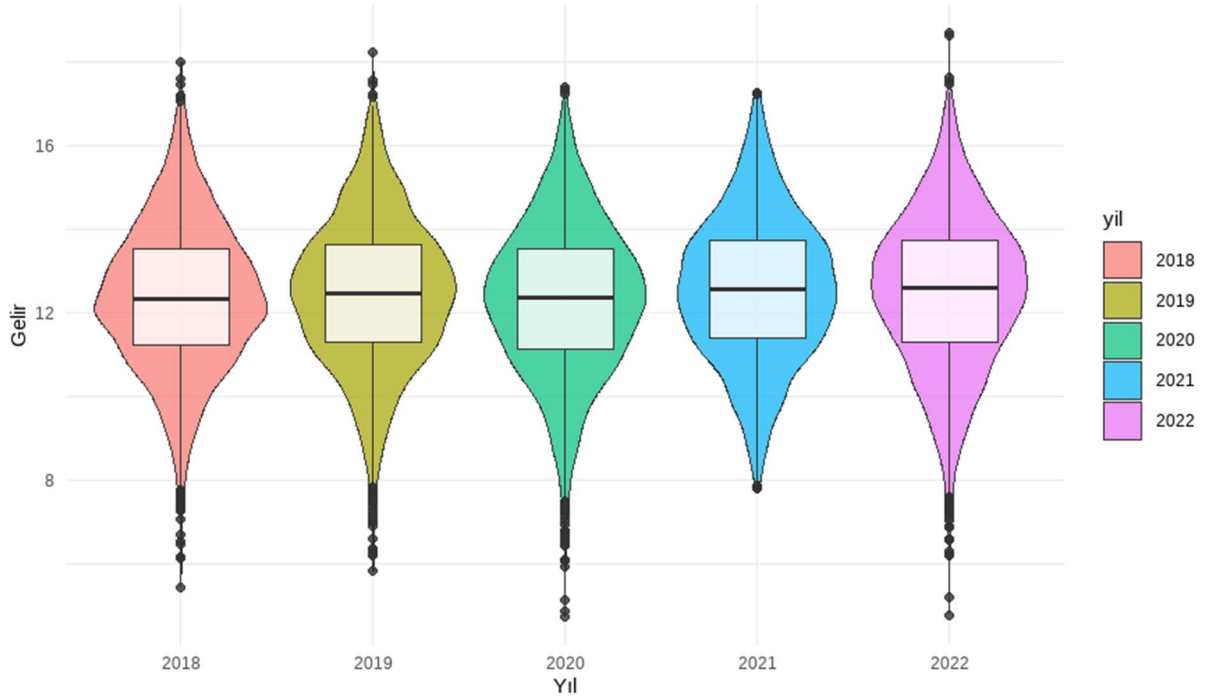
➤ Güncel Veri Setinin Dağılımının İncelenmesi

Verinin dağılımının detaylı analizi, farklı yıllardaki gelir değişimlerini, trendleri ve potansiyel anomalileri anlamak için gerçekleştirilmektedir. Bu görselleştirme yöntemi, yalnızca mevcut durumu değil, aynı zamanda zaman içindeki değişimleri de kapsamlı bir şekilde sunmaktadır. Aşağıdaki kod bloğu, *ggplot2* paketini kullanarak *long_tarim_veri_3_2021* veri çerçevesindeki *log_gelir* değişkeninin yıllara göre dağılımını görselleştirmektedir.

```
ggplot(long_tarim_veri_3_2021, aes(x = yil, y = log_gelir, fill = yil)) +
  geom_violin(alpha = 0.7) +
  geom_boxplot(width = 0.5, fill = "white", alpha = 0.8) +
  theme_minimal() +
  labs(title = "Yıllara Göre Log Hasılatın Görselleştirmesi",
       y = "Log Hasılat",
       x = "Yıl")
```

Grafik 21. Örnek 1 - Yıllara Göre Gelir Dağılımı

Yıllara Göre Gelirin Görselleştirilmesi



Kodun açıklaması:

- **Görselleştirme Fonksiyonu Belirleme:** `ggplot(long_tarim_veri_3_2021, aes(x = yıl, y = log_gelir, fill = yıl))` ile bir ggplot nesnesi oluşturulmaktadır. `aes` (aesthetics), x eksenini için `yıl` ve y eksenini için `log_gelir` değişkenlerini tanımlayan argümandır. `fill = yıl` argümanı, tamamen görsellik amaçlı farklı yılları farklı renklerle doldurmak için kullanılmaktadır.
- **Violin Grafiği Ekleme:** `geom_violin(alpha = 0.7)` ile, verinin dağılımını gösteren violin grafikler eklenmektedir. `alpha = 0.7` saydamlık derecesini ayarlamaktadır. Böylece grafikte saydamlık oluşmakta alttaki katmanlar görülebilmektedir.
- **Kutu Grafiği Ekleme:** `geom_boxplot(width = 0.5, fill = "white", alpha = 0.8)` ile, dağılıma ilişkin kutu grafikler eklenmektedir. `width = 0.5` kutu grafiklerinin genişliğini belirlemektedir. `fill = "white"` kutuların içini beyaz renkle doldurmakta, `alpha = 0.8` saydamlık derecesini belirlemektedir.
- **Tema Uygulama:** `theme_minimal()` ile grafik için minimal bir tema seçilmektedir.
- **Başlık ve Eksen Etiketleri Ekleme:** `labs(title = "Yıllara Göre Log Hasılatın Görselleştirilmesi", y = "Log Hasılat", x = "Yıl")` ile grafik başlığı, y eksenini ve x eksenini için etiketler belirlenmektedir.

Özetle, her bir yılın gelir dağılımı farklı renkte bir violin grafiği ile gösterilmiştir. Violin grafiklerinin genişliği, veri yoğunluğunun fazla olduğu gelir seviyelerini göstermektedir. Eğer bir violin grafiği belirli bir noktada daha genişse, o noktadaki gelir değerleri daha sık görülmektedir. Her bir violin grafiğinin içinde, medyan (orta çizgi), çeyrekler (kutunun üst ve alt kenarları) ve aykırı değerler (kutunun dışında kalan

noktalar) yer almaktadır. Bu, verinin dağılımının merkezi eğilim ve yayılım ölçüleri⁷² hakkında bilgi vermektedir. Örneğin, violin grafiğinde üst veya alt kısım daha genişse, o yönde bir çarpıklık olduğu ve verinin o alanda yoğunlaştığı şeklinde yorumlanmaktadır. Kutu grafiklerinin dışında kalan tekil noktalar aykırı değerleri temsil etmektedir. Eğer bir yıl için çok sayıda aykırı değer varsa, bu yılın gelir dağılımında bazı gözlemlerin diğerlerinden önemli ölçüde farklı olduğunu göstermektedir. Aykırı değerlerin sayısı ve yoğunluğu, bu gözlemlerin analize etkisinin değerlendirilmesinde önemlidir. Veri setinde son olarak yapılan aykırı değer atma işlemi ile 330 kişi müdahale grubu, 3044 kişi kontrol grubu olarak analize devam edilmiştir.

```
yardim_sayisi_3 <- tarim_veri_3_2021 %>%
  group_by(yardim) %>%
  summarize(count = n())

yardim_sayisi_3
## # A tibble: 2 × 2
##   yardim count
##   <fct>   <int>
## 1 0       3044
## 2 1        330
```

Adım 6: Varsayımların Kontrolü

Veri setinden sonra yönetime karar verme aşamasına geçilmektedir. Çalışmada hem zaman hem de grup bazında bir etki değerlendirme yapılması ön görülmüştür. Bu kapsamda Farkların Farkı yönteminin uygulanması mümkün görünmektedir; ancak öncelikle bu yöntemin en önemli varsayımı olan “Paralel Eğilim” varsayımının kontrol edildikten sonra modelin kurulması gerekmektedir.

► Paralel Eğilim Varsayımının Kontrolü

Aşağıdaki kod ile yapılan işlemlerde *long_tarim_veri_3_2021* adlı veri çerçevesi kullanılarak, yardımdan yararlanan ve yararlanmayan gruplar arasında yıllık gelir ortalamalarının ve bu ortalamaların güven aralıklarının karşılaştırmalı bir analizi gerçekleştirilmektedir. Bu analiz, farklı gruplar arasında zaman içindeki eğilimlerin benzerliğini ve dolayısıyla paralel eğilim varsayımını test etmek için kullanılmaktadır.

```
plot_tarim_veri <- long_tarim_veri_3_2021 %>%
  mutate(yardim = factor(yardim, labels = c("Yardım Almayan", "Yardım Alan")),
         yil = factor(yil, labels = c("2018", "2019", "2020", "2021", "2022"))) %>%
  group_by(yardim, yil) %>%
  summarize(ort_gelir = mean(log_gelir, na.rm = TRUE),
            se_gelir = sd(log_gelir, na.rm = TRUE) / sqrt(n()),
            upper = ort_gelir + (1.96 * se_gelir),
            lower = ort_gelir - (1.96 * se_gelir))
## 'summarise()' has grouped output by 'yardim'. You can override using the
## 'groups' argument.
plot_tarim_veri
## # A tibble: 10 × 6
## # Groups:   yardim [2]
##   yardim    yil ort_gelir se_gelir upper lower
##   <fct>    <fct>   <dbl>   <dbl> <dbl> <dbl>
## 1 Yardım Almayan 2018    12.2  0.0305  12.3  12.2
```

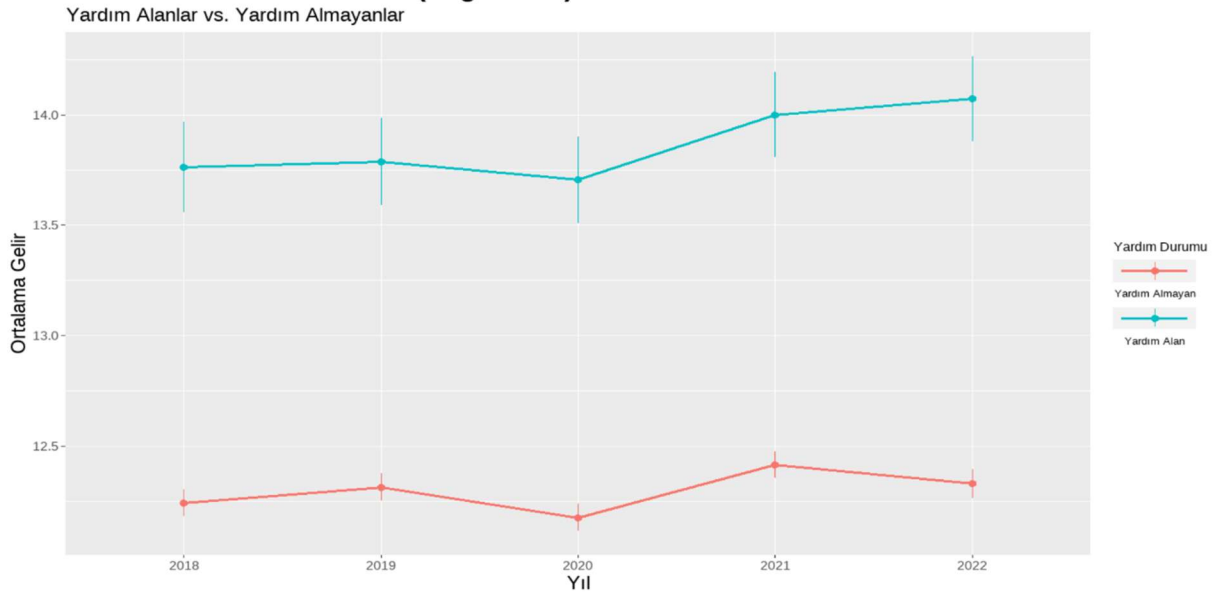
⁷² Merkezi eğilim ve yayılım ölçüleri, bir veri setinin özelliklerini özetlemek için kullanılan istatistiksel terimlerdir. Bu ölçüler, veri setinin genel yapısını ve dağılımını anlamada yardımcı olmaktadır. Örneğin ortalama, mod, medyan merkezi ölçüm; varyans, standart sapma yayılım ölçüleridir.

```
## 2 Yardım Almayan 2019 12.3 0.0310 12.4 12.3
## 3 Yardım Almayan 2020 12.2 0.0319 12.2 12.1
## 4 Yardım Almayan 2021 12.4 0.0303 12.5 12.4
## 5 Yardım Almayan 2022 12.3 0.0327 12.4 12.3
## 6 Yardım Alan 2018 13.8 0.104 14.0 13.6
## 7 Yardım Alan 2019 13.8 0.101 14.0 13.6
## 8 Yardım Alan 2020 13.7 0.0997 13.9 13.5
## 9 Yardım Alan 2021 14.0 0.0979 14.2 13.8
## 10 Yardım Alan 2022 14.1 0.0986 14.3 13.9
```

- **mutate** fonksiyonu, **yardım** ve **yıl** sütunlarını etiketli faktörlere dönüştürür. Bu, grafikteki grupları ve eksen etiketlerini daha anlaşılır hale getirir.
- **group_by** fonksiyonu, veriyi **yardım** ve **yıl** sütunlarına göre gruplandırır, böylece her bir yardım/yıl kombinasyonu için istatistikler ayrı ayrı hesaplanabilir.
- **summarize** fonksiyonu, her grupta logaritmik gelirin ortalamasını (**ort_gelir**), standart hatasını (**se_gelir**) ve %95 güven aralığının üst ve alt sınırlarını hesaplar.

```
ggplot(plot_tarim_veri, aes(x = yıl, y = ort_gelir, color = yardım)) +
  geom_pointrange(aes(ymin = lower, ymax = upper), size = 0.5) +
  geom_line(aes(group = yardım)) +
  labs(x = "Yıl",
       y = "Ortalama Gelir")
```

Grafik 22. Örnek 1 - Paralel Eğilim Grafiği
Yıllara Göre Ortalama Gelir (Logaritmik)



- **geom_pointrange** fonksiyonu, her yıl için hesaplanan ortalamaları ve güven aralıklarını puan ve çizgi olarak çizer.
- **geom_line** fonksiyonu, aynı yardım grubundaki yıllar arasındaki bağlantıyı çizgi ile gösterir, bu da zaman içindeki trendleri gözlemlemeyi kolaylaştırır.

Görsel analiz ile verinin logaritması alındıktan sonra grafikten de anlaşılacağı üzere büyük ölçüde ortak bir trende sahip oldukları görüldüğünden paralel eğilim varsayımı kabul edilerek yöntemin uygulamasına geçilmektedir.

Öte yandan, paralel eğilim varsayımının doğrulanması ekonometrik model aracılığıyla da kontrol edilebilir. Bu çerçevede, müdahale öncesi dönemde paralel eğilim varsayımının ihlal edilip edilmediği oluşturulacak model yardımı ile test edilmektedir.

Söz konusu varsayımın kontrolü için oluşturulacak modelde bağımlı değişken olarak yine sonuç değişkeni kullanılmalıdır. Görselleştirme ile yapılan kontrolde de sonuç değişkeninin kullanıldığı hatırlanacaktır. Oluşturulacak modelin bağımsız değişkenlerini ise müdahale öncesi döneme ait yıllar ile müdahale ve kontrol grubunu birbirinden ayıran kukla değişken oluşturmaktadır.

$$\log(Yatırım) = \beta_0 + \beta_1 yıl + \beta_2 grup + \beta_3 (yıl * grup) + \epsilon$$

Oluşturulan modelde etkileşim katsayısına (β_3) dikkat edilmelidir. Zira, belirtilen katsayı, yılların bağımlı değişken üzerindeki etkisinin müdahale ve kontrol grupları açısından farklılaşıp farklılaşmadığını belirlemeyi sağlar. Bu çerçevede, eğer müdahale öncesinde yıllar itibarıyla bağımlı değişkenin eğilimi müdahale ve kontrol grubu açısından farklılaşıyorsa, başka bir deyişle etkileşim katsayısı istatistiksel olarak anlamlı ise o zaman söz konusu gruplar arasında, müdahale öncesi dönemde bir farklılık ortaya çıkmaktadır. Bu nedenle paralel eğilim varsayımının ihlal edildiği sonucuna ulaşılmaktadır. Öte yandan, etkileşim katsayısının istatistiksel olarak anlamsız olması, müdahale ve kontrol gruplarının müdahale öncesi dönemde benzer eğilim sergilediği ve paralel eğilim varsayımının desteklendiği anlamına gelmektedir.

Paralel eğilim varsayımının ihlal edilip edilmediğinin kontrol edilmesine yönelik bu yaklaşım, görselleştirmeyle birlikte de kullanılabilir. Böylelikle, ekonometrik modelden elde edilen sonuçlar görselleştirmeyle de desteklenerek paralel eğilim varsayımının kontrolüne ilişkin daha bütüncül ve tutarlı bir değerlendirme yapılabilmektedir.

Belirtilen açıklamalar ışığında ekonometrik model yardımıyla kontrol yapılabilmesi için öncelikle müdahale öncesindeki yıllar filtrelenmelidir.

```
veri_did_pre <- veri_did %>% filter(year < 2025)
```

Bir sonraki aşamada, bağımlı değişkenin sonuç değişkeni, bağımsız değişkenlerin ise müdahale öncesi döneme ait yıllar ile müdahale ve kontrol gruplarını birbirinden ayıran kukla değişkenin olduğu model oluşturularak tahmin edilmelidir.

```
model_kontrol <- lm(log_investment ~ year + treat + (year * treat),
                    data = veri_did_pre)
summary(model_kontrol)
```

Örnek olarak oluşturulan modelde “year” isimli değişken müdahale öncesi yılları, “treat” isimli değişken ise müdahale ve kontrol gruplarını ayıran kukla değişkeni temsil etmektedir. Buna ek olarak, modelin bağımlı değişkeni “log_investment” isimli logaritmik dönüşümü yapılmış yatırım harcaması oluşturmaktadır.

```

Coefficients:
##          Estimate   Std. Error t value Pr(>|t|)
## (Intercept) 31.593709   8.158047   3.873  0.00011 ***
## year        -0.009968   0.004033  -2.472  0.01350 *
## treat        -9.911949  11.537221  -0.859  0.39034
## year:treat  0.004935  0.005703  0.865  0.38691
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

Oluşturulan modele ilişkin etkileşim katsayısının (year:treat) kontrol edilmesi gerekmektedir. Bu çerçevede, örnek uygulamadaki etkileşim katsayısına ait p-değeri (0,38691), yüzde 5 (0,05) anlamlılık düzeyinden büyük olduğu için boş hipotez reddedilememekte ve müdahale öncesi dönemde yılların bağımlı değişken üzerindeki etkisinin, müdahale ve kontrol grupları arasında istatistiksel olarak anlamlı bir farklılık göstermediği sonucuna ulaşılmaktadır.⁷³ Böylelikle, paralel eğilim varsayımı ihlal edilmediğinden Farkların Farkı yöntemiyle etki değerlendirme aşamasına geçilebilecektir.

Öte yandan, eğer gerçekleştirilen tahmin sonrasında etkileşim katsayısına ait p-değeri, yüzde 5 olarak belirlenen anlamlılık düzeyinden küçük olsaydı, bu durumda boş hipotez reddedilecekti. Bu durumda, müdahale öncesi dönemde yılların bağımlı değişken üzerindeki etkisinin müdahale ve kontrol grupları arasında farklılaştığı ve bu bakımdan paralel eğilim varsayımının ihlal edildiği sonucuna ulaşılabilecekti. Böyle bir durumun söz konusu olması halinde eşleştirme yapıldıktan sonra paralel eğilim varsayımı yeniden kontrol edilmeli ve söz konusu varsayımın geçerliliği doğrulandıktan sonra etki değerlendirme aşamasına geçilmelidir. Bununla birlikte, eğer eşleştirme yapıldıktan sonra da paralel eğilim varsayımı yine ihlal ediliyorsa, bu durumda Farkların Farkı yöntemi kullanılmamalıdır.⁷⁴

Bu örneğimizde görsel analiz sonrası, paralel eğilim varsayımının geçerli olduğu görüldüğünden yöntemin uygulanmasına geçilmektedir.⁷⁵

Adım 7: Farkların Farkı ile Tahmin

Etki değerlendirme yöntemi olarak Farkların Farkı kullanılmıştır. Farkların Farkı yöntemi, müdahalenin iki gruptaki etkisini karşılaştırarak, müdahalenin neden olduğu olası değişiklikleri netleştirmeyi amaçlamaktadır. Öncelikle 2021 ve 2022 yıllarına ait veriler veri setinden filtrelenmektedir. Ardından 2021 müdahalenin uygulanmadığı, 2022 uygulandığı yıl olarak ele alınmaktadır.

```

filtered_tarim_veri <- long_tarim_veri %>%
  filter(yil %in% c("2021", "2022"))

```

Kurulacak model, müdahalenin (yardımın) etkisini, müdahale yapılmayan grupla karşılaştırarak gelir üzerindeki etkisini ölçmeye çalışmaktadır. Bu etki, grup ve zaman

⁷³ Gerçekleştirilen hipotez testinde, boş hipotez, “müdahale öncesi dönemde yılların bağımlı değişken üzerindeki etkisi müdahale ve kontrol grupları arasında anlamlı biçimde farklılaşmamaktadır” şeklindedir.

⁷⁴ Eşleştirme yaparken dikkat edilecek hususlar için bakınız “3.2.1. Farkların Farkı Yönteminde İzlenmesi Gereken Adımlar: Adım 3: Varsayımların Kontrolü”

⁷⁵ Varsayımın ihlaline karşı Farkların Farkı yöntemi ile elde edilen modelin gücünü test eden robust analysis başlıklı çalışmalar yazında mevcuttur. Elde çeşitli yıllara ait verilerin olması bu tür testlerin yapılmasını kolaylaştırmaktadır.

(yil_num) değişkenlerinin sabit etkisi ile müdahalenin zamana bağlı değişen etkisini gösteren etkileşim (yardim * yil_num)⁷⁶ terimi açısından değerlendirmektedir.

```
model_did <- lm(log_gelir ~ yardım + yil_num +  
  (yardim * yil_num),  
  data = long_tarim_veri)  
tidy(model_did)
```

- **Kesişim (Intercept):** Modelin kesme noktası (sabit terimi), program uygulanmamış olsaydı ne olurdu sorusunun yanıtını vermektedir. Hesaplamada yardım almayan ve *yil_num* 0'a denk olan grubun ortalama geliri tahmin edilmektedir. Çıkan sonuçlarda istatistiki olarak anlamlılık tespit edilmiştir. Burada yapılan anlamlılık testi $H_0: \beta = 0$ yokluk hipotezini test etmektir. Hipotezin test değeri için t istatistiği ve bu istatistiğe ait bir olasılık (p-value) değeri hesaplanmaktadır. Bu çalışmanın sonucunda kesişim katsayısı ($2e-16 < p=0,05$) için yokluk hipotezi red edilmektedir. Bu sebeple anlamlı olduğu ifade edilmektedir. Bağımlı değişkenin öngörülen değeri olarak yorumlanmaktadır. Bağımsız değişkenin olmadığı durumda y, yani bağımlı değişken için öngörülen beklenen değerdir.
- **yardim1:** Yardımdan yararlanan grubun, yararlanmayan gruba göre ortalama gelirindeki artış miktarını göstermektedir. Müdahale grubunda yer alan kişilerin ortalamalarının yararlanmayan gruba göre daha yüksek olduğu görülmektedir. Aynı zamanda t istatistik değeri için hesaplanan p olasılık değerine⁷⁷ bakıldığında istatistiki olarak bu ortalama değer anlamlıdır.
- **yil_num1:** Çalışmada 2021 yılı için "0", 2022 yılı için "1" kategorik değeri kullanıldığından çalışmada 2022 yılının bağımlı değişken üzerindeki etkisi yorumlanmaktadır. İstatistiki olarak anlamlı olmadığı görülmektedir. Diğer bağımsız değişkenler sabit olduğunda (ceteris paribus) zaman 1 birim ilerlediğinde bağımlı değişkende 0.04143 birim değişme olacağı öngörülmektedir. Ancak bu katsayı tek başına istatistiki olarak anlamlı değildir (bkz p olasılık değeri = $2.39e-1$).
- **yardim1:yil_num1:** Bu değişken Farkların Farkı modelinde, asıl yorumlanması gereken değişkendir ve tahmin edilen bu etkileşim teriminin katsayısı gelir üzerindeki etkiyi göstermektedir. Bu değişkenin katsayısı için hesaplanan p olasılık değeri 0.0544 çıkmıştır. %95 anlamlılık düzeyinde (0,05 hata payı referans alındığında) hipotezi red etmek mümkün olmamakta, dolayısıyla istatistiki olarak anlamlı olmamaktadır. Fakat risk payı yükseltirirse yani %90 anlamlılık düzeyinde (0,10 hata payı) hipotez test edilirse anlamlı kabul edilebilmektedir. Hata düzeyini artırmak, tahmin edilen örneklem güven aralığını genişletmek yani tahminlerin kesimliliğini azaltmaktır. Fakat p olasılık değeri için %5 hata payına yakın bir değer çıkmış olması sebebiyle bu çalışmada %90 anlamlılık düzeyinde anlamlılığı kabul edilmiştir. Gerçek verilerle yapılan çalışmalarda beklenen, olabildiğince risk payı düşük olacak şekilde hipotezlerin doğrulunu test etmektir. Analizlerde tek bir model tahmini ile sınırlı kalınmamasının sebeplerinden biri de bu şekilde çeşitli durumlarla karşılaşma durumunun gerçek ekonomik verilerde sıkça karşılaşılan bir durum olmasından kaynaklıdır.

⁷⁶ Yardımın ve zamanın etkileşim terimi, müdahalenin (yardımın) zamanla değişen etkisini ölçer. Eğer etkileşim terimi istatistiksel olarak anlamlıysa, bu müdahalenin belirli bir zaman etkisi olduğunu gösterir.

⁷⁷ İstatistiksel bir testin sonuçlarını yorumlamada kullanılan olasılık değerleridir.

- **Belirtme Katsayısı (Multiple R-squared):** R-kare değeri, bağımsız değişkenlerin bağımlı değişkeni açıklama yüzdesini göstermektedir. Çalışmada 0.06731, yani model, veri setindeki varyansın yaklaşık %6.7'sini açıklamaktadır. Oransal bakıldığından düşük çıkmıştır. Bu bağımlı değişkenin açıklaması için ek bağımsız değişkenlere ihtiyaç olduğunu göstermektedir. Bu örnek bir senaryo olduğundan bu şekilde yorumlanarak bırakılmıştır.
- **F-istatistiği:** F istatistiği, modelin tüm katsayılarının sıfıra eşit olup olmadığını ($H_0: \beta_1 = \beta_2 = \dots = \beta_k = 0$) test etmektedir. Modelin anlamlı olup olmadığını ifade etmektedir. Çalışmalarda sadece katsayıların değil, modelinde anlamlı olması gerekmektedir. F istatistiği için hesaplanan p olasılık değerini yorumlamak pratiktir. Çıktıdan görüldüğü üzere, p olasılık değeri $2.2e-16$ olup yokluk hipotezi red edilmektedir. Bu durum en bir katsayının sıfırdan farklı olduğu ve modelin istatistiki olarak anlamlı olduğu şeklinde yorumlanmaktadır.

```
## # A tibble: 4 × 5
##   term      estimate std.error statistic  p.value
##   <chr>      <dbl>    <dbl>    <dbl>    <dbl>
## 1 (Intercept)  12.3    0.0158    780. 0
## 2 yadiml      1.52    0.0504    30.3 8.09e-196
## 3 yil_numl     0.0414  0.0352     1.18 2.39e- 1
## 4 yadiml:yil_numl 0.217  0.113     1.92 5.44e- 2
## summary(model_did, robust=T)
##
## Call:
## lm(formula = log_gelir ~ yadiml + yil_numl + (yadiml * yil_numl),
##     data = long_tarim_veri)
##
## Residuals:
##   Min     1Q   Median     3Q    Max
## -7.5598 -1.1055  0.0599  1.1598  5.9637
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   12.28909   0.01575  780.140 <2e-16 ***
## yadiml         1.52376   0.05037  30.252 <2e-16 ***
## yil_numl       0.04143   0.03522   1.176  0.2395
## yadiml:yil_numl 0.21671   0.11263   1.924  0.0544 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.738 on 16866 degrees of freedom
## Multiple R-squared:  0.06731, Adjusted R-squared:  0.06714
## F-statistic: 405.7 on 3 and 16866 DF, p-value: < 2.2e-16
```

Çalışmanın başında veri setinde “alan” adlı bağımsız değişken olduğundan bahsedilmiş olmasına rağmen, bu değişkenin dahil edildiği regresyon modelinde hem değişkenin katsayısının anlamsız olması hem de modelin anlamsız olması sebebiyle rehberde sonuçlarına yer verilmemiştir.

Sonuç olarak, yapılan etki değerlendirmede uygulanan yardım programının çiftçilerin gelirleri üzerinde %21,7 oranında bir etkisi bulunmaktadır. Ancak iktisadi açıdan düşünüldüğünde gelirin sadece devlet yardımından etkilenmeyecektir. Bu sebeple modele bağımsız değişken eklenerek katsayıların ve modelin anlamlılığının yanı sıra belirtme katsayısının da göz önünde bulundurulması gerekmektedir. Örneğin, bu örnekte “alan” bağımsız değişkeni geliri açıklayan değişken olarak düşünülmüş ve modele dahil edildiğinde belirtme katsayısı %25 çıkmıştır; ancak katsayının anlamsız çıkması, aynı zamanda etki katsayısının da anlamsız çıkması sebebiyle gelirdeki değişimi

açıklayabilecek farklı değişkenlerin modele dahil edilmesi gerektiği açıktır. Örneğin, alan değişkeni yerine bu alanın ekimin kullanılan makine teçhizatın kalitesi, bölge, iklim, pazara erişim, hammadde fiyatları gibi birçok kategoride değişkenle çalışma tasarlanabilir.

Örnek 2. Ters Olasılıklı Ağırlıklandırma (IPW) Yöntemi

Ters Olasılıklı Ağırlıklandırma yönteminin etki değerlendirme çalışmasında nasıl uygulanacağı bu örnek ile uygulamalı olarak anlatılmıştır. Bu çerçevede, 1.250 firmaya ilişkin örnek bir veri seti⁷⁸ oluşturulmuştur. Veri setinde yer alan firmaların; 2023 yılına ait dolar cinsinden satış hasılatlarını “hasılat” değişkeni, faaliyet sürelerini “faaliyet süresi” değişkeni, devlet yardımından yararlanma durumlarını “yardım” değişkeni, dolar cinsinden yatırım tutarlarını “yatırım” değişkeni ve firmalarda hibrit çalışma yapılıp yapılmamasını ise “hibrit” değişkeni temsil etmektedir. Yardım değişkeninin ‘1’ değerini alıyor olması firmanın ilgili devlet yardımından yararlandığını, ‘0’ değeri tam aksini ifade etmektedir. Benzer şekilde “hibrit” değişkeninin ‘1’ olması da ilgili firmada hibrit çalışma yönteminin geçerli olduğunu temsil etmektedir.

```
# A tibble: 10 × 6
  sıra hasılat faaliyet_suresi yardım hibrit yatırım
  <int> <dbl>      <dbl> <int> <int> <dbl>
1 1 248491.    47.5 0 1 5416.
2 2 250347.    30.0 1 0 7823.
3 3 251356.    32.0 1 1 9093.
4 4 247068.    42.7 1 0 4691.
5 5 250536.    29.4 0 0 7114.
6 6 250633.    23.4 1 0 8126.
7 7 249282.    25.8 1 1 6862.
8 8 249317.    36.2 0 1 6160.
9 9 249294.    25.4 0 0 5989.
10 10 248887.    27.4 1 1 6272.
```

Örnekte imalat sanayii sektöründe faaliyet gösteren 1.250 firmanın, stratejik olarak tanımlanan ürünlerin üretimini sağlamaya yönelik gerçekleştirecekleri yatırım harcamalarını desteklemek üzere 2022 yılında yeni bir devlet yardımı uygulanmaya başlanmıştır. 2022 yılı başından itibaren teknoloji yoğun ve stratejik nitelikli ürün üreten söz konusu firmaların, makine ve teçhizat başta olmak üzere yatırım harcamalarına yönelik nakdi destek sağlanmaya başlanmıştır. Verilen devlet yardımı ile birlikte, yardımdan yararlanan firmaların yatırım harcamalarında artış sağlanması hedeflenmiştir. 2023 yılında söz konusu devlet yardımının etkisinin değerlendirilmesi için bir çalışma başlatılmıştır.

Araştırma ekibi bağımlı ve bağımsız değişkeni kullanarak doğrudan ortalama müdahale etkisi tahmin etmeye çalışmıştır. Bu çerçevede, tahmin edilecek modelin bağımlı değişkeni yatırım, bağımsız değişkeni ise yardım olmaktadır. Zira, devlet yardımından yararlanıp yararlanmamanın firmaların yatırım harcamaları üzerindeki etkisi araştırılmak istenilmektedir.

⁷⁸ Veri setinin adı “yardım_verisi” olup Başkanlığımız internet sitesinden erişilebilmektedir.

```

Call:
lm(formula = yatırım ~ yardım, data = yardım_verisi)

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  6167.74    52.61  117.23  <2e-16 ***
yardım      1781.51    72.79   24.47  <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1286 on 1248 degrees of freedom
Multiple R-squared:  0.3243,    Adjusted R-squared:  0.3238
F-statistic: 599 on 1 and 1248 DF, p-value: < 2.2e-16

```

Gerçekleştirilen tahmin sonucunda devlet yardımından yararlanmayanlara kıyasla devlet yardımından yararlanan firmaların yatırım harcamalarının 2023 yılında 1.781,51 dolar artış gösterdiği hesaplanmıştır. Yardımdan yararlanma durumunu temsil eden katsayının p olasılık değeri (2e-16), %5 olarak belirlenen hata payından (0,05) düşük olduğundan tahmin edilen katsayının %95 güven düzeyinde anlamlı olduğu sonucuna ulaşılmıştır. Benzer şekilde modelin anlamlılığına ilişkin p-değeri (2,2e-16) de %5 olarak belirlenen hata payından (0,05) düşük olduğundan tahmin edilen modelin de anlamlı olduğu sonucuna ulaşılmıştır.

Bununla birlikte ekipte yer alan araştırmacılardan biri, veri setinde yer alan firmaların farklı karakteristik özellikleri olduğunu ve bu özelliklerin de etki değerlendirme çalışmasında dikkate alınması gerektiği uyarısında bulunmuştur. Bunun üzerine ekip, IPW yöntemiyle ortalama müdahale etkisini yeniden hesaplamaya ve elde edilen sonucu ilk tahminin sonucu ile karşılaştırmaya karar vermiştir.

Ters Olasılıklı Ağırlıklandırma ile etki değerlendirme üç aşamadan oluşmaktadır:

1. Öncelikle eğilim skorları oluşturulmaktadır. Başka bir deyişle gözlem birimlerinin müdahaleye maruz kalma olasılıkları hesaplanmaktadır.
2. Sonrasında hesaplanan eğilim skorları IPW formülü ile bir ağırlığa dönüştürülmektedir.
3. Hesaplanan ters olasılıklı ağırlıklar son aşamada regresyon modeline dahil edilerek nedensel etki hesaplanmaktadır.

İlk aşamada bahsedilen eğilim skorunun hesaplanması için kullanılan farklı yöntemler olmasına rağmen yazında en sık kullanılan yöntemlerden biri lojistik regresyondur.⁷⁹ Bu çerçevede, öncelikle lojistik regresyon yönteminden yararlanılarak eğilim skorları oluşturulmuştur.

```

model_tahmin_yardım <- glm(yardım ~ hasılat + faaliyet_süresi,
                           family = binomial(link = "logit"),
                           data = yardım_verisi)
yardım_verisi_eğilim_skoru <- augment_columns(model_tahmin_yardım, yardım_verisi,
                                              type.predict = "response") %>%
  rename(eğilim = .fitted)

```

Uygulama kapsamında öncelikle **glm()** fonksiyonu kullanılarak firmaların yardımdan yararlanma durumuna ilişkin eğilim skorları hesaplanmaktadır. Bu çerçevede, yardımdan yararlanma durumu üzerinde belirleyici olan “hasılat” ve “faaliyet süresi”

⁷⁹ Diskriminant analizi, probit model ve kümeleme analizi.

değişkenleri kullanılmaktadır. Daha sonra hesaplanan eğilim skorları ana veri seti (yardım_verisi) içerisine yeni bir sütun olarak eklenmektedir. Bunun için `augment_columns()` fonksiyonu kullanılmaktadır.

```
yardım_verisi_eğilim_skoru %>%
  select(sıra, faaliyet_süresi, yardım, hibrit, yatırım, eğilim) %>%
  slice(1023, 816)
# A tibble: 2 × 6
  sıra faaliyet_süresi yardım hibrit yatırım eğilim
  <int>      <dbl> <int> <int>   <dbl> <dbl>
1 1018      28.1   0     1  3349. 0.215
2  348      34.4   1     0 4591. 0.218
```

Veri setinde yer alan firmalar için eğilim skorları incelendiğinde 1018'inci ve 348'inci sırada yer alan gözlemler dikkat çekmektedir. Zira, her iki firmanın da eğilim skoru aynı olmasına rağmen diğer değişkenler bakımından özellikleri birbirinden farklılaşmaktadır. Örneğin 1018'inci sıradaki firmanın faaliyet süresi 28 yıl iken, 348'inci sıradaki firmanın faaliyet süresi 34 yıldır. Ayrıca, kontrol grubunda yer alan firmada hibrit çalışma modeli uygulanırken, müdahale grubunda yer alan firmada hibrit çalışma söz konusu değildir. Aynı zamanda müdahale grubundaki firmanın yatırım harcaması tutarı 4.591 dolar iken kontrol grubundaki firmanın yatırım harcaması 3.349 dolardır. Bununla birlikte, her iki firmanın da devlet yardımından yararlanma ihtimali %22 olarak hesaplanmıştır.

Eğer eşleştirme yöntemiyle ortalama müdahale etkisi tahmin edilecek olsaydı beklenen ve beklenmeyen gözlemlerin eşleştirilmeye çalışılacağı bu çıkan olasılık değerinden anlaşılmaktadır. Bu sebeple IPW yerine sadece eşleştirme yapıldığında, doğru eşleştirmenin yapılıp yapılmadığının incelenmesinde fayda vardır. Bu sebeptendir ki yazında çeşitli şekillerde eşleştirme yöntemleri mevcuttur. Avantaj ve dezavantajlarına göre deneme-yanılma usulü yöntemlerin uygulanması gerektiği bundan ötürü Rehber'in çeşitli kısımlarında vurgulanmaktadır.

Söz konusu örnek açısından düşünüldüğünde faaliyet süresi birbirine benzeyen devlet yardımından yararlanmış ve yararlanmamış firmaların eşleştirilmesi gerekecektir. Bu aşamada belirtilen bu hususa benzer bir uygulama IPW ile yapılacaktır. Devlet yardımından yararlanma olasılığı düşük olan ve fiili olarak devlet yardımından yararlanmamış firmaların nispeten daha küçük bir ağırlığa sahip olması gerekmektedir. Bunun yanında, devlet yardımından yararlanma olasılığı düşük olan ancak fiili olarak devlet yardımından yararlanan firmaların da nispeten daha yüksek bir ağırlığa sahip olması gerekmektedir. Bu çerçevede, Ters Olasılıklı Ağırlıklandırma yöntemi veri setindeki gözlemlerin hesaplanan skorlarına bir ağırlık atanmasına imkân tanımaktadır.

Ters olasılıklı ağırlıkların hesaplanmasında kullanılabilecek farklı formüller bulunmakla birlikte, ortalama müdahale etkisinin hesaplanmasına yönelik olarak örnekte şu (6) no'lu formül kullanılmıştır⁸⁰:

$$IPW = \left(\frac{Müdahale}{Eğilim} \right) + \left(1 - \frac{Müdahale}{1 - Eğilim} \right) \quad (6)$$

⁸⁰ Austin, P. and Stuart, E. A. (2015) Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies, *Statistics in Medicine*, 34(28).

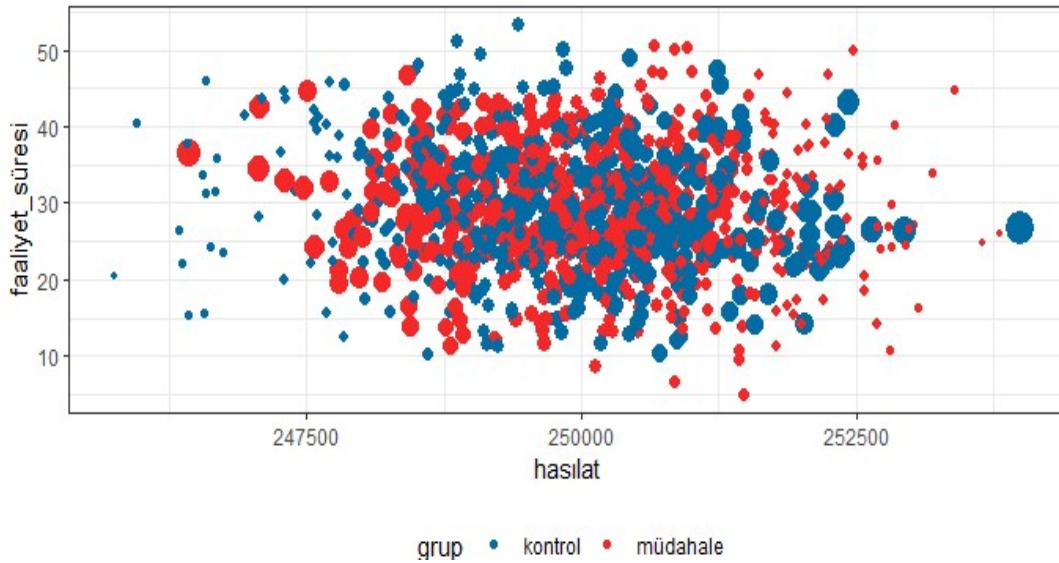
Belirtilen formül kullanılarak tahmin edilen eğilim skorları ağırlık değerine dönüştürülmüş ve veri setinin içerisine aşağıdaki kod bloğuyla eklenmiştir. Ardından veri seti içerisinde yer alan ve aynı eğilim skoruna sahip olan 1018'inci ve 348'inci firmaların ağırlıkları yeniden karşılaştırılmıştır.

```
yardım_verisi_ipw <- yardım_verisi_eğilim_skoru %>%
mutate(ipw = (yardım / eğilim) + ((1 - yardım) / (1 - eğilim)))

yardım_verisi_ipw %>%
select(sıra, faaliyet_süresi, yardım, hibrit, yatırım,
eğilim, ipw) %>%
slice(1018, 348)
# A tibble: 2 × 7
  sıra faaliyet_süresi yardım hibrit yatırım eğilim ipw
<int>      <dbl>    <int> <int>    <dbl> <dbl> <dbl>
1 1018        28.1      0     1 3349. 0.215 1.27
2  348        34.4      1     0 4591. 0.218 4.60
```

Bu yöntemin avantajı veri setinde yer alan tüm gözlemlere bir ters olasılık ağırlığı verilmesi ve buna bağlı olarak veri setinden herhangi bir gözlemin atılmasına gerek kalmamasıdır. Grafik 23'te yer alan her bir noktanın büyüklüğü, veri setindeki her bir gözlemin/firmanın kendi IPW değerini temsil etmektedir.

Grafik 23. Örnek 2 - Saçılım Grafiği



Hesaplanan IPW değerlerinden yararlanılarak karıştırıcıların sonuç üzerindeki etkisi tahmin edilmektedir. Bunun için regresyon modelinden yararlanılmıştır.

```
model_ipw <- lm(yatırım ~ yardım,
data = yardım_verisi_ipw,
weights = ipw)
tidy(model_ipw)
# A tibble: 2 × 5
  term      estimate std.error statistic p.value
<chr>      <dbl>    <dbl>    <dbl>    <dbl>
1 (Intercept) 6535.    53.0     123.      0
2 yardım      1079.    74.9     14.4 1.08e-43
```

Gerçekleştirilen tahmin sonucunda devlet yardımından yararlanmayan firmalara kıyasla devlet yardımından yararlanan firmaların ortalama yatırım harcamasının 1.079

dolar arttığı hesaplanmıştır. Yardım katsayısına ilişkin p-değeri (1.08e-43), hata payı olarak seçilen %5'ten küçük olduğu için tahminin istatistiki olarak %95 güven düzeyinde anlamlı olduğu sonucuna ulaşılmaktadır.

“ipw” Paketi ile Ters Olasılıklı Ağırlıkların Hesaplanması

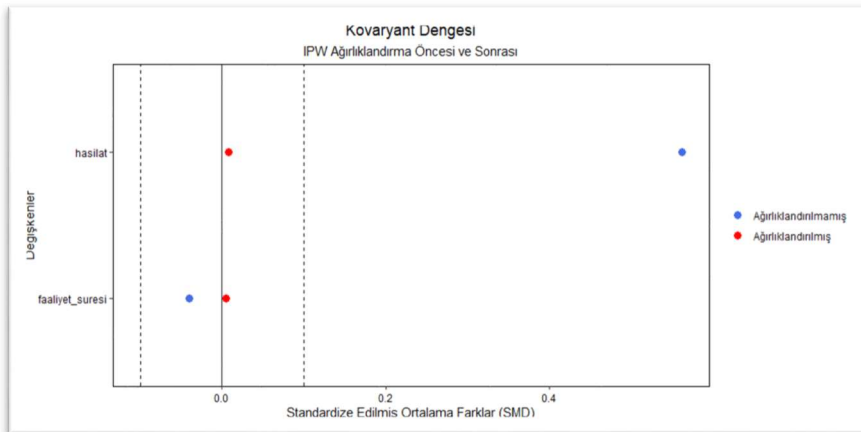
Buraya kadar anlatılan ters olasılıklı ağırlıkların R programında farklı bir paket yardımıyla tahmin edilmesi de mümkündür. Bu çerçevede, söz konusu firmalar için devlet yardımdan yararlanma olasılığını, lojistik regresyon yöntemiyle eğilim skorunu tahmin etmek ve sonrasında manuel olarak söz konusu eğilim skorlarına ağırlık formülü tanımlamak yerine “ipw” paketindeki “ipwpoint ()” fonksiyonu kullanılabilir.

```
install.packages("ipw")  
library(ipw)
```

Paket yüklendikten sonra “ipwpoint ()” fonksiyonu içerisindeki ilgili argümanlar tanımlanarak ters olasılıklı ağırlıkların tahmin işlemi gerçekleştirilmiştir.

```
weights_ipwpoint <- ipwpoint(  
  exposure = yardım,  
  family = "binomial",  
  link = "logit",  
  denominator = ~ hasılat + faaliyet_suresi,  
  data = as.data.frame(yardım_verisi))
```

Fonksiyon içerisinde yer alan “denominator” argümanı karıştırıcıları tanımlamak için kullanılmaktadır. Bu nedenle, söz konusu argüman için seçilecek değişkenlerin karıştırıcı değişken olmasına dikkat edilmelidir. Ayrıca, “exposure” argümanı da sonuç değişkenini tanımlamak için kullanılmaktadır.



Ters olasılıklar oluşturulduktan sonra eşleştirme yönteminde olduğu gibi bu yöntemde de seçilen değişkenler bakımından denge kontrolü yapılmalıdır. Yukarıda yer verilen grafikten de görüleceği üzere, seçilen değişkenler bakımından müdahale ve kontrol gruplarının SMD değerleri arasındaki fark sıfıra oldukça yaklaşmıştır. Bu çerçevede, söz konusu grupların daha karşılaştırılabilir hale geldiği değerlendirilmektedir. Bu bakımdan, eşleştirme yönteminde olduğu gibi IPW yönteminde de denge kontrolünün yapılması önem arz etmektedir.

Ters olasılıklı ağırlıklar söz konusu paket yardımıyla hesaplandıktan sonra ortalama müdahale etkisinin tahmini gerçekleştirilecektir.

```
yardım_verisi_ipwpoint <- yardım_verisi %>%
  mutate(ipw = weights_ipwpoint$ipw.weights)

model_ipwpoint <- lm(yatırım ~ yardım,
  data = yardım_verisi_ipwpoint,
  weights = ipw)

tidy(model_ipwpoint)
```

Dikkat edilecek olursa öncelikle tahmin edilen ters olasılıklı ağırlık değerleri yardım verisinin içerisine eklenerek “yardım_verisi_ipwpoint” adında yeni bir veri seti oluşturulmaktadır. Sonrasında veri seti kullanılarak devlet yardımından yararlanmanın firmaların yatırım harcamaları üzerindeki ortalama etkisi tahmin edilmektedir.

```
# A tibble: 2 × 5
  term      estimate std.error statistic p.value
<chr>      <dbl>    <dbl>    <dbl>    <dbl>
1 (Intercept)  6535.    53.0     123.0    0
2 yardım      1079.    74.9     14.4 1.08e-43
```

Gerçekleştirilen tahmin sonucunda devlet yardımından yararlanmayan firmalara kıyasla devlet yardımından yararlanan firmaların ortalama yatırım harcamasının 1.079 dolar arttığı ve yardım katsayısına ilişkin tahminin istatistiki olarak %95 düzeyinde anlamlı olduğu sonucuna ulaşılmıştır.

Son olarak bu başlık altında tahmin edilen üç modelin sonuçlarının karşılaştırmasına yer verilecektir.

```
texreg::screenreg(list(model_naive, model_ipw, model_ipwpoint),
  custom.model.names = c("Hatalı Model", "IPW", "IPW_Point"))
```

```
=====
      Hatalı   Model IPW   IPW_Point
-----
(Intercept) 6167.74 *** 6534.53 *** 6534.53 ***
            (52.61) (52.97) (52.97)
yardım      1781.51 *** 1079.44 *** 1079.44 ***
            (72.79) (74.88) (74.88)
-----
R^2          0.32      0.14      0.14
Adj. R^2      0.32      0.14      0.14
Num. obs.    1250      1250      1250
=====
*** p < 0.001; ** p < 0.01; * p < 0.05
```

Tabloda “Hatalı” olarak isimlendirilen sütun, araştırma ekibi tarafından başlangıçta tahmin edilen ve karıştırıcıları dikkate almayan modelin sonuçlarını yansıtmaktadır. Burada dikkat edilmesi gereken husus tahmin sonucu her ne kadar anlamlı çıkmış olsa da karıştırıcılar dikkate alınmadığı için tahmin edilen ortalama müdahale etkisinin sapmalı olmasıdır. Bununla birlikte, Ters Olasılıklı Ağırlıklandırma yöntemiyle gerçekleştirilen her iki modelin sonuçları da benzer şekilde istatistiksel olarak anlamlı çıkmıştır. Ayrıca her iki modelde de karıştırıcıların etkisi dikkate alınarak ortalama müdahale etkisi tahmin edilmiş ve karıştırıcılardan kaynaklanan sapma azaltılmıştır. Böylelikle karıştırıcıların etkisinin dikkate alınmadığı durumda 1.781,51

dolar olarak hesaplanan devlet yardımının ortalama etkisi karıştırıcılar dikkate alındığında 1.079,44 dolar olarak tahmin edilmiştir.

Örnek 3. Regresyon Süreksizliği Tasarımı

“İş Sağlığı ve Güvenliği” Yardım Programı

Madencilik sektöründe yer alan firmaların iş sağlığı ve güvenliği koşullarının iyileştirilmesine yönelik rehberlik faaliyetleri kapsamında “İş Sağlığı ve Güvenliği” adı altında bir yardım programı uygulanmıştır. İçeriğinde rehberlik faaliyetleri gerçekleştirilecek olan yardımın hedef alt grubu; çok tehlikeli sınıfta yer alan, 50 den az çalışanı bulunan, küçük ölçekli firmalar olarak belirlenmiştir. Yardım programına yapılan başvurular içerisinde kabul edilme şartı puanlama sistemine göre gerçekleştirilmiştir. Yardımın tasarımında baştan bir eşik değer (skor, puan vb.) belirlenmiştir. Bu sebeple tasarımı belli olan durumlarda verinin analizi için Rehber’de yer alan yöntemlerden Regresyon Süreksizliği Tasarımı (RST) kullanılarak etki değerlendirme analizi gerçekleştirilmiştir.

Yardımın değerlendirme kriterleri; son 2 yıl içerisinde; işyerinde iş sağlığı ve güvenliği hususunda tespit edilmiş olan mevzuata aykırılık sayısı, meydana gelmiş olan iş kazası sayısı, ramak kala olay sayısı ve işyerinde işin durdurulmasına yönelik idari yaptırım uygulanıp uygulanmaması göstergeleri olup her kriter üzerinden firmanın aldığı toplam puan, 100 tam puan üzerinden düşülmüştür.

Örnekte programın hedef alt grup özellikleri göz önünde bulundurularak, bu firmanın taşıdığı özelliklere yönelik sentetik *isgtutor.xlsx* adlı veri seti kullanılmıştır. Örneğin, her firmaya ait; çalışan sayısı, mevzuata aykırılık sayısı, iş kazası sayısı, ramak kala olay sayısı, idari yaptırım uygulanıp uygulanmamasına ilişkin verilen puanlar toplanmış ve sonucunda eşik değer olan 70 puanın altında kalan firmalar programa dahil edilmiş, diğerleri dahil edilmemiştir. Firmalara ait başlangıç puanları ve program sonrası puanlarındaki değişiklikler değerlendirilerek, programın etkisi ölçülmesi amaçlanmıştır.

RST, eşik değer etrafında değişen etkilerin incelenmesine olanak tanımaktadır. Belirli bir eşik değer varlığına dayalı olarak programa katılım koşulu olması, regresyon temelli süreksizlik analizini gerektirmektedir. Bu yöntem, program etkisinin eşik değer altında ve üstünde olduğu durumları tespit etmeyi hedeflemektedir. Eşik değer çevresindeki bireylerin birbirlerine en yakın özelliklere sahip olanlar olduğu ve programa dahil olması ve olmaması arasında bir fark olmadığı varsayılmaktadır.

Bireylerin hangi gruba atanacakları, eşik değer altında veya üstünde olma durumlarına göre belirlenmektedir. Süreksizliğin varlığı ilk aşamada grafiklerle görsel olarak belirlenebilmektedir. Etkinin hesaplanması için eşik değer etrafında hangi gözlemlerin kullanılacağını gösteren bant genişliği hesaplanmaktadır. Bu bant içinde, eşik değer altında ve üstünde yer alan bireyler için regresyon modelleri oluşturulmakta ve iki grup arasında bir süreksizlik olup olmadığı incelenmektedir. Bu süreksizlik, uygulanan programın etkisi olarak yorumlanmaktadır.

Analiz için gerekli R paketleri için gerekli paketler aşağıda yer almaktadır. Söz konusu paketlerden daha önce yüklenip çalıştırılmayanlar var ise öncelikle

install.packages(...) yazılarak yüklenmesi ve çalıştırılması gerekmektedir. Ardından "library" fonksiyonunun kullanılması uygun olacaktır.

```
library(tidyverse)
library(here)
library(broom)
library(scales)
library(modelsummary)
library(MatchIt)
library(readr)
library(rdrobust)
library(rddensity)
library(dplyr)
```

Çalışma R üzerinde yapılacağından "isgtutor.xlsx" adlı veri setinin aktarımı sağlanmış ve veri setine ilişkin Tablo 17 'de bilgiler sunulmuştur.

```
• rehberlikrs <- readxl::read_xlsx("../rehber_md/data/isgtutor.xlsx")
```

Tablo 17. Örnek 3- "İş Sağlığı ve Güvenliği" Analiz Değişkenleri

Değişken Adı	Değişken Açıklaması
<i>id</i>	İşletme numarası
<i>ilk_skor</i>	Program öncesi İşletmenin skoru (nicel değişken)
<i>bitis_skoru</i>	Program sonrası İşletme skoru (nicel değişken)
<i>rehberlik</i>	Programdan yararlanıp yararlanmama durumu (0,1 kategorik değişken)

Veri setinde yer alan ilk altı firmanın değerlerine verinin türlerini ve nasıl tutulduğunu anlamak amacıyla aşağıdaki şekilde bakılabilir.

```
head(rehberlikrs)
##   id ilk_skor rehberlik rehberlik_text bitis_skoru
## 1  1  92.4   FALSE    No tutor    78.1
## 2  2  72.8   FALSE    No tutor    58.2
## 3  3  53.7   TRUE     Tutor     62.0
## 4  4  98.3   FALSE    No tutor    67.5
## 5  5  69.7   TRUE     Tutor     54.1
## 6  6  68.1   TRUE     Tutor     60.1
```

Eşik değer, genellikle belirli bir puan üzerinden sınıflandırma yapılmasını sağlamaktadır. Bu tür eşik puanları belirlemek, genellikle bir karar noktasıdır ve bu nokta, neyin kabul edilebilir veya kabul edilemez olduğunu belirlemede kullanılmaktadır. Çalışmada belirlenen eşik değer 70 puandır. Sırasıyla, belirlenen eşik değere göre grup atamaları gerçekleştirilmiş ardından grup atamalarında hata olup olmadığı grafiklerle test edilmiştir (bkz. Grafik 24).

Hata olmasından kasıt 70'in altında puanı olan bir firmanın müdahale grubu yerine kontrol grubunda yer alması durumudur.

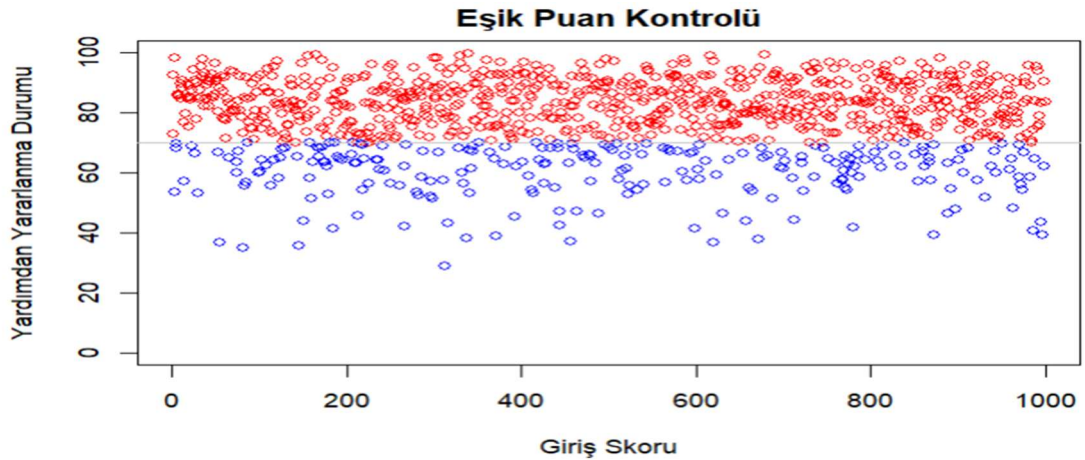
```
esikdeger <- 70
esikdeger <- kodu ile eşik değerin ataması gerçekleştirilmektedir.
plot(rehberlikrs$ilk_skor,
     col = ifelse(rehberlikrs$rehberlik == "TRUE", "cyan", "orange"),
     main = "Eşik Puan Kontrolü",
     xlab = "Giriş Skoru",
     ylab = "Yardımdan Yararlanma Durumu",
     ylim = c(0,100)) +
abline(h = esikdeger, col = "gray")
```

- **plot()**: Bu fonksiyon, *rehberlikrs* isimli veri çerçevesindeki *ilk_skor* sütununu x eksenine yerleştirerek bir grafik çizmektedir.
- **col**: Bu argüman, **plot()** fonksiyonunda kullanılan bir parametredir ve grafikteki veri noktalarının renklerini belirlemek için kullanılmaktadır.
- **ifelse()**: Bu fonksiyon, *rehberlikrs* veri çerçevesindeki *rehberlik* sütununu **"TRUE"** veya **"FALSE"** olarak değerlendirerek, **col** argümanı içinde renklendirme işlemi için kullanılmıştır.

Başka bir ifadeyle,

- **ifelse(rehberlikrs\$rehberlik == "TRUE", "blue", "red")**: Bu kısım, *rehberlikrs* veri çerçevesindeki *rehberlik* sütununu kontrol etmektedir. Eğer *rehberlik* sütunundaki değer **"TRUE"** yani yardımdan yararlanmış ise, o veri noktası mavi (**"blue"**) renkte, değilse kırmızı (**"red"**) renkte gösterilmektedir.
- **(main)** : Bu argüman grafik başlığını,
- **(xlab)** : Bu argüman x ekseninin etiketini,
- **(ylab)** : Bu argüman y ekseninin etiketini,
- **(ylim)** : Bu argüman y eksen aralığını belirlemektedir.
- **abline(h = esikdeger, col = "green")**: Bu kısım, **abline()** fonksiyonunu kullanarak grafiğe yatay bir çizgi eklemektedir. **h = esikdeger** argümanı yatay çizgiyi belirtir. Çizginin y ekseninde belirtilen *threshold* (eşik değeri) değerine denk gelecek şekilde oluşturulmasını sağlamaktadır. **col = "green"**, çizginin rengini yeşil olarak belirlemektedir.

Grafik 24. Örnek 3 - Dağılım Nokta Grafiği



Grafikte hata durumu tespit edilmemiştir. Ancak seçilen eşik değere göre gözlemlerin o eşik değeri etrafında dengeli dağılması beklenmektedir. Dengeli dağılmama durumları aşağıdaki durumlardan kaynaklı ortaya çıkabilir:

1. **Gözlemlerde Ani Değişiklikler:** Eşik değere yakın veya üstündeki firmaların davranışlarında veya performanslarında ani ve belirgin değişiklikler olmuş olabilir.
2. **Gözlemlerin Yoğunluğu:** Eşik değere yakın olan firmaların sayısının diğerlerine göre belirgin şekilde farklılık gösteriyor olması,
3. **Kovaryanslarda Anormallikler:** Eşik değerin üstündeki firmaların, başka değişkenlerle ilişkileri incelendiğinde diğer firmalardan farklı bir desen göstermesi veya bu değişkenlerle beklenmeyen ilişkiler içinde olması,
4. **Zaman Serisinde Anormallikler:** Programa katılım veya performansın zamanla anormal bir şekilde değişmesi veya eşik değerine yakın olduğu bir dönemde beklenmeyen bir değişiklik olması.

Dengeli dağılıma ilişkin veri görselleştirmesinin yanı sıra *McCrary Yoğunluk Testi* ile olası değişiklik/anormallik durumu test edilebilmektedir. Bu test, eşik değeri etrafında yayılan verinin yoğunluğunu test etmekte ve anormal bir yoğunluğa sahip olup olmadığını değerlendirerek eşik değerin etkisini sınamaktadır. Aşağıda yer alan R kodu, regresyon yoğunluk testini uygulamak içindir.

```
test_density <- rddensity(rehberlikrs$ilk_skor, c = 70)
summary(test_density)
```

- **test_density <- ...:** Elde edilen yoğunluk testi sonuçları **test_density** adlı bir nesneye aktarılmaktadır.
- **rddensity(rehberlikrs\$ilk_skor):** **rddensity()** fonksiyonunu kullanarak RST analizi için yoğunluk testini başlatmaktadır. “**rehberlikrs\$ilk_skor**” ifadesi, analizde kullanılacak olan “**running variable**⁸¹” (araştırmamızda eşik değeri) olarak belirtilen **ilk_skor** değişkenini temsil etmektedir. **c = 70**, eşik değeri belirtmektedir.
- **summary(test_density):** “**test_density**” nesnesinin özetini göstermek içindir. Bu özet, RST analizinin yoğunluk testi sonuçlarına dair sonuçları içermektedir.

```
##
## Manipulation testing using local polynomial density estimation.
##
## Number of obs =      1000
## Model =          unrestricted
## Kernel =         triangular
## BW method =       estimated
## VCE method =      jackknife
##
## c = 70           Left of c       Right of c
## Number of obs    237            763
## Eff. Number of obs 208            577
## Order est. (p)    2              2
## Order bias (q)    3              3
## BW est. (h)       22.444         19.966
##
## Method           T              P > |T|
## Robust           -0.5521        0.5809
## Warning in summary.CJMrddensity(test_density): There are repeated observations.
## Point estimates and standard errors have been adjusted. Use option
## massPoints=FALSE to suppress this feature.
##
## P-values of binomial tests (H0: p=0.5).
##
## Window Length / 2    <c    >=c    P>|T|
```

⁸¹ Bir modelin veya sistemin performansını değerlendirmek için kullanılan değişkenlerdir.

## 0.900	20	21	1.0000
## 1.800	36	38	0.9076
## 2.700	47	57	0.3776
## 3.600	62	78	0.2047
## 4.500	75	98	0.0941
## 5.400	87	120	0.0259
## 6.300	105	151	0.0048
## 7.200	119	173	0.0019
## 8.100	129	200	0.0001
## 9.000	136	227	0.0000

Hipotez:

H_0 : Eşik değeri etrafında manipülasyona yönelik anlamlı bir yoğunluk yoktur. (Yokluk Hipotezi)

H_a : Eşik değeri etrafında manipülasyona yönelik anlamlı bir yoğunluk vardır. (Alternatif Hipotez)

Sonuçta, p-olasılık değeri **0.5809** olduğu için %95 güven düzeyinde yokluk hipotezi red edilemeyecektir. Bu sebeple dağılımın dengeli dağıldığı ifade edilebilmektedir. Ayrıca t-istatistiğine dayanarak, eşik altında ve üstünde kalanlar arasında ortalamalar açısından farklılık olup olmadığına bakıldığında, eşik değeri altındaki 105. gözlemden sonra p olasılık değerine göre örnekleme bir farklılık olduğu görülmektedir. Ancak analizde bizim için önemli olan eşik değeri etrafında dağılan iki grup açısından etki değerlendirme yapılmasıdır. Bu sebeple seçilecek örnekleme bu değerleri göz önünde bulundurulmalıdır.

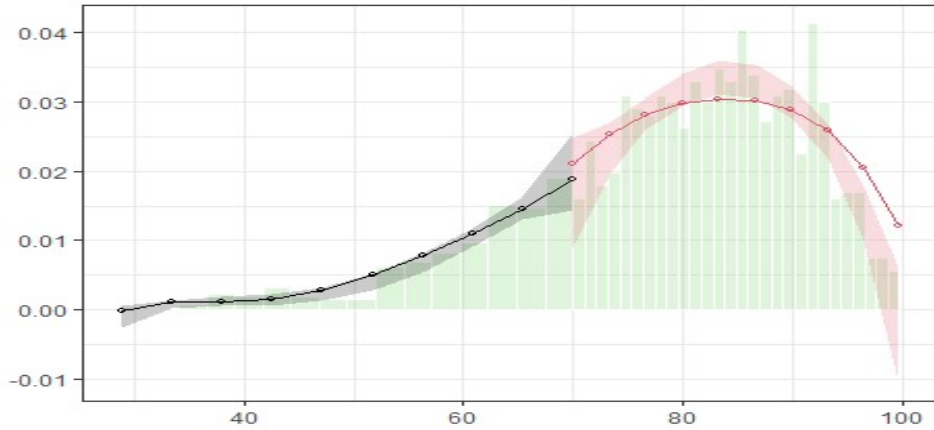
Yoğunluk testine ilişkin yukarıdaki sonuçların yanı sıra grafikte görselleştirmesi de sağlanmıştır. Verinin belirli bir noktada veya aralıkta yoğunlaşp yoğunlaşmadığını anlamak için grafik de önemli bir araçtır.

```
plot_density_test <- rdplotdensity(rdd = test_density,
                                   X = rehberlikrs$ilk_skor,
                                   type = "both")
```

Kodda, RST yoğunluk testi sonuçlarını görselleştirmek için `rdplotdensity()` fonksiyonu kullanılmaktadır.

- **rdd**: RST yoğunluk testi verilerinin aktarıldığı argümadır.
- **X**: RST analizinde kullanılan ilk puan verisidir. (`rehberlikrs$ilk_skor`).
- **type**: Çizim türünü belirten parametredir. **“both”** parametresi, hem kütleli yoğunluk fonksiyonunu (mass density function) hem de pencere yoğunluk fonksiyonunu (window density function) içeren bir grafik çizilmesini sağlamaktadır.

Grafik 25. Örnek 3 - Yoğunluk Grafiği



Grafik 25’te görüldüğü üzere güven aralıkları çakışmaktadır. Güven aralıklarının çakışması, veri setindeki gruplar arasında ortalamalar açısından istatistiksel olarak anlamlı bir farklılık bulunmadığını göstermektedir.

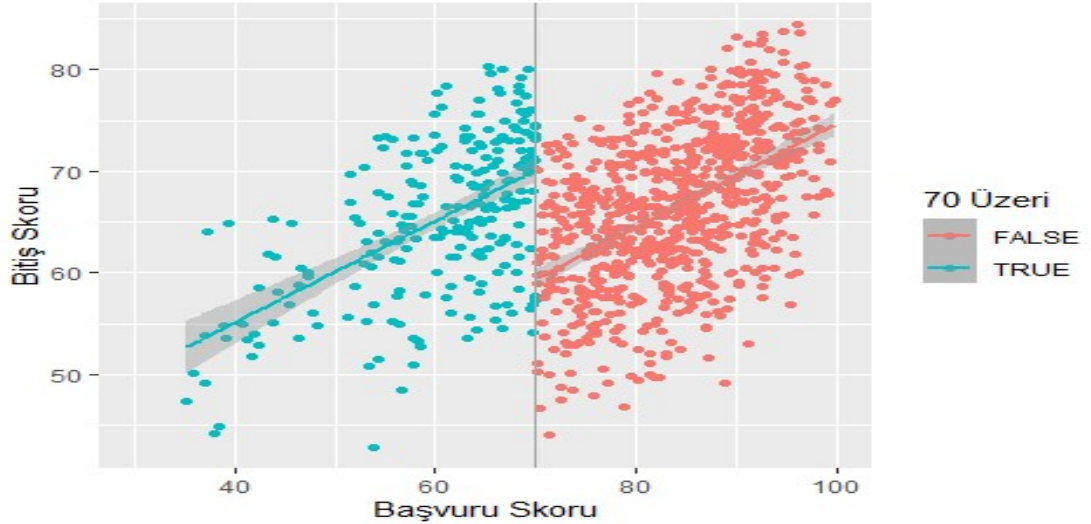
Başlangıçta grafiksel yöntemlerin kullanımı, potansiyel süreksizlikleri görsel olarak belirlemek için önemlidir. Programın etkisi değerlendirilirken, belirli bir eşik değeri etrafındaki bant genişliği hesaplanmaktadır. Bu bant genişliği, analize dahil edilecek örnekleme belirlemektedir. Ardından, iki grup olarak eşik değeri altında ve üstünde yer alan iki örneklem için regresyon modeli tahmin edilmektedir. Süreksizliğin kontrolüne yönelik grafik için R kodu aşağıda yer almaktadır:

```
ggplot(rehberlikrs, aes(x = ilk_skor, y = bitis_skoru, color = rehberlik)) +  
  geom_point() +  
  geom_smooth(data = filter(rehberlikrs, ilk_skor <= 70), method = "lm") +  
  geom_smooth(data = filter(rehberlikrs, ilk_skor > 70), method = "lm") +  
  labs(x = "Başvuru Skoru", y = "Bitiş Skoru", color = "70 Üzeri") +  
  xlim(30, 100) + geom_vline(xintercept = 70, col = "darkgray")
```

- **ggplot()** : Bu fonksiyon ile **aes()** argümanı içinde **ilk_skor** ve **bitis_skoru** değişkenleri x ve y eksenlerine atanır. **color** argümanı, **ilk_skor**’un 70’den büyük olup olmasına göre ayırım yapmaktadır.
- **geom_point()** : Bu fonksiyon ile saçılım grafiği (scatter plot) çizilmektedir. Her bir nokta, **ilk_skor** ve **bitis_skoru** değerlerinin düzlemdeki yerini göstermektedir.
- **geom_smooth()** : Bu fonksiyon, **filter()** ile alt gruplara ayrılarak, **method = "lm"** (lineer regresyon) ile her bir alt grup için ayrı ayrı regresyon eğrileri oluşturmaktadır. Bir alt grup **ilk_skor**’un 70’den küçük veya eşit olduğu durumu temsil ederken, diğer alt grup **ilk_skor**’un 70’den büyük olduğu durumu temsil etmektedir.
- **labs()** : Bu fonksiyon ile, x eksen etiketi “**Başvuru Skoru**” olarak, y eksen etiketi ise “**Bitiş Skoru**” olarak atanmaktadır. **color** etiketi “70 Üzeri” olarak belirtilmiştir.
- **scale_color_manual()**: Renk skalası manuel olarak ayarlanabilmektedir. Eşik değeri altında ve üstünde olan gözlemler için sırasıyla “**darkorange**” ve “**green**” renkleri atanmıştır.
- **xlim()** : Bu fonksiyon ile x eksen sınırları belirlenmiş olup sadece 30 ile 100 arasında puan alan gözlemlerin gösterimi sağlanmaktadır.

- **geom_vline()**: *xintercept* argümanı 70 eşik değeri olarak belirtilmiştir. X ekseninde 70 değerine denk gelen dikey bir çizgi (*xintercept*) çizdirilmiş ve bu çizgi “darkgray” renkle belirtilmiştir. Bu çizgi ile 70 eşik değerine göre iki grup arasındaki ayrım görsel olarak vurgulanmaktadır.
- Bu grafik; *ilk_skor* değişkenine göre, *bitis_skoru* değişkeninin nasıl değiştiğini göstermekte olup görülen regresyon eğrileri, *ilk_skor* değişkenine göre (değeri 70’in altında olanlar ve 70’in üstünde olanlar) iki farklı grup için çizilmiştir. Süreksizliğin görselleştirilmesi sağlanmıştır.

Grafik 26. Örnek 3 - Süreksizliğin Tespiti



Grafik 26’da yeşil noktalar programdan yararlananların başlangıç skorlarının 70 puan altında olduğunu temsil ederken; turuncu noktalar, programdan yararlanmayanların başlangıç skorlarının 70 puan üstünde kaldığını ifade etmektedir. Eğer programdan yararlanan grup programa dahil edilmeseydi, karşı-olgusallık durumu sebebiyle turuncu gruba ait regresyon doğrusunun yer aldığı noktadan doğrusal ilerleyeceği varsayılmaktadır. Fakat arada oluşan süreksizlik yani farklılık etki olarak tanımlanmakta ve yardıma atfedilmektedir.

Süreksizliğin varlığı tespit edildikten sonra, programın etkisini belirlemek adına, istatistiksel anlamlılığın test edilmesi ve süreksizlik miktarının belirlenmesi programın gerçek etkisini değerlendirmek için kritik önem taşımaktadır. Bu aşamada regresyon modeli (7) numaralı eşitlik ile gösterilmektedir:

$$\text{Bitiş Skoru} = \beta_0 + \beta_1 * \text{İlk_Skor} + \beta_2 * \text{Rehberlik} + \varepsilon \quad (7)$$

Eşik değeri içeren ± 10 bant aralığındaki gözlemler için regresyon modeli tahmin edilmiştir. Daha önce ifade edildiği üzere, eşleştirme yapmadan, iki grubun homojen dağıldığının varsayımı veri setinin daraltılması ile gerçekleştirilmektedir. Bu gözlemler içerisinde aykırı değer olma ihtimali yoktur.

Eşik değerin, +10 ve -10 bant genişliği içinde yer alan veriler üzerinden tahmin edilen regresyon koduna aşağıda yer verilmiştir:

```
model_bw_10 <- lm(bitis_skoru ~ ilk_skor + rehberlik,  
  data = filter(rehberlikrs,  
    ilk_skor >= -10 &  
    ilk_skor <= 10))
```

ilk_skor değeri -10 ile +10 arasında olan gözlemleri içeren bir alt küme oluşturmaktadır. Bu alt küme üzerinde *bitis_skoru* bağımlı değişkenini *ilk_skor* ve *rehberlik* bağımsız değişkenleriyle ilişkilendiren bir lineer regresyon modeli oluşturulmaktadır.

- *filter()*: Bu fonksiyon, *ilk_skor* değeri -10 ile +10 arasında olan gözlemleri *rehberlikrs* veri setinden seçmektedir.
- *lm()*: Bu fonksiyon, *filter()* fonksiyonu ile daraltılmış veri setindeki *bitis_skoru* değişkenini *ilk_skor* ve *rehberlik* değişkenleriyle ilişkilendiren bir regresyon modeli oluşturmaktadır.

Modele ilişkin; regresyon katsayıları, standart hataları, p-değerleri ve diğer istatistikleri içeren tablo aşağıdaki şekildedir;

```
tidy(model_bw_10)  
## # A tibble: 3 × 5  
##   term          estimate std.error statistic  p.value  
##   <chr>          <dbl>    <dbl>    <dbl>    <dbl>  
## 1 (Intercept)    33.18      8.64      3.83 1.43e-04  
## 2 ilk_skor        0.388     0.114      3.40 7.45e-4  
## 3 rehberlikTRUE    9.27      1.31      7.09 6.27e-12
```

- **(Intercept)**: Sabit değer **33.18** olarak hesaplanmış olup diğer bağımsız değişkenler sıfırken, bağımlı değişkenin beklenen değerinin 33.18 ortalama değer olduğu şeklinde yorumlanmaktadır.
- **ilk_skor**: Değişkenin katsayısı **0.388** olarak değer almıştır. *ilk_skor* değişkenindeki her bir birimlik artışın, bağımlı değişken üzerinde ortalama 0.388 birimlik bir artışa sebep olduğu görülmektedir.
- **rehberlikTRUE**: *rehberlik* değişkeninin kategorik bir değişken olduğunu ve burada “TRUE” durumunun etkisini göstermektedir. Bu durumda, *rehberlik* değişkeni **TRUE** (*programdan faydalanıldığında*) olduğunda, bağımlı değişkende ortalama **9.27** birimlik etki olduğunu göstermektedir.
- **Statistic**: Her bir katsayının t istatistiğini temsil etmektedir.
- **p-value**: p olasılık değeri, bağımsız değişkenlerin bağımlı değişken üzerindeki etkilerinin istatistiki olarak anlamlı olup olmadığını değerlendirmek için kullanılmaktadır.

Kurulan modele ilişkin özet istatistiklerin bir arada izlenebilmesi amacıyla; **modelssummary** paketi kullanılabilmektedir.

```
modelssummary(list("Bandwidth = 10" = model_bw_10))
```

- **list("Bandwidth = 10" = model_bw_10):** Bu kısım, bant aralığı 10 olan ve model_bw_10 olarak adlandırılan modele ilişkin sonuçları listelemektedir.
- **modelssummary(...):** Bu fonksiyon ile bir veya daha fazla modelin özet bilgilerini içeren bir tablo oluşturulmaktadır. Fonksiyonun içine, bir liste içinde bulunan modellerin adlarını ve nesnelerini içeren bir argüman yazılmaktadır.

Tablo 18. Örnek 3- Regresyon Süreksizliği Tasarımının Tahmin Sonuçları

	Bandwidth = 10
(Intercept)	60.377
	(0.752)
ilk_skor	0.388
	(0.114)
rehberlikTRUE	9.273
	(1.309)
Num.Obs.	404
R2	0.162
R2 Adj.	0.158
AIC	2663.5
BIC	2679.5
Log.Lik.	-1327.755
F	38.891
RMSE	6.47

Sonuç olarak, hem model hem katsayılar %95 güven düzeyinde anlamlı çıkmıştır. Uygulanan yardım programına katılan firmaların iş sağlığı ve güvenliği ile ilgili ortalama 9.273 birim puanlarını yükseltebileceği ve yardımın etkisinin pozitif olduğu şeklinde yorum yapılabilir.

5.102 SAYILI CUMHURBAŞKANLIĞI KARARNAMESİ KAPSAMINDA DEVLET YARDIMLARI HAKKINDA ETKİ DEĞERLENDİRMESİ

Önceki bölümlerde, etki değerlendirmenin ne anlama geldiği ve nasıl uygulanabileceği ile ilgili bilgiler paylaşılmıştır. Bu bölümde ilgili Kararname kapsamında teklif edilen, uygulanmasına devam edilen ve uygulanması tamamlanan veya sonlandırılan yardımlar hakkında hazırlanması gereken etki değerlendirme raporları hakkında bilgiler sunulmuştur.

Bilindiği gibi Kararnamenin usul ve esaslarını düzenleyen Yönetmelik'in "Etki değerlendirmenin amacı ve türleri" başlıklı 8 inci maddesinin 2 nci fıkrasında uygulayıcı kurumların yardımlar ile ilgili olarak Başkanlığa sunması gereken belgeler açıklanmıştır. Söz konusu belgeler arasında *Ön Etki Değerlendirme Raporu*, *Uygulama Dönemi Etki Değerlendirme Raporu* ve *Uygulama Sonrası Etki Değerlendirme Raporu* bulunmaktadır. Söz konusu raporların formatı Başkanlığın resmî internet sitesinde yayımlanmıştır.⁸²

5.1. Ön Etki Değerlendirme

Uygulama öncesi değerlendirmelerde ilk defa uygulanacak bir programın olası fayda veya riskleri için *yapısal ekonomik modeller (structural economic models)* ve çeşitli *simülasyon teknikleri* uygulanır. Bu yaklaşım, elde bireysel davranış ve piyasalar hakkındaki varsayımlar varken, amaçlanan politika değişikliklerinin sonuçlarını tahmin etmeye çalışır. Burada uygulayıcı kurumun programı tasarlarken bir teori üzerinden programı uygulamaya koyması beklenir. Ancak program öncesinde de takip edilen anahtar göstergeler için elde bulunan ya da toplanan veri mümkündür.

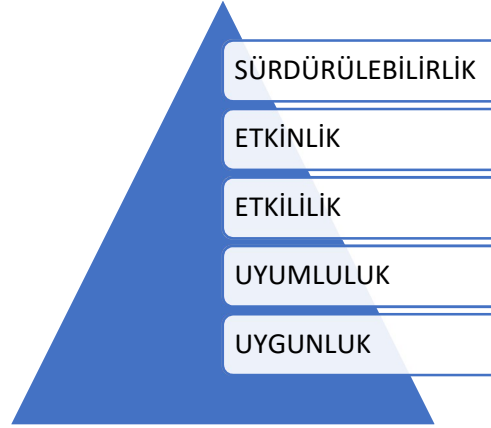
Program öncesi değerlendirmeye bir örnek olarak 2012 yılında Ranjeeta Thomas'ın araştırması verilebilir. Thomas araştırma kapsamında Nikaragua'da çocukların okullaşmasını artırmak ve koruyucu sağlık hizmetlerine erişimi iyileştirmek amaçlı uygulanan şartlı nakit transferi programı *Red de Proteccion Social*'in farklı düzeylerde sunulacak nakit transferleri senaryolarında ortaya çıkaracağı sonuçları tahmin etmek için bir program öncesi değerlendirme çalışması yapmıştır. Bu çalışmada, ilk olarak değerlendirmede kullanılan ekonomik model programın gerçekleşen çıktılarıyla kıyaslanıp uygunluğu test edilmiş, sonrasında simülasyonlar uygulanarak iki farklı politika senaryosunun olası etkileri tahmin edilmiştir. İlk senaryoda, mevcut program dahilinde nakit transferine hak kazanma koşulları değiştirilmeden sunulan nakit destek miktarı yüzde 75 oranında düşürülmüştür. Yapılan değerlendirme sonucunda, okullaşma ve koruyucu sağlık hizmeti kullanımındaki artışın daha düşük oranda nakit transferiyle de sağlanabileceği tespit edilmiştir. İkinci senaryoda ise, yalnızca gıda/sağlık yardımı ile bu yardıma hak kazanma koşulu olan sağlık eğitimlerine katılma şartı değiştirilmemiş, ancak okula gitme yardımı ve şartı kaldırılmıştır. Bu senaryonun değerlendirmesi ise küçük çocukların okullaşmasındaki artışın orijinal senaryoya yani programın mevcut uygulamasına kıyasla daha düşük gerçekleştiğini ortaya çıkarmıştır (Thomas, 2012).

⁸² Detaylar için bakınız <https://www.sbb.gov.tr/devlet-yardimlariyla-ilgili-belge-ve-yayinlar/>

Etki değerdendirmerinin aksine, bu yöntemler, uygulanan yardımların gerçek etkilerini ölçmek yerine potansiyel yani gelecekteki etkilerini simüle etmek için kullanılır. Bu tür değerdendirme, yardımın muhtemel etkilerini karşılaştırmak ve performans gösterge hedeflerini daha gerçekçi belirlemek, maliyetleri, getiri oranlarını ve diğer ekonomik parametreleri tahmin etmek için son derece faydalı olabilir.⁸³

Ex-ante kapsamında genellikle Şekil 8’deki kriterler bazında değerdendirmeler yapılmakta, bu kriterlere ilişkin standart sorular cevaplanmaya çalışılmaktadır.⁸⁴

Şekil 8. Değerdendirme Kriterleri



- Uygunluk: Müdahale doğru şeyleri yapıyor mu? Uygunluğun değerdendirilmesi, müdahalenin doğru olanı yapıp yapmadığının anlaşılması konusunda değerdendiren ekibe yardımcı olur. Müdahalenin hedefleri ve uygulanması ile yararlanıcı ve paydaşların ihtiyaçları ve önceliklerinin ne kadar uyumlu olduğunu değerdendirmeye yarar. Böylelikle hedef kitlenin müdahaleyi faydalı ve değerli bulup bulmadığını araştırmış olur.
- Uyumluluk: Özellikle diğer politikalar ile ne kadar örtüşüyor? Önerilen amaçlar ve faaliyetler, sosyo-ekonomik analizle mantıklı bir şekilde ilişkilendirilmiş mi? Bunlar birbirleriyle tutarlı mı ve bölgesel, ulusal, makro-bölgesel ve küresel düzeyde amaçlar ile ne kadar bağlantılı?
- Etkililik: Müdahale hedeflerine ulaşmayı başarıyor mu? Etkililik analizi; performans değerdendirmesi niteliğinde bir yöntem olup belirlenen hedeflere ne oranda ulaşıldığını anlamak için yapılır. Bir müdahalenin planlanan sonuçlara ulaşp ulaşmadığı, bunun hangi süreçle gerçekleştirildiği, bu süreçte hangi faktörlerin belirleyici olduğu ve istenmeyen etkilerin olup olmadığı hakkında fikir verebilir.
- Etkinlik: Kaynaklar ne kadar iyi kullanılıyor? Müdahalenin kaynakları ekonomik ve zaman yönden analiz edilmesidir. Burada ‘ekonomik’ kelimesinden kasıt girdilerin yani fonların, doğal kaynakların, zamanın, uzmanlığın çıktıya ve etkiye sonuca diğer benzer alternatiflere göre en maliyet-etkin şekilde ulaşmasıdır.

⁸³ Impact Evaluation in Practice v2 p.14

⁸⁴ CNMC_Methodological Guide On The Evaluation Of State Aid p14

- Sürdürülebilirlik: Beklenen sonuçlar zaman içinde kalıcı olacak mı? Yararlanıcı devlet yardımı almadığı takdirde ne olacak?

Sonuç olarak, bu değerlendirme bir yardımın uygulanmadan önceki potansiyel etkilerini değerlendirmek için önemli bir araçtır ve ön etki değerlendirme raporu, Kararname ve Yönetmelikte belirtildiği gibi uygulayıcı kurumlar tarafından teklif edilen yardımlar için hazırlanması gereken raporlardan biridir. Yönetmelik'in "Teklife konu Devlet yardımları" başlıklı 6 ncı maddesinin 1 inci fıkrasında,

"...İlk defa yürürlüğe konulmak üzere hazırlanan veya yürürlükte bulunmakla birlikte 4/1/1961 tarihli ve 213 sayılı Vergi Usul Kanunu gereğince tespit ve ilan edilen yeniden değerlendirme oranından fazla ödenek artışı veya gelir kaybına yol açan ya da Devlet yardımının türü, oranı, süresi, miktarı, yararlanıcıları gibi yardımın asli unsurlarında değişiklik öngören taslaklar; 7 nci maddeye uygun olarak teklif edilir."

hükmü yer almaktadır.

Anılan hükümden anlaşılacağı üzere, sadece ilk defa teklif edilen bir yardım için değil, aynı zamanda hâlihazırda uygulanan bir yardımın asli unsurlarında bir değişiklik teklif edilmesi durumunda da ön etki değerlendirme raporunun hazırlanması gerekmektedir.

Başlangıç aşamasında yapılan gözlemler, elde bulunan veri, uluslararası yükümlülükler, sektör temsilcilerinin görüşleri veya talepleri vb. nedenlerle yardım fikri ortaya çıkabilir. Bu tür durumlarda, söz konusu fikrin uygulayıcı kurum tarafından sistematik bir şekilde çalışılmasının ve muhtemel yardım teklifinin bu sistematik değerlendirme süzgecinden geçirilerek uygulanabilirliğine karar verilmesinin ardından teklifin hazırlanması gerekmektedir.

Ön etki değerlendirme raporunun, Yönetmeliğin "Tekliflerin yapılması" başlıklı 7 nci maddesinin 1 inci fıkrasının (b) bendinde geçen *Devlet Yardımı Teklif Formu* içeriğine göre değerlendirilmesinde yarar bulunmaktadır. Çünkü etki değerlendirme planlanırken ortaya çıkabilecek muhtemel sorulara teklif formunda yer verilmiştir. Ayrıca ön etki değerlendirme formunun "Etki Değerlendirme" başlıklı kısmında da Rehber de belirtilen hususlara göre detaylar sunulmalıdır.

Yardımanın etkin, etkili, verimli, sürdürülebilir ve tutarlı olacak şekilde tasarlanabilmesi ve anlaşılır olması için söz konusu soruların en uygun şekilde, riskleri ve varsayımları ortaya koyarak gerçeğe yakın cevaplanması gerekmektedir.

Tablo 19, bahse konu *Ön Etki Değerlendirme Raporunda* yer alan "Devlet Yardımı ve Etkileri" tablosudur. Uygulayıcı kurumlara yardım ile ilgili bir mantıksal çerçeve çizmek amaçlı oluşturulmuştur. Yardımın nasıl bir planlama ile yararlanıcılara ve ulusal ekonomiye fayda sağlayacağını gösteren özet bir tablodur.

Söz konusu tablonun ilk sütununda Yönetmeliğin "Devlet yardımı türleri" başlıklı 5 inci maddesi kapsamında yardımın türünün ve bütçesinin belirtilmesi beklenmektedir.

İkinci sütunda, yardımdan yararlanmaya ilişkin koşul ve kriterler ile uygulanacak faaliyetler ve zaman planlaması gibi yardımın uygulanma sürecine dair bilgilerin sunulması gerekmektedir.

Üçüncü sütunda, hedef kitle içerisinde başvurması beklenen yararlanıcı sayısı ya da hedeflenen başvuru sayısı ile bu yararlanıcılar için hedeflenen ortalama yardım tutarları yazılacaktır. Burada amaç, kaynağın dağıtımında ne tür bir plan yapıldığını anlamaya çalışmaktır. Örneğin, yardımı planlanırken öncesinde yapılacak bir anket çalışması yolu ile hedef kitle ve yapılması gereken ortalama yardım tutarları hakkında da bilgi edinilebilir.

Dördüncü sütun olan yardımın yararlanıcılara etkisi ise özellikle çıktı bazında öngörülen etkilerdir. Burada oran, miktar, imtiyaz ya da çeşitli sınıflandırmalar şeklinde öngörülere yer verilmelidir.

Son sütunda ise yararlanıcılara sağlanan yardımın, ilgili sektör veya alandaki sonuca katkısı belirtilecektir. Örneğin, yararlanıcının belirli bir ürünün üretimini artırmaya yönelik uygulanan bir yardımın, o ürünün toplam üretimine katkısının yazılması gerekmektedir.

Ön etki değerlendirme raporunda kurum tarafından öngörülen tahminlerin yazılması gerekmektedir. Mevcut ya da geçmiş veri kullanılarak veya hedef kitleyi ilgilendiren kurumlardan temin edilerek hesaplanabilir. Bu tahminlere yönelik varsayımlar ya da riskler ise tablonun altında yer alan iki satırda detaylandırılacaktır.

Tablo 19. Devlet Yardımı ve Etkileri Tablosu

YARDIM	YARDIM KAPSAMINDA ÖNGÖRÜLEN FAALİYETLER	YARDIMDAN YARARLANANLAR	YARDIMIN YARARLANICILARA ETKİSİ	YARDIMIN EKONOMİYE ETKİSİ
Devlet yardımının katkı türü ve toplam miktarı nedir? <i>Örneğin: Yıllık 1 Milyar TL tutarında nakdi yardım. Bu kapsamda ayrıca varsa diğer katkılar da belirtilecektir (Örn. uzman personel temini)</i>	Devlet yardımı kime, hangi koşullarda ve hangi amaçla sağlanacaktır? Yardımla ilgili yapılacak ilave faaliyetler (tanıtım, eğitim vb.) nelerdir? <i>Örneğin: Yardımla ilgili eğitim ve tanıtım yapılması, başvuru duyurusunun yapılması, başvuruların alınması, başvuruların değerlendirilerek uygun görülen firmalara yardımın sağlanması</i>	Devlet yardımı sağlanması öngörülen firma sayısı ve firma başına öngörülen yardım tutarı nedir? <i>Örneğin: 4.000 firmaya yıllık ... başvuru üzerinden ... TL tutarında yardım sağlanması</i>	Devlet yardımından yararlanan firmanın bu yardımdan elde edeceği fayda nedir? <i>Örneğin: Yardım sağlanan firmaların ... TL tutarında üretim veya ihracat yapması, dekar başına ... üretimini %... oranında artırması gibi</i>	Devlet yardımının ulusal ekonomiye sağlayacağı fayda nedir? <i>Örneğin: Türkiye'nin toplam üretiminin/ihracatının ... TL/Dolar ya da %..oranında artması</i>
Varsayımlar	Devlet yardımı teklifinin, hangi koşulların varlığı halinde yukarıda belirtilen etkiyi doğuracağı belirtilir. Söz konusu varsayımların belirlenmesinde idari ve hukuki altyapının mevcut durumu ile olası değişiklikler, ulusal ve küresel beklentiler vb. hususlar dikkate alınır.			
Riskler	Varsayımlardan hangilerinin gerçekleşmemesi halinde yardımın beklenen etkiyi doğurmayacağı ve nedenleri belirtilir.			

Başlangıç aşamasında dikkat edilmesi gereken bir diğer önemli husus da *Ön Etki Değerlendirme Raporunda* yer alan performans göstergeleridir (bkz. Tablo 20). Performans göstergesi, bir programın ya da politikanın başarılı olup olmadığına karar vermek için kullanılacak sonuç ölçüleridir/kriterleridir.

Tablo 20. Devlet Yardımının Performans Göstergeleri

<i>Performans Göstergesi</i>	<i>Mevcut Durum</i>	<i>yıl</i>	<i>yıl</i>	<i>yıl</i>
<i>Örneğin: Yardımdan yararlananların yıllık ilave ortalama ihracat artış oranı</i>	<i>0 veya 100 (ölçülmüş değer)</i>	<i>%20 veya 120</i>	<i>2. yıl/mevcut durum sütunu</i>	<i>3. yıl/mevcut durum sütunu</i>
<i>Örneğin: Yardımdan yararlananların ilave kalıcı istihdam oranı</i>	<i>0 veya 100 (ölçülmüş değer)</i>	<i>%20 veya 120</i>	<i>2. yıl/mevcut durum sütunu</i>	<i>3. yıl/mevcut durum sütunu</i>
<i>Örneğin: Yardımdan yararlananlarca ilave ekilen dönüm/dönüm başına rekolte</i>	<i>0 veya 100 (ölçülmüş değer)</i>	<i>%20 veya 120</i>	<i>2. yıl/mevcut durum sütunu</i>	<i>3. yıl/mevcut durum sütunu</i>

Tablo 20’de yer alan performans göstergeleri örnek olarak verilmiştir. Yardımın amacını ortaya koyan sonuç ve çıktıların performans göstergesi olarak yazılması beklenmektedir. Diğer bir deyişle, performans göstergelerinin, yardımın yararlanıcılara etkisini ölçecek şekilde belirlenmesi gerekir. Bu bilgi aynı zamanda Tablo 22’de “Yardımanın Yararlanıcılara Etkisi” sütununa da yazılacaktır. Öte yandan, performans göstergelerine dair tahminler yıl bazında talep edilecektir. Üç yıl örnek olarak konulmuştur. Temel göstergeler seçildiği zaman, programın başarısı bakımından ihtiyaç duyulan açık hedefler oluşturulmuş olur.

Uygulayıcı kurum lüzumu halinde daha uzun yıllar için de performans göstergelerinin yıllık etkisini tahmin edebilir. Bu aşama, seçilen ana göstergeler üzerinde programın beklenen etkilerinin belirlenmesi anlamına gelmektedir. Uygulamaya geçen yardımlar için yapılacak etki değerlendirme analizlerinde bu göstergeler kullanılarak sonuç değişkeni değerlendirilecektir. Performans göstergeleri Şekil 9’da belirtilen özellikleri taşımalıdır.

Şekil 9. Performans Göstergelerinde Olması Gereken Özellikler



Spesifik olması, performans göstergelerinin belirli bir hedefe veya amaca yönelik olması anlamına gelir. Örneğin, Yükseköğretim mezunu gençlerin istihdamını artırmak amacıyla yeni bir aktif işgücü piyasası programı uygulanmaya başlanmış ve program kapsamında istihdam edilen her genç için ödenmesi gereken sosyal güvenlik primi devlet tarafından karşılanmaktadır. Böyle bir programın amacı gençlerin istihdamını artırtmak

ise bu durumda gençlerin spesifik olarak tanımlandığı bir performans göstergesi belirlemek gerekmektedir.

Ölçülebilir olması, ilerlemeyi takip edebilmek için ölçülebilir bir hedefe sahip olmak demektir. Bu sayede hedefin ne kadarının gerçekleştiğini göstermektedir ve hedefe ne oranda yaklaşıldığının belirlenmesine yardımcı olmaktadır.

Atfedilebilir olması, yardımın amacından uzaklaşmadan ana faydayı sağladığı düşünülen uyumlu bir gösterge olması demektir.

Gerçekçi olması, doğrudan yardımın hedeflediği çıktıdan doğacak faaliyetlerle uyumlu olması demektir. Görülme olasılığı daha düşük hedefler belirlemek ilerleyen dönemlerde yardımın başarısını izleme konusunda zorluklar oluşturabilmektedir.

Son olarak hedeflenmiş olması gerekmektedir. Bu, belirlenen performans göstergesi için belirlenmiş periyodik zamanlar üzerinden takibi anlamına gelmektedir. Bu durum işin tamamlandığı anlamına gelmez; fakat belli zamanlarda da hedef kontrolünün yapılması, duruma ilişkin alınacak aksiyonları kolaylaştıracaktır.

Yardım teklifinin uygun bulunması ve yürürlüğe konulması durumunda, uygulayıcı kurum her yıl Başkanlığa sunacağı izleme raporlarıyla birlikte gerçekleşme bilgilerini iletacaktır. Bu nedenle ön etki değerlendirmede planlamanın ve yardımın yararlanıcılara etkisi konusunda öngörülerin sağlıklı yapılması önem arz etmektedir.

Ön Etki Değerlendirme Raporu ve yardım teklif formu hazırlanırken aşağıdaki adımların takip edilmesinde yarar bulunmaktadır.

Adım 1. Başlangıç ve Ekibin Oluşturulması

Etki değerlendirme, bir ekip tarafından yapılması gereken ve süreklilik gerektiren bir çalışmadır. Bu nedenle ön etki değerlendirmeye ilgili olarak sorumlu ve yetkili bir ekip oluşturulmasında yarar bulunmaktadır.

Ekipte yönetici (veya baş değerlendirici), araştırmacı ve analist olmalıdır. Ekibin ön etki değerlendirme raporu ve yardım teklif formu formatlarını dikkatlice inceleyerek, çalışma takvimi, veri toplama yöntemi, eğer anket gerekiyorsa hizmet alımı, etki değerlendirme yöntemi, paydaşlar vb. konusunda istişare ederek karar alması gerekmektedir.

Adım 2. Paydaşların Belirlenmesi

Yardım teklifi bir sorunun çözümü ve ihtiyacın karşılanması için gündeme gelmektedir. Bu nedenle, konuyla ilgili olabilecek tüm paydaşların katkılarının alınması gerekmektedir. Paydaşlar belirlenirken her bir paydaşın niçin tasarlanan yardımla ilgili bir paydaş olduğunun teklif formunda açıklanması gerekir. Bu yaklaşım, konu itibarıyla bazı önemli paydaşların unutulması riskini de azaltabilecektir.

Adım 3. Paydaşlarla İstişare

Esasında süreç içerisinde paydaşlarla başka istişareler de yapılacaktır. Ancak ilk istişare planlama açısından önemlidir. Bu istişarede, sorunları, sorunların olası nedenlerini, yardım fikrini, yardımın tasarımını, beklenen çıktı ve sonuçlarını, hedeflerini, hedef kitlesini, mali yükünü, saha çalışması yöntemini, etki değerlendirme yöntemini, performans göstergelerini vb. birçok hususu görüşebilecektir.

Paydaşlarla yapılacak ilk istişare yardımı teklif etmeyi düşünen kuruma birçok konuda fikir verecek, paydaşların desteğini sağlayacak ve olabilecek mükerrerlikleri engelleyebilecektir. Paydaş istişaresinin daha etkili olabilmesi için yardım fikri olan kurumun istişareler öncesinde tüm konularda hazırlık yapmış ve taslak planını hazırlamış olması gerekir. Hazırlık sürecinde izleyeceği yöntemin tüm detayları hakkında hazırlıklı olması, bunlarla ilgili ön bilgilendirme yaparak istişareyi yapması daha uygun olacaktır.

Adım 4. Etki Değerlendirme Yönteminin Belirlenmesi

Önceki bölümlerde de açıklandığı gibi yardımın özelliği ve veri gibi çeşitli kriterler gözetilerek etki değerlendirme yöntemine karar verilmesi gerekmektedir. Öte yandan, etkinin tespit edilmesi ve uygulama sonrası olabilecek farklılıkların açıklanarak uygun tedbirlerin geliştirilebilmesi açısından karma yöntemin (nicel ve nitel yöntemin birlikte) kullanılmasında yarar bulunmaktadır.

Adım 5. Veri Toplama Aracı

Etki değerlendirmede verinin olabildiğince önce toplanmaya başlanması uygun olacaktır. Yardımın yararlanıcılar üzerindeki etkisinin ölçülmesi hedeflenmektedir. Bu nedenle yararlanıcılar ve benzer özellikte yararlanmayanlarla ilgili uygulama öncesi, uygulama sırası ve uygulama sonrası veri toplamak gerekecektir.

Günümüzde birçok veri kaynağı kullanılmaktadır. Bunlar arasında anketler, odak grubu toplantıları, idari kayıtlar, istatistikler vb. ön planda yer almaktadır. Bunlardan uygun olanlarının araştırılarak belirlenmesinde yarar bulunmaktadır. Anket yönteminin seçilmesi durumunda hizmet alımı için teknik şartnamenin hazırlanarak ihale yapılması gerekebilecektir. İdari kayıtların kullanılmasının öngörülmesi halinde ise, veri başka kurum veya kurumlardan temin edilecekse ilgili kurum veya kurumlarla ön görüşme yapılarak verinin toplanabileceğinin garanti altına alınması gerekir. Verinin ilgili kurumlardan sağlanamayacağının anlaşılması durumunda en pratik şekilde toplanabileceği yol seçilerek devam edilmesi uygun olacaktır.

Adım 6. Soruların Tasarlanması

Ön etki değerlendirmede amaç, sorunları ve nedenlerini, uygulanabilecek yardımı, hedef kitleyi ve bu kitleye muhtemel etkilerini tespit etmek için veri toplayıp, bu veriden yararlanarak oluşturulacak yardımın muhtemel etkisini tahmin etmektir. Veri bu kapsamda oluşturulacak araştırma sorularına cevaben aranacaktır. Özellikle anket yönteminin kullanıldığı durumlarda soruların belirlenmesi önemli bir aşamadır. Bu nedenle sorularla ilgili çalışmaların etki değerlendirmenin amacı gözetilerek titizlikle yürütülmesinde fayda vardır.

Adım 7. Anketörlerin Eğitimi

Anket veya mülakat ile veri toplanan durumlarda anketörlere önemli görev ve sorumluluklar düşmektedir. Verinin amacına uygun bir şekilde toplanabilmesi için anketörlerin anket yapılanlara ilave sorular sorması veya açıklamalar yapması gerekebilmektedir. Özellikle nitel veri toplanan durumlarda anketörlere daha büyük görevler düşebilmektedir. Bunlara ek olarak, anketörlerin toplanan veriyi muhafaza etmesi, etik kurallara sadık davranması vb. sorumlulukları bulunmaktadır. Bu ve benzeri durumlarla ilgili olarak anketörlerin önceden eğitilmesi, veri toplamada karşılaşılabilecek risklere karşı önemli bir tedbir olabilecektir.

Adım 8. Örneklem Seçimi

Her yardımın bir hedef kitlesi bulunmaktadır. Hedef kitlenin tamamında etkiyi ölçmek maliyetli olacağından kitleyi temsil edecek bir örneklem seçilerek ölçüm yapılması gerekecektir. Bazı etki değerlendirme yöntemlerinde bunu en başta yapıp daha sonra grubun bir kısmına yardım sağlanıp bir kısmına sağlanmadığından bunu sağlamak kolaydır. Ancak bazı yöntemlerde ise bu başlangıç aşamasında imkânsızdır.

Örneklemin sayı olarak ne kadar olması gerektiği, etki değerlendirmenin türüne göre farklılık arz edebilmektedir. Örneklemini sayıca fazla tutmak genel etkinin görülmesi açısından daha sağlıklı sonuçlar vermektedir. Ancak zaman ve maliyet önemli bir kısıtlayıcı faktör olacağından örneklem seçiminde bu hususun da dikkate alınmasında yarar bulunmaktadır.⁸⁵

Sonuç olarak, ön etki değerlendirmede anket durum ve ihtiyaç tespiti niteliği de taşıdığından, paydaşlarla yapılan istişarelerin de katkısıyla hedef kitleden örneklem seçilmelidir.

Adım 9. Veri Toplama

Örneklem seçiminin yapılmasının ardından veri toplanabilir. Verinin toplanması aşamasında anket yapılıyorsa anketörlerin zorluk yaşaması veya istenilen şekilde veri toplanamaması durumunda etki değerlendirmeden sorumlu ekibin hızlı bir şekilde sorunun çözümü için bir yol bulması uygun olacaktır. Bu nedenle veri toplanması süreci sıkı takip edilmelidir.

Adım 10. Veri Analizi

Toplanan veri analiz edilirken, hedeflenen amaç ve gereklilikler göz önünde bulundurulmalıdır. Yardıma ihtiyaç olup olmadığını, nasıl bir yardıma ihtiyaç olduğunu, odaklanılması gereken alanı, hedef kitleyi, performans göstergelerini ve *Devlet Yardımı ve Etkileri Tablosunun* doldurulması için gerekli olan veriye odaklanılmalıdır.

Ayrıca, toplanan veri analiz edilerek ve uygun etki değerlendirme yöntemi kullanılarak performans göstergelerinde önümüzdeki yıllara ilişkin hedefler oluşturulmalıdır.

⁸⁵ Örneklem seçiminde dikkat edilmesi gereken hususlar hakkında detaylı bilgi için: White ve Raitzer (2017)

Adım 11. Ön Rapor

Veri yukarıda belirtilen çerçevede analiz edildikten sonra ön raporu oluşturma aşamasına geçilebilecektir. Ön raporda etki değerlendirmenin sadece sonuçları değil, yukarıda sayılan adımlar çerçevesinde yapılan işlemlerin de özetlenmesi gerekmektedir. Böylece bulguların dayanakları ve yapılan çalışmaların çerçevesi görülebilecektir.

Adım 12. Ön Rapor Hakkında Paydaşların Görüşünün Alınması

Ön rapor, taslak rapor niteliğinde olup paydaşlarla görüşülerek olabilecek görüşlerin alınması ve nihai hale getirilmesinde yarar bulunmaktadır. Bu istişarenin sebebi, bulguların sahiplenilmesi ve çalışmaya olabilecek genel katkıların alınmasıdır. Bu sürecin hızlı ve sağlıklı bir şekilde yapılabilmesi için planlamanın önceden yapılarak paydaşların bilgilendirilmesinde yarar bulunmaktadır.

Adım 13. Nihai Rapor

Paydaşlarla ön rapor hakkında yapılan istişareler sonrasında olabilecek ekleme ve değişikliklerin ardından rapora son şekli verilecektir.

Adım 14. Devlet Yardımı Teklif Formunun ve Ön Etki Değerlendirme Raporunun doldurulması ve belgelerin Başkanlığa iletilmesi

Nihai rapor ile yardıma ihtiyaç olduğu sonucuna ulaşırsa, ilgili kurum tarafından *Devlet Yardımı Teklif Formu* ve *Ön Etki Değerlendirme Raporu* yapılan çalışmaya uygun olarak eksiksiz doldurulur. Söz konusu belgeler doldurulduktan sonra Cumhurbaşkanı karar taslağı ile birlikte Başkanlığa resmî yazıyla iletilmelidir. Yönetmelik'in 7 nci maddesinin 2 nci fıkrasına göre, Cumhurbaşkanı karar taslağının, yardımın amacı, kapsamı, unsurları, yardım oranı ve/veya miktarı, süresi, uygulayıcı kurumun yetki ve sorumlulukları gibi yardıma ilişkin asli unsurları içermesi esastır.

Ön Etki Değerlendirme Raporuna dayanak teşkil eden nihai rapor ve hesaplamada kullanılan verinin de yazı ekinde gönderilmesi beklenmektedir.

5.2. Uygulama Dönemi Etki Değerlendirme

Teklif edilen bir yardımın onaylanıp uygulamaya konulmasının ardından geçen her takvim yılı için *Uygulama Dönemi Etki Değerlendirme Raporu* hazırlanarak takip eden yılın ilk üç ayı sonunda yani Mart ayı sonuna kadar *Yıllık İzleme Raporu* ekinde Başkanlığa sunulur.

Uygulama döneminde yapılacak etki değerlendirmede amaç, uygulamanın bir etki doğurup doğurmadığını anlamak ve yardım uygulanmadan önce yapılan ön etki değerlendirme sonuçları ile mukayese etmektir. Uygulayıcı kurumların değerlendirmeyi yapabilmeleri için üç aylık süreleri bulunduğundan bu sürenin etkin bir şekilde kullanılmasında yarar bulunmaktadır.

Bu süreçte toplanan bilgiler, mevcut en iyi kanıtlarla uygulamaya rehberlik edebilir ve yardımlarda yapılacak revizelerde kullanılabilir. Bulunan sonuçlara göre uygulama aynı şekilde devam ettirilebilir, sonuçlandırılabilir veya revize edilerek etkinliği artırılabilir.

Yardım hakkında kanıtlara dayalı hareket etmek uygulayıcılar için doğru yolda olunduğunun bir göstergesi olacaktır. Uygulama dönemi etki değerlendirme için Başkanlığın resmî internet sayfasında yayımlanan rapor formatı bir önceki takvim yılına ait gerçekleşme verisiyle hazırlanacaktır.

Ön etki değerlendirme raporunda kullanılan yöntemle belirlenen performans göstergeleri seçilen örneklem üzerinden yeniden hesaplanır. Söz konusu veri yardımın uygulandığı tüm süre boyunca veya belirli zaman dilimlerinde toplanabilir. Bu detayların oluşturulan ekip⁸⁶ tarafından raporun sonuç ve değerlendirme kısmında belirtilmesi faydalı olacaktır. Ayrıca sunulan ilk yıllık izleme raporunda performans göstergelerinin gerçekleşen değerlerine ilişkin bildirim yapılmasının yanı sıra izleyen yıllar için eklenen yeni performans hedefleri de aynı rapora eklenmesi mümkündür.

Amaç uygulama sırasında faaliyetlerin etkisinin tespit edilmeye çalışılmasıdır. Raporda yer alacak yönlendirici değerlendirmeler, bir yardımın nasıl çalıştığını ve ilerlediğini gösterir ve öngörülemeyen ve amaçlanmayan sonuçları ve zorlukları vurgulayabilir. Bu nedenle bu bilgi ve değerlendirmelerinde bu raporda belirtilmesi beklenmektedir.

Yapılacak 3 aylık çalışmada sürenin etkin bir şekilde kullanılabilmesi için Ön Etki Değerlendirme Raporu hazırlanırken oluşturulan adımlara göre hareket edilmesi faydalı olacaktır. Özellikle veri analizine kadar olan adımlarla ilgili hem belirlenmiş hem de yeniden belirlenecek konular göz önünde bulundurularak eksiksiz hareket edilirse sonraki adımların gerçekleştirilmesi daha pratik ve sistematik olacaktır.

5.3. Uygulama Sonrası Etki Değerlendirme

Uygulama sonrası etki değerlendirme izlenecek yol itibarıyla uygulama dönemi etki değerlendirmeye benzetilmektedir. Aradaki temel fark, uygulama dönemi etki değerlendirme yıllık olarak yapılıp bir önceki yılın etkisini araştırırken, uygulama sonrası etki değerlendirme, yardımın uygulamaya başlanmasından tamamlanıncaya kadar olan dönemdeki toplam etkinin araştırılmasıdır. Uygulama sonrası etki değerlendirme raporu hazırlanırken de uygulama dönemi etki değerlendirme raporunun hazırlanmasındaki sürecin takip edilmesi yararlı olacaktır.

Uygulama sonrası etki değerlendirme yardımın nihai sonuçlarına ve etkilerine geriye dönük bir bakış sağlar. Başlangıç (temel), orta (uygulama) ve son (çıktı ve sonuç) şeklinde bir olay örgüsü oluşturur.

⁸⁶ Personelin yetkinliklerinin olması durumunda görevlendirilecek personel birden fazla rolü üstlenebileceğinden üç kişiden az bir ekip de olabilir. Uygun olması halinde ön etki değerlendirmeyi yapan ekibin bu değerlendirmeyi yapması da sağlanabilir. Böylece birikimden istifade edilebilecektir.

Yardımanın uygulanması tamamlandıktan sonra, bir nihai rapor, hedefe ulaşıp ulaşılmadığının anlaşılmasına yardımcı olur. Sonuçları ve etkiyi görmek için yapılan bir dizi faaliyetin tamamının değerlendirilmesine imkân sağlar ve gelecekte uygulanacak yardımlar için görüş ve önerilere zemin hazırlar. Ayrıca çıkan sonuçlar yardımdan yararlananlar ve temsil ettikleri bölge, sektör ile bütçeyi takip eden birimler için de faydalı ve ikna edici olacaktır.

SONUÇ

Başarılı etki değerlendirme, yararlanıcıların müdahaleye yanıt verdiği mekanizmaları tam olarak değerlendirebilmelidir. Bu mekanizmalar, bağımsız mekanizmalar olmayıp diğer mevcut politikalarla, sektörlerle bağlantılar içerebilir. Mevcut politikalarla bağlantı, politika yapıcılarının bir müdahalenin etkilerini anlamalarına yardımcı olup benzer müdahalelerin gelecekteki etki değerlendirme sürecine rehberlik etmesi açısından önemlidir. İyi tasarlanmış bir etki değerlendirmenin faydaları bu nedenle uzun vadelidir ve dış geçerliliği vardır. Bu sayede bulgular başka durum ve zamanlarda benzer kitleleri kapsayacak şekilde genelleştirilebilir. Bunun, müdahalenin ölçeklenebilirliği ve tekrarlanması açısından önemli sonuçları olacaktır. Ancak zaman içindeki değişiklikler, farklı politika kararları, müdahale grubu seçimleri vb. etkenlerin yardımın etkisini değiştirebileceği de göz önünde tutulmalıdır.

Uygulamanın ayrıntılı olarak raporlanması, büyük ölçüde bir müdahalenin ne kadar iyi uygulandığına, izlendiğine ve değerlendirildiğine bağlı olacaktır. Yardımın nasıl uygulandığına dair eksiksiz bilgiye sahip olmanın önemi, değerlendirme sonuçları sınırlı veya olumsuz bir etki gösterdiğinde ortaya çıkacaktır. Bu sayede uygulamadaki sorunları yardımın tasarımındaki sorunlardan ayırmamız kolaylaşacaktır.

Özetle değerlendirme sonuçlarını doğru bir şekilde yorumlayabilmek için etki değerlendirme, sürekliliği olan bir izleme ve değerlendirme çerçevesine oturtulmalıdır. Bu kapsamda Rehberin etki değerlendirme konusunda uygulayıcı kurumlara faydalı olması amaçlanmıştır ve daha da kapsamlı bir bakış açısı sunmak için güncellenerek ikinci versiyonu hazırlanmıştır.

Bu versiyonda alana ilişkin örnek uygulamalar, katkı sağlayacağı düşünülen öneriler aktarılmaya çalışılmıştır. Etki değerlendirme ile alakalı hazırlanacak raporlarda başlangıç noktası bir kaynak olarak hizmet etmesi amaçlanmıştır. Rehber hakkında her türlü görüş ve öneriler dygm_iard@sbb.gov.tr adresine iletilir.

KAYNAKÇA

- Abadie, A. (2005). "Semiparametric Difference-in-Differences Estimators." Review of Economic Studies, 72(1)
- Abadie, A. and Imbens, G.W. (2006), Large Sample Properties of Matching Estimators for Average Treatment Effects. Econometrica
- Abadie, A., Diamond A, Hainmueller J (2010) Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program. Journal of the American Statistical Association, <https://web.stanford.edu/~jhain/Paper/JASA2010.pdf>.
- Abadie, A., (2021) Using Synthetic Controls: Feasibility, Data Requirements, and Methodological Aspects, <https://pubs.aeaweb.org/doi/pdfplus/10.1257/jel.20191450>.
- Adato, Michelle (2011) "Combining Quantitative and Qualitative Methods for Program Monitoring and Evaluation: Why Are MixedMethod Designs Best?", Prem Notes, May 2011, Number 9, <https://openknowledge.worldbank.org/server/api/core/bitstreams/ccf8e788-6d8e-58cd-8e28-178bc69f77e8/content>.
- Aditya, K., & Subash, S. (2019). Propensity Score Matching. V. Nikam, A. Jhahria, & S. Pal içinde, *Quantitative Methods For Social Sciences* (s. 199-210).
- Andrews, D. W. K. (2003), End-Of-Sample Instability Tests, Econometrica, Vol. 75, No. 6, <https://users.ssc.wisc.edu/~behansen/718/Andrews2003b.pdf>.
- Angrist, J. D., & Pischke, J.-S. (2009) "Mostly harmless econometrics: An empiricist's companion", Princeton University Press
- Austin, P.C. (2009), Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. Statist. Med., 28
- Austin, P.C. (2011), Optimal caliper widths for propensity-score matching when estimating differences in means and differences in proportions in observational studies. Pharmaceut. Statist., 10
- Austin, P. C. (2011). An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies. Multivariate Behavioral Research, 46(3), 399–424
- Austin, P. and Stuart, E. A. (2015) Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies, Statistics in Medicine, 34(28).
- Austin, P. (2024). Propensity Score Analysis With Baseline and Follow-Up Measurements of the Outcome Variable. <https://onlinelibrary.wiley.com/doi/10.1002/pst.2436> adresinden alındı.
- Arkhangelsky, D. ve diğerleri (2019), Synthetic Difference in Differences, https://www.researchgate.net/publication/329945694_Synthetic_Difference_in_Differences.
- Ayanoğlu, Kamil, Düzyol, M.Cüneyd, İlater, Niyazi ve Yılmaz, Cevdet (1996) "Kamu Yatırım Projelerinin Planlanması ve Analizi" İktisadi Sektörler ve Koordinasyon Genel Müdürlüğü, Devlet Planlama Teşkilatı.
- Aydos, M., Toker, Ç. D. (2021). Eğitim Seviyelerine Göre Genç İşsizliğinin Ekonometrik Analizi: Türkiye Örneği. İzmir İktisat Dergisi 36 (1). <https://dergipark.org.tr/en/download/article-file/1020796>

- BIS (2011) "Guidance on Evaluating the Impact of Interventions on Business", August 2011, Department for Business, Innovation and Skills, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/212318/11-1085-guidance-evaluating-interventions-on-business.pdf.
- Bertrand, M., Duflo, E. ve Mullainathan, S., (2004), "How Much Should We Trust DiD Estimates?", *The Quarterly Journal of Economics*, 119(1), 249–275.
- Bilinski, Alyssa and Laura A. Hatfield, "Seeking evidence of absence: Reconsidering tests of model assumptions," arXiv:1805.03273 [stat], May 2018.
- Bjarkefur, Kristoffer, Luiza Cardoso de Andrade, Benjamin Daniels, and Maria Ruth Jones. 2021. *Development Research in Practice: The DIME Analytics Data Handbook*. Washington, DC: World Bank. <https://documents1.worldbank.org/curated/en/176451623903242202/pdf/Development-Research-in-Practice-The-DIME-Analytics-Data-Handbook.pdf>
- Borusyak, Kirill and Jann Spiess, "Revisiting Event Study Designs: Robust and Efficient Estimation," *arXiv:2108.12419 [econ]*, 2021. Ek olarak bkz. Gardner, John, "Two-stage differences in differences," *Working Paper*, 2021.
- Caetano, C. ve Callaway, B., (2024), "Difference-in-Differences when Parallel Trends Holds Conditional on Covariates," Papers 2406.15288, arXiv.org.
- Caliendo, M., & Kopeinig, S. (2008). Some practical guidance for the implementation of propensity score matching. *Journal of Economic Surveys*, 22(1), 31–72.
- Callaway, Brantly and Pedro H. C. Sant'Anna, "Difference-in-Differences with multiple time periods," *Journal of Econometrics*, 2021, 225 (2), 200–230.
- Card, D., & Krueger, A. B., "Minimum Wages and Employment: A Case Study of the Fast Food Industry in New Jersey and Pennsylvania", *American Economic Review*, Vol. 84, No. 4, pp. 772-793, September 1994.
- Cappelleri, Joseph C, M Trochim, William (2015), "Regression Discontinuity Design", <https://www.billtrochim.net/research/2015/2015%20-%20Cappelleri%20and%20Trochim%20-%20Regression%20Discontinuity%20Designs.pdf>.
- Cattaneo, Matias D. , Idrobo, Nicolás, Titiunik Roico, "a Practical Introduction to Regression Discontinuity Designs", Cambridge University Press, USA, 2019.
- Cavallo, A. (2016). "Scraped Data and Sticky Prices" *Review of Economics and Statistics*. doi: 10.1162/REST_a_00652.
- Chan, C. *What is a ROC Curve and How to Interpret It*. Ekim 2005 tarihinde <https://www.displayr.com/>: <https://www.displayr.com/what-is-a-roc-curve-how-to-interpret-it/>
- Chen G., Cai Z. ve Taylor P.A., (2024), "Through the lens of causal inference: Decisions and pitfalls of covariate selection", *Apert Neuro*. 2024;4:10.52294/001c.124817.
- Coady, D., Grosh, M., & Hoddinott, J. (2004). Targeting of Transfers in Developing Countries. *WORLD BANK REGIONAL AND SECTORAL STUDIES*. <https://openknowledge.worldbank.org/server/api/core/bitstreams/ede11a32-56d0-5486-bd05-f179dc231ee5/content>
- Conley, T. G. ve Taber, C. R. (2011), *Inference With "Difference In Differences" With A Small Number Of Policy Changes*, *The Review of Economics and Statistics*, https://users.ssc.wisc.edu/~lhong22/Teaching_Econ705/EventStudy/Conley_Taber_2011.pdf.
- Copestake, James, Morsink, Marlies and Remnant, Fiona (2019) "Attributing Development Impact: the Qualitative Impact Protocol Case Book", Rugby, UK, Practical Action Publishing, <http://dx.doi.org/10.3362/9781780447469>.

- Craig, P. (2015), Synthetic controls: a new approach to evaluating interventions, <https://whatworksscotland.ac.uk/wp-content/uploads/2015/03/Synthetic-controls-working-paper.pdf>.
- Cunningham, S., (2021), Causal Inference: The Mixtape, Yale University Press.
- Çelik, H. vd., (2020), Nitel Veri Analizi ve Temel İlkeleri, Eğitimde Nitel Araştırmalar Dergisi – ENAD, Cilt 8 / Sayı 1.
- Çukurova Kalkınma Ajansı (2021). 2010-2016 Mali Destek Programları Etki Değerlendirme Raporu, https://www.cka.org.tr/uploads/document_center_v/cka-etki-analizi-raporu-01032022-87.pdf.
- “DAC Criteria For Evaluating Development Assistance”, <https://www.oecd.org/dac/evaluation/39119068.pdf>.
- David S. Lee and Thomas Lemieux (2010), “Regression Discontinuity Designs in Economics”, <https://www.aeaweb.org/articles?id=10.1257/jel.48.2.281>.
- de Chaisemartin, C., & d’Haultfoeuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, 110(9), 2964–2996
- de Chaisemartin, C.D., Ciccia, D., d’Haultfoeuille, X., & Knau, F. (2025). Difference-in-Differences Estimators When No Unit Remains Untreated.
- Detle, Holger and Martin Schumann, “Difference-in-Differences Estimation Under NonParallel Trends,” Working Paper, 2020.
- Dehejia, R. H., & Wahba, S. (2002). Propensity score-matching methods for nonexperimental causal studies. *The Review of Economics and Statistics*, 84(1)
- Elçi, Ş. (2014). Ar-Ge ve İnovasyon Destek Programlarının Değerlendirilmesi ve Etki Analizi. https://www.researchgate.net/publication/314239891_Ar-Ge_ve_Inovasyon_Destek_Programlarinin_Degerlendirilmesi_ve_Etki_Analiz.
- European Commission, “Common methodology for State aid evaluation”, Brussels, 28.5.2014 SWD(2014) 179 final, https://ec.europa.eu/competition/state_aid/modernisation/state_aid_evaluation_methodology_en.pdf.
- Fan, S. (2010). Covariate. *In Encyclopedia of research design* . <https://methods.sagepub.com/ency/edvol/encyc-of-research-design/chpt/covariate>
- Farrell, M. H. “Robust inference on average treatment effects with possibly more covariates than observations”, *Journal of Econometrics* Volume 189, Issue 1, November 2015, Pages 1-23, s.2.
- Frechtling, Joy and Sharp, Laure (1997) “User Friendly Handbook for Mixed Evaluations”, <https://www.nsf.gov/pubs/1997/nsf97153/start.htm>.
- Garrett Grolemond ve Hadley Wickham. 2022. Hands-On Programming with R. RStudio Education. Erişim tarihi: 30 Temmuz 2025. <https://rstudio-education.github.io/hopr/>
- Genç, M., Atasoy, Y. (2010). Ar-Ge Harcamaları ve Ekonomik Büyüme İlişkisi: Panel Veri Analizi. *Bilgi Ekonomisi ve Yönetim Dergisi*, Cilt: 5 (2), 27-34.
- GEI (2022) “Impact Evaluation”, Global Evaluation Initiative, <https://www.betterevaluation.org/methods-approaches/themes/impact-evaluation>.
- Gertler, Paul J., Sebastian Martinez, Patrick Premand, Laura B. Rawlings, and Christel M. J. Vermeersch (2016) “Impact Evaluation in Practice”, second edition. Washington, DC: Inter-American Development Bank and World Bank. doi:10.1596/978-1-4648-0779-4IEG (2012) .
- Gertler, P., Martinez, S., Rawlings, L. B., Premand, P., ve Vermeersch, C. M. J.,(2016), *Impact Evaluation in Practice: Second Edition*, s. 136.

- Ghanem, D. Sant'Anna, P. H. C. ve Wüthrich, K., (2023), "Selection and Parallel Trends", <https://psantanna.com/files/gsw2022.pdf>, s. 3.
- Glewwe, P. and Todd, P. (2022) "Impact Evaluation in International Development: Theory, Methods, and Practice" Washington, DC, World Bank.
- Grubbs, F. E., (1969), "Procedures for Detecting Outlying Observations in Samples", *Technometrics*, Vol. 11, No. 1, 1-21.
- Güneş, C., Ünlü, M., & Birecikli, Ş. (2016). *Türkiye'de Sağlık Hizmetleri Talebinin Sayma Veri Modelleriyle İncelenmesi: İçsellik Sorunu*. <https://dergipark.org.tr/tr/download/article-file/230721> adresinden alındı.
- Josse, Julie. 2020. "Data Analysis with Missing Values in R." R-miss-tastic Tutorials. Erişim: 30 Temmuz 2025. https://rmissstastic.netlify.app/tutorials/josse_bookdown_dataanalysismissingr_2020
- Hadley Wickham, Mine Çetinkaya-Rundel. 2023. R for Data Science (2e). Erişim tarihi: 30 Temmuz 2025. <https://r4ds.hadley.nz/>
- Hagemann, A. (2023), *Permutation Inference with a Finite Number of Heterogeneous Clusters*, The Review of Economics and Statistics, <https://arxiv.org/pdf/1907.01049>.
- Hahn, J. ve Shi, R. (2017), *Synthetic Control and Inference*, *Econometrics*, <https://www.mdpi.com/2225-1146/5/4/52>.
- Ham, D. W. ve Miratrix, L., (2024), "Benefits and costs of matching prior to a Difference in Differences analysis when parallel trends does not hold", arXiv:2205.08644v5.
- Hansen, B. B. (2004). Full Matching in an Observational Study of Coaching for the SAT. *Journal of the American Statistical Association*, 99(467)
- Hansen, B. E. (2024). *Jackknife Standard Errors For Clustered Regression*, <https://users.ssc.wisc.edu/~behansen/papers/tcauchy.pdf>; Hansen, B. E. (2025), *Standard Errors For Difference-In-Difference Regression*, *Journal of Applied Econometrics*, https://users.ssc.wisc.edu/~behansen/papers/jae_25.pdf.
- Hayat, E., Özden, A. S. (2021). Genelleştirilmiş Sıralı Lojistik Regresyon Analizi ile Bireylerin Mutluluk Düzeylerine Etki Eden Faktörlerin Belirlenmesi. *Uluslararası Toplum Araştırmaları Dergisi* 18 (40). <https://dergipark.org.tr/tr/download/article-file/1620468>.
- Heckman, J. J., Ichimura, H., & Todd, P. E. (1997). Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme. *The Review of Economic Studies*, 64(4)
- Hedges, L. V., (2024), "Interpretation of the Standardized Mean Difference Effect Size When Distributions Are Not Normal or Homoscedastic", *Educ Psychol Meas.* 2024 Oct 6;85(2):245–257.
- Heinrich, C., Maffioli, A., Vazquez, G. (2010). A Primer for Applying Propensity-Score Matching. Office of Strategic Planning and Development Effectiveness, Inter-American Development Bank. <https://publications.iadb.org/publications/english/viewer/A-Primer-for-Appling-Propensity-Score-Matching.pdf>.
- Hill, Jennifer. 2008. "Discussion of research using propensity-score matching: Comments on A critical appraisal of propensity-score matching in the medical literature between 1996 and 2003 by Peter Austin, *Statistics in Medicine*." *Statistics in medicine* 27(12):2055–2061.
- Ho, D. E., Imai, K., King, G. and Stuart, E. A. (2007). Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. *Political Analysis* 15 199–236.

- Iacus, S. M., King, G., & Porro, G. (2012). Causal inference without balance checking: Coarsened exact matching. *Political analysis*, 20(1)
- Ibragimov, R. and Müller, U. K. (2016), *Inference with Few Heterogeneous Clusters*, *The Review of Economics and Statistics*,
<https://www.princeton.edu/~umueller/BehrensFisher.pdf>.
- IEG (2012) “World Bank Group Impact Evaluations: Relevance and Effectiveness”, June 2012, Independent Evaluation Group, World Bank, IFC, MIGA,
<https://openknowledge.worldbank.org/server/api/core/bitstreams/efaebc2a-37e1-5991-9d78-38ecbac0f8e2/content>.
- ILO, 2018, Guide on Measuring Decent Jobs for Youth Monitoring, evaluation and learning in labour market programmes, Note 6. A Step-by-Step Guide to Impact Evaluation, https://www.ilo.org/wcmsp5/groups/public/---ed_emp/documents/instructionalmaterial/wcms_627314.pdf.
- Imbens, G. W. (2004). Nonparametric estimation of average treatment effects under exogeneity: A review. *Review of Economics and Statistics* 86 4–29.
- Imbens, G. W. ve Wooldridge, J. M. (2009). "Recent Developments in the Econometrics of Program Evaluation." *Journal of Economic Literature*, 47(1), 5–86, s. 15.
- Imbens, G. W., (2004), “Nonparametric Estimation of Average Treatment Effects Under Exogeneity: A Review”, *The Review of Economics and Statistics*, 86(1): 4–29.
- Kang, Hyun. 2013. “The Prevention and Handling of the Missing Data.” *Korean Journal of Anesthesiology* 64 (5): 402–406.
<https://pmc.ncbi.nlm.nih.gov/articles/PMC3668100/>
- Khandker, S. R., Koolwal, G. B., Samad, H. A. (2010). *Handbook on Impact Evaluation Quantitative Methods and Practices*. The World Bank, Washington, D. C.
<https://openknowledge.worldbank.org/server/api/core/bitstreams/67f37dac-345d-57db-8289-244ad8c60c83/content>.
- King, G. and Nielsen, R. (2019), Why Propensity Scores Should Not Be Used for Matching, *Political Analysis*, 27, 4, Pp. 435-454.
- Kreif, N., Grieve, R., Hangartner, D., Turner A., Nikolova S., Sutton, M., Examination of the Synthetic Control Method for Evaluating Health Policies with Multiple Treated Units, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5111584/>.
- Kwak, S. K. ve Kim, J. H., (2017), “Statistical data preparation: management of missing values and outliers”, *Korean J Anesthesiol*. 2017 Jul 27;70(4):407–411.
- Lee, S. (2025, Mart). *McFadden's R-squared: Top 5 Stats for Model Evaluation*. Ekim 2025 tarihinde <https://www.numberanalytics.com/blog/mcfaddens-r-squared-top-5-stats-model-evaluation>
- Lee, J., & Little, T. D. (2017). A practical guide to propensity score analysis for applied clinical research. *Behaviour Research and Therapy*, 98, 76–90.
- Marin Kukoč, Bruno Škrinjarić ve Josip Juračak, 2020. *The Impact Assessment of the EU Pre-Accession Funds on Agriculture and Food Companies: The Croatian Case*. EIZ.
- Morgan, S. L. and Harding, D. J. (2006). Matching estimators of causal effects: Prospects and pitfalls in theory and practice. *Sociological Methods & Research* 35 3–60.
- Morten O. Nielsen, and Matthew D. Webb, “Cluster-robust inference: A guide to empirical practice,” *Journal of Econometrics*, 2022, Volume 232, Issue 2, February 2023, Pages 272-299 ile Rambachan, Ashesh and Jonathan Roth,

- “Design-Based Uncertainty for QuasiExperiments,” November 2022. arXiv:2008.00602 [econ, stat].
- Nichols, Austin (2022) “Causal inference with observational data”, *The Stata Journal* (2007) 7, Number 4, pp. 507–541, Washington, DC, Urban Institute.
- O’Brien, R. M., (2007), “A Caution Regarding Rules of Thumb for Variance Inflation Factors”, *Quality & Quantity*, Volume 41, pages 673–690.
- OECD (1991) “Principles for Evaluation of Development Assistance”, OECD, Paris, France, <https://www.oecd.org/dac/evaluation/2755284.pdf>.
- OECD (2007) “OECD Framework for the Evaluation of SME and Entrepreneurship Policies and Programmes”, https://www.oecd-ilibrary.org/industry-and-services/oecd-framework-for-the-evaluation-of-sme-and-entrepreneurship-policies-and-programmes_9789264040090-en.
- OECD (2009) “Impact Evaluations and Development”, NONIE Guidance on Impact Evaluation, <https://www.oecd.org/development/evaluation/dcdndep/47466906.pdf>.
- OECD (2021) “Applying Evaluation Criteria Thoughtfully”, OECD Publishing, Paris, <https://doi.org/10.1787/543e84ed-en>.
- Pattanayak, C. W., Rubin, D., & Zell, E. (2011, Ekim). Propensity Score Methods for Creating Covariate Balance in Observational Studies. *Cardiologia*, 64(10), 897-903. <https://www.revespcardiol.org/en-propensity-score-methods-for-creating-articulo-S1885585711004749> adresinden alındı
- Pokropek, Artur (2016),” Introduction to instrumental variables and their application to large-scale assessment data”, <https://eric.ed.gov/?id=EJ1182248>.
- Porter, D., & Gujarati, D. (2009). *Basic Econometrics*.
- Polat, B., Aktakke, N. (2017). Etki Değerlendirme Yöntemleri. *Kalkınma Programlarının Değerlendirmesinde Etki Analizi Yöntemleri ve Türkiye’de Uygulamalar*, 19-49.
- Pötscher, B. M. and Preinerstorfer, D. (2023), *Valid Heteroskedasticity Robust Testing*, *Econometric Theory*, <https://www.cambridge.org/core/journals/econometric-theory/article/valid-heteroskedasticity-robust-testing/AE47F17F50EDF86FE607EE4645640304> ; Pötscher, B. M. and Preinerstorfer, D. (2024), *A Necessary And Sufficient Condition For Size Controllability Of Heteroskedasticity Robust Test Statistics*, <https://arxiv.org/pdf/2412.17470>. Pre-test with Caution: Event-study Estimates After Testing for Parallel Trends,” *American Economic Review: Insights*, 2022, 4 (3), 305–322.
- R-Bloggers. 2016. “Outlier Detection with Mahalanobis Distance.” R-Bloggers. Erişim: 30 Temmuz 2025. <https://www.r-bloggers.com/2016/12/outlier-detection-with-mahalanobis-distance/>
- Ravenfo. 2021. “Aykırı Değer Analizi.” Ravenfo Blog. Erişim: 30 Temmuz 2025. <https://ravenfo.com/2021/02/11/aykiri-deger-analizi/>
- Roth, Sant’Anna, Bilinski ve Poe, “What’s Trending in Difference-in-Differences? A Synthesis of the Recent Econometrics Literature”, January 2023, arXiv:2201.01194v3.
- Rosenbaum, P. R. ve Rubin, D. B., (1983), “The Central Role of the Propensity Score in Observational Studies for Causal Effects”, *Biometrika*, Vol. 70, No. 1 (Apr., 1983), pp. 41-55.
- Rosenbaum, P. R. ve Rubin, D. B., (1985), “Constructing a Control Group Using Multivariate Matched Sampling Methods That Incorporate the Propensity Score”, *The American Statistician*, Vol. 39, No. 1, pp. 33-38.

- Rosenbaum, P. R. (1989). Optimal Matching for Observational Studies. *Journal of the American Statistical Association*, 84(408)
- Rosenbaum, P. R. (2002). *Observational Studies*, 2nd ed. Springer, New York.
- Rubin, D. B., (1980), “Randomization Analysis of Experimental Data: The Fisher Randomization Test Comment”, *Journal of the American Statistical Association*, Vol. 75, No. 371 (Sep., 1980), pp. 591-593.
- Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, 66(5), 688–701
- Rubin, D. B., (2001), “Using Propensity Scores to Help Design Observational Studies: Application to the Tobacco Litigation”, *Health Services & Outcomes Research Methodology* 2:169–188, 2001.
- Ryu, Ho Jin. 2023. dlookr: Tools for Data Diagnosis, Exploration, Transformation, and Reporting. R package version 0.6.1. <https://CRAN.R-project.org/package=dlookr>
- Ryu, Choonghyun. n.d. dlookr. Erişim: 30 Temmuz 2025. <https://choonghyunryu.github.io/dlookr/>
- Sant’Anna, P. v. (2025). *Difference-in-Differences Designs: A Practitioner’s Guide*.
- Scholz, Vera (2021) “The Qualitative Impact Protocol”, <https://www.intrac.org/wpcms/wp-content/uploads/2019/05/QUIP.pdf>.
- Sills, E.O. ve diğerleri, (2015), <https://journals.plos.org/plosone/article/file?id=10.1371/journal.pone.0132590&type=printable>.
- Simister, Nigel, James, Dan and Lunin, Letisha (2017) “Qualitative Analysis”, <https://www.intrac.org/wpcms/wp-content/uploads/2017/01/Qualitative-analysis.pdf>.
- STDF (2021) “Guidelines for the Evaluation of Projects Funded by the Standards and Trade Development Facility”, Standards and Trade Development Facility, https://standardsfacility.org/sites/default/files/STDF_214_Evaluation_Guidelines_2021_Final.pdf.
- Staffa, S. J., & Zurakowski, D. (2018). *Five steps to successfully implement and evaluate propensity score matching in clinical research studies*.
- Stock, J., & Watson, M. (tarih yok). *Introduction to Econometrics*. <https://users.ssc.wisc.edu/~mchinn/stock-watson-econometrics-3e-lowres.pdf>
- Stuart, E. A. (2010). Matching methods for causal inference: A review and a look forward. *Statistical Science*, 25(1), 1–21.
- The World Bank. (2010). *Handbook on Impact Evaluation Quantitative Methods and Practices*. <https://documents1.worldbank.org/curated/en/650951468335456749/pdf/520990PUB0EPI1101Official0Use0Only1.pdf>
- Thomas, R. (2012) “Conditional cash transfers to improve education and health: An ex-ante evaluation of red de protección social, Nicaragua” *Health Economics*, 21(10), 1136-1154.
- Thompson, S. (2017), *Impact Evaluations using Administrative Data: How the Synthetic Control Method Opens Doors to Otherwise Impossible Impact Evaluations*, <https://www.summitllc.us/blog/impact-evaluations-using-administrative-data-how-the-synthetic-control-method-opens-doors-to-otherwise-impossible-impact-evaluations>
- Thulin, Måns. 2024. *Modern Statistics with R*. 2nd ed. Chapman & Hall/CRC Press. Erişim tarihi: 30 Temmuz 2025. <https://www.modernstatisticswithr.com/>
- Tierney, Nicholas, Di Cook, and Thomas Lin Pedersen. 2021. *naniar: Data Structures, Summaries, and Visualisations for Missing Data*. R package version 0.6.1.

- <https://CRAN.R-project.org/package=naniar> Tierney, Nicholas. n.d. naniar. Erişim: 30 Temmuz 2025. <https://naniar.njtierney.com/>
- Totik Doğan, N., Püren, S., Karakaya, Y., & Alparslan, Z. (2023). An Introduction to Propensity Score Analysis: Checklist for Clinical. <https://avesis.cu.edu.tr/yayin/4294aa64-20ee-49b7-b19f-06dd5d87ebe7/an-introduction-to-propensity-score-analysis-checklist-for-clinical-researches>
- Trevor French. 2023. R for Data Analysis – Data Types. Erişim tarihi: 30 Temmuz 2025. <https://trevorfrench.github.io/R-for-Data-Analysis/plc3-data-types.html>
- Tukey, J. W., Exploratory Data Analysis, Addison-Wesley Publishing Company Reading, Mass. — Menlo Park, Cal., London, Amsterdam, Don Mills, Ontario, Sydney 1977.
- UNDP (2019) “The Revised UNDP Evaluation Policy), United Nations Development Programme, https://popp.undp.org/UNDP_POPP_DOCUMENT_LIBRARY/Public/AC_Evaluation.pdf.
- UNEG (2013). Impact Evaluation in UN Agency Evaluation Systems: Guidance on Selection, Planning and Management. Guidance Document. United Nations Evaluation Group.
- United States Agency for International Development -USAID (2006), Collecting and Using Data for Impact Assessment, https://pdf.usaid.gov/pdf_docs/Pnadn197.pdf.
- United States Agency for International Development, https://usaidlearninglab.org/system/files/resource/files/mod7_decision_tree_for_selecting_evaluation_design.pdf.
- Van Buuren, Stef. 2018. Flexible Imputation of Missing Data. 2nd ed. Boca Raton, FL: Chapman & Hall/CRC. <https://stefvanbuuren.name/fimd/>
- Van Buuren, Stef, and Karin Groothuis-Oudshoorn. 2011. “mice: Multivariate Imputation by Chained Equations.” Journal of Statistical Software 45 (3): 1–67. <https://CRAN.R-project.org/package=mice>
- VanderWeele, T. J., ve Hernán, M. A., (2013), “Causal Inference Under Multiple Versions of Treatment”, J Causal Inference. 2013 May 29;1(1):1–20.
- Verheuve, N., Witteman, J., Vlaanderen M., Synthetic Control Method for Dutch Policy Evaluationutton, M., , <https://link.springer.com/article/10.1007/s10645-022-09417-5>.
- Villabona, J. (2022), Tax incentives as carrots for solar energy adoption, <https://repository.urosario.edu.co/bitstreams/8059b275-8698-4b7f-ba96-54d9e2f9b2c2/download>.
- Yang, J., Webster-Clark, M., Lund, J., Sandler, R., Dellon, E., & Stürmer, T. (2019). Propensity score methods to control for confounding in observational cohort studies: a statistical primer and application to endoscopy research. Gastrointest Endosc. doi:10.1016/j.gie.2019.04.236. Epub 2019
- Yıldırım, F. (2014). Karşı-Olgusal Etki Analizi. *Anahtar Dergisi* (302), 45-48.
- Yıldız, S. B. (2013). Küçük ve Orta Ölçekli İşletmelerin Desteklenmesinde Kosgeb’in Rolü: Bir Etki Değerlendirmesi Araştırması. Süleyman Demirel Üniversitesi Sosyal Bilimler Enstitüsü İşletme Anabilim Dalı (Doktora Tezi). Yükseköğretim Kurulu Ulusal Tez Merkezi Veri Tabanından erişilebilir. (Tez No. 339280).
- Zhou, Y., Aryal, S. ve Bouadjenek, M. R., (2024), “A Comprehensive Review of Handling Missing Data: Exploring Special Missing Mechanisms”, arXiv:2404.04905v1.
- Westat J. F.(2002) , The 2002 User-Friendly Handbook for Project Evaluation, p.43-53.

- Wickham, Hadley, Mine Çetinkaya-Rundel, and Garrett Grolemund. 2023. R for Data Science (2e), “Missing values.” <https://r4ds.hadley.nz/missing-values>
- White, Howard and Raitzer, David A. (2017) “Impact Evaluation of Development Interventions: A Practical Guide”, Asian Development Bank <https://www.adb.org/sites/default/files/publication/392376/impact-evaluation-development-interventions-guide.pdf>.
- Wooldridge, J. M., Econometric Analysis of Cross Section and Panel Data, The MIT Press, Cambridge, Massachusetts, London, England, 2010.
- World Bank (2010), <https://openknowledge.worldbank.org/entities/publication/f89faa3e-3aba-5b06-ab9c-5fc4dce9be59>.
- Wooldridge, Jeffrey M, “Two-Way Fixed Effects, the Two-Way Mundlak Regression, and Difference-inDifferences Estimators,” Working Paper, 2021, pp. 1–89.
- World Bank (2006), Outline of Principles of Impact Evaluation <https://www.oecd.org/dac/evaluation/dcdndep/37671602.pdf>.
- World Bank Group (2018), Toolkit for Impact Evaluation of Public Credit Guarantee Schemes for SMEs, <https://www.summitllc.us/blog/impact-evaluations-using-administrative-data-how-the-synthetic-control-method-opens-doors-to-otherwise-impossible-impact-evaluations>
- İnternet: <https://www.resmigazete.gov.tr/eskiler/2007/04/20070403-13.htm>
 chibueze99. 2019. “MissingR.” RPubs. Erişim: 30 Temmuz 2025.
<https://rpubs.com/chibueze99/MissingR>
 rpatel40. 2021. “Handling Missing Data in R.” RPubs. Erişim: 30 Temmuz 2025.
https://rpubs.com/rpatel40/handling_missing_data_in_R
 Justjooz. 2019. “Missing Data and Data Imputation.” RPubs (November 21).
 Erişim tarihi: 30 Ekim 2025. https://www.rpubs.com/justjooz/miss_data
 yuzaR Data Science. 2021. “Handling Missing Data in R.” YouTube (video).
 Erişim: 30 Temmuz 2025. <https://www.youtube.com/watch?v=Akb401i32Oc>
 Built In. n.d. “Mahalanobis Distance.” Erişim: 30 Temmuz 2025.
<https://builtin.com/data-science/mahalanobis-distance>
 maulikpatel. 2016. “Mahalanobis Distance.” RPubs. Erişim: 30 Temmuz 2025.
<https://rpubs.com/maulikpatel/228329>
 EZBioCloud Help Center. n.d. “Box Plot.” Erişim: 30 Temmuz 2025.
<https://help.ezbiocloud.net/box-plot/>
 OpenIntro Project. 2023. Exploratory Data Analysis. Erişim tarihi: 30 Temmuz 2025. <https://openintro-ims.netlify.app/exploratory-data-analysis>
 DataCamp. 2024. Introduction to Statistics in R. Erişim tarihi: 30 Temmuz 2025.
<https://app.datacamp.com/learn/courses/introduction-to-statistics-in-r>

EKLER

EK 1. Farkların Farkı Yöntemini Uygulayacak Araştırmacılar için Kontrol Listesi⁸⁷

A. Herkes aynı zamanda mı müdahaleye maruz kalmaktadır?

Eğer herkes aynı zamanda müdahaleye maruz kalıyorsa ve panel veri dengeliyse⁸⁸ iki yönlü sabit etkiler (TWFE) modelleri kullanılarak yapılan tahminin, kolayca yorumlanabilir tahminler sağladığı değerlendirilmektedir. TWFE modelleri, bağımlı değişkenin grup ve zaman sabit etkileri ile birlikte müdahale ile ilişkilendirildiği bir yapıdır. Bu sayede, gruplar arasındaki sabit farklılıklar ve zaman içindeki genel eğilimler kontrol edilebilir ve müdahalenin doğrudan etkisi daha net bir şekilde anlaşılabilir.

Eğer herkes aynı zamanda müdahaleye maruz kalmıyorsa, kademeli müdahale zamanlaması için “heterojenlik açısından sağlam” bir tahmin aracının kullanılmasının değerlendirilmesi önerilmektedir. Bu çerçevede, söz konusu makalenin (bkz. dipnot 82) 3.3 numaralı başlığı altında bahsedilen Callaway ve Sant’Anna⁸⁹ (CS) tahmincisi ile veri atama⁹⁰ (imputation) tahmincisi yaklaşımlarının kullanılabileceği değerlendirilmektedir.

CS tahmincisi yaklaşımının statik ya da dinamik TWFE modeline kıyasla iki temel avantajı bulunmaktadır. Bunlardan ilki müdahale etkilerinin gelişigüzel heterojenliği altında bile duyarlı tahminler sağlamasıdır. Duyarlı ifadesi ile hem yaklaşımın negatif ağırlıklandırmayı önlediği hem de gruplar arasındaki etkilerin ağırlıklandırılmasının en küçük kareler (EKK) metodu ile belirlenmek yerine araştırmacı tarafından belirlendiği kastedilmektedir. İkinci avantajı ise gözlemlenemeyen potansiyel sonuçların çıkarımı için hangi birimlerin kontrol grubunda kullanılıyor olduğunu tam anlamıyla şeffaflaştırmasıdır.

Veri atama (imputation) tahmincisi yaklaşımları spesifik olarak, henüz müdahaleye maruz kalmamış birimler ve dönemler için gözlemleri kullanarak bir TWFE regresyon denklemi uyarlamaktadır. Bu, regresyondan tahmin edilen değeri kullanarak, müdahaleye maruz kalmış her bir birim için, hiç müdahale edilmemiş olma potansiyellerini tahmin etmektedir. Müdahaleye maruz kalmış her birim için müdahale etkisinin bir tahminini sağlamakta ve bireysel düzeydeki bu tahminler ortalama müdahale etkisi (ATT) gibi özet parametrelerin tahminlerini oluşturmak için toplanabilmektedir. Bu yaklaşımlar, paralel eğilim varsayımının tüm gruplar ve dönemler için geçerli olduğu ve

⁸⁷ Roth, Sant’Anna, Bilinski ve Poe, “What’s Trending in Difference-in-Differences? A Synthesis of the Recent Econometrics Literature”, January 2023, *arXiv:2201.01194v3*.

⁸⁸ Dengeli panel veri, bir araştırma veya analizde kullanılan panel veri setinin her bir gözleminin her bir zaman birimine ve her bir gözlem birimine karşılık geldiği durumu ifade etmektedir. Yani, tüm gözlemler için zaman serisi ve birimler (örneğin, bireyler, ülkeler, firmalar) eksiksiz ve tamdır.

⁸⁹ Callaway, Brantly and Pedro H. C. Sant’Anna, “Difference-in-Differences with multiple time periods,” *Journal of Econometrics*, 2021, 225 (2), 200–230.

⁹⁰ Borusyak, Kirill and Jann Spiess, “Revisiting Event Study Designs: Robust and Efficient Estimation,” *arXiv:2108.12419 [econ]*, 2021. Ek olarak bkz. Gardner, John, “Two-stage differences in differences,” Working Paper, 2021 ile Wooldridge, Jeffrey M, “Two-Way Fixed Effects, the Two-Way Mundlak Regression, and Difference-in-Differences Estimators,” Working Paper, 2021, pp. 1–89.

herhangi bir öngörünün/beklentinin olmadığı durumlarda geçerli tahminler sağlamaktadır.

B. Paralel eğilim varsayımı geçerli midir?

Eğer cevap “evet” ise bunun nasıl olduğu fonksiyonel biçim seçiminin sebebinin de içerecek şekilde açıklanmalıdır. Eğer sebep, (yarı) rassal müdahale zamanlaması ise, daha etkili bir tahmincinin kullanılması düşünülmelidir. Zira, kademeli müdahale zamanlamasının söz konusu olduğu durumlarda, genelleştirilmiş paralel eğilim varsayımının genellikle müdahale zamanlamasının rassal ya da yarı-rassal olduğu tartışılarak gerekçelendirilmektedir.

Eğer paralel eğilim varsayımı geçersizse, aşağıdaki aşamaların değerlendirilmesi önerilmektedir:

- Paralel eğilim varsayımı ortak değişkenlere bağlı olarak daha makul olacaksa, bu durumda ortak değişkenlere bağlı bir yöntemin düşünülmesi önerilmektedir. Önerilen yöntemlerden biri “en yakın komşu” eşleştirmesi iken diğeri ise ters olasılıklı ağırlıklandırma (Inverse Probability Weighting-IPW).
- Bir grafik yardımıyla paralel eğilim varsayımının geçerliliğinin incelenmesi önerilmektedir. Eğer ortak bir müdahale tarihi varsa ve koşulsuz bir paralel eğilim varsayımı kullanılıyorsa o zaman müdahaledeki ilerlemeleri ve gecikmeleri içeren bir TWFE modeli kullanılabilir. Eğer böyle bir durum söz konusu değilse müdahale öncesindeki dönemde paralel eğilim söz konusu olsa bile tahmin edilen katsayılar hatalı olabilecektir.
- Eğer grafik yardımıyla paralel eğilim varsayımının geçerliliği test edilecekse bu çalışmaya, ilgili alternatiflere ve/veya eşdeğerlik testlerine karşı ön testin gücünün belirlenmesi/tespit edilmesi hususunun da eklenmesi önerilmektedir. Örneğin Roth’un⁹¹ araştırmacının paralel eğilimin ihlali varsayımının altında güç analizi yapmak ve ön testlerden kaynaklanan olası sapmaları hesaplamak için geliştirdiği araçlardan yararlanılabileceği değerlendirilmektedir. Bilinski ve Hatfield⁹² ile Dette ve Schumann⁹³ ise boş ve alternatif hipotezlerin rollerini tersine çevirerek düşük güç sorununu çözmeye yardımcı olan ön teste yönelik “eşdeğerlik” yaklaşımlarını önermektedirler.
- Paralel eğilim varsayımının muhtemel ihlallerine karşı sonuçların sağlamlığını gösteren resmi duyarlılık analizlerinin raporlanması önerilmektedir.

C. Bir süper popülasyondan örneklenmiş çok sayıda müdahale ve kontrol kümesi var mı?

Eğer cevap “evet” ise, küme düzeyinde küme-sağlam (cluster-robust) yöntemler kullanılması önerilmektedir. FF yöntemine ilişkin temel modelde sonsuz bir süper

⁹¹ “Pre-test with Caution: Event-study Estimates After Testing for Parallel Trends,” *American Economic Review: Insights*, 2022, 4 (3), 305–322

⁹² Bilinski, Alyssa and Laura A. Hatfield, “Seeking evidence of absence: Reconsidering tests of model assumptions,” *arXiv:1805.03273 [stat]*, May 2018.

⁹³ Dette, Holger and Martin Schumann, “Difference-in-Differences Estimation Under NonParallel Trends,” *Working Paper*, 2020.

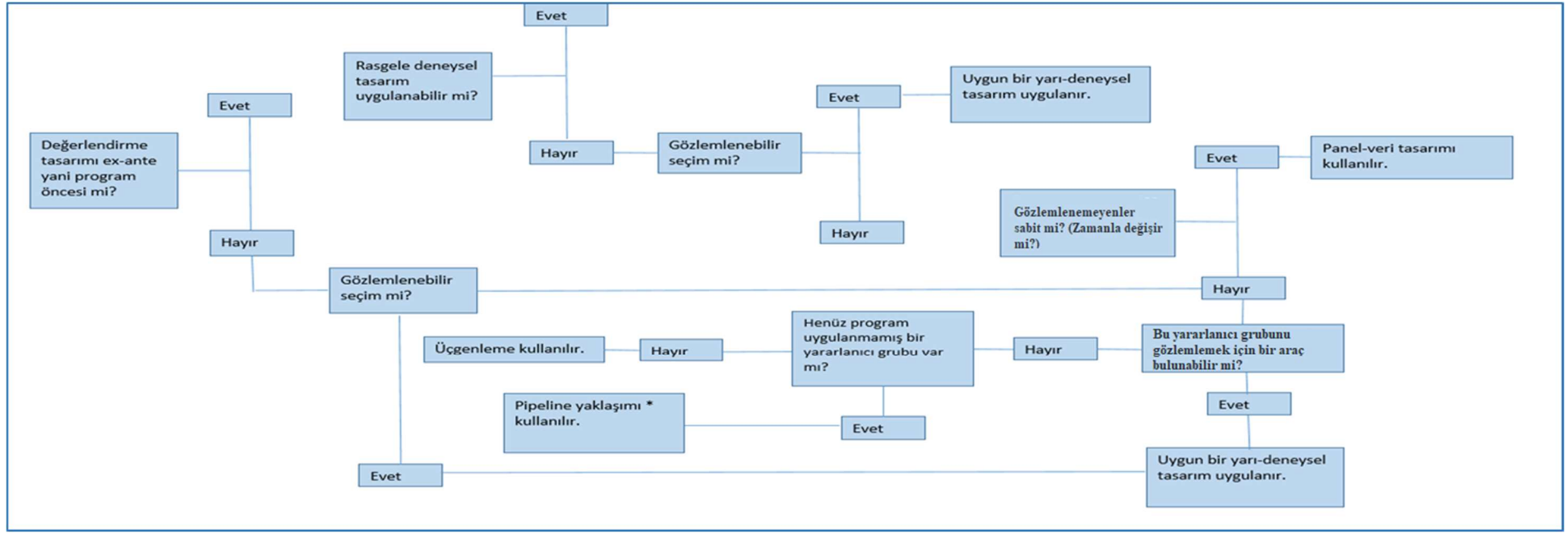
popülasyondan örneklenmiş çok sayıda bağımsız olarak çekilen küme olduğu varsayılmaktadır. Ancak, bu varsayımla ilgili iki kavramsal zorluk bulunmaktadır. Öncelikle, birçok durumda kümelerin süper popülasyonunun ne olduğu belirsizdir. İkinci olarak birçok ortamda uygun kümeleme düzeyinin ne olduğunu belirlemek zordur. Bu çerçevede, uygun kümeleme düzeyinden emin olunmadığı durumda, müdahalenin bağımsız olarak atandığı düzeyde kümeleme yapmanın iyi bir genel kural olduğu belirtilmektedir.

Müdahaleye maruz kalmış az sayıda kümenin olması durumunda alternatif çıkarım yöntemlerinden birinin kullanılması önerilmektedir. Bu yöntemlerden biri Fisher Rastgeleleştirme Testleri (Fisher Randomization Tests), başka bir deyişle permütasyon testleridir. Söz konusu testlerdeki temel mantık, verinin bazı istatistiklerinin (örneğin farkların farkı tahmincisinin t-istatistiği) hesaplanması ve sonrasında (küme seviyesinde) müdahale atamasının farklı permütasyonları altında söz konusu istatistiklerin yeniden hesaplanması şeklindedir. Permütasyon testlerinin faydalı bir kıyaslama olabileceği değerlendirilmektedir. Zira, müdahale rassal bir şekilde atanmış olsa bile müdahale etkisinin olmadığı şeklindeki boş hipotez reddedilemez. Bu durum, diğer güçlü varsayımlar olmadan veride bir etkinin olduğuna dair güçlü bir kanıt bulunmadığını göstermektedir. Kademeli müdahale zamanlamasının olduğu durumlarda, bir birimin müdahaleye maruz kalma zamanının olabildiğince rassal olduğunun varsayılmasının daha makul olabileceği değerlendirilmektedir.

Eğer süper popülasyon tasavvur edilemiyorsa başka bir deyişle süper popülasyon kavramsallaştırılamıyorsa çıkarım için tasarım temelli bir gerekçe düşünülmelidir. Tasarım temelli yaklaşımda, verideki birimler sabit ve müdahale ataması stokastik/rassal olarak kabul edilmektedir. Böylelikle, süper popülasyonun kavramsallaştırılmasına gerek kalmamakta ve uygun kümeleme seviyesi, müdahalenin atanma şekline göre belirlenmektedir. Son dönemlerde gerçekleştirilen çalışmalara⁹⁴ ait sonuçların kümelerin içinden çekildiği bir süper popülasyonu kavramsallaştırmanın zor olması durumunun aslında bir sorun olmayabileceği ve eğer süper popülasyonu yoksa ve belirsizlik müdahalenin stokastik/rassal atamasından kaynaklanıyorsa “olağan” yaklaşımın geçerliliğini koruyacağı değerlendirilmektedir.

⁹⁴ Morten O. Nielsen, and Matthew D. Webb, “Cluster-robust inference: A guide to empirical practice,” *Journal of Econometrics*, 2022, Volume 232, Issue 2, February 2023, Pages 272-299 ile Rambachan, Ashesh and Jonathan Roth, “Design-Based Uncertainty for QuasiExperiments,” November 2022. *arXiv:2008.00602 [econ, stat]*.

EK 2. Nicel Etki Değerlendirme Karar Ağacı⁹⁵



*: Pipeline yaklaşımı, müdahaleyi (deney grubu) zaten deneyimlemiş olanların sonuçları ile henüz programa katılmamış birimlerin olduğu kontrol grubu karşılaştırılır. Ancak bu yaklaşımın geçerli olabilmesi için hem müdahale hem de kontrol gruplarının benzer özelliklere sahip olması önemlidir

⁹⁵ Değerlendirme müdahaleden önce tasarlanıyorsa (program öncesi), deneysellik sorusunu sorarak ile başlıyoruz. Eğer müdahale grubu rasgele seçilirse, o zaman kitleden seçilen rastgele bir örneklem, geçerli bir kontrol grubu olur ve programdan yararlanmamış olmaları şartıyla tasarım kurgulanabilir. Bu yaklaşım, belirli birimleri hedeflemenin mümkün olmadığı anlamına gelmez. Rastgele dağıtım, toplam nüfusun bir alt grubuna, örneğin en fakir bölgelerden birine olabilir. Rastgeleleştirme mümkün değilse, tüm seçim kriterleri gözleniyor mu sorusu sorulur. Eğer öyleyse, seçim yanlılığını ortadan kaldıracak bir dizi regresyona dayalı yaklaşım vardır. Seçim kriterleri gözlemlenmiyorsa ve zamanla değişmeyecekleri (sabit) düşünülüyorsa, o zaman panel veri kullanılması seçim yanlılığı etkilerini ortadan kaldıracaktır. Burada panel veri için bir referans (baseline) çizgisi gereklidir. Çalışma, tam olarak aynı birimler (bireyler, haneler, firmalardan oluşan bir panel) için bilgi elde etmenin mümkün olmayacağı şekilde ex-post yapılırsa ve seçim gözlemlenemeyenler (unobservables) tarafından belirlenirse, o zaman varsayılan gözlemlenemeyenleri gözlemek için bazı yollar aranmalıdır. Bu mümkün değilse, henüz programdan faydalanmamış yararlanıcılar varsa bir boru hattı yaklaşımı (pipeline approach) kullanılabilir. Yukarıda belirtilen prosedürlerden hiçbirisi mümkün değilse, seçim yanlılığı sorunu ele alınamaz. Etki değerlendirme, makul ilişkilendirme yoluyla bir argüman oluşturmak için büyük ölçüde müdahale teorisine ve üçgenlemeye dayanmalıdır.

EK 3. Değerlendirme Tasarımı için Karar Ağacı Süreci

1. Değerlendirme gerekli mi?	HAYIR A. doğru zaman değil B. Bir değerlendirmenin ele alabileceği hiçbir soru yok C. Paydaşlardan talep yok EVET A. Paydaşlardan bilgi talebi var B. Bir değerlendirmenin ele alabileceği sorular var
2. Bir performans değerlendirmesine mi yoksa bir etki değerlendirmesine mi ihtiyaç var?	Performans değerlendirmesi: Sorular, programın ne kadar iyi uygulandığı, hedeflerine ulaşp ulaşmadığı ve hedef kitlenin tüm kesimlerine ulaşılıp ulaşılmadığı ile ilgilidir. Etki Değerlendirmesi: Sorular, programın amaçlanan sonuçlarına ve etkilerine ulaşp ulaşmadığı ve bunların müdahalenin etkilerine atfedilebilir atfedilemeyeceği ile ilgilidir.
PERFORMANS DEĞERLENDİRMESİ	
3. İş bildirimi (Statement of Work-SOW) gerçekçi ve başarılı olabilir mi? Not: Statement of Work-SOW, belirli bir projenin gereksinimlerini açıklayan bir belgedir. Yapılacak işin kapsamı, zaman çizelgesi, çalışma yeri, bütçe vb. hüküm ve koşullarını tanımlar.	HAYIR A. Sorular ölçmek için karmaşık ve belirsizdir. B. Bütçe yetersiz. C. Zaman çizelgesi gerçekçi değil. EVET A. Bir performans değerlendirmesinde ele alınabilecek açıkça tanımlanmış sorular vardır. B. Önerilen bütçe ve zaman çizelgeleri gerçekçi.
4. Gerçek dünyadaki kısıtlamalar nelerdir? Kısıtlamaların tümü ele alınabilir mi veya değerlendirmeyi gerçekleştirmeyi imkânsız kılan kısıtlamalar var mı?	A. Bütçe kısıtları B. Zaman kısıtları C. Veri kısıtları D. Politik ve kurumsal kısıtlar
5. Program döngüsünün hangi aşamasında değerlendirme devreye alınır?	A. Planlama ve tasarım aşaması. B. Program başlangıcı. C. Program uygulaması sırasında D. Programın tamamlanmasıyla.

	E. Program birkaç yıl çalıştıktan sonra.
6. Hangi tür performans değerlendirme çalışması önerilmektedir?	i. Gelişimsel değerlendirme (Developmental evaluation) ii. Tanı Değerlendirmeleri (Diagnostic studies) iii. Kıyaslama verileri toplama (Benchmark data collection) iv. Hızlı geri bildirim çalışmaları (Rapid feedback studies) v. Katılımcı değerlendirme (Participatory evaluation) vi. Güçlendirme değerlendirmesi (Empowerment evaluation) vii. Dahili program tamamlama raporu (Internal program completion report) viii. Denetim ve uyum raporları (Audit and compliance reports) ix. Sürdürülebilirlik değerlendirmesi (Sustainability assessment)
7. Hangi araç ve teknikler kullanılacak?	Teori –Tabanlı Metotlar (Theory-based methods) Nicel Metotlar (Quantitative methods) Nitel Metotlar (Qualitative methods) Karma Metotlar (Mixed-methods)
ETKİ DEĞERLENDİRMESİ	
8. Müdahale basit mi (simple), karmaşık mı (complex) değerlendirilecek?	Basit: tasarımların tartışılmasıyla devam edin Karmaşık: Hem bağımsız hem de birbirine bağlı olarak hareket edebilen bir dizi bileşenden oluşur. Bu, müdahaleyi zorlaştırır. Dolayısıyla "aktif bileşenleri" tanımlayıp ve hangi bileşen veya bileşen kombinasyonlarının daha önemli olduğundan emin olarak tasarımlar değerlendirilmeye çalışılır.
9. İş bildirimi (Statement of Work-SOW) gerçekçi ve başarılabılır mı?	HAYIR A. Sorular ölçmek için karmaşık ve belirsizdir. B. Bütçe yetersiz. C. Zaman çizelgesi gerçekçi değil. EVET A. Bir performans değerlendirmesinde ele alınabilecek açıkça tanımlanmış sorular vardır.

	B. Önerilen bütçe ve zaman çizelgeleri gerçekçi.
10. Program döngüsünün hangi aşamasında değerlendirme devreye alınır?	A. Planlama ve tasarım aşaması. B. Program başlangıcı. C. Program uygulaması sırasında D. Programın tamamlanmasıyla. E. Program birkaç yıl çalıştıktan sonra. Not: Program değerlendirmesi ne kadar geç yapılırsa, araştırma tasarım seçenekleri yelpazesi o kadar sınırlı olur ve yüksek düzeyde bir metodolojik titizlik sağlamakta bir o kadar zorlaşır. Değerlendirme, programda geç başladığında, temel verileri yeniden oluşturmak için geriye dönük bir tasarım kullanma olasılığı göz önünde bulundurulmalıdır.
11. Gerçek dünyadaki kısıtlamalar nelerdir? Kısıtlamaların tümü ele alınabilir mi veya değerlendirmeyi gerçekleştirmeyi imkânsız kılan kısıtlamalar var mı?	A. Bütçe kısıtları B. Zaman kısıtları C. Veri kısıtları D. Politik ve kurumsal kısıtlar
12. Ne tür bir değerlendirme tasarımı önerilir?	A. Deneysel tasarım (experimental design) B. Yarı deneysel tasarım (quasi-experimental design) C. Deneysel olmayan tasarım (non-experimental design)
Deneysel Tasarım (Experimental Design)	
13. Yararlanıcılar nasıl seçilecek?	i. Kitle içinde seçilebilir olanların içinden gelişigüzel seçim (Ör. Kura ile) ii. Her p'inci girişi alarak bir listeden seçim (burada $p = N/n$; N = kitle (popülasyon)'deki toplam birimler ve n ise seçilecek birim sayısıdır). iii. İki aşamalı seçim. 1. aşamada tüm uygun birimler seçilir ve listelenir ve 2. aşamada yararlanıcılar rasgele seçilir. iv. İstatistiksel olarak uygun örneklem metodolojilerinden biri (ör. Kota, sistematik, kümeleme, tabakalı vb.)

14. Temel değerlendirme tasarımını geliştirmek için hangi teknikler kullanılacak?	A. Gelişmiş teknikler kullanılmayacaktır. B. Değerlendirmeyi bir değişim teorisine (theory of change) ve bir program mantığı modeline (logic model) dayandırmak. C. Süreç analizi (Process analysis) Ç. Bağlamsal analiz (Contextual analysis) D. Karma yöntem yaklaşımları E. Mevcut ikincil verilerin (secondary data) kullanımı F. Üçgenleme
15. Değerlendirmenin genel kalitesini değerlendirmek ve iç ve dış geçerliliğe yönelik tehditleri değerlendirmek için herhangi bir kalite değerlendirme prosedürü kullanılacak mı?	HAYIR Sistematik kalite değerlendirmesi kullanılmayacaktır. EVET Yalnızca kurumun kendi genel yönergeleriyle değerlendirilir ya da, Ayrıntılı bir kontrol listesi oluşturarak değerlendirilir.
Yarı-Deneysel Tasarımlar (Quasi-experimental designs)	
16. Ne tür eşleşen kontrol grubu kullanılacak?	İstatistiksel olarak eşleşen kontrol grubu (örneğin, Eğilim Skoru Eşleştirmesi kullanılarak) Regresyon süreksizliği tasarımı ile yargısal olarak seçim eşleştirme
17. Bir temel data (veri) olacak mı?	i. Değerlendirme, projenin başlatılması sırasında başlar ve hem program hem de kontrol grubu için temel verileri toplamak üzere anketler yapılır. ii. Birincil mevcut durum verileri yalnızca program (müdahale) grubu için toplanır. iii. Değerlendirme, geç başlarsa mevcut durum koşulları, daha önce yürütülen araştırmalardan elde edilen ikincil veriler kullanılarak tahmin edilir. iv. Mevcut durum verileri toplanmaz ve etki, program (müdahale) ve kontrol gruplarının son test (post-test) kontrollsından değerlendirilir.
18. Yarı deneysel tasarımlardan hangisi kullanılacaktır?	A. Regresyon Süreksizliği Tasarımı B. Farkın Farkı C. Eşleştirme (Matching) (bir değişken veya birden çok değişken)

19. Temel değerlendirme tasarımını geliştirmek için hangi teknikler kullanılacak?	A. Gelişmiş teknikler kullanılmayacaktır. B. Değerlendirmeyi bir değişim teorisine (theory of change) ve bir program mantığı modeline (logic model) dayandırmak. C. Süreç analizi (Process analysis) Ç. Bağlamsal analiz (Contextual analysis) D. Karma yöntem yaklaşımları E. Mevcut ikincil verilerin (secondary data) kullanımı F. Üçgenleme
20. Değerlendirmenin genel kalitesini değerlendirmek ve iç ve dış geçerliliğe yönelik tehditleri değerlendirmek için herhangi bir kalite değerlendirme prosedürü kullanılacak mı?	HAYIR Sistematik kalite değerlendirmesi kullanılmayacaktır. EVET Yalnızca kurumun kendi genel yönergeleriyle değerlendirilir ya da, Ayrıntılı bir kontrol listesi oluşturarak değerlendirilir.
Deneyisel Olmayan Tasarımlar (non-experimental Designs)	
21. Hangi deneyisel olmayan tasarım kullanılacak?	i. Kontrol grubu olmadan proje grubunun ön test-son test karşılaştırması ii. Tek bir durum için art arda üç ön test-son test karşılaştırmasının yapıldığı tek durum karşılaştırma (single-case comparison) tasarımı⁹⁶ iii. Herhangi bir temel veri olmadan ve kontrol grubu olmadan, programın sonlarında proje grubunun hızlı analizi (rapid analysis). Bu analiz genellikle çok az zaman ve küçük bir bütçe ile proje yerlerini ziyaret etmeyi içerir. iv. Boylamsal (longitudinal) çalışmalar, panel çalışmaları, vaka çalışmaları (case studies) , katılımcı gözlemi, katılımcı grup teknikleri gibi tekniklerin kullanıldığı daha güçlü nitel veya karma yöntem tasarımları.
22. Temel değerlendirme tasarımını geliştirmek için hangi teknikler kullanılacak?	A. Gelişmiş teknikler kullanılmayacaktır. B. Değerlendirmeyi bir değişim teorisine (theory of change) ve bir program mantığı modeline (logic model) dayandırmak.

⁹⁶ Her faz için program başlatılıp sonlandırılarak üç kez incelenebilir veya üç benzer durum test edilebilir.

	C. Süreç analizi (Process analysis) Ç. Bağlamsal analiz (Contextual analysis) D. Karma yöntem yaklaşımları E. Mevcut ikincil verilerin (secondary data) kullanımı F. Üçgenleme
23. Değerlendirmenin genel kalitesini değerlendirmek ve iç ve dış geçerliliğe yönelik tehditleri değerlendirmek için herhangi bir kalite değerlendirme prosedürü kullanılacak mı?	HAYIR Sistemati kalite değerlendirmesi kullanılmayacaktır. EVET Yalnızca kurumun kendi genel yönergeleriyle değerlendirilir ya da,ayrıntılı bir kontrol listesi oluşturarak değerlendirilir.

Kaynak: Amerika Birleşik Devletlerinin bağımsız kuruluşlarından “USAID” Kalkınma Ajansı dokümanından çevrilmiştir. (https://usaidlearninglab.org/system/files/resource/files/mod7_decision_tree_for_selecting_evaluation_design.pdf)