



HACETTEPE ÜNİVERSİTESİ
EĞİTİM BİLİMLERİ ENSTİTÜSÜ

Eğitim Bilimleri Anabilim Dalı
Eğitimde Ölçme ve Değerlendirme Programı

ÇOK KATEGORİLİ PUANLANAN MADDELERDE DEĞİŞEN MADDE VE ADIM
FONKSİYONU BELİRLEME YÖNTEMLERİNİN İNCELENMESİ

Yasemin KUZU

Doktora Tezi

Ankara, 2021

Liderlik, arařtırma, inovasyon, kaliteli eęitim ve deęiřim ile

Daha ileriye... En İyiyeye...



HACETTEPE ÜNİVERSİTESİ
EĞİTİM BİLİMLERİ ENSTİTÜSÜ

Eğitim Bilimleri Anabilim Dalı
Eğitimde Ölçme ve Değerlendirme Programı

ÇOK KATEGORİLİ PUANLANAN MADDELERDE DEĞİŞEN MADDE VE ADIM
FONKSİYONU BELİRLEME YÖNTEMLERİNİN İNCELENMESİ

INVESTIGATION OF DIFFERENTIAL ITEM AND DIFFERENTIAL STEP
FUNCTIONING PROCEDURS IN POLYTOMOUS ITEMS

Yasemin KUZU

Doktora Tezi

Ankara, 2021

Kabul ve Onay

Eđitim Bilimleri Enstitüsü M¼d¼rl¼đ¼ne,

Yasemin KUZU'nun hazırladıđı “Çok Kategorili Puanlanan Maddelerde Deđişen Madde ve Adım Fonksiyonu Belirleme Yöntemlerinin İncelenmesi” başlıklı bu çalıřma j¼rimiz tarafından **Eđitim Bilimleri Ana Bilim Dalı, Eđitimde Ölçme ve Deđerlendirme Bilim Dalında Doktora Tezi** olarak kabul edilmiřtir.

J¼ri Bařkanı	Prof. Dr. H¼lya KELECİOđLU
J¼ri Üyesi (Danıřman)	Prof. Dr. Selahattin GELBAL
J¼ri Üyesi	Prof. Dr. Cem Oktay G¼ZELLER
J¼ri Üyesi	Doç. Dr. Ayfer SAYIN
J¼ri Üyesi	Dr. Öğr. Üyesi Sevda ÇETİN

Bu tez Hacettepe Üniversitesi Lisansüstü Eđitim, Öğretim ve Sınav Yönetmeliđi'nin ilgili maddeleri uyarınca yukarıdaki j¼ri üyeleri tarafından 28/05/2021 tarihinde uygun gör¼lm¼ř ve Enstitü Yönetim Kurulunca / / tarihi itibarıyla kabul edilmiřtir.

Prof. Dr. Selahattin GELBAL
Eđitim Bilimleri Enstitüsü M¼d¼r¼

Öz

Bu çalışmanın amacı çok kategorili maddelerde değişen madde (DMF) ve adım fonksiyonu (DAF) belirleme yöntemlerini çeşitli koşullarda karşılaştırmaktır. Bu bağlamda PISA 2018 kapsamında öğrencilere yönelik hazırlanan “Bilgi ve İletişim Teknolojilerine Aşinalık Anketi” içerisindeki okulda dijital cihazları kullanma sıklığıyla ilgili maddelerin (IC011) yanıtlandığı Kazakistan, Türkiye ve ABD verileri üzerinde çalışılmıştır. Çok kategorili DMF analizleri için Mantel test, Liu- Agresti istatistiği, Cox β ve poly-SIBTEST yöntemleri; DAF analizleri için ise Ardışık Kategori Log Odds Oranı (AC-LOR) ve Kümülatif Kategori Log Odds Oranı (CU-LOR) yöntemleri kullanılmıştır. İlişkisel tarama modelinde yürütülen bu çalışmada “kategori birleştirme kuralı, odak grup- referans grup, odak grup örneklem büyüklüğü, odak grup:referans grup örneklem oranı, DMF belirleme yöntemi” şeklinde koşullar oluşturulmuştur. Koşulların oluşturulmasında SAS ve R yazılımlarından; analizler için ise poly-SIBTEST yöntemi için SIBTEST; diğer yöntemler için DIFAS programlarından faydalanılmıştır. Analizler sonucunda yüksek düzey DMF/DAF sergileyen madde/adım sayısının; çok kategorili DMF yöntemlerine göre küçük örnekleme, DAF yöntemlerine göre ise büyük örnekleme daha fazla olduğu belirlenmiştir. Bununla birlikte yöntemlerin büyük örnekleme daha tutarlı sonuçlar verdiği tespit edilmiştir. Adımlar boyunca zıt işaretli DAF bulduran maddelerde DMF değerinin daha düşük olduğu, dolayısıyla DMF sergilemeyen bir maddede DAF analizi yapılmamasının hatalı sonuçlar verebileceği belirtilmiştir. Araştırma kapsamında oluşturulan kategori birleştirme koşullarının sonuçlar üzerinde sistematik bir etkisinin olmadığı görülmekle birlikte birleştirilen kategorilerin işaretlenme sıklığının sonuçları farklılaştığı düşünülerek sonraki çalışmalarda bu durumun incelenmesi önerilmiştir.

Anahtar sözcükler: çok kategorili değişen madde fonksiyonu, değişen adım fonksiyonu, ardışık yaklaşım, kümülatif yaklaşım, AC-LOR, CU-LOR

Abstract

The aim of this study is to compare the polytomous differential item functioning (DIF) and differential step functioning (DSF) detection methods in polytomously scored items. In this context, it was studied Turkey, Kazakhstan and the USA samples responding to items (IC011) related to frequency of using digital devices in "ICT Familiarity Questionnaire" which was administered to students in within the scope of PISA 2018. Mantel test, Liu-Agresti statistics, Cox β and poly-SIBTEST methods for multi-category DIF analysis; Adjacent Category Log Odds Ratio (AC-LOR) and Cumulative Category Log Odds Ratio (CU-LOR) methods were used for DSF analyzes. The research model is relational screening and conditions such as "category combination rule, focus group-reference group, focus group sample size, sample rating of focus group: reference group, DIF investigation method" were created. In the creation of the conditions, from SAS and R software; in analysis, SIBTEST for poly-SIBTEST method; DIFAS programs were used for other methods. As a result of the analysis, the number of item / steps with high levels of DIF / DSF; it was observed that it was more in the small sample compared to the multi-category DIF methods and in the large sample compared to the DSF methods. However, it was determined that the methods appeared more consistent results in a large sample. Throughout the steps, it was stated that the DIF effect was lower for the substances with DIF of the opposite signs. Therefore, if DSF analysis is not done for items that has no DIF, incorrect results may be obtained. Although, it is seen that the category combining conditions created within the scope of the study do not have a systematic effect on the results, considering choosing frequency of the combined categories differentiated the results, it is suggested that this situation should be examined in future studies.

Keywords: multi-category differential item function, differential step function, adjacent approach, cumulative approach, AC-LOR, CU-LOR

Teşekkür

Öncelikle doktora öğrenimim boyunca bilimsel ve bireysel her konuda bana destek olan, tez çalışmamın her aşamasında yol gösteren, önüme çıkan engelleri aşmamda öncülük eden, doktora sürecinin bana kazandırdığı en değerli insan canım hocam, danışmanım Prof. Dr. Selahattin GELBAL'a sonsuz şükranlarımı sunarım.

Çok kıymetli fikirleri ve nazik eleştirileriyle tezimi şekillendiren, sorularımı içtenlikle yanıtlayan çok değerli hocalarım Prof. Dr. Hülya Kelecioğlu, Prof. Dr. Cem Oktay GÜZELLER, Dr. Sevda ÇETİN ve Doç. Dr. Ayfer SAYIN başta olmak üzere derslerdeki ve sosyal hayatlarındaki anlayışlı tutumlarından dolayı bölümümüzdeki bütün hocalarıma, doktora öğrenimim süresinde Yurtiçi Doktora Bursu kapsamında desteklerini sunan TÜBİTAK'a teşekkürlerimi sunarım.

Tezimin şekillenmesinde değerli fikirlerini esirgemeyen ve yazdıklarımı noktasına, virgülüne kadar bıkmadan tekrar tekrar okuyan canım arkadaşım Dr. Sakine GÖÇER ŞAHİN'e ayrıca teşekkür ederim.

“Önce derslerinizi yapın, başka iş yapmanızı istemiyorum.” diyen, evimizin çınarı, Osmanlı kadını babaannem Afife İÇER'e; “Kız çocuğunu da mı okutuyorsun.” Sözlerini kulak ardı ederek “Hepsi benim evladım, okuyabildiği kadar okusun.” diyen canım babam Mehmet İÇER'e; bizi okuma ruhuyla büyüten canım annem Nesrin İÇER'e; gece gündüz demeden çocuklarıma annelik, babalık, oyun arkadaşlığı yapan fedakâr diğer canım annem Nezaket KUZU ve canım babam Ümüt KUZU'ya; hayatım boyunca desteğini tüm kalbimle hissettiğim başta canım abim Musti'm olmak üzere tüm kardeşlerime; çocuklarıma amcalıktan öte arkadaş olan KUZU kardeşlerime; varlıklarıyla her daim motivasyon kaynağım olan ve hayatıma disiplin katan canım kızlarım Aysima Ece ve Simay Gülce'ye; veeee en önemlisi düştüğümde kaldıran, başardığımda onurlandıran, benimle üzüldü benimle sevinen, doktora hayatım boyunca lojistik ve psikolojik desteğini esirgemeyen canım cancağzım biricik eşim, hayat arkadaşım Okan KUZU'ya minnettarım...

Yasemin KUZU

Canım kızlarım Aysima Ece ve Simay Gülce'ye...

İçindekiler

Öz.....	ii
Abstract	iii
Teşekkür.....	iv
Tablolar Dizini	viii
Şekiller Dizini.....	ix
Simgeler ve Kısaltmalar Dizini	x
Bölüm 1 Giriş.....	1
Problem Durumu	1
Araştırmanın Amacı ve Önemi	4
Araştırma Problemi	5
Alt problemler.....	5
Sayıtlar	6
Sınırlılıklar	6
Tanımlar	7
Bölüm 2 Araştırmanın Kuramsal Temeli ve İlgili Araştırmalar	8
Değişen Madde Fonksiyonu.....	8
Çok kategorili maddeler için değişen madde fonksiyonu belirleme yöntemleri.	10
Değişen Adım Fonksiyonu	15
Değişen adım fonksiyonu belirleme yöntemleri.	17
İlgili Araştırmalar	19
Bölüm 3 Yöntem.....	23
Araştırmanın Türü	23
Çalışma Grubu	23
Veri Toplama Aracı	24
Boyutluluğun incelenmesi.	27
İncelenen Koşullar.....	29
Kategori birleştirme kuralı.	29
Odak grup – Referans grup.....	30
Örneklem büyüklüğü ve odak grup referans grup örneklem oranı.....	30
Çok kategorili DMF/DAF belirleme yöntemleri.....	31
Verilerin Analizi	31

Çok Kategorili DMF Analizleri	31
DAF Analizleri	32
Bölüm 4 Bulgular ve Yorum	33
Alt Problem 1'e İlişkin Bulgular.....	33
Alt Problem 2'ye İlişkin Bulgular.....	48
Alt Problem 3'e İlişkin Bulgular.....	62
Alt Problem 4'e İlişkin Bulgular.....	79
Alt Problem 5'e İlişkin Bulgular.....	90
Alt Problem 6'ya İlişkin Bulgular.....	91
Tartışma	94
Çok kategorili DMF belirleme yöntemlerinin örneklem büyüklüğü ve koşullara göre incelenmesi	94
DAF belirleme yöntemlerinin örneklem büyüklüğü ve koşullara göre incelenmesi	95
Çok kategorili DMF/DAF belirleme yöntemlerinden elde edilen sonuçların birlikte incelenmesi ..	98
Bölüm 5 Sonuç ve Öneriler	100
Sonuçlar	100
Çok kategorili DMF belirleme yöntemlerine ilişkin sonuçlar.....	100
DAF belirleme yöntemlerine ilişkin sonuçlar	100
Çok kategorili DMF ve DAF belirleme yöntemlerinden birlikte elde edilen sonuçlar.	101
Öneriler.....	102
Uygulayıcılar için öneriler	102
Sonraki araştırmalara yönelik öneriler.	102
Kaynaklar	104
Ekler	113
EK-A: "Bilgi ve İletişim Teknolojilerine Aşinalık Anketi" içindeki IC011 kodlu maddeler	113
EK-B: Maddelere İlişkin DMF Miktarının Örneklem Büyüklüğüne Göre Karşılaştırılması	114
EK-C: Madde Adımlarına İlişkin DAF Miktarının Örneklem Büyüklüğüne Göre Karşılaştırılması ..	120
EK-Ç: AC-LOR ve CU-LOR Yöntemlerinin Maddeleri Sınıflamasına İlişkin Benzerlik Oranları	126
EK-D: Etik Komisyonu Onay Bildirimi	127
EK-E: Etik Beyanı.....	128
EK-F: Doktora Tez Çalışması Orijinallik Raporu.....	129
EK-G: Dissertation Originality Report	130
EK-H: Yayımlama ve Fikrî Mülkiyet Hakları Beyanı.....	131

Tablolar Dizini

Tablo 1 <i>Mantel Test Olasılık Tablosu</i>	11
Tablo 2 <i>β Değerinin Yorumlanmasında Kullanılan Kritik Değerler</i>	15
Tablo 3 <i>Veri Toplama Aracına İlişkin Betimsel İstatistikler ve Cronbach Alfa Değerleri</i>	24
Tablo 4 <i>Veri Toplama Aracının Maddelerine İlişkin Betimsel İstatistikler ve Puan Kategori Dağılımı</i> ..	26
Tablo 5 <i>Veri Toplama Aracının Faktör Analizine Uygunluğuna İlişkin Veriler</i>	27
Tablo 6 <i>Açımlayıcı Faktör Analizi Sonucunda Elde Edilen Faktörler ve Açıklanan Varyans Miktarları</i>	28
Tablo 7 <i>Açımlayıcı Faktör Analizi Sonucunda Maddelere İlişkin Faktör Yük Değerleri</i>	28
Tablo 8 <i>Araştırmanın Amacı Kapsamında Oluşturulan Kategori Birleştirme Koşulları</i>	29

Şekiller Dizini

Şekil 1. Tek biçimli DMF ve Tek biçimli olmayan DMF	9
Şekil 2. Odak grup örneklem büyüklüğü 200 iken M4'e ait DMF miktarının değişimi.....	34
Şekil 3. Odak grup örneklem büyüklüğü 200 iken Madde 9'a ait DMF miktarının değişimi.....	39
Şekil 4. Odak grup örneklem büyüklüğü 200 iken Madde 10'a ait DMF miktarının değişimi.....	44
Şekil 5. Odak grup örneklem büyüklüğü 1000 iken M4'e ait DMF miktarının değişimi.....	49
Şekil 6. Odak grup örneklem büyüklüğü 1000 iken Madde 9'a ait DMF miktarının değişimi.....	53
Şekil 7. Odak grup örneklem büyüklüğü 1000 iken Madde 10'a ait DMF miktarının değişimi.....	58
Şekil 8. Odak grup örneklem büyüklüğü 200 iken M4'e ait DAF miktarının değişimi.	63
Şekil 9. Odak grup örneklem büyüklüğü 200 iken Madde 9'a ait DAF miktarının değişimi.	69
Şekil 10. Odak grup örneklem büyüklüğü 200 iken Madde 10'a ait DAF miktarının değişimi.	74
Şekil 11. Odak grup örneklem büyüklüğü 1000 iken M4'e ait DAF miktarının değişimi.	80
Şekil 12. Odak grup örneklem büyüklüğü 1000 iken Madde 9'a ait DAF miktarının değişimi.	83
Şekil 13. Odak grup örneklem büyüklüğü 1000 iken Madde 10'a ait DAF miktarının değişimi.	87

Simgeler ve Kısaltmalar Dizini

ABD : Amerika Birleşik Devletleri

ATM : Aşamalı Tepki Modeli

DAF : Değişen Adım Fonksiyonu

DIF : Differential Item Functioning

DMF : Değişen Madde Fonksiyonu

DSF : Differential Step Functioning

KAZ : Kazakistan

KPM : Kısmi Puan Modeli

TUR : Türkiye

Bölüm 1

Giriş

Problem Durumu

Teknoloji ve iletişim araçlarının hızla gelişmesiyle dünya bir küreselleşme sürecine girmiş ve ülkeler arasında siyasal, toplumsal ve kültürel etkileşim artmıştır. Bu etkileşim modern çağın gerektirdiği becerilere ayak uyduran bireyler yetiştirmeyi daha da önemli kılmaktadır. Ülkelerin aynı ölçüte dayanarak, özellikle eğitim alanında dünyadaki konumunu görme, eğitim politikalarını değerlendirme ve bu politikalara şekil verme amacıyla uluslararası geniş ölçekli sınavlara (PIRLS, PISA, TIMSS vb.) katılım her geçen gün artmaktadır. Ülkelerin kendi eğitim politikaları sonucunda öğrencilerin edindikleri bilgi ve beceri düzeylerini belirleyerek diğer ülkelerdeki öğrencilerle kıyaslama, eğitim politikalarının güçlü ve zayıf yönlerini belirleme ve eğitim reformu gerçekleştirirken ölçütler oluşturmada PISA uygulamasından elde edilen sonuçlardan büyük ölçüde etkilendikleri görülmektedir (Ababneh ve diğerleri, 2016; Baird ve diğerleri, 2016; Simola, 2005; Tienken, 2016). Dünya çapında bu kadar önemsenen PISA, Ekonomik İşbirliği ve Kalkınma Örgütü (OECD) tarafından rutin olarak üçer yıllık dönemler halinde, öğrencilerin kazanmış oldukları bilgi ve becerileri değerlendiren bir çalışmadır. Uygulamada matematik, fen ve okuma olmak üzere toplamda üç alandan soru sorulmakla birlikte her uygulamada bu alanlardan biri ağırlıklı alan olarak belirlenmektedir. Örneğin ilk uygulama olan PISA 2000 uygulamasında okuma alanı; ikinci uygulamada (PISA 2003) matematik ve üçüncü uygulamada (PISA 2006) fen alanı ağırlıklı alan olmuştur. Bu döngü bu şekilde devam etmiş olup PISA 2018 uygulamasında okuma becerileri ağırlıklı alan olarak karşımıza çıkmıştır. 15 yaş grubundaki öğrencilere hitap eden bu uygulamada ayrıca öğrencilerin motivasyonları, kendileri hakkındaki görüşleri, öğrenme biçimleri, okul ortamları ve aileleri ile ilgili veriler toplanmaktadır (MEB, 2020). Bu bağlamda bir yandan öğrencilerin analiz etme, akıl yürütme ve etkin iletişim kurabilme becerileri ölçülürken; diğer yandan öğrenci, veli ve okul yönetimine uygulanan anketlerle de öğrenci düzeyini olumlu ya da olumsuz yönde etkileyen faktörler araştırılmaktadır. Son yıllarda teknolojiye yaşanan büyük gelişmelerle birlikte PISA uygulamasında bilgisayar tabanlı uygulamaya başlanmış; PISA 2018 uygulamasına katılan 79 ülkenin 70'inde bilgisayar tabanlı test, 9'unda ise kâğıt kalem testi uygulanmıştır (OECD, 2019). Bilgisayar tabanlı testin uygulanması,

öğrencilerde bilgi ve iletişim teknolojilerine aşinalığın ne düzeyde olduğu sorusunu beraberinde getirmiştir. Bu duruma açıklık getirmek üzere öğrencilerin bilgisayar kullanma durumlarını ve bilgisayar kullanımına yönelik tutumlarını belirlemek üzere; bilgi ve iletişim teknolojilerinin ulaşılabilirliği ve kullanımına yönelik “Bilgi ve İletişim Teknolojilerine Aşinalık Anketi” uygulanmıştır. Alanyazında bilgi ve iletişim teknolojileri kullanımının öğrencilerin eğitim çıktıları üzerinde olumlu/olumsuz etkisi olduğu bildirilmektedir (Leino, 2014; Luu & Freeman, 2011; Tezoh, 2015).

Görüldüğü üzere bu sınavların ve sınavlardan elde edilen puanların amacına hizmet etmesi büyük önem taşımaktadır. Diğer bir ifadeyle yapılan yorumların doğru sonuçları; test puanlarının ise bireylerin gerçek puanlarını yansıtması için geçerli ölçmelere ihtiyaç duyulmaktadır. Test puanlarına dayalı yapılan yorumları ya da alınan kararları destekleyen teori ve kanıtların bir derecesi (AERA ve diğerleri, 2014) olan geçerlik, ölçme araçlarında bulunması gereken en önemli özelliklerden biridir. Testler, ölçülen yapıyla ilgisi olmayan değişkenleri dikkate almaksızın tüm bireyler için aynı doğrulukla ölçme yapmalıdır (Sireci & Rios, 2013). Herkes için aynı şeyi ifade etmeyen diğer bir ifadeyle amacına hizmet etme derecesi gruplara veya ülkelere göre değişen bir test ile bu ülkeleri veya grupları aynı ölçek düzeyine yerleştirip kıyaslamak yanıltıcıdır. Bu açıdan test maddelerinin ölçtüğü özelliğin bireyler, gruplar, ülkelere göre değişmez olması gerekir. Maddelerin değişmezliği, maddelerin yanıtlanma olasılıklarının aynı özellikteki gruplara göre değişmemesidir. Madde ve test yanlılığı geçerliği etkileyen en önemli tehdit unsurları olarak karşımıza çıkmaktadır (Clauser & Mazor, 1998).

Yanlılık, testteki bir maddeye aynı yetenek düzeyindeki bireyler tarafından verilen yanıtların ölçülmek istenen özellikten farklı sebeplere bağlı olarak değişkenlik göstermesidir. Bu durum ölçme sonuçlarına ölçülmek istenen özellik dışında başka değişkenlerin karıştığı anlamına gelmektedir. Diğer bir ifadeyle yanlılık ölçülmek istenen yapıyla ilgili olmayan unsurların, belirlenen gruptaki bireyler için sistematik olarak daha iyi ya da daha kötü sonuçlar doğurmasıdır (Wood, 2011). Eğitimde ve psikolojide uygulanan testlerde yapıyla ilgisiz unsurlar ise çoğunlukla cinsiyet, din, etnik köken, sosyoekonomik durum, ırk, farklı dile uyarlamada çeviri problemleri, ölçme uygulamasının yapılma koşulları, aynı psikolojik yapının kültürlere göre farklılık göstermesi vb. şekillerde görülmektedir (Allalouf ve diğerleri, 1999; Dancer ve diğerleri, 1994; Feng, 2008; Henderson, 1999; Kabasakal, 2014; Özberk, 2016; Sireci

& Khaliq, 2002; Smith, 2009; Taylor & Lee, 2012; Woods, 2011). Psikolojik ölçmelerde bu unsurların tamamen kontrol altına alınması imkânsızdır. Ancak bazı istatistiksel yöntemler yardımıyla belirlenen gruplardan herhangi birine avantaj ya da dezavantaj oluşturmasının önüne geçmek mümkündür. Bir testte herhangi bir gruba avantaj ya da dezavantaj oluşturan maddelerin belirlenerek testin bu maddelerden arındırılması ya da revize edilmesi, yapılan ölçme ve değerlendirmenin doğru sonuçlar vermesi adına son derece önemlidir.

Bir testte yanlılık içeren maddelerin tespitinde ilk olarak maddelerde değişen madde fonksiyonu (DMF) olup olmadığı belirlenmelidir. DMF, farklı alt gruplarda yer alan aynı yetenek düzeyindeki bireylerin bir maddeyi doğru cevaplama olasılığının gruplara göre farklılık göstermesidir (Embretson & Reise, 2000; Hambleton ve diğerleri, 1991). Alanyazında DMF çalışmaları incelendiğinde ilk olarak iki kategorili maddeler üzerinde çalışıldığı, son yıllarda ise performansa dayalı değerlendirmenin yaygınlaşmasıyla iki kategorili maddelerle birlikte çok kategorili maddelerde de DMF tespitine odaklanıldığı görülmektedir. İki kategorili maddelerin aksine çok kategorili maddelerde yanıt kategorilerinin sayısı nedeniyle DMF farklı biçimler alabilmektedir. Çok kategorili bir maddede yanıt kategorisi sayısının bir eksiği kadar adım fonksiyonu tanımlanmaktadır. Tanımlanan bu adım fonksiyonlarının ölçülen özelliklerin cinsiyet, ırk, sosyo-ekonomik düzey gibi demografik özelliklere göre farklılık göstermemesi beklenir. Çok kategorili maddelerde belirli bir adımda ölçülen özelliklerde görülen gruplar arası fark değişen adım fonksiyonu (DAF) olarak tanımlanmaktadır (Penfield, 2007). Bu bağlamda madde düzeyinde tek bir değişmezlik hesaplamak yerine, çok kategorili maddenin her bir puan düzeyinin değişmezliği göz önünde bulundurulmalıdır. Bu yaklaşım, değişmezlik ihlalinin hangi puan düzeylerinden kaynaklandığını belirleme imkânına sahiptir. Testte yer alan maddelerin bazılarının bir grup (örneğin kadınlar) bazılarının da diğer grubun (örneğin erkekler) lehine olduğu durumda DMF miktarının ve yönünün benzer olması durumunda testteki toplam DMF miktarı birbirini etkisiz hale getirecektir (Wyse, 2013). Böyle bir durumda DMF'nin öğrencilerin üzerine bir etkisi olmasa da test geliştiricilere test geliştirme sürecinin kalitesi ve bundan sonra bu sürecin nasıl ilerlemesi gerektiği hakkında da bilgi vermeyecektir. Maddelerde DAF tespiti ise işaret veya büyüklüklerine göre ters yönde olan DMF etkilerinin tespit edilememesi yönünden ve madde revizyonunda hangi kategoriler üzerinde durulacağına karar verilmesi adına oldukça önemlidir. Diğer yandan DMF

yöntemleriyle tespit edilen DMF'li maddelerde DMF'nin madde düzeyinde mi yoksa adım düzeyinde mi olduğunun bilinmesi, maddenin çıkarılması ya da revize edilmesinde yol göstericidir. Geçerleme sürecinde DAF çerçevesinin kullanılması test geliştiren kişiye gruplar arasındaki farklılıklara yol açan madde içerik ve bileşenleri hakkında güçlü araçlar sunmaktadır. Bu bağlamda uluslararası düzeyde yapılan PISA 2018 uygulamasında çok kategorili maddelerin DMF ve DAF analiziyle istatistiksel olarak incelenmesi ve yöntemlerin DMF/DAF'li olarak işaretledikleri maddelerin karşılaştırılması amaçlanmıştır.

Araştırmanın Amacı ve Önemi

Bu çalışmanın amacı çok kategorili maddelerde DMF ve DAF belirleme yöntemlerini karşılaştırmaktır. Bu bağlamda PISA 2018 kapsamında öğrencilere yönelik hazırlanan "Bilgi ve İletişim Teknolojilerine Aşinalık Anketi" içerisindeki okulda dijital cihazları kullanma sıklığıyla ilgili maddelerin (IC011) yanıtlandığı Kazakistan, Türkiye ve ABD verileri üzerinde çalışılmıştır. Ülkelerin seçiminde öncelikle PISA 2018 sonuçlarına göre ağırlıklı alanda ülke sıralamaları incelenmiş ve ülkeler düşük, orta ve yüksek düzey olarak üç gruba ayrılmıştır. Aynı zamanda ülkelerin ekonomik durumları da göz önünde bulundurularak, her iki durum (başarı ve ekonomik düzey) için de düzeyi düşük olan gruptan ilgili anketi yanıtlayan Kazakistan, orta düzey gruptan Türkiye ve üst düzey gruptan ABD tercih edilmiştir. Çok kategorili DMF analizleri için aşağıda detaylı olarak anlatılan Mantel test, Liu- Agresti istatistiği, Cox β ve poly-SIBTEST yöntemleri; DAF analizleri için ise Ardışık Kategori Log Odds Oranı (AC-LOR) ve Kümülatif Kategori Log Odds Oranı (CU-LOR) yöntemleri kullanılmıştır.

Çok kategorili maddelerde DMF tespiti iki kategorili maddelerdeki DMF tespitine göre daha karmaşıktır. Çok kategorili maddelerde değişmezlik göz önüne alındığında, değişmezlik biçimi puan düzeylerinde farklılık gösterebilir. Öyle ki; bir puan düzeyinde değişmezlik sağlanamazken diğer puan düzeylerinde sağlanabilir; maddede değişmezliğin sağlanamadığı durumlarda, bir puan düzeyinde referans grubun lehine, başka bir puan düzeyinde ise odak grubu lehine DMF görülebilir (Penfield, Alvarez ve diğerleri, 2008).

Çok kategorili maddelerde yaygın olarak kullanılan DMF belirleme yöntemleri değişmezliğin madde düzeyinde incelenmesine dayanmaktadır (Penfield & Lam, 2000). Mantel'in ki-kare istatistiği (Mantel, 1963), Genelleştirilmiş Mantel Haenszel

(GMH) istatistiđi (Somes, 1986) gibi yaklařımlarda tm puan dzeyinde toplanan genel deđiřmezlik lldđ iin tek bir DMF indeksi verilmektedir. Bu durumda DMF'nin hangi puan dzeyinden kaynaklandığı belirlenememektedir. Dolayısıyla, yapılan DMF analizlerinden sonra uzmanların bir araya gelerek yrttđ; olası DMF nedenlerini belirleme alıřmaları ve madde revizyonu sreci daha az verimli gemekte, bu durum hem zaman hem de ekonomik olarak daha zorlayıcı hale gelmektedir.

DAF analizleri, ok kategorili maddelerde kapsamlı bir DMF analizi iin nemli bir bileřendir. Son yıllarda arařtırmacılar ok kategorili maddelerde deđiřmezlik formunu incelerken birok sebep gstererek, tek bir toplam puan dzeyi yerine her bir puan dzeyinin dikkate alınması gerektiđini belirtmektedirler (Gattamorta & Penfield, 2012). Bu sebeplerden birincisi, DAF etkisi ok kategorili bir maddenin adımlarında iřaret veya byklk olarak deđiřtiđinde, poly-SIBTEST, Standart Ortalama Farkı (SMD) gibi birok omnibus DMF yntemlerinin nispeten dřk g gstermesidir (Penfield & Algina, 2003; Wang & Su, 2004). İkincisi, omnibus DMF yntemleri tm adımlarda toplanmıř DAF'yi temsil eden bir deđer verir ve bu nedenle sadece bir adımın byk bir DAF miktarına sahip olması veya DAF'nin adımlar boyunca zıt iřaretili olması durumunda belirli adımlarda byk DAF miktarlarının gzden kamasıdır. rneđin; bir adımda .3'lk, bařka bir adımda -.3'lk DAF gsteren bir maddenin sonu itibariyle analizlerde tespit edilen DMF miktarı sıfır (0) olacaktır. Dolayısıyla her adım iin DAF'nin hesaplanması nemli bilgilerin fark edilmesini sađlayacaktır. Son olarak byle bir yaklařımla deđiřmezliđin ihlalinden hangi puan dzeylerinin sorumlu olduđu anlařılabilecek, dolayısıyla DMF'nin olası nedenleri hakkında bilgi sahibi olunacaktır.

Arařtırma Problemi

lkelere gre (TUR-KAZ, TUR-ABD) ok kategorili DMF/ DAF yntemleriyle elde edilen DMF/DAF miktarları; deđiřen kategori birleřtirme kuralı, odak grup rneklem byklđ (200, 1000) ve odak grup: referans grup rneklem oranlarında (2:1,1:1,1:2,1:3) farklılık gstermekte midir?

Alt problemler. Arařtırma problemine ait alt problemler ařađıda sunulmuřtur:

1. Odak grup rneklem byklđ 200 (kk) olduđunda, lke karřılařtırmalarında ok kategorili DMF yntemleriyle elde edilen DMF

- miktarları; deęişen kategori birleřtirme kuralı ve odak grup:referans grup örneklem oranlarında farklılık göstermekte midir?
2. Odak grup örneklem büyüklüęü 1000 (büyük) olduęunda, ülke karřılařtırmalarında çok kategorili DMF yöntemleriyle elde edilen DMF miktarları; deęişen kategori birleřtirme kuralı ve odak grup:referans grup örneklem oranlarında farklılık göstermekte midir?
 3. Odak grup örneklem büyüklüęü 200 (küçük) olduęunda, ülke karřılařtırmalarında DAF yöntemleriyle elde edilen DAF miktarları; deęişen kategori birleřtirme kuralı ve odak grup:referans grup örneklem oranlarında farklılık göstermekte midir?
 4. Odak grup örneklem büyüklüęü 1000 (büyük) olduęunda, ülke karřılařtırmalarında DAF yöntemleriyle elde edilen DAF miktarları; deęişen kategori birleřtirme kuralı ve odak grup:referans grup örneklem oranlarında farklılık göstermekte midir?
 5. Ülke karřılařtırmalarında; deęişen kategori birleřtirme kuralı ve odak grup:referans grup örneklem oranlarında çok kategorili DMF yöntemleriyle elde edilen DMF miktarları; odak grup örneklem büyüklüęüne göre farklılık göstermekte midir?
 6. Ülke karřılařtırmalarında; deęişen kategori birleřtirme kuralı ve odak grup:referans grup örneklem oranlarında DAF yöntemleriyle elde edilen DMF miktarları; odak grup örneklem büyüklüęüne göre farklılık göstermekte midir?

Sayıtlılar

Katılımcıların; PISA 2018 uygulamasına aynı kořullar altında katıldıkları kabul edilmiřtir.

Sınırlılıklar

Bu çalıřma;

1. PISA 2018 uygulamasında yer alan “Bilgi ve İletişim Teknolojilerine Ařinalık Anketi” içerisindeki okulda dijital cihazları kullanma sıklığıyla ilgili maddelere (IC011) verilen öğrenci yanıtlarıyla,

2. PISA 2018 uygulamasında elde edilen Kazakistan, Türkiye ve ABD verileriyle,
3. Yapılan karşılařtırmalarda kullanılan yöntemlerle sınırlıdır.

Tanımlar

DMF: Farklı alt gruplarda yer alan aynı yetenek düzeyindeki bireylerin bir maddeyi doğru cevaplama olasılığının gruplara göre farklılık göstermesidir.

DAF: Çok kategorili bir maddenin belli bir adımındaki ölçme özelliklerinde görülen gruplar arası farktır.

Bölüm 2

Araştırmanın Kuramsal Temeli ve İlgili Araştırmalar

Değişen Madde Fonksiyonu

Bir testteki maddelerin yanlılığının araştırılmasında ilk adım maddelerde DMF olup olmadığını belirlemektir. Bir testte olması beklenen durum, testte ölçülmek istenen özellik bakımından aynı düzeyde olan bireylerin bir maddeyi doğru yanıtlama olasılıklarının da aynı olmasıdır (Dorans & Holland, 1993). DMF, farklı alt gruplarda yer alan aynı yetenek düzeyindeki bireylerin bir maddeyi doğru cevaplama olasılığının gruplara göre farklılık göstermesidir (Embretson & Reise, 2000; Hambleton ve diğerleri, 1991). Diğer bir ifadeyle DMF, maddenin ölçmeyi amaçladığı yetenek düzeyi için yapılan bir karşılaştırmada farklı gruplardaki bireylerin maddeyi doğru yanıtlama olasılıklarındaki değişiklikleri açıklamaktadır (Zumbo, 1999). DMF analizleri madde yanlılığını belirlemede ön koşul olup aynı zamanda geçerlik için bir kanıttır (Embretson, 2007). Yanlı bir maddede DMF varlığı kesin olmakla birlikte bir maddedeki DMF varlığı o maddenin yanlı olduğu anlamına gelmemektedir. DMF varlığı saptanan bir maddenin yanlı olup olmadığına ancak uzman görüşüyle karar verilebilir (Zumbo & Gelin, 2005), dolayısıyla bu işlem temel olarak niteldir (Ellis & Raju, 2003; Furlow ve diğerleri, 2009; Sireci & Allalouf, 2003).

DMF çalışmalarında dezavantajlı grup (odak) ile avantajlı grup (referans) olmak üzere iki grup üzerinde çalışılır (Osterlind, 1983). DMF analizlerinde amaç aynı yetenek düzeyinde fakat farklı gruplardaki bireylerin bir maddeyi doğru yanıtlama olasılıklarını incelemek olduğundan, öncelikle odak ve referans gruplar ölçülen yetenek bakımından aynı yetenek düzeyinde eşitlenir. DMF içermeyen bir madde her iki grup için de aynı ölçme özelliklerini göstermelidir. Aksi takdirde maddenin DMF'li olduğuna karar verilerek hangi tür DMF içerdiği incelenir.

DMF, tek biçimli ve tek biçimli olmayan DMF olmak üzere iki farklı şekilde ele alınmaktadır. DMF gösteren bir madde için, bir grup her yetenek düzeyinde diğer gruptan daha düşük veya daha yüksek performans gösteriyorsa burada tek biçimli DMF söz konusudur. Tek biçimli DMF, yetenek düzeyi ile grup üyeliği arasında madde performansında etkileşim olmadığı durumlarda ortaya çıkmaktadır. Dolayısıyla tek biçimli DMF olduğunda gruplar arasında yalnızca madde güçlük parametresi farklılık gösterir. Kısacası gruplar arasındaki performans farklılığı kesiksiz bütün yetenek alanı

boyunca tek biçimli olduğunda, tek biçimli DMF'nin varlığından bahsedilmektedir (Penfield & Lam, 2000). Tek biçimli olmayan DMF'de ise; DMF gösteren bir madde için, bir grup diğer gruba göre bazı yetenek düzeylerinde daha yüksek bazı yetenek düzeylerinde ise daha düşük performans göstermektedir. Yani yetenek düzeyi farklılaştıkça, odak ve referans grupların maddeyi doğru yanıtlama olasılıkları da farklı düzey ve yönlerde farklılık göstermektedir (Osterlind & Everson, 2009). Diğer bir deyişle grup üyeliği ile yetenek düzeyi arasında etkileşim olduğu durumlarda tek biçimli olmayan DMF'nin varlığından bahsedilebilir (Camilli & Shepard, 1994). Tek biçimli olmayan DMF'de alt gruplar arasında madde ayırt ediciliği farklıyken güçlük parametresi için aynı şey söylenemez (Turhan, 2006).



Şekil 1. Tek biçimli DMF ve Tek biçimli olmayan DMF.

Şekil 1'de görüldüğü gibi tek biçimli olmayan DMF için gruplara ait madde karakteristik eğrileri birbirini keserken; tek biçimli DMF'de böyle bir çakışma söz konusu değildir.

DMF'nin varlığı bir testin geçerliği için olumsuz bir tehdit unsuru olmakla birlikte DMF gösteren bir maddenin hemen ölçekten çıkarılması doğru değildir (Zumbo & Gelin, 2005). Öncelikle DMF gösteren maddeler istatistiksel analizler sonucunda DMF düzeyi bakımından A (ihmal edilebilir), B (orta düzeyde) ve C (yüksek düzeyde) kategorilerine ayrılır (Zieky, 1993). A kategorisindeki maddelerin hepsi ve B kategorisindeki maddelerin önemli olanları testte kullanılmalı; C kategorisindeki maddeler ise testte ölçülen özellik için çok önemli değilse kullanılmamalıdır (Zwick, 2012). Ancak alanyazında farklı DMF yöntemlerinin belirlediği maddeler ve bu maddelerin DMF düzeylerinin farklı olmasından dolayı sadece bir yöntemle göre analiz yapılması ve sonuçların bir yöntemle göre yorumlanmasının uygun olmadığı belirtilmektedir (Karami & Nodoushan, 2011). Dolayısıyla eğer bir madde birden fazla

yönteme göre DMF içeriyorsa, bu maddenin DMF'li olduğuna yönelik daha fazla gerekçe elde edilmiş olur.

Alanyazında birçok DMF belirleme yönteminden bahsedildiği, ayrıca bu yöntemlerin farklı kaynaklarda farklı şekillerde sınıflandırıldığı görülmektedir. Örneğin; Camilli ve Shepard (1994); DMF belirleme yöntemlerini, madde tepki kuramına dayalı yöntemler, olasılık tablosuna dayalı yöntemler ve varyans analizine dayalı yöntemler olmak üzere üç sınıfta ele almıştır. Diğer yandan Zumbo (2007) ise DMF belirleme yöntemlerinin Çok Boyutlu Yöntemler, Madde Tepki Kuramına Dayalı Yöntemler ve Olasılık Tablosuna Dayalı Yöntemler şeklinde sınıflandırılabilirliğini belirtmiştir. Bu sınıflamalardan farklı olarak yöntemlerin Klasik Test Kuramı ve Madde Tepki Kuramına Dayalı Yöntemler şeklinde iki başlık altında incelendiği de görülmektedir (Ellis & Raju, 2003). Klasik Test Kuramına (KTK) dayalı yöntemlerde odak ve referans grupların eşleştirilmesi ham puanlar üzerinden yapılırken; Madde Tepki Kuramına dayalı yöntemlerde grupların eşleştirilmesinde yetenekler kullanılmaktadır (Emenogu & Childs, 2003). KTK'ya dayalı yöntemler incelendiğinde çoğunlukla Mantel-Haenszel (MH), Varyans Analizi (ANOVA), lojistik regresyon (LR), Korelasyona Dayalı Yöntemler, Simultaneous Item Bias Test (SIBTEST) ve ki-kare istatistikleri kullanılırken, MTK'ya dayalı olarak daha çok Lord'un ki-karesi, Olabilirlik Oranı yöntemleri ve Raju'nun alan ölçümleri yöntemleri ile karşılaşılmaktadır (Camilli & Shepard, 1994; Ellis & Raju, 2003; Holland & Wainer, 1993).

Çok kategorili maddeler için değişen madde fonksiyonu belirleme yöntemleri. Çok kategorili maddeler için kullanılan DMF belirleme yöntemleri daha karmaşık yapıda olması nedeniyle (Maydeu-Olivares ve diğerleri, 1994) alanyazında daha az kullanılmakla birlikte bir bakıma iki kategorili puanlanan maddelerde kullanılan DMF belirleme yöntemlerinin uzantısı şeklindedir. Bu bağlamda geliştirilen çok kategorili DMF yöntemlerine örnek olarak Mantel Test (Mantel, 1963), poly-SIBTEST (Chang ve diğerleri, 1996), Genelleştirilmiş Mantel-Haenszel (Zwick ve diğerleri, 1993), Lojistik Diskriminant Fonksiyon Analizi (Miller & Spray, 1993), Liu Agresti istatistiği (Liu & Agresti, 1996), HW1 ve HW3 istatistiği (Welch & Hoover, 1993) ve Cox's β istatistiği (Camilli & Congdon, 1999) istatistikleri verilebilir. Bu çalışmada yalnızca araştırma sorularını yanıtlamada kullanılan yöntemlere ilişkin detaylı bilgilere yer verilmiştir. Bu yöntemlerin tercih edilmesinde; yöntemlerin parametrik yöntemlerin gerektirdiği katı varsayımlardan bağımsız olması ve bu yöntemlerle görece daha

küçük örneklerde analiz yapılabilmesi (Penfield & Algina, 2003) etkili olmuştur. Nitekim parametrik yöntemlerde parametre tahminlerinin yeterli gücü ve kararlılığı sağlaması için nispeten daha büyük örneklemelere ihtiyaç duyulmaktadır (French & Miller, 1996). Diğer yandan parametrik yöntemler model veri uyumuna dayanmakta; bu uyumun sağlanmaması Tip I hata oranının artırmaktadır (Bolt, 2002). Parametrik olmayan yöntemlerin seçilmesinde diğer bir sebep ise parametrik yöntemlerin daha zaman alıcı olması ve analizlerin maddenin her bir adımı için tek tek yapılmasını gerektirmesidir (Penfield, 2007).

Mantel test. Mantel test istatistiği, MH testinin bir uzantısı olup sıralama ölçeği düzeyindeki değişkenler üzerindeki eşleşen gruplar arasındaki ilişkiyi saptamak için geliştirilmiştir (Mantel, 1963). Daha önce ikili puanlanan maddeler üzerinde hesaplanan bu istatistiğin 90'lı yıllarda çok kategorili maddeler üzerinde DMF belirlemek amacıyla kullanılması önerilmiş (Dorans & Holland, 1993; Zwick ve diğerleri, 1993) ve farklı çalışmalarda kullanıldığı görülmüştür (Chang ve diğerleri, 1996; Welch & Hoover, 1993). Mantel test ile DMF analizinde bir serbestlik derecesinde ki-kare dağılımına ilişkin istatistik ile yokluk hipotezi test edilir. Yöntemin DMF analizlerinde kullanılmasında madde yanıt kategorilerinin sıralı, odak ve referans grupları için madde ortalamalarının karşılaştırılabilir olduğu varsayılır (Zwick ve diğerleri, 1993).

Mantel test analizi için öncelikle veriler, J yanıt kategori sayısı, K ise eşleştirme değişkeninin düzeylerinin sayısı olmak üzere üç boyutlu $2 \times J \times K$ matrisi şeklinde düzenlenir. Tablo 1'de gösterildiği gibi bütün K düzeyleri için $2 \times J$ olasılık tablosu oluşturulabilir.

Tablo 1

Mantel Test Olasılık Tablosu

Grup	Madde puanı					Toplam
	y_1	y_2	y_3	...	y_j	
Referans	n_{R1k}	n_{R2k}	n_{R3k}	...	n_{Rjk}	n_{R+k}
Odak	n_{F1k}	n_{F2k}	n_{F3k}	...	n_{Fjk}	n_{F+k}
Toplam	n_{+1k}	n_{+2k}	n_{+3k}	...	n_{+jk}	$n_{++k} (N_k)$

Mantel test olasılık tablosunda $y_1, y_2, y_3, \dots, y_j$ değerleri maddeden elde edilebilecek olası J puanlarını, n_{RTk} ve n_{FTk} değerleri ise sırasıyla referans ve odak grupları temsil etmek üzere, eşleştirme değişkeninin k. düzeyinde olan ve madde puanı y_j olan bireylerin sayısını göstermektedir. "+" ise belirli bir indeks üzerindeki

toplama göstermektedir. Örneğin n_{R+k} k. düzeydeki referans grubu üyelerinin toplam sayısını vermektedir. Bu bağlamda Mantel testi analizine ait denklem (1) şu şekildedir (Zwick ve diğerleri, 1993):

$$Mantel \chi^2 = \frac{(\sum_k F_k - \sum_k E(F_k))^2}{\sum_k Var(F_k)}, \quad (1)$$

F_k : odak grubu için eşleştirme değişkeninin k. düzeyindeki puanların toplamı

$E(F_k)$: beklenen odak grup puanları

$Var(F_k)$: odak grup puanlarının varyansı

Mantel test istatistiği bir serbestlik derecesinde ki-kare dağılımına sahiptir. Bu test sonucunda yokluk hipotezinin reddedilmesi, ölçülen özellik bakımından benzer düzeyde olan odak ve referans grup üyelerinin söz konusu maddedeki performanslarının farklılaştığını, dolayısıyla maddenin DMF içerdiğini göstermektedir. Mantel test istatistiği için kritik değer; .05 anlamlılık düzeyinde 3,84, .01 anlamlılık düzeyinde ise 6,63 olarak belirtilmiştir (Penfield, 2013). Mantel test istatistiğinin pozitif olması DMF'nin referans grup, negatif olması durumunda ise odak grup lehine olduğu anlamına gelmektedir. Mantel test istatistiğine dayanarak DAF büyüklüklerinin A, B ve C şeklinde kategorilere ayrılarak yorumlanmasına ilişkin herhangi bir ölçüt bulunmamaktadır. Bu nedenle Mantel yöntemine dayanarak yüksek veya düşük düzeyde DMF şeklinde yorum yapmak mümkün olmadığından, DMF büyüklükleri ancak birbiri ile kıyaslanarak değerlendirilmektedir.

Liu Agresti istatistiği. Liu Agresti istatistiği, diğer Mantel-Haenszel tabanlı yöntemler kadar yaygın olmasa da çok kategorili maddelerde DMF analizi için önerilen bir yöntemdir (Penfield & Algina, 2003). Liu ve Agresti (1996), sıralı yanıt değişkenlerinin tüm K düzeyleri için $\hat{\alpha}_{MH}$ 'nin bir genellemesi olarak common odds oranının bir tahmin edicisini tanımlamıştır. Bu yöntem bir odds oranının tahmin edicisi olduğu için, maddede bulunan yanıt kategorilerinin sayısından nispeten bağımsız olarak bir etki büyüklüğü hesaplayarak, iki ve çok kategorili maddelerde etki büyüklüklerinin karşılaştırılmasına olanak sağlayabilmektedir (Penfield & Algina, 2003). Liu Agresti istatistiğini hesaplamak için öncelikle odds oranının hesaplanması gerekmektedir. Tablo 1 kullanılarak j. adımda referans grubun odak gruba göre başarı odds oranı şu şekilde tahmin edilir:

$$\hat{\alpha}_j = \frac{\sum_{k=1}^K A_{jk} D_{jk} / N_{jk}}{\sum_{k=1}^K B_{jk} C_{jk} / N_{jk}}. \quad (2)$$

A_{jk} / B_{jk} : j. adımda başarılı/başarısız olan referans grup üyelerinin sayısı

C_{jk} / D_{jk} : j. adımda başarılı/başarısız olan odak grup üyelerinin sayısı

N_{jk} : $A_{jk} + B_{jk} + C_{jk} + D_{jk}$

Liu Agresti istatistiği, denklem (2)'nin K eşleştirme değişkeninin tüm düzeyleri için ayrı ayrı hesaplanıp toplanmasıyla elde edilir. Öyle ki;

$$\hat{\psi}_{LA} = \frac{\sum_{k=1}^K \sum_{j=1}^J A_{jk} D_{jk} / N_{jk}}{\sum_{k=1}^K \sum_{j=1}^J B_{jk} C_{jk} / N_{jk}} \quad (3)$$

dir. $\hat{\alpha}_{MH}$ ile aynı ölçekte bir tahminci üretmek için;

$$\hat{\alpha}_{LA} = \frac{1}{\hat{\psi}_{LA}} \quad (4)$$

dönüşümü yapılarak doğal logaritması alınır. Eğer $\ln(\hat{\alpha}_{LA}) = 0$ ise DMF yok, $\ln(\hat{\alpha}_{LA}) > 0$ ise referans grup lehine, $\ln(\hat{\alpha}_{LA}) < 0$ ise odak grup lehine DMF olduğu şeklinde yorumlanır (Penfield & Algina, 2003).

Cox's β istatistiği. Cox's β istatistiği matematiksel olarak Mantel teste eşit fakat kavramsal olarak farklı bir yaklaşım (Cox, 1958) olup, verilerin β parametresine sahip merkezi olmayan çok değişkenli hipergeometrik dağılımdan geldiğini varsaymaktadır. β değeri Mantel test olasılık tablosu kullanılarak şu şekilde hesaplanmaktadır (Camilli & Congdon, 1999). τ_k ve ζ_k^2 çok kategorili bir maddede sırasıyla k. toplam puan düzeyi için hipergeometrik ortalama ve varyans olmak üzere;

$$\tau_k = \frac{n_{R+k} n_{Jk}}{N_k}, \quad (5)$$

$$\zeta_k^2 = \frac{n_{R+k} n_{F+k}}{N_k} \text{Var}(j), \quad (6)$$

dir. Bu bağlamda β değeri ve varyansı;

$$\hat{\beta} = \frac{\sum_k \sum_j J(n_{Rjk} - \tau_{jk})}{\sum_k \zeta_k^2} \quad (7)$$

$$\text{var}(\hat{\beta}) = (\sum_k \zeta_k^2)^{-1} \quad (8)$$

şeklinde hesaplanmaktadır. β değerinin sıfırdan anlamlı şekilde farklılaşması maddenin DMF gösterdiği anlamına gelmektedir. Diğer yandan (Mantel, 1963) β değerinin yorumlanmasında standart normal dağılım gösteren z istatistiğini önermiştir. Z istatistiğinin hesaplanması şu şekildedir:

$$z^2 = \frac{\hat{\beta}^2}{\text{var}(\hat{\beta})} \quad (9)$$

z^2 veya $|z|$ değerinin büyük olması maddede yüksek düzeyde DMF olduğunu göstermektedir.

Poly-SIBTEST. Çok kategorili maddelerde DMF tespiti için kullanılan poly-SIBTEST istatistiği, iki kategorili maddelerde kullanılan SIBTEST'in genişletilmiş halidir ve parametrik olmayan bir modeldir (Chang ve diğerleri, 1996). SIBTEST yönteminde maddeler iki alt gruba ayrılır. Birincisi DMF'li olduğu düşünülen maddelerden oluşan "çalışılan alt test", ikincisi ise test ile ölçülmek istenen özellikleri ölçtüğü, dolayısıyla DMF'li olmadığı düşünülen maddelerden oluşan "geçerli alt test"tir (Narayanan & Swaminathan, 1996). Referans ve odak gruptaki bireylerin yeteneklerini temsil eden toplam puan "geçerli alt test" ten elde edilir. Dolayısıyla SIBTEST yöntemiyle DMF varlığı incelenen bir madde, eşleştirme değişkeninin içinde yer almaz. Toplam puan ve kategori sayısına göre eşleştirilen bireylerin performansları her bir "çalışılan test" için K grupta karşılaştırılır (Henderson, 1999). Bu çalışmada her bir madde için "çalışılan alt test" sadece tek maddeden (maddenin kendisi) oluşturulmuş; diğer maddeler "geçerli alt test"te yer almıştır.

SIBTEST yöntemi maddedeki DMF varlığının yanı sıra DMF miktarını da belirten bir etki büyüklüğü (β) sunmaktadır. Her geçerli alt test puan düzeyinde DMF'li olduğu düşünülen maddede beklenen grup farkı olarak tanımlanan β etki büyüklüğünün kestirimi şu şekilde tanımlanmaktadır:

$$\hat{\beta} = \sum_{k=0}^{n_m} p_k (\bar{Y}_{Rk} - \bar{Y}_{Fk}) \quad (10)$$

k : geçerli alt testteki puan düzeylerinin sayısı,

n_m : geçerli alt testteki maksimum puan düzeyi,

p_k : odak grupta geçerli alt test puanı k olan bireylerin oranı,

\bar{Y}_{Rk} : referans grup için k . geçerli alt test puan düzeyinde şüpheli maddenin ortalama madde puanı,

\bar{Y}_{Fk} : odak grup için k . geçerli alt test puan düzeyinde şüpheli maddenin ortalama madde puanı,

Roussos and Stout (1996) tarafından SIBTEST etki büyüklüğü olarak önerilen β etki büyüklüğü indeksi, ikili ve çok kategorili maddelerde poly-SIBTEST DMF indeksini yorumlamak için de kullanılmaktadır (Henderson, 2001). β değerinin negatif

olması odak grup lehine, pozitif olması ise referans grup lehine DMF varlığı anlamına gelmektedir. Bu bağlamda β etki büyüklüğünün yorumlanmasında kullanılan kritik değerler Tablo 2’de verilmiştir (Roussos & Stout, 1996).

Tablo 2

β Değerinin Yorumlanmasında Kullanılan Kritik Değerler

Kritik değerler	Yorum
$ \beta < 0,059$	DMF yoktur ya da ihmal edilebilir düzeydedir. (A düzey)
$0,059 \leq \beta < 0,088$	DMF vardır ve orta düzeydedir (B düzey)
$ \beta \geq 0,088$	DMF vardır ve yüksek düzeydedir (C düzey)

Tablo 2’de β etki büyüklüğünün yorumlanmasında A, B ve C düzeyleri için kritik değerler verilmiştir. Bu çalışmada DMF etkisinin yüksek olduğu C düzeyine giren maddelere odaklanılmış, dolayısıyla poly-SIBTEST istatistiği sonucunda $|\beta| \geq 0,088$ olan maddeler DMF’li olarak işaretlenmiştir.

Değişen Adım Fonksiyonu

Değişen adım fonksiyonu, çok kategorili bir maddenin belirli bir adımındaki ölçülen özelliklerde ortaya çıkan “gruplar arası farkı” tanımlamak için kullanılan kapsamlı bir yaklaşımdır (Penfield, 2007). Madde boyutunda tek bir istatistik veren DMF analizlerinin aksine, DAF analizlerinde maddede yer alan adım sayısı kadar istatistik elde edilir. Dolayısıyla değişen adım fonksiyonu, değişen madde fonksiyonunun madde içindeki DMF etkilerine odaklanan bir alt kümesi olarak görülebilir.

DAF için Penfield (2008) tarafından iki boyutlu bir taksonomi önerilmiştir. DAF taksonomisindeki birinci boyut, DAF’nin yaygınlığı ile ilgilidir. Örneğin, tüm madde adımlarında orta ile büyük arası DAF’nin mevcut olması durumunda, DAF yaygındır (pervasive); aksi halde yaygın değildir (non-pervasive). Yaygın DAF’nin varlığı, DMF’nin nedeninin madde düzeyinde etkisini gösterdiği anlamına gelmektedir. Yaygın olmayan DAF ise madde düzeyinde değil sadece bir ya da birkaç adımda etkisini gösteren, dolayısıyla daha kolay izole edilebilecek DMF’nin varlığını göstermektedir. Örneğin; öğrencilere belirli bir konuyla ilgili verilen kompozisyon yazma ödevinde yaygın DAF görülebilir. Bu durumda yaygın DAF’nin varlığı değişmezlik eksikliğinden sorumlu olan faktörün, maddenin kendisinden kaynaklandığı anlamına gelecektir. Eğer ödevde sadece ‘dil bilgisi kurallarına uyma’ ile ilgili adımda DAF görülmüşse yaygın olmayan DAF’den bahsedilir ve sorun maddeden değil sadece söz konusu

adımdan kaynaklanmıştır. Dolayısıyla madde revizyonunda araştırmacı özellikle bu adımla ilgilenecektir. DAF taksonomisinin ikinci boyutu; sabit (DAF, bir grup için aynı büyüklükte avantajlıdır), yakınsak (DAF, bir grup için değişen büyüklüklerde avantajlıdır), ırsak (bir adımda referans grubun, başka adımda odak grubun lehinedir) şeklinde bir DAF örüntüsü tanımlamaktadır.

Çok kategorili bir maddede DAF'nin değerlendirilmesi, (r maddedeki puan düzeylerinin sayısı olmak üzere) maddeyi $J = r - 1$ adım fonksiyonuna ayırarak başlar. Her adım fonksiyonu, her bir puan düzeyinden art arda daha yüksek bir puan düzeyine ilerleme veya "adım atma" olasılığını tanımlamaktadır. Maddenin adım fonksiyonlarından birinde veya daha fazlasında gruplar arası fark olması durumunda maddenin DAF'li olduğu sonucuna varılır. DAF analizi farklı yaklaşımlar kullanılarak gerçekleştirilebilir. Aşamalı Tepki Modeli (ATM) (Cohen ve diğerleri, 1993) ve Kısmi Puan Modeli (KPM) (Penfield, Myers ve diğerleri, 2008) kapsamında MTK tabanlı yaklaşımlar ile lojistik regresyon (French & Miller, 1996) bunlara örnek olarak verilebilir. Bu çalışmada DAF tespiti için, alanyazında kullanılan en yaygın DAF yöntemlerinden Ardışık Kategori Log Odds Oranı (AC-LOR) ve Kümülatif Kategori Log Odds Oranı (CU-LOR) yöntemleri tercih edilmiştir. Bu amaçla Penfield (2008)'in aynı gözlenen puana sahip odak ve referans grup üyelerinin j . adımda başarılı olma olasılıklarını karşılaştırdığı olasılık oran yaklaşımı kullanılmıştır. Buna göre sınava girenler, olası puan değerleri $k = 1, 2, 3, \dots, K$ olan bir testin ham toplam puanlarına göre puan gruplarına ayrılır. Bu bağlamda referans grubun j . adımdaki başarı olasılığının odak grubun başarı olasılığına oranı şu şekilde hesaplanmaktadır:

$$\hat{\alpha}_j = \frac{\sum_{k=1}^K A_{jk} D_{jk} / N_{jk}}{\sum_{k=1}^K B_{jk} C_{jk} / N_{jk}} \quad (11)$$

A_{jk} : j . adımda başarılı olan referans grup üyelerinin sayısı

B_{jk} : j . adımda başarısız olan referans grup üyelerinin sayısı

C_{jk} : j . adımda başarılı olan odak grup üyelerinin sayısı

D_{jk} : j . adımda başarısız olan odak grup üyelerinin sayısı

N_{jk} : $A_{jk} + B_{jk} + C_{jk} + D_{jk}$

Bu değer, iki kategorili maddeler için Mantel-Haenszel olasılık oranına eşdeğerdir, öyle ki her adım bir ikili madde olarak değerlendirilmektedir (Gattamorta & Penfield, 2012). $\hat{\alpha}_j$ 'nin doğal logaritması $\hat{\lambda}_j$ ile gösterilmektedir. $\hat{\lambda}_j$ 'nin sıfır olması DAF

olmadığı, negatif bir $\hat{\lambda}_j$ değeri odak grup lehine DAF olduğu, pozitif bir $\hat{\lambda}_j$ değeri ise referans grup lehine DAF olduğu anlamına gelmektedir. $\hat{\alpha}_j$ değerinin ölçek ve yön bakımından, iki kategorili maddelere uygulanan Mantel-Haenszel ortak log-olasılık oranı ile tutarlı olması nedeniyle, $\hat{\lambda}_j$ 'nin yorumlanmasında ETS sınıflandırma şeması (Zieky, 1993) kullanılmaktadır. Öyle ki; $|\hat{\lambda}_j| < 0,43$ değeri düşük bir DAF etkisi, $0,43 \leq |\hat{\lambda}_j| \leq 0,64$ değeri orta düzey bir DAF etkisi, $|\hat{\lambda}_j| > 0,64$ değeri ise büyük bir DAF etkisi olduğunu göstermektedir (Penfield, 2007; Penfield, Alvarez ve diğerleri, 2008). Yapılan çalışmalarda düşük düzeyde bir DAF'nin, ihmal edilerek orta veya yüksek bir DAF içeren maddelerin incelenebileceği belirtilmiştir (Penfield, Alvarez ve diğerleri, 2008). Bu çalışmada incelenen koşul sayısının fazla olması nedeniyle sadece büyük DAF'ler işaretlenmiştir.

Değişen adım fonksiyonu belirleme yöntemleri. Alanyazında değişen adım fonksiyonunu belirlemek için kullanılan birçok yöntem bulunmaktadır. Örneğin; MTK tabanlı yaklaşımlarını ele almak gerekirse Cohen ve diğerleri (1993)'in aşamalı tepki modeli altındaki yaklaşımı; referans ve odak grupları için elde edilen madde gerçek puan fonksiyonları arasındaki işaretli ve işaretsiz alanın yanı sıra bir maddenin J adımlarının her biri için Lord'un χ^2 'sine (Lord, 1980) dayalı olarak madde parametrelerinin karşılaştırılmasını dikkate almaktadır. Diğer yandan Penfield (2008) kısmi puan modeli altında madde ve adım düzeyinde değişmezliği modellemek üzere hem parametrik hem de parametrik olmayan yöntemler sunmaktadır. MTK tabanlı bu yaklaşımlar iki kategorili maddelerdeki DMF örneklerine paralel olarak adım fonksiyonlarındaki gruplar arası farklılıkları incelemektedir. Yani çok kategorili bir maddenin J adımları iki kategorili maddeler gibi düşünülerek adımların her birinde ayrı bir analiz gerçekleştirilmektedir (Gattamorta, 2009).

DAF analizinde kullanılan yöntemler arasında Dorans ve Kulick (1986) tarafından önerilen standartlaştırma yöntemi ve Shealy ve Stout (1993) tarafından önerilen SIBTEST yöntemi, odak ve referans grupların bir maddedeki ortalama puanları arasındaki ağırlıklı farkı kullanmaktadır. Her iki yöntem de bir etki büyüklüğü vermektedir. Diğer yandan French ve Miller (1996), çok kategorili bir maddenin her adımında bir olasılık oranı testi kullanarak DMF yokluk hipotezinin birden çok test edildiği, lojistik regresyona dayalı bir yaklaşım önermektedir. Çok kategorili maddeler

için lojistik regresyon, belirli bir madde ile ilişkili adım fonksiyonlarının her biri için iki kategorili bağımlı değişkenler için kullanılan lojistik regresyonla aynıdır.

DAF yöntemleri kullanılarak yapılan araştırmaların çoğunda MTK tabanlı iki temel adım fonksiyonu kullanılmaktadır. Bunların ilki Genelleştirilmiş Kısmi Puan Modeli (GKPM) (Muraki, 1992) altında ardışık kategori yaklaşımı; ikincisi ise ATM (Samejima, 1997) altında kümülatif yaklaşımdır.

Ardışık kategori yaklaşımı. Çok kategorili maddelerde DAF analizi yapılırken GKPM ile tutarlı olan ardışık kategori yaklaşımı kullanılarak J adım fonksiyonlarının her biri tanımlanmaktadır. Bu yaklaşım altında j. adım fonksiyonu j-1 puan düzeyinden j puan düzeyine başarılı bir şekilde ilerleme olasılığını ifade etmektedir. Örneğin 4 kategorili bir maddede (0, 1, 2, 3) 3 adım bulunmaktadır. 1. adım 0 puandan 1 puana, 2. adım 1 puandan 2 puana, 3. adım ise 2 puandan 3 puana ilerlemeyi göstermektedir. 3. adım için ardışık kategori yaklaşımı altında denklem 11'de kullanılan semboller şu şekilde kullanılmalıdır:

$$\hat{\alpha}_j = \frac{\sum_{k=1}^K A_{jk} D_{jk} / N_{jk}}{\sum_{k=1}^K B_{jk} C_{jk} / N_{jk}} \quad (12)$$

A_{jk} / B_{jk} : sırasıyla 3 ve 2 alan referans grubu üyelerinin sayısı,

C_{jk} / D_{jk} : sırasıyla 3 ve 2 alan odak grubu üyelerinin sayısı,

N_{jk} : $A_{jk} + B_{jk} + C_{jk} + D_{jk}$

Ardışık kategori yaklaşımı altında 3. adımdaki bir DAF etkisi için sadece 2 ve 3 puan alan bireyler dikkate alınmakta 1 veya daha düşük puan alan bireyler analiz dışı bırakılmaktadır. Dolayısıyla 3. adımdaki bir DAF etkisi referans grup üyelerine karşı odak grup için 2 puandan 3 puana geçişin daha zor olduğunu göstermektedir.

Kümülatif kategori yaklaşımı. Çok kategorili maddelerde DAF analizi yapılırken ATM ile tutarlı olan kümülatif kategori yaklaşımı kullanılarak J adım fonksiyonlarının her biri tanımlanmaktadır. Bu yaklaşım altında j. adım fonksiyonu 0, 1, ..., j-1 puan düzeyinden j, ..., J puan düzeyine başarılı bir şekilde ilerleme olasılığını ifade etmektedir. Örneğin 4 kategorili bir maddede 1. adım 0 puandan 1, 2 veya 3 puana, 2. adım 0 veya 1 puandan 2 veya 3 puana, 3. adım ise 0, 1 veya 2 puandan 3 puana ilerlemeyi göstermektedir. 3. adım için kümülatif kategori yaklaşımı altında denklem 11'de kullanılan semboller şu şekilde kullanılmalıdır:

$$\hat{\alpha}_j = \frac{\sum_{k=1}^K A_{jk} D_{jk} / N_{jk}}{\sum_{k=1}^K B_{jk} C_{jk} / N_{jk}}. \quad (13)$$

A_{jk} : 3 alan referans grubu üyelerinin sayısı,

B_{jk} : 0, 1 veya 2 alan referans grubu üyelerinin sayısı,

C_{jk} : 3 alan odak grubu üyelerinin sayısı,

D_{jk} : 0, 1 veya 2 alan odak grubu üyelerinin sayısı,

N_{jk} : $A_{jk} + B_{jk} + C_{jk} + D_{jk}$

Kümülatif kategori yaklaşımı altında 3. adımdaki bir DAF etkisi referans grup üyelerine karşı odak grup için 0, 1 veya 2 puandan 3 puana geçişin daha zor olduğunu göstermektedir. Dolayısıyla kümülatif yaklaşım altında DAF analizinde toplamda bütün puanlar dikkate alınmaktadır ve maddeler kümülatif ve ardışık kategoriler yaklaşımları altında farklı şekilde ikiye ayrıldığından, A_{jk} , B_{jk} , C_{jk} , D_{jk} ve N_{jk} değerleri iki yaklaşım arasında farklılık göstermektedir. Bu nedenle adım düzeyi parametrelerinin yorumlanmasında adım fonksiyonunu tanımlamakta kullanılan yaklaşımı bilmek oldukça önemlidir.

İlgili Araştırmalar

Bu bölümde alanyazında DAF yöntemlerinin kullanıldığı ya da DMF ve DAF yöntemlerinin birlikte ele alındığı çalışmalara yer verilmiştir.

Penfield (2007), çok kategorili maddelerde DMF'nin değerlendirilmesinde kullanılan istatistiklerin madde düzeyinde tek bir ölçüm değişmezliği indeksi verdiğini, dolayısıyla bu durumun yanıltıcı olabileceğini öne sürmüştür. Bu bağlamda çalışmada çok kategorili maddelerin her adımındaki değişmezliği inceleyen, değişen adım fonksiyonu olarak açıklanan bir çerçeve sunulmuştur. Simulatif veri üzerinde, ikili maddelerde sıklıkla kullanılan Mantel Haenszel tabanlı parametrik olmayan bir DAF yöntemi uygulanmıştır. Çalışmanın sonucunda çok kategorili maddenin adımları boyunca DAF düzeyi büyüklük veya işaret bakımından değiştiğinde, DAF yönteminin madde düzeyinde istatistik veren DMF yöntemine göre daha güçlü ve doğru sonuçlar verdiği görülmüştür.

Penfield, Alvarez ve diğerleri (2008)'in çalışmasında çok kategorili maddelerde DMF'nin yorumlanmasını iyileştirmeye dönük DAF taksonomisi sunulmuştur. Buna göre çok kategorili bir maddenin adımları boyunca DAF etkileri birkaç farklı form alabilir

ve farklı formlar, DMF tespitinin hassasiyeti ve DMF nedenlerinin yorumu için farklı sonuçlara sebep olabilir. Bu bağlamda çalışmada DAF formları sınıflandırılmış, istatistiklerin yorumlanması objektif hale getirilmiştir.

Penfield (2008)'in çalışmasında ATM, Sürekli Oran Modeli (SOM) ve GKPM tarafından belirlenen modellere dayalı üç tane parametrik olmayan DAF yöntemi açıklanmaktadır. Bu bağlamda üç DAF yönteminin istatistiksel özelliklerini çeşitli koşullar altında karşılaştırmak için bir simülasyon çalışması yapılmıştır. Çalışmada her biri dörder yanıt kategorisinden oluşan 16 çok kategorili maddeden oluşan yapay bir test üretilmiş ve ilk madde çalışılan madde olarak belirlenmiştir. Yani sadece ilk maddede DAF analizi yapılmıştır. Odak ve referans grupların ortalamaları, yanıtları üretmekte kullanılan model, çalışılan maddenin ayırıcılık parametresi ve DAF tipi manipüle edilerek toplam 60 koşul içeren bir araştırma deseni oluşturulmuştur. Çalışmanın sonucunda ATM'ye dayanan DAF yönteminin, en düşük genel yanlılık seviyesini korurken, en yüksek kararlılık seviyesini verdiği görülmüştür.

Penfield (2010), çalışmasında net ve global DMF ayırımını yapmış ve konuyu gerçek veri üzerinde tartışmıştır. Çalışmada 18.248 Kanadalı öğrenciye 30 çok kategorili madde içeren standartlaştırılmış matematik testi uygulanmasıyla elde edilen veriler kullanılmıştır. Örneklem kadın (referans) ve erkek (odak) gruplara ayrılmıştır. Maddelerin DMF gösterip göstermediğine bakmak için Simultaneous Step Level test (SSL) ve Mantel'in ki-kare testleri, DMF içerdiği belirlenen maddelerin DAF analizi için ise bir olasılık oran yaklaşımı kullanılmıştır. Analiz sonuçları ATM ve KPM modelleri altında yorumlanmıştır.

Miller ve diğerleri (2010) öğretmenin öğretim uygulamalarının matematiğin 9. sınıf değerlendirmesi üzerindeki etkisini araştırmak için DMF ve DAF analizlerini birleştirerek bir çalışma yürütmüştür. Bu bağlamda Ontario'nun 9. Sınıf Matematik Sınavı'nın iki formundan elde edilen veriler kullanılmıştır. Öğrenciler öğretmenlerinin kullandıkları öğretim yöntemine göre gruplandırılmıştır. Maddelerin DMF analizleri için Mantel ki-kare, Liu Agresti ve Cox's β yöntemleri kullanılmıştır. Analizler sonucunda orta ve yüksek düzeyde DMF gösterdiği tespit edilen maddeler DAF analizine tabi tutulmuştur. Analiz sonuçları maddeler hakkında bilgi vermenin ötesinde kullanılan öğretim yöntemi hakkında da bilgilendirme yapmaktadır.

Gattamorta ve Penfield (2012), DAF analizi yapmak için kullanılan yaklaşımlardan ardışık kategori ve kümülatif kategori yaklaşımlarını karşılaştırmıştır. Eyalet çapında bilişsel engelli öğrencileri değerlendirmek için hazırlanmış 5. ve 8. sınıf Fen testlerinin kullanıldığı çalışmada 16 çok kategorili madde üzerinde çalışılmıştır. Sonuçlar, özellikle DAF etkilerinin ihmal edilebilir olduğu durumlarda, iki yaklaşımın genellikle tutarlı sonuçlar verdiğini, önemli DAF etkileri mevcut olduğunda ise, iki yaklaşımın zaman zaman farklılaştığını ortaya koymuştur.

Gattamorta ve diğerleri, (2012)'de çok kategorili maddelerde madde ve adım düzeyinde değişmezlik etkilerini eş zamanlı olarak değerlendirmek için KPM kullanılmıştır. Bir simülasyon çalışmasının yanı sıra bu metodolojinin uygulamadaki kullanımını göstermek için, uyarlanmış büyük ölçekli bir testin uygulamasından elde edilen büyük bir ampirik örnek sunulmuştur. Çalışmanın gerçek verileri Kanada'da ulusal çapta uygulanan "the School Achievement Indicators Program (SAIP)" geniş ölçekli başarı testinden elde edilmiştir. Testin Fransızca formunu alan öğrenciler referans grubu (N = 2.867), İngilizce formunu alan öğrenciler ise odak grup (N = 4,652) olarak belirlenmiştir. Matematik alanında cebir, ölçüm, geometri ve veri yönetimi becerilerini değerlendiren 30 çok kategorili madde üzerinde parametrik olmayan analizler gerçekleştirilmiştir. Diğer yandan 30 çok kategorili madde üretilerek WINSTEPS yardımıyla farklı koşullarda parametrik yöntemlerle analizler yapılmıştır. Çalışmanın sonucunda parametrik ve parametrik olmayan yöntemlerin DAF'nin varlığı, yokluğu ve/veya büyüklüğü konusunda tutarlı oldukları belirtilmiştir.

Akour ve diğerleri, (2015), çalışmalarında test diliyle ilgili ölçme değişmezliğini araştırmak için net ve global DMF olmak üzere iki tür DMF ile birlikte DAF yöntemlerini kullanmışlardır. Çalışmada PISA 2006'da çok kategorili maddelerin bulunduğu fen testine ait veriler kullanılmıştır. Çalışmanın örneklemini Fransızca ve İngilizce'nin referans dilleri temsil ettiği, farklı test dillerini uygulayan dört ülke oluşturmaktadır. Maddelerin DMF gösterip göstermediğine bakmak için SSL test (global DMF) ve Mantel'in ki-kare testi (net DMF), DMF içerdiği belirlenen 6 maddenin DAF analizi için ise bir olasılık oran yaklaşımı kullanılmıştır. Bulgular, birçok maddenin her iki DMF tipini gösterdiğini, ancak çoğu durumda sonuçların iki referans dil için tutarsız olduğunu göstermiştir. Ayrıca, DAF analizi, DMF etkisinin doğası ve yeri ile ilgili net DMF analizinden daha fazla bilgi vermiştir.

Benítez ve diğerleri, (2015), çalışmalarında maddelerin farklı dillere göre neden DMF gösterdiğini araştırmak amacıyla DMF analizinden elde edilen nicel sonuçları ve uzman değerlendirmesinden elde edilen nitel bulguları birleştiren karma yöntem kullanmışlardır. Çalışmanın nicel kısmında PISA 2006 öğrenci anketinde yer alan yedi ölçeğe ait toplam 38 çok kategorili maddeye yanıt veren İspanya (referans grup) ve ABD (odak grup) örneklemleri üzerinde çalışılmıştır. Çok kategorili DMF analizi için DAF ve Sıralı Lojistik Regresyon yöntemleri kullanılarak her iki yöntem için de DMF'li olduğu tespit edilen maddeler uzman görüşüne sunulmuştur. Çalışmanın sonucunda maddelerdeki alışılmış çeviri sorunlarının arkasında yatan potansiyel yanlılık kaynakları hakkında bilgiler verilmiştir.

Ayodele (2017), çalışmasında çok kategorili maddelerde, puan kategori sayısının DAF ve DMF yöntemlerinin güç ve Tip I hatası üzerindeki etkisini araştırmıştır. Simülatif veri ile yapılan çalışmada değişen koşullarda analizler DAF yöntemlerinden AC-LR ve CU-LOR; DMF yöntemlerinden ise Mantel (ki kare) test, Liu Agresti, GMH, SSL test ile yapılmıştır. Değişen koşullar arasında örneklem büyüklüğü oranı, puan kategori sayısı, veri üretme modeli, etki ve DAF modeli bulunmaktadır. Bu bağlamda 3 ve 4 kategorili 20 madde üretilmiş ve bunlardan 2 madde üzerinde koşullar değiştirilmiştir. Diğer yandan SSL testinde karşılaştırmalı Tip I hatalarını ayarlama istatistiksel yöntemlerden en etkili olanı araştırmak için Dunn-Bonferroni, Benjamini ve Hochberg ve Holm's yöntemleri incelenmiştir. Çalışmanın sonuçları puan kategori sayılarının değiştirilmesinin DAF / DMF belirleme yöntemleri üzerinde bir etkisi olmadığını göstermiştir. Gücü etkileyen en önemli faktörün ise DAF modeli olduğu sonucuna ulaşılmıştır. DMF yöntemleri içerisinde en iyi performansı gösteren yöntemin SSL test olduğu, hem DAF hem de DMF'yi aynı anda tespit etmekte kullanılabileceği belirtilmiştir.

DMF ve DAF analizlerinin birlikte yapıldığı çalışmalar incelendiğinde alanyazında oldukça sınırlı çalışma bulunduğu görülmektedir. Yapılan çalışmalarda istatistiklerin daha çok simülasyon veri üzerinde yapıldığı, gerçek veri üzerinde çalışıldığında ise mevcut duruma odaklanıldığı dikkatleri çekmektedir. Bu çalışmada gerçek veri üzerinde koşullar manipüle edilerek yöntemlerin karşılaştırılmasıyla alanyazındaki bu eksikliğin giderilmesine katkı sağlanması amaçlanmaktadır.

Bölüm 3

Yöntem

Bu bölümde araştırmanın türü, çalışma grubu, verilerin elde edilmesi, veri toplama aracı, araştırma kapsamında incelenen koşullar ve verilerin analizine yer verilmiştir.

Araştırmanın Türü

PISA 2018 uygulamasında aşağıda belirtilen maddeler üzerinde çok kategorili DMF/DAF belirleme yöntemlerinin çeşitli koşullarda karşılaştırıldığı bu araştırma ilişkisel tarama modelinde bir araştırmadır.

Çalışma Grubu

Bu çalışmada PISA 2018 kapsamında öğrencilere uygulanan “Bilgi ve İletişim Teknolojilerine Aşinalık Anketi” içerisindeki okulda dijital cihazları kullanma sıklığıyla ilgili maddelere yanıt veren Türkiye, Kazakistan ve Amerika Birleşik Devletleri (ABD) örneklemi üzerinde çalışılmıştır. Ülkelerin seçiminde öncelikle PISA 2018 sonuçlarına göre okuma becerileri alanında (ağırlıklı alan) ülke sıralamaları incelenmiş ve ülkeler düşük, orta ve yüksek düzey olarak üç gruba ayrılmıştır. Ayrıca ülkelerin ekonomik durumları da göz önünde bulundurulmuştur. İlgili anketin PISA 2018 uygulamasına katılan ülkelerin tümüne uygulanmamış olması da hesaba katılarak her iki durum (başarı ve ekonomik düzey) için de düzeyi düşük olana gruptan ilgili anketi yanıtlayan ve OECD ülkesi olmayan Kazakistan (69. sırada), orta düzey gruptan OECD ülkesi olan Türkiye (40. sırada) ve üst düzey gruptan yine OECD ülkesi olan ABD (13. sırada) tercih edilmiştir.

İlgili ülkelerden PISA 2018 uygulamasına katılan birey sayısı Kazakistan için 19.507, Türkiye için 6.890 ve ABD için 4838 olduğu görülmektedir (OECD, 2019) Ülkelere ait verilerin analize uygun olup olmadığı incelenmiş, gerekli veri temizleme işlemleri yapılmıştır. İncelenen anket için 1 ve daha fazla maddeyi yanıtlamayan birimler veri setinden çıkarılmıştır. Sonrasında tek yönlü uç değer incelemesi için tüm bireylerin z puanları hesaplanmış, kutu grafikleri ve histogram grafiği oluşturulmuştur. (-3,3) aralığı dışında kalan z puanları (Raykov & Marcoulides, 2008), histogram grafiğinde normal dağılım eğrisinin dışında kalan değerler ve kutu grafiğinde sınırların dışında kalan değerler (Tukey, 1977) uç değer kabul edilerek bu veriler silinmiştir. Çok

değişkenli uç değer hesaplaması için ise Mahalanobis Uzaklığı yöntemi kullanılmış; hesaplanan Mahalanobis Uzaklık değeri, serbestlik derecesi bağımsız değişken sayısı olan χ^2 tablo değeriyle karşılaştırılmıştır. Bu bağlamda .01 manidarlık düzeyinde Mahalanobis Uzaklık değeri χ^2 tablo değerinin üzerinde olan veriler uç değerler olarak kabul edilmiş ve veri setinden çıkarılmıştır (Büyüköztürk, 2018). Tüm veri temizleme işlemleri sonucunda toplamda Kazakistan verisinde 10.991 (%60,63), Türkiye verisinde 3.997 (%22,05) ve ABD verisinde 3.140 (%17,32) kişi üzerinde çalışılmıştır.

Veri Toplama Aracı

Araştırmanın verilerine PISA 2018 verilerinin açıklandığı, OECD'nin resmi internet adresinden (<https://www.oecd.org/pisa/data/2018database>) ulaşılmıştır. Bu bağlamda, PISA 2018 kapsamında öğrencilere yönelik hazırlanan “Bilgi ve İletişim Teknolojilerine Aşinalık Anketi” içerisindeki okulda dijital cihazları kullanma sıklığıyla ilgili maddelerin (IC011) yanıtlandığı Kazakistan, Türkiye ve ABD verileri üzerinde çalışılmıştır. Söz konusu ankette okulda dijital cihazları kullanma sıklığıyla ilgili 10 madde bulunmaktadır. Maddeler 1(hemen hemen hiç), 2 (ayda 1-2 kez), 3 (haftada 1-2 kez), 4 (hemen hemen her gün), 5 (her gün) olmak üzere 5’li dereceleme ölçeği tipindedir. Anketten elde edilen puanlar 10-50 arasında değişmekte olup yüksek puanlar katılımcının okulda dijital cihazları kullanma sıklıklarının yüksek, düşük puanlar ise okulda dijital cihazları kullanma sıklıklarının düşük düzeyde olduğu anlamına gelmektedir. IC011 kodlu ankete ilişkin maddeler Ek A’da sunulmuştur.

Veri toplama aracına ilişkin betimsel istatistikler ve anketin güvenilirliğine ilişkin Cronbach alfa değerleri hesaplanarak Tablo 3’te verilmiştir.

Tablo 3

Veri Toplama Aracına İlişkin Betimsel İstatistikler ve Cronbach Alfa Değerleri

	Ülkeler	N	Ortalama	SS	Basıklık	Çarpıklık	Cronbach α
IC011	KAZ	10.991	25,01	10,93	-,60	,43	,95
	TUR	3.997	16,39	6,14	-,44	,79	,85
	ABD	3.140	23,81	8,23	-,54	,34	,85

Tablo 3’e göre “Bilgi ve İletişim Teknolojilerine Aşinalık Anketi” nde okulda dijital cihazları kullanma sıklığına ilişkin en yüksek ortalama Kazakistan verisine aittir (\bar{x} = 25,01). ABD verilerine ait ortalama (\bar{x} = 23,81) Kazakistan ortalamasına yakın olup ikinci sıradadır. Türkiye ise ülkeler arasında en düşük ortalamaya sahiptir (\bar{x} = 16,39). İç tutarlılık anlamında güvenilirlik katsayısı veren Cronbach alfa değerleri

incelendiğinde Kazakistan örnekleme için ,95; Türkiye ve ABD örnekleme için ise ,85 olduğu görülmektedir. George ve Mallery (2003)'e göre güvenilirlik katsayısını yorumlamada kullanılan kriterler şu şekildedir: $\alpha > .90$ ise mükemmel, $\alpha > .80$ ise iyi, $\alpha > .70$ ise kabul edilebilir, $\alpha > .60$ ise tartışılabilir, $\alpha > .50$ ise düşük düzeyde güvenilir; $\alpha < .50$ ise güvenilir değildir. Dolayısıyla ölçeğin Kazakistan verisi için mükemmel düzeyde, Türkiye ve ABD verisi için iyi düzeyde güvenilir olduğunu söylemek mümkündür.

Diğer yandan ülkeler bazında her bir maddeye ilişkin betimsel istatistikler ve puan kategori dağılımları hesaplanarak Tablo 4'te sunulmuştur.

Tablo 4

Veri Toplama Aracının Maddelerine İlişkin Betimsel İstatistikler ve Puan Kategori Dağılımları

MADDELER	Ülkeler	Ortalama	SS	Basıklık	Çarpıklık	Madde toplam korelasyonu	PUAN KATEGORİ DAĞILIMLARI (%)				
							1	2	3	4	5
M1	KAZ	2,71	1,43	-1,30	,20	,63	29,7	17,1	20,0	19,0	14,2
	TUR	1,78	1,16	,41	1,27	,43	61,9	13,2	13,1	8,6	3,3
	USA	2,65	1,52	-1,52	-1,48	,33	38,2	9,8	16,1	21,1	14,9
M2	KAZ	2,38	1,31	-,92	,52	,79	35,8	20,6	21,6	13,7	8,2
	TUR	1,39	,74	2,19	1,79	,58	74,5	13,9	9,9	1,8	
	USA	2,52	1,33	-1,17	-1,09	,53	32,0	18,5	23,9	16,6	8,9
M3	KAZ	2,79	1,30	-1,06	,11	,79	21,9	19,5	27,6	19,3	11,6
	TUR	2,20	1,15	-,64	,57	,52	37,3	22,8	25,9	10,3	3,6
	USA	3,11	1,27	-,96	-,93	,59	15,6	13,9	29,0	26,4	15,1
M4	KAZ	2,54	1,32	-1,06	,33	,84	30,3	20,1	24,0	16,4	9,2
	TUR	1,56	,88	1,27	1,47	,65	66,0	16,9	13,1	3,5	,5
	USA	2,09	1,26	-,80	-,63	,62	47,9	17,3	17,7	12,4	4,8
M5	KAZ	2,22	1,29	-,73	,69	,78	42,6	18,9	19,7	12,0	6,8
	TUR	1,36	,73	2,82	1,97	,64	77,5	11,3	9,4	1,9	
	USA	1,84	1,21	-,30	,01	,51	61,2	10,6	14,8	9,7	3,7
M6	KAZ	2,17	1,27	-,66	,72	,77	43,5	19,1	19,9	11,5	6,0
	TUR	1,34	,69	2,87	1,97	,60	77,4	12,5	8,9	1,3	
	USA	1,63	,99	,81	,93	,46	64,8	15,6	12,1	6,6	,9
M7	KAZ	2,62	1,27	-,99	,24	,82	25,5	21,3	27,2	17,1	8,8
	TUR	1,96	1,13	-,23	,89	,49	48,8	19,9	20,5	8,0	2,8
	USA	2,25	1,26	-1,01	-,93	,60	40,6	17,4	23,0	14,2	4,8
M8	KAZ	2,49	1,30	-1,00	,38	,84	31,4	20,7	24,0	15,5	8,5
	TUR	1,52	,86	1,98	1,63	,59	67,9	16,7	12,0	2,7	,7
	USA	2,85	1,33	-1,20	-1,15	,56	22,3	17,7	25,5	22,3	12,2
M9	KAZ	2,54	1,30	-1,03	,34	,84	29,6	20,9	24,4	16,2	8,8
	TUR	1,51	,86	2,31	1,69	,64	68,1	17,0	11,4	2,6	,9
	USA	2,53	1,26	-1,09	-1,02	,66	28,2	22,3	24,5	17,9	7,1
M10	KAZ	2,55	1,32	-1,06	,33	,84	29,7	20,7	23,7	16,4	9,5
	TUR	1,78	1,04	,66	1,21	,61	55,4	20,8	16,5	5,0	2,3
	USA	2,34	1,26	-1,00	-,90	,63	35,5	22,1	21,4	15,0	5,9

Tablo 4'e göre "Bilgi ve İletişim Teknolojilerine Aşinalık Anketi" nde okulda dijital cihazları kullanma sıklığına ilişkin maddelerde tüm ülkeler için en yüksek ortalama M3 maddesinde ($\bar{X}_{KAZ} = 2,79$, $\bar{X}_{TUR} = 2,20$, $\bar{X}_{ABD} = 3,11$); en düşük ortalama ise M6 maddesinde ($\bar{X}_{KAZ} = 2,17$, $\bar{X}_{TUR} = 1,34$, $\bar{X}_{ABD} = 1,63$) elde edilmiştir. Bununla birlikte madde ortalamalarının Kazakistan ve ABD verileri için çoğunlukla 2'nin üzerinde, Türkiye verisi için ise çoğunlukla 2'nin altında olduğu görülmektedir. Bu durumda PISA 2018 uygulamasına Türkiye'den katılan öğrencilerin okulda dijital cihazları kullanma durumlarının düşük düzeyde olduğu söylenebilir. Diğer yandan maddelere ilişkin puan kategori dağılımları incelendiğinde maddelerin çoğunluğunda Türkiye verisinin yarısından fazlasının 1. kategoride yoğunlaştığı, özellikle 4. ve 5. kategorinin oldukça az işaretlendiği görülmektedir. M2, M5 ve M6 maddelerinde 5. kategorinin hiç işaretlenmemiş olması; M4, M7, M8, M9 maddelerinde ise 5. kategoriye işaretleyenlerin %1'in altında olması oldukça dikkat çekmektedir. Kazakistan ve ABD verileri incelendiğinde ise dağılımın tüm kategori düzeylerine yayıldığı görülmektedir. Her üç ülke için de 1. kategori en fazla, 5. kategori en az oranda işaretlenmiştir.

Boyutluluğun incelenmesi. Ölçeğin boyutluluğunu incelemek amacıyla açımlayıcı faktör analizi yapılmıştır. Ükelere ait verilerin örneklem büyüklüğü Kaiser-Meyer-Olkin (KMO) katsayısı ile; verinin dağılımı ise Barlett Küresellik Testi ile kontrol edilmiş, sonuçlar Tablo 5'te sunulmuştur.

Tablo 5

Veri Toplama Aracının Faktör Analizine Uygunluğuna İlişkin Veriler

		KAZAKİSTAN	TÜRKİYE	ABD
KMO		,942	,872	,877
	Yaklaşık ki kare değeri	101454,496	16161,504	10587,332
Barlett Küresellik Testi	Serbestlik derecesi	45	45	45
	p	,000	,000	,000

Tablo 5'e göre ülke verilerine ilişkin KMO katsayıları ,87-,94 aralığında değer almıştır. Kaiser (1970)'a göre KMO değeri 0 ile 1 arasında değer almaktadır ve bu değer 1'e yaklaşması örneklem büyüklüğünün faktör analizi için uygun olduğu anlamına gelmektedir. Diğer yandan Barlett Küresellik Testi sonuçları incelendiğinde ki-kare değerinin üç ülke için de istatistiksel olarak anlamlı ($\chi^2_{KAZ(45)} = 101454,496$, $\chi^2_{TUR(45)} =$

16161,504, ($\chi^2_{ABD(45)} = 10587,332; p < ,01$) olduğu; dolayısıyla verilerin açımlayıcı faktör analizi için uygun olduğu görülmektedir. Kazakistan, Türkiye ve ABD verileri ile yapılan açımlayıcı faktör analizi sonuçları Tablo 6'da sunulmuştur.

Tablo 6

Açımlayıcı Faktör Analizi Sonucunda Elde Edilen Faktörler ve Açıklanan Varyans Miktarları

			Özdeğer	Açıklanan varyans
IC011	KAZ	1. Faktör	7,01	70,09
	TUR	1. Faktör	6,279	62,791
	ABD	1. Faktör	5,004	50,044
		2. Faktör	1,023	10,229

Açımlayıcı faktör analizi sonucunda Kazakistan ve Türkiye verisi için öz değeri 1'in üzerinde tek bileşen olduğu, dolayısıyla tek boyutlu olduğu görülmektedir. ABD verisi için ise öz değeri 1'in üzerinde iki bileşen olduğu tespit edilmiştir. İki bileşenden birincisinin özdeğeri ile ikinci bileşenin özdeğeri arasındaki fark yaklaşık 4 kat olup maddelerin faktör yükleri de incelendiğinde tüm maddelerin birinci boyutta daha fazla faktör yükü gösterdiği görülerek ölçeğin tek boyutlu olduğuna karar verilmiştir. Maddelerin faktör yüküne ilişkin sonuçlar Tablo 7'de verilmiştir.

Tablo 7

Açımlayıcı Faktör Analizi Sonucunda Maddelere İlişkin Faktör Yük Değerleri

Maddeler	FAKTÖR YÜKLERİ		
	KAZ	TUR	ABD
IC011 M1	,69	,66	,49
M2	,83	,79	,67
M3	,82	,70	,70
M4	,88	,85	,77
M5	,83	,86	,71
M6	,82	,83	,67
M7	,86	,72	,76
M8	,88	,84	,70
M9	,88	,85	,78
M10	,88	,80	,78

Tablo 7'de Türkiye, Kazakistan ve ABD verileri ile yapılan açımlayıcı faktör analizi sonucunda maddelere ilişkin elde edilen faktör yük değerleri verilmiştir. Genel olarak faktör yük değerlerinin ,49 ile ,88 arasında değiştiği görülmektedir.

Araştırma kapsamında incelenen maddeler: Ayodele (2017) çalışmasında 20 maddelik bir test üretmiş bunlardan 2 madde üzerinde manipülasyon yaparak araştırma

sorularını incelemiştir. Bu çalışmada da araştırma koşullarının fazla olması sebebiyle yorumlamak üzere üç madde seçilmiştir. Maddelerin seçiminde psikometrik özellikler dikkate alınmış, üç ülke için madde toplam korelasyonu ve faktör yükü yüksek olan; diğer bir ifadeyle ölçeği temsil gücü en yüksek olan maddeler tercih edilmiştir. Tablo 7 incelendiğinde her üç ülke için bulunduğu faktöre en büyük yük veren ve madde toplam test korelasyonları en yüksek olan maddelerin M4, M9 ve M10 olduğu görülmektedir. Bu bağlamda çalışmada M4, M9 ve M10 maddelerine ilişkin çok kategorili DMF ve DAF analiz sonuçları raporlanarak yorumlanmıştır.

İncelenen Koşullar

Bu bölümde araştırmada manipüle edilen koşullardan bahsedilmiştir.

Kategori birleştirme kuralı. Öncelikle mevcut durumda 5'li dereceleme ölçeği tipinde (1-5 şeklinde) kodlanmış olan maddeler DIFAS 5.0 programının çalışma prensipleri gereği (0-4) şeklinde kodlanmıştır (Penfield, 2013). Sonrasında madde kategori sayısını değiştirmek amaçlandığından kategori birleştirme yoluna gidilmiştir. Birleştirilen kategorilerin birbirine yakın (ardışık gelen) kategoriler olmasına dikkat edilerek kategori birleştirmede olabilecek tüm kombinasyonlar ele alınmıştır (Gelin & Zumbo, 2003; Göçer-Şahin ve diğerleri, 2016). Araştırmanın amacına yönelik oluşturulan kategori birleştirme koşulları Tablo 8'de gösterilmiştir.

Tablo 8

Araştırmanın Amacı Kapsamında Oluşturulan Kategori Birleştirme Koşulları

		Kodlama öncesi	Yeni kategoriler	Açıklama
Üç kategorili	1. Koşul	(1,2)	0	(1 ile 2) ve (4 ile 5) birleştirilmiştir.
		3	1	
		(4,5)	2	
	2. Koşul	1	0	(2 ile 3) ve (4 ile 5) birleştirilmiştir.
		(2,3)	1	
		(4,5)	2	
	3. Koşul	(1,2)	0	(1 ile 2) ve (3 ile 4) birleştirilmiştir.
		(3,4)	1	
		5	2	
Dört kategorili	4. Koşul	(1,2)	0	(1 ile 2) birleştirilmiştir.
		3	1	
		4	2	
		5	3	
	5. Koşul	1	0	(2 ile 3) birleştirilmiştir.
		(2,3)	1	

		4	2	
		5	3	
	6. Koşul	1	0	(4 ile 5) birleştirilmiştir.
		2	1	
		3	2	
		(4,5)	3	
	7. Koşul	1	0	(3 ile 4) birleştirilmiştir.
		2	1	
		(3,4)	2	
		5	3	
Beş kategorili	8. Koşul	1	0	DIFAS 5.0 programının koşulları nedeniyle yeniden kodlanmıştır.
		2	1	
		3	2	
		4	3	
		5	4	

Tablo 8'e göre verilerin analizinde üç kategorili veri için üç, dört kategorili veri için dört ve beş kategorili veri için bir tane olmak üzere toplamda sekiz kategori birleştirme koşulu elde edilmiştir.

Odak grup – Referans grup. DMF/DAF analizleri için iki farklı odak grup- referans grup oluşturulmuştur. Bunlardan birincisinde ekonomisi ve başarısı bakımından düşük düzeyde bulunan Kazakistan (referans) ile ekonomisi ve başarısı bakımından orta düzeyde bulunan Türkiye (odak) örnekleme karşılaştırılmıştır. İkinci grupta ise ekonomisi ve başarısı bakımından orta düzeyde bulunan Türkiye (odak) ile ekonomisi ve başarısı bakımından yüksek düzeyde bulunan ABD (referans) örnekleme karşılaştırılmıştır.

Örneklem büyüklüğü ve odak grup referans grup örneklem oranı. Araştırmada incelenen koşullardan bir diğeri ise odak grup örneklem büyüklüğüdür. DMF çalışmalarında örneklem büyüklüğü oldukça önemlidir. Öyle ki örneklemin çok küçük olması zayıf parametre kestirimine, dolayısıyla DMF varlığının tespit edilememesine; örneklemin çok büyük olması ise DMF tespitinde aşırı duyarlılığa neden olabilmektedir (Ayodele, 2017). Bu sebeple bu çalışmada farklı örneklem büyüklüklerinde çalışılarak en doğru kararı verme eğilimine gidilmiştir. Çok kategorili maddelerle yapılan çalışmalar incelendiğinde 40 odak grup, 400 referans grup olmak üzere en az 440 kişilik veri ile; yaygın olarak ise 100 ila 2000 kişilik veri ile çalışıldığı görülmektedir (Ankenmann ve diğerleri, 1999; Elosua & Wells, 2013; Gonzalez-Roma ve diğerleri, 2006; Meade & Lautenschlager, 2004; Wood, 2011). Bu çalışmada odak grup örneklem büyüklüğü 200 (küçük) ve 1000 (büyük) olmak üzere iki farklı büyüklükte ele alınmıştır. Bununla birlikte

(odak grup):(referans grup) örneklem oranı 2:1, 1:1, 1:2 ve 1:3 olacak şekilde dört koşulda incelenmiştir.

Çok kategorili DMF/DAF belirleme yöntemleri. Araştırmada çok kategorili DMF belirleme yöntemlerinden Mantel test, Liu Agresti, Cox's β , poly-SIBTEST; DAF belirleme yöntemlerinden ise AC-LOR ve CU-LOR ile analizler gerçekleştirilmiştir.

Verilerin Analizi

Çok Kategorili DMF Analizleri. Maddelerin DMF gösterip göstermediğini belirlemek amacıyla parametrik olmayan çok kategorili DMF belirleme yöntemlerinden Mantel test, Liu Agresti istatistiği, Cox's β istatistikleri için DIFAS 5.0 programından faydalanılmıştır. Öncelikle veriler ilgili programın çalışma prensipleri gereği en küçük değer 0'dan başlayacak şekilde yeniden kodlanmış (1=0, 2=1, 3=2, 4=3, 5=4) ve analizlerde eşleştirme değişkeni olarak toplam puan kullanılmıştır. Çok kategorili DMF analizleri için; kategori birleştirme kuralları (8), odak grup- referans grup (2), odak grup örneklem büyüklüğü (2), odak grup: referans grup örneklem oranı (4), DMF belirleme yöntemi (4) olmak üzere $8*2*2*4*4 = 512$ hücreli bir desen oluşturulmuştur. Analiz çıktıları Ek-2 de verilmiştir. Mantel test sonuçları yorumlanırken .05 düzeyinde Tip I hata olasılığı için kritik değer 3,84 iken; .01 düzeyinde Tip I hata olasılığı için kritik değer 6,63 olarak kabul edilmiştir. Diğer yandan Liu Agresti istatistiği yorumlanırken analiz çıktılarındaki standartlaştırılmış Liu Agresti Kümülatif Ortak Log-Odds Oranı (LOR Z) değeri kullanılmıştır. Bu değer 2'den büyük ya da -2'den küçük olması durumunda maddede DMF varlığından bahsedilmiştir. Liu Agresti istatistiğinin pozitif olması referans grup, negatif olması odak grup lehine DMF olduğu anlamına gelmektedir. DIFAS program çıktılarından elde edilen diğer bir istatistik ise Cox's β istatistiğidir. Cox's β tablo değerinin kendi standart hatasına bölünmesiyle elde edilen Cox Z değerinin 2'den büyük ya da -2 den küçük olması durumunda maddede DMF varlığından söz edilmiştir. Cox's β değeri yaklaşık olarak standart normal dağılım göstermektedir. Bu değer pozitif olması referans grup lehine, negatif olması ise odak grup lehine DMF olduğu anlamını taşımaktadır (Penfield, 2013).

Çok kategorili DMF analizlerinden sonuncusu poly-SIBTEST yöntemiyle yapılmıştır. SIBTEST programı kullanılarak yapılan analizin sonuçları yorumlanırken β değeri ele alınmış ve $|\beta| \geq 0,088$ değerleri DMF'li (C düzeyi) olarak işaretlenmiştir (Roussos & Stout, 1996).

DAF Analizleri. Maddelerin DAF gösterip göstermediğini belirlemek amacıyla CU-LOR ve AC-LOR istatistikleri için DIFAS 5.0 programından faydalanılmıştır. Her bir maddenin her adımı için hesaplanan DAF değerleri incelenmiştir. Bu bağlamda analiz çıktılarından elde edilen $\hat{\lambda}_j$ değerleri yorumlanarak adımları yüksek DAF gösteren maddeler her iki yöntem için ayrı ayrı işaretlenmiştir. Yüksek DAF gösteren maddelerin işaretlenmesinde $|\hat{\lambda}_j| > 0,64$ ölçütü dikkate alınmıştır (Penfield, 2007; Penfield, Alvarez ve diğerleri, 2008).

Seçilen maddelere ilişkin çok kategorili DMF ve DAF analizi sonuçları grafikler yardımıyla Bulgular kısmında sunulmuştur. Grafiklerin yorumlanmasını kolaylaştırmak amacıyla her yönteme ait kritik değerler kesik çizgilerle belirtilmiştir. Yöntemlere ilişkin çıktıların yorumlanmasında, elde edilen DMF/DAF'lerin hangi grup lehine olduğunu belirlemek adına mutlak değer alma yoluna gidilmemiş; hesaplanan negatif/pozitif DMF/DAF değerleri orijinal haliyle sunulmuştur. Cox's B istatistiğinin yorumlanmasında, kendi standart hatasına bölünerek elde edilen Cox Z istatistiği; Liu Agresti istatistiğinin yorumlanmasında ise yine kendi standart hatasına bölünerek elde edilen LOR Z istatistiği kullanılmıştır. Bu bağlamda grafikte kesik çizgilerle belirtilen kritik değerler; Cox Z ve LOR Z için ± 2 , Mantel test için 6,63 ve poly-SIBTEST için ise $\pm 0,088$ 'dir. Cox B ve Liu Agresti istatistiklerinde $[-2,2]$ aralığı dışında; poly-SIBTEST istatistiğinde $-(0,088), (0,088)$ aralığı dışında, Mantel test istatistiğinde ise 6,63 çizgisinin üzerinde kalan değerler yüksek düzey DMF anlamına gelmektedir. Diğer yandan DAF analizleri için bu değer $\pm 0,64$ olup, $[-(0,64), (0,64)]$ aralığı dışında kalan değerler madde adımının ilgili koşullarda yüksek düzey DAF sergilediğini göstermektedir. DMF ve DAF grafiklerinde kritik değerlerden uzaklaşıldıkça DMF/DAF düzeyi artmakta olup bu uzaklaşma negatif yönde ise DMF/DAF odak grup lehine, pozitif yönde ise referans grup lehinedir.

Bölüm 4

Bulgular ve Yorum

Bu bölümde araştırma problemlerine yönelik yapılan çok kategorili DMF/DAF analizlerinden elde edilen sonuçlar grafikler yardımıyla sunulmuş ve yorumlanmıştır. Çok kategorili DMF analizleri için Cox's B, Liu Agresti, MT ve poly-SIBTEST yöntemleri; DAF analizleri için ise AC-LOR ve CU-LOR yöntemleri kullanılmıştır. Bulgular her bir alt problem altında ilgili maddeler için ayrı ayrı grafikler halinde sunulmuştur. Analizler tüm maddeler üzerinde yapılmış olmakla birlikte incelenen koşul sayısının fazla olması nedeniyle sadece M4, M9 ve M10'a ait bulgulara yer verilmiştir. Sonrasında maddelerin üçü birlikte ele alınmış, alanyazınla desteklenerek yorumlanmıştır.

Alt Problem 1'e İlişkin Bulgular

Odak grup örneklem büyüklüğü 200 (küçük) olduğunda, ülke karşılaştırmalarında çok kategorili DMF yöntemleriyle elde edilen DMF miktarları; değişen kategori birleştirme kuralı ve odak grup:referans grup örneklem oranlarında farklılık göstermekte midir?

Odak grup örneklem büyüklüğü 200 (küçük) olduğunda, değişen koşullarda çok kategorili DMF yöntemlerine (Mantel test, Liu- Agresti istatistiği, Cox β , poly-SIBTEST) göre Türkiye-ABD ve Türkiye-Kazakistan karşılaştırmalarından elde edilen sonuçlar M4 için Şekil 2'de, M9 için Şekil 3'te ve M10 için Şekil 4'te verilmiştir.



Şekil 2. Odak grup örneklem büyüklüğü 200 iken M4'e ait DMF miktarının değişimi.

Şekil 2’de örneklem büyüklüğü 200 iken M4’e ait DMF miktarının, farklı örneklem büyüklüğü oranları, DMF belirleme yöntemleri, kategori birleştirme kuralları ve ülkelere göre dağılımı yer almaktadır. M4’ün Türkiye-ABD karşılaştırmasında çeşitli koşullar altında elde edilen DMF miktarının değişimi incelendiğinde;

Örneklem büyüklüğü oranı 2:1 iken Cox’s B yönteminden elde edilen DMF miktarı son iki koşulda bir miktar azalma gösterse de koşullar bazında genel olarak benzer düzeyde olduğu söylenebilir. Cox’s B yönteminden elde edilen DMF miktarı -2,78 ile -1,97 arasında; Liu Agresti yönteminden elde edilen DMF miktarı -2,69 ile -2,11 arasında; MT yönteminden elde edilen DMF 3,87 ile 7,73 arasında ve son olarak poly-SIBTEST yönteminden elde edilen DMF ise -0,22 ile 0,01 arasında değişmektedir. Liu Agresti ve Cox’s B yöntemlerinden elde edilen DMF miktarının koşullar bazında kendi içinde genel olarak benzer düzeyde seyrettiği söylenebilir. Yöntemlerden elde edilen DMF miktarları Cox’s B yöntemine göre koşul 3 haricinde tüm koşullarda; Liu Agresti yöntemine göre tüm koşullarda ve MT yöntemine göre koşul 1, 2, 3 ve 7’de C düzeyindedir. Poly-SIBTEST yönteminden elde edilen DMF miktarı bazı koşullarda yüksek bazılarında ise düşük olup koşul 3, 7 ve 8 haricinde tüm koşullarda C düzeyindedir. Dolayısıyla bu yöntemden elde edilen DMF miktarına ait değişimin koşullar arasında sistematik bir örüntü sergilemediği söylenebilir.

Örneklem büyüklüğü oranı 1:1 iken Cox’s B, Liu Agresti ve MT yöntemlerinden elde edilen DMF miktarı sistematik bir örüntü sergilemese de DMF miktarının koşullar bazında ani bir düşme ya da artış göstermediği söylenebilir. Cox’s B ve Liu Agresti yöntemlerine göre tüm koşullarda C düzeyinde DMF hesaplanırken; MT yöntemine göre koşul 1 ve 2’den elde edilen DMF miktarları kritik değer altındadır. Poly-SIBTEST yönteminden elde edilen DMF miktarının ise koşul 7 de bir miktar artış gösterdiği, diğer koşullarda ise örneklem büyüklüğü oranı 2:1’den elde edilen DMF miktarları ile benzer düzeyde olduğu söylenebilir. Bu yöntemde göre koşul 3 haricinde tüm koşullarda C düzeyinde DMF elde edilmiştir. Cox’s B, Liu Agresti, MT ve poly-SIBTEST yöntemlerinden elde edilen DMF miktarları sırasıyla, -3,43 ile -2,16; -3,29 ile -2,19; 4,68 ile 11,76 ve -0,3 ile -0,07 arasında değişmektedir.

Örneklem büyüklüğü oranı 1:2 iken; Cox's B ve Liu Agresti yöntemlerinden elde edilen DMF miktarının bir önceki örneklem büyüklüğü oranına kıyasla bir miktar artış gösterdiği ancak bu değişimin koşullar bazında düzenli bir örüntü sergilemediği görülmektedir. MT yöntemine göre elde edilen DMF miktarı, bir önceki örneklem büyüklüğü oranına kıyasla özellikle koşul 2 ve koşul 5'te ani bir artış göstermiş, diğer koşullarda ise biraz artmış ancak koşullar bazında bu değişimlerin yine sistematik bir örüntü sergilemediği belirlenmiştir. Cox's B, Liu Agresti ve MT yöntemlerine göre tüm koşullarda C düzeyinde DMF elde edilmiştir. Poly-SIBTEST yönteminden elde edilen DMF miktarı ne bir önceki örneklem büyüklüğü oranına göre ne de koşullar arasında sistematik bir örüntü çizmiştir. Cox's B yönteminden elde edilen DMF miktarı -4.57 ile -2.81 arasında; Liu Agresti yönteminden elde edilen DMF miktarı -4.37 ile -2.83 arasında; MT yönteminden elde edilen DMF 7.89 ile 20.87 arasında ve son olarak poly-SIBTEST yönteminden elde edilen DMF ise -0.37 ile -0.04 arasında değişmekte değişmekte olup koşul 3 ve koşul 4 dışındaki bütün koşullarda C düzeyindedir.

Son olarak örneklem büyüklüğü oranı 1:3 iken, bütün koşullarda ve bütün yöntemlerinden elde edilen DMF miktarları bir önceki örneklem büyüklüğü oranından elde edilen DMF miktarına göre azalma gösterdiği belirlenmiştir. Ortaya çıkan bu değişimler koşullar bazında sistematik bir örüntü sergilememiştir. Cox's B yönteminden elde edilen DMF miktarı -4.07 ile -2.02 arasında; Liu Agresti yönteminden elde edilen DMF miktarı -3.84 ile -1.99 arasında; MT yönteminden elde edilen DMF 4.07 ile 16.46 arasında ve son olarak Poly-SIBTEST yönteminden elde edilen DMF ise -0.27 ile -0.02 arasında değişmektedir. Cox's B ve Liu Agresti yöntemlerinden elde edilen DMF sonuçları bütün koşullarda C düzeyinde iken MT ve poly-SIBTEST yöntemlerinden elde edilen sonuçlar koşulların çoğunda C düzeyindedir.

M4'ün Türkiye-Kazakistan karşılaştırmasında;

Örneklem büyüklükleri oranı 2:1 iken Cox's B yönteminden elde edilen DMF miktarının koşul 2 ve koşul 5'te C düzeyinde olduğu diğer koşullarda ise düşük ve benzer düzeyde olduğu söylenebilir. Cox's B yönteminden elde edilen DMF miktarının -2.40 ile -0.38 arasında değiştiği görülmektedir. Liu Agresti yönteminden elde edilen DMF miktarının ise koşul 2'de C düzeyinde olduğu, diğer koşullarda kritik değer altında

seyrettiği ve genel olarak koşullar arasında sistemli bir örüntü sergilemediği söylenebilir. Liu Agresti yönteminden elde edilen DMF miktarları -2.43 ile -0.37 arasında değişmektedir. MT yönteminden elde edilen DMF miktarının koşul 2 dışında genel olarak benzer düzeyde seyrettiği; koşul 2 ve koşul 5'ten elde edilen DMF'lerin kendi aralarında benzer ve nispeten yüksek olmakla birlikte kritik değer altında olduğu görülmektedir. MT yönteminden elde edilen DMF miktarı 0.14 ile 5.75 arasında değişmektedir. Grafikte, Poly-SIBTEST yöntemine göre koşul 3'te bu örneklem büyüklüğü oranı için herhangi bir değer olmadığı görülmektedir. Poly-SIBTEST yöntemi bu koşul için kestirim yapamamıştır. Genel olarak değerlendirildiğinde, Poly-SIBTEST yönteminden elde edilen DMF miktarının koşullar arasında düzenli bir örüntü sergilemediği söylenebilir. Maddenin koşul 1 ve 5 dışındaki tüm koşullarda C düzeyinde DMF sergilediği görülmekte ve bu DMF'nin koşul 4'te pozitif yönde, diğer koşullarda ise negatif yönde olması dikkat çekmektedir. Poly-SIBTEST yönteminden elde edilen DMF miktarları -0.42 ile 0.17 ranjında değişmektedir.

Örneklem büyüklüğü oranı 1:1 iken bütün yöntemlerde neredeyse bütün koşullarda elde edilen DMF miktarının bir önceki örneklem büyüklüğü oranına göre azalma gösterdiği görülmektedir. Cox's B yönteminden elde edilen DMF miktarı -0.28 ile 0.46 arasında Liu Agresti yönteminden elde edilen DMF miktarı ise -0.29 ile 0.42 arasında değişmektedir. MT yönteminde ise koşul 2 ve koşul 5 dışındaki bütün koşullarda bu örneklem büyüklüğü oranında elde edilen DMF miktarlarının bir önceki örneklem büyüklüğü oranından elde edilen DMF miktarı ile benzer olduğu ve bu benzerliğin koşullar bazında değişmediği ve genel olarak DMF'nin 0 düzeyinde seyrettiği söylenebilir. Cox's B, Liu Agresti ve MT yöntemlerine göre hiçbir koşulda C düzeyinde DMF yoktur. Poly-SIBTEST yönteminden elde edilen DMF miktarının koşullar arasında ve bir önceki örneklem büyüklüğü oranına kıyasla düzenli bir örüntü sergilemediği söylenebilir. Bu yöntemde koşul 3 haricinde tüm koşullarda C düzeyinde DMF mevcuttur. Poly-SIBTEST yönteminden elde edilen DMF miktarlarının -0.36 ile 0.1 arasında değiştiği görülmektedir.

Örneklem büyüklüğü oranı 1:2 iken MT dışındaki bütün yöntemlerden elde edilen DMF miktarının değişiminin, farklı koşullara ve bir önceki örneklem büyüklüğü oranına göre sistematik bir örüntü sergilemediği görülmektedir. Cox's B ve Liu Agresti

yöntemlerinden elde edilen DMF miktarı 0.06 ile 0.90 arasında değişmektedir. MT yönteminden elde edilen DMF miktarı ise bir önceki örneklem büyüklüğü oranından elde edilen DMF miktarı ile oldukça benzer ve 0'a yakın bir düzeydedir. Poly-SIBTEST yönteminden elde edilen DMF miktarının -0.03 ile 0.14 arasında değiştiği görülmektedir. Cox's B, Liu Agresti ve MT yöntemlerinden elde edilen DMF sonuçları hiçbir koşulda C düzeyinde olmayıp yalnızca poly-SIBTEST yönteminden elde edilen sonuçlar birkaç koşulda C düzeyinde DMF göstermiştir.

Son olarak örneklem büyüklüğü oranı 1:3 iken, Cox's B, Liu Agresti ve Poly-SIBTEST yöntemlerinden elde edilen DMF miktarları bir önceki örneklem büyüklüğü oranından elde edilen DMF miktarına göre bazı koşullarda artış bazı koşullarda azalma göstermiş olup düzenli bir örüntü sergilememiştir. MT yönteminden elde edilen DMF miktarının bir değişim göstermediği belirlenmiştir. MT yöntemi, koşullar değiştiğinde de benzer sonuçlar üretmiştir. Bu örneklem büyüklüğü oranında yöntemlerin ve koşulların tamamı için elde edilen DMF miktarları kritik değer altında olup, hiçbir yöntem ve koşulda C düzeyinde DMF bulunmamaktadır. Cox's B ve Liu Agresti yöntemlerinden elde edilen DMF miktarları sırasıyla -1.09 ile 0.12 ve -0.98 ile 0.11 arasında; Poly-SIBTEST yönteminden elde edilen DMF miktarı ise -0.08 ile 0.09 arasında değişmektedir.

M4'ün içerdiği DMF miktarı ülkelere göre değerlendirildiğinde, Cox's B, Liu Agresti ve MT yöntemlerinde bütün koşullarda; Poly-SIBTEST yönteminde ise birçok koşulda Türkiye-ABD karşılaştırmasında ortaya çıkan DMF miktarının Türkiye-Kazakistan karşılaştırmasında ortaya çıkan DMF'den çok yüksek olduğu dikkat çekmektedir. Türkiye-ABD karşılaştırmasında elde edilen DMF'nin çoğunlukla C düzeyinde ve negatif yönde olduğu görülmektedir. Dolayısıyla elde edilen DMF odak grup yani Türkiye lehinedir. Türkiye-Kazakistan karşılaştırmasında ise maddelerin Cox's B, Liu Agresti ve MT yöntemlerine göre çoğunlukla C düzeyinde DMF içermediği; poly-SIBTEST yöntemine göre ise bazı koşullarda Türkiye bazı koşullarda ise Kazakistan lehine C düzeyinde DMF olduğu görülmektedir. DMF'nin koşullar bazında yön değiştirmesi oldukça dikkat çekmektedir.



Şekil 3. Odak grup örneklem büyüklüğü 200 iken Madde 9'a ait DMF miktarının değişimi.

Şekil 3'te örneklem büyüklüğü 200 iken madde 9'a ait DMF miktarının, farklı örneklem büyüklüğü oranları, DMF belirleme yöntemleri, kategori birleştirme kuralları ve ülkelere göre dağılımı bulunmaktadır. Madde 9'un Türkiye-ABD karşılaştırmasında çeşitli koşullar altında elde edilen DMF miktarının değişimi incelendiğinde;

Örneklem büyüklüğü oranı 2:1 iken koşul 2'de Cox's B, Liu Agresti ve MT yöntemlerinden elde edilen DMF'lerin nispeten daha düşük, diğer koşullardaki DMF'nin ise birbirine benzer olmakla birlikte tüm koşullarda C düzeyinde olduğu görülmektedir. Poly-SIBTEST yönteminden elde edilen DMF ise koşullara göre düzenli bir örüntü çizmese de ilk üç koşulda benzer ve nispeten daha düşük olmakla birlikte yine tüm koşullarda C düzeyindedir. Bu örneklem büyüklüğü oranında Cox's B yönteminden elde edilen DMF miktarı 2.60 ile 3.51 arasında; Liu Agresti yönteminden elde edilen DMF miktarı 2.49 ile 3.68 arasında; MT yönteminden elde edilen DMF 6.77 ile 12.34 arasında ve son olarak Poly-SIBTEST yönteminden elde edilen DMF ise 0.3 ile 0.43 arasında değişmektedir.

Örneklem büyüklüğü oranı 1:1 iken Cox's B ve Liu Agresti yöntemlerinden elde edilen DMF miktarının bir önceki örneklem büyüklüğünden elde edilen DMF miktarına kıyasla azalma gösterdiği; koşullar bazında ise ilk beş koşulda benzer düzeylerde seyrettiği görülmektedir. Cox's B yönteminden elde edilen DMF miktarı 2.76 ile 3.87 arasında; Liu Agresti yönteminden elde edilen DMF miktarı 2.73 ile 3.87 arasında değişmektedir. MT ve Poly-SIBTEST yöntemlerinden elde edilen DMF sonuçlarının ise bir önceki örneklem büyüklüğü oranına göre bazı koşullarda azaldığı bazı koşullarda ise arttığı görülmektedir. Koşullar bazında ise yine sistematik bir örüntü sergilemediği belirlenmiştir. MT yönteminden elde edilen DMF 7.64 ile 15 arasında ve son olarak Poly-SIBTEST yönteminden elde edilen DMF ise 0.24 ile 0.5 arasında değişmektedir. Tüm yöntem ve koşullarda elde edilen DMF pozitif yönde ve C düzeyindedir.

Örneklem büyüklüğü oranı 1:2 iken Cox's B ve Liu Agresti yöntemlerinden elde edilen DMF miktarının bir önceki örneklem büyüklüğü oranına kıyasla bir miktar azalma meydana gelmiş olup; bu değişimin koşullar bazında düzenli bir örüntü sergilemediği görülmektedir. Örneklem büyüklüğü oranının 1:1 olduğu durumda ortaya çıkan sonuçlara benzer olarak, MT ve poly-SIBTEST yöntemlerinden elde edilen DMF sonuçlarının bir

önceki örneklem büyüklüğü oranına göre bazı koşullarda azaldığı bazı koşullarda ise arttığı görülmektedir. Koşullar bazında ise yine sistematik bir örüntü sergilenmediği belirlenmiştir. Cox's B yönteminden elde edilen DMF miktarı 2.62 ile 3.81 arasında; Liu Agresti yönteminden elde edilen DMF miktarı 2.51 ile 3.75 arasında; MT yönteminden elde edilen DMF 6.83 ile 14.51 arasında ve son olarak poly-SIBTEST yönteminden elde edilen DMF ise 0.21 ile 0.54 arasında değişmektedir. Bütün yöntemlerden elde edilen DMF sonuçları C düzeyinde olup pozitif yönde yani ABD lehinedir.

Son olarak örneklem büyüklüğü oranı 1:3 iken, bütün yöntemlerden elde edilen DMF miktarları bir önceki örneklem büyüklüğü oranından elde edilen DMF miktarına göre artış göstermiş olup; koşullar bazında düzenli bir örüntü olacak şekilde değişim göstermemiştir. Cox's B yönteminden elde edilen DMF miktarı 3.09 ile 4.45 arasında; Liu Agresti yönteminden elde edilen DMF miktarı 2.81 ile 4.21 arasında; MT yönteminden elde edilen DMF 9.54 ile 19.83 arasında ve son olarak Poly-SIBTEST yönteminden elde edilen DMF ise 0.27 ile 0.62 arasında değişmektedir. Yine bütün yöntemlerden elde edilen DMF sonuçları C düzeyinde olup pozitif yönde yani ABD lehinedir.

Madde 9'un Türkiye-Kazakistan karşılaştırmasında ortaya çıkan DMF miktarının değişimi incelendiğinde;

Örneklem büyüklüğü oranı 2:1 iken Cox's B, Liu Agresti ve MT yöntemlerinden elde edilen DMF miktarının koşullar bazında düzenli bir örüntü sergilemediği, Cox's B ve Liu Agresti yöntemleri için özellikle koşul 3 ve 4'teki DMF miktarlarının diğer koşullardakinden daha yüksek ve C düzeyinde olduğu görülmektedir. Her üç yöntemden elde edilen DMF miktarı örüntüsünün benzer olmakla birlikte MT yönteminden elde edilen tüm DMF miktarları kritik değer altındadır. M4'te olduğu gibi Poly-SIBTEST koşul 3'te bu örneklem büyüklüğü oranı için herhangi bir değer olmadığı görülmektedir. Poly-SIBTEST yöntemi bu koşul için kestirim yapamamıştır. Poly-SIBTEST yönteminden elde edilen DMF'lerin koşullar bazında düzenli bir örüntü sergilemese de DMF'lerin koşul 4 haricinde benzer düzeyde olduğu söylenebilir. Cox's B yönteminden elde edilen DMF miktarının 1.29 ile 2.58 arasında; Liu Agresti yönteminden elde edilen DMF miktarının 1.41 ile 2.36 arasında; MT yönteminden elde edilen DMF miktarının 1.67 ile 6.67 arasında

ve son olarak Poly-SIBTEST yönteminden elde edilen DMF miktarının ise 0 ile 0.41 arasında deęiřtięi görölmektedir.

Örneklem büyüklüęü oranı 1:1 iken Cox's B ve Liu Agresti yöntemlerinden elde edilen DMF miktarının bir önceki örneklem büyüklüęü oranından elde edilen DMF miktarına kıyasla daha benzer düzeylerde seyrettięi görölmektedir. MT yönteminde ise kořulların birçoęunda elde edilen DMF miktarı, bir önceki örneklem büyüklüęü oranına kıyasla artış göstermiř, ancak bu artış C düzeyinde DMF'ye neden olmamıřtır. MT yönteminden elde edilen DMF miktarının kořullar bazında benzer düzeyde olduęu söylenebilir. Son olarak Poly-SIBTEST yöntemlerinden elde edilen DMF sonuçlarının ise bir önceki örneklem büyüklüęü oranına göre azaldıęı görölmektedir. Kořullar bazında ise en fazla azalma kořul 5'te meydana gelmiř, ilk üç kořul kendi içinde, son üç kořul ise kendi içinde benzer düzeyde DMF içermektedir. Bu yöntemle göre kořul 2 ve 5'ten elde edilen DMF miktarları kritik deęerin altındadır. Cox's B yönteminden elde edilen DMF miktarının 1.48 ile 2.48 arasında; Liu Agresti yönteminden elde edilen DMF miktarının 1.46 ile 2.37 arasında; MT yönteminden elde edilen DMF miktarının 2.87 ile 6.36 arasında ve son olarak Poly-SIBTEST yönteminden elde edilen DMF miktarının ise -0.04 ile 0.16 arasında deęiřtięi görölmektedir.

Örneklem büyüklüęü oranı 1:2 iken Cox's B ve Liu Agresti yöntemlerinden elde edilen DMF miktarlarının kritik deęer etrafında seyrettięi; MT ve Poly-SIBTEST yöntemlerinden elde edilen DMF sonuçlarının ise bir önceki örneklem büyüklüęü oranına göre ilk üç kořulda azaldıęı son üç kořulda ise arttıęı görölmektedir. Özellikle kořul 6'da Poly-SIBTEST yönteminden elde edilen DMF miktarının 0'dan -0.25'e doęru ani bir ivmelenmesi gözlenmiřtir. Kořul 3, 4, 5, 7 ve 8'de pozitif yönde C düzeyinde olan DMF, kořul 6'da yön deęiřtirmiřtir. Cox's B yönteminden elde edilen DMF miktarının 1.68 ile 2.12 arasında; Liu Agresti yönteminden elde edilen DMF miktarının 1.59 ile 2.02 arasında; MT yönteminden elde edilen DMF miktarının 2.82 ile 12.36 arasında ve son olarak Poly-SIBTEST yönteminden elde edilen DMF miktarının ise -0.22 ile 0.23 arasında deęiřtięi görölmektedir.

Son olarak örneklem büyüklüęü oranı 1:3 iken, bütün yöntemlerden elde edilen DMF miktarları bir önceki örneklem büyüklüęü oranından elde edilen DMF miktarına göre

artış göstermiş olup; koşullar bazında düzenli bir örüntü olacak şekilde değişim göstermemiştir. Cox's B yönteminden elde edilen DMF miktarının 2.88 ile 3.9 arasında; Liu Agresti yönteminden elde edilen DMF miktarının 2.6 ile 3.55 arasında; MT yönteminden elde edilen DMF miktarının 8.28 ile 15.04 arasında ve son olarak Poly-SIBTEST yönteminden elde edilen DMF miktarının ise 0.16 ile 0.45 arasında değişmektedir. Dört yöntemden elde edilen DMF miktarı bütün koşullarda C düzeyinde olup, Kazakistan lehinedir.

Madde 9'un içerdiği DMF miktarı ülkelere göre değerlendirildiğinde, birçok koşulda Türkiye-ABD karşılaştırmasında elde edilen DMF'nin miktarının Türkiye-Kazakistan karşılaştırmasında ortaya çıkan DMF'den yüksek olduğu dikkat çekmektedir. İşaretler göz önünde bulundurulduğunda ise Türkiye-ABD karşılaştırmasında madde 9'un ABD lehine; Türkiye-Kazakistan karşılaştırmasında ise Kazakistan lehine olduğu görülmektedir. Bununla birlikte Türkiye-Kazakistan karşılaştırmasında örneklem büyüklüğü 1:2 iken koşul 6'da poly-SIBTEST yöntemine göre elde edilen DMF'nin Türkiye lehine olması dikkat çekmektedir.



Şekil 4. Odak grup örneklem büyüklüğü 200 iken Madde 10'a ait DMF miktarının değişimi.

Şekil 4'te örneklem büyüklüğü 200 iken madde 10'a ait DMF miktarının, farklı örneklem büyüklüğü oranları, DMF belirleme yöntemleri, kategori birleştirme kuralları ve ülkelere göre dağılımı yer almaktadır. Madde 10'un Türkiye-ABD karşılaştırmasında çeşitli koşullar altında elde edilen DMF miktarının değişimi incelendiğinde;

Örneklem büyüklüğü oranı 2:1 iken Cox's B yönteminden elde edilen DMF miktarının farklı koşullar arasındaki değişiminin düzenli olmadığı görülmektedir. Bu koşuldaki DMF'lerin -2.85 ile -0.83 arasında değiştiği görülmektedir. Liu Agresti yönteminden elde edilen DMF'ler -2.59 ile -0.87 arasında değişmekte olup; koşullar arasında bu değişim düzenli bir örüntü çizmemiştir. MT yöntemlerinden elde edilen DMF miktarının koşul 2 dışında, genel olarak benzer düzeyde seyrettiği söylenebilir. Koşul 2 hariç tutulduğunda MT yönteminden elde edilen DMF'lerin 0 ile 5 arasında değiştiği görülmektedir. Cox's B yöntemine göre koşul 2 ve 5'te; Liu Agresti ve MT yöntemlerine göre ise sadece koşul 2'de C düzeyinde DMF tespit edilmiştir. Poly-SIBTEST yönteminden elde edilen DMF miktarları da genel olarak benzer düzeylerde seyretmekle birlikte koşul 4 haricinde tüm koşullarda C düzeyinde DMF elde edilmiştir. Bu yöntemden elde edilen DMF'ler bu örneklem büyüklüğü oranında -0.2 ile -0.05 arasında değişmektedir.

Örneklem büyüklüğü oranı 1:1 iken Cox's B ve Liu Agresti yöntemlerinden elde edilen DMF miktarının bir önceki örneklem büyüklüğü oranına kıyasla genel olarak bir miktar artış gösterdiği ancak bu değişimin koşullar bazında düzenli bir örüntü sergilemediği görülmektedir. Bu örneklem büyüklüğü oranında Cox's B yönteminden elde edilen DMF'ler -3.02 ile -2.02; Liu Agresti yönteminden elde edilen DMF'ler ise -2.96 ile -2.09 ranjında değişmekte; madde tüm koşullarda C düzeyinde DMF içermektedir. MT yöntemine göre elde edilen DMF miktarı, bir önceki örneklem büyüklüğü oranına kıyasla artmıştır. MT yönteminden elde edilen DMF miktarı 4.08 ile 9.15 arasında değişmektedir. Poly-SIBTEST yönteminden elde edilen DMF miktarı bir önceki örneklem büyüklüğü oranından elde edilen DMF miktarına göre genel olarak biraz azalma göstermiştir. Bu örneklem büyüklüğü oranında, -0.3 ile 0 arasında değişmektedir. MT ve poly-SIBTEST yöntemlerinden elde edilen sonuçlar bazı koşullarda C düzeyinde DMF göstermiştir.

Örneklem büyüklüğü oranı 1:2 iken Cox's B, Liu Agresti ve MT yöntemlerinden elde edilen DMF miktarı bir önceki örneklem büyüklüğü oranına göre ani bir artış; Poly-SIBTEST yönteminde ise bir miktar bir artış göstermiştir. Bu örneklem büyüklüğü oranında tüm yöntem ve koşullarda C düzeyinde DMF elde edilmiştir. Cox's B yönteminden elde edilen DMF'ler -4.13 ile -3.57; Liu Agresti yönteminden elde edilen DMF'ler -3.93 ile -3.53; MT yönteminden elde edilen DMF'ler 12.78 ile 17.05; son olarak poly-SIBTEST yönteminden elde edilen DMF'ler -0.4 ile -0.1 ranjında değişmektedir.

Son olarak örneklem büyüklüğü oranı 1:3 iken, son iki koşul hariç bütün koşullarda Cox's B, Liu Agresti, MT ve Poly-SIBTEST yöntemlerinden elde edilen DMF miktarları bir önceki örneklem büyüklüğü oranından elde edilen DMF miktarına göre azalma göstermiş olup yine de tüm yöntem ve koşullarda C düzeyindedir. Son iki koşulda ise bütün yöntemlerden elde edilen DMF'ler bir miktar artış göstermiştir. Bu örneklem büyüklüğü oranında Cox's B yönteminden elde edilen DMF'ler -4.08 ile -3.22; Liu Agresti yönteminden elde edilen DMF'lerin ise -4.09 ile -3.18; MT yönteminden elde edilen DMF'ler 10.37 ile 16.63; son olarak Poly-SIBTEST yönteminden elde edilen DMF'ler -0.3 ile -0.1 arasında değişmektedir. Cox's B ve Liu Agresti yöntemlerinde koşulların benzer sonuçlar ürettiği; MT ve Poly-SIBTEST yöntemlerinde ise sonuçların koşullar bazında farklılaştığı söylenebilir.

Madde 10'un Türkiye-Kazakistan karşılaştırmasında;

Örneklem büyüklükleri oranı 2:1 iken Cox's B ve Liu Agresti yöntemlerinden elde edilen DMF miktarlarının kritik değer etrafında seyrettiği ve koşul 6'da en yüksek değerini aldığı görülmektedir. Cox's B yönteminden elde edilen DMF'nin -2.28 ile -1.48; Liu Agresti yönteminden elde edilen DMF'nin ise -2.65 ile -1.64 arasında değiştiği görülmektedir. MT yönteminden elde edilen DMF miktarının sadece koşul 6'da C düzeyinde DMF sergilediği; bunun dışında koşullar bazında genel olarak benzer olduğu söylenebilir. Bu örneklem büyüklüğünde MT yönteminden elde edilen DMF miktarları genel olarak 2.05 ile 7.84 arasında değişmektedir. Grafikte, yine poly-SIBTEST yönteminin koşul 3'te kestirim yapamamasından dolayı bu koşulda herhangi bir değer olmadığı görülmektedir. Genel olarak değerlendirildiğinde, poly-SIBTEST yönteminden elde edilen DMF miktarının -0.4 ile 0 arasında değiştiği ve koşullar arasında düzenli bir örüntü olmadığı söylenebilir. Bu

yönteme göre madde, koşul 4 ve 5 haricinde tüm koşullarda C düzeyinde DMF sergilemektedir.

Örneklem büyüklüğü oranı 1:1 iken bütün yöntemlerde koşullar bazında farklılaşmalar artmıştır. Buna göre bazı koşullarda elde edilen DMF miktarında bir önceki örneklem büyüklüğü oranına göre azalma bazılarında ise artış gösterdiği görülmektedir. MT yöntemine göre son üç koşulda hesaplanan DMF miktarındaki ani artış oldukça dikkat çekmektedir. Bu örneklem büyüklüğü oranında Cox's B yönteminden elde edilen DMF miktarı -2.43 ile -1.23; Liu Agresti yönteminde -2.46 ile -1.32; MT yönteminden elde edilen DMF miktarı 1.52 ile 15.63; Poly-SIBTEST yönteminden elde edilen DMF ise -0.6 ile 0.1 arasında değişmektedir. Bütün yöntemlerden elde edilen DMF miktarının koşullar arasında ve bir önceki örneklem büyüklüğü oranına kıyasla düzenli bir örüntü sergilemediği söylenebilir.

Örneklem büyüklüğü oranı 1:2 iken yöntemlerden elde edilen DMF miktarının değişiminin, farklı koşullara ve bir önceki örneklem büyüklüğü oranına göre sistematik bir örüntü sergilemediği görülmektedir. MT yönteminden elde edilen DMF miktarı ise bir önceki örneklem büyüklüğü oranından elde edilen DMF miktarı ile oldukça benzer düzeydedir. Bir önceki örneklem büyüklüğü oranında ani artış gösteren son üç koşuldaki DMF miktarlarında bir azalma görülmektedir. Cox's B yönteminden elde edilen DMF miktarı -3.51 ile -1.86 arasında; Liu Agresti yönteminden elde edilen DMF miktarı -3.80 ile -1.86 arasında; MT yönteminden elde edilen DMF miktarı 3.4 ile 11.49 arasında ve son olarak poly-SIBTEST yönteminden elde edilen DMF miktarı ise -0.30 ile 0.05 arasında değişmektedir.

Son olarak örneklem büyüklüğü oranı 1:3 iken, Cox's B, Liu Agresti ve Poly-SIBTEST yöntemlerinden elde edilen DMF miktarları bir önceki örneklem büyüklüğü oranından elde edilen DMF miktarına göre azalma göstermişken; MT yönteminden elde edilen DMF miktarlarının son üç koşulda ani bir ivmeyle arttığı belirlenmiştir. Son üç koşulda elde edilen DMF miktarı C düzeyinde ve oldukça yüksektir. MT dışındaki yöntemlerden elde edilen DMF miktarlarının farklı koşullar arasında sistemli olarak değişim göstermediği görülmektedir. Cox's B yönteminden elde edilen DMF miktarının -3.02 ile -1.59 arasında; Liu Agresti yönteminden elde edilen DMF miktarının -2.96 ile -

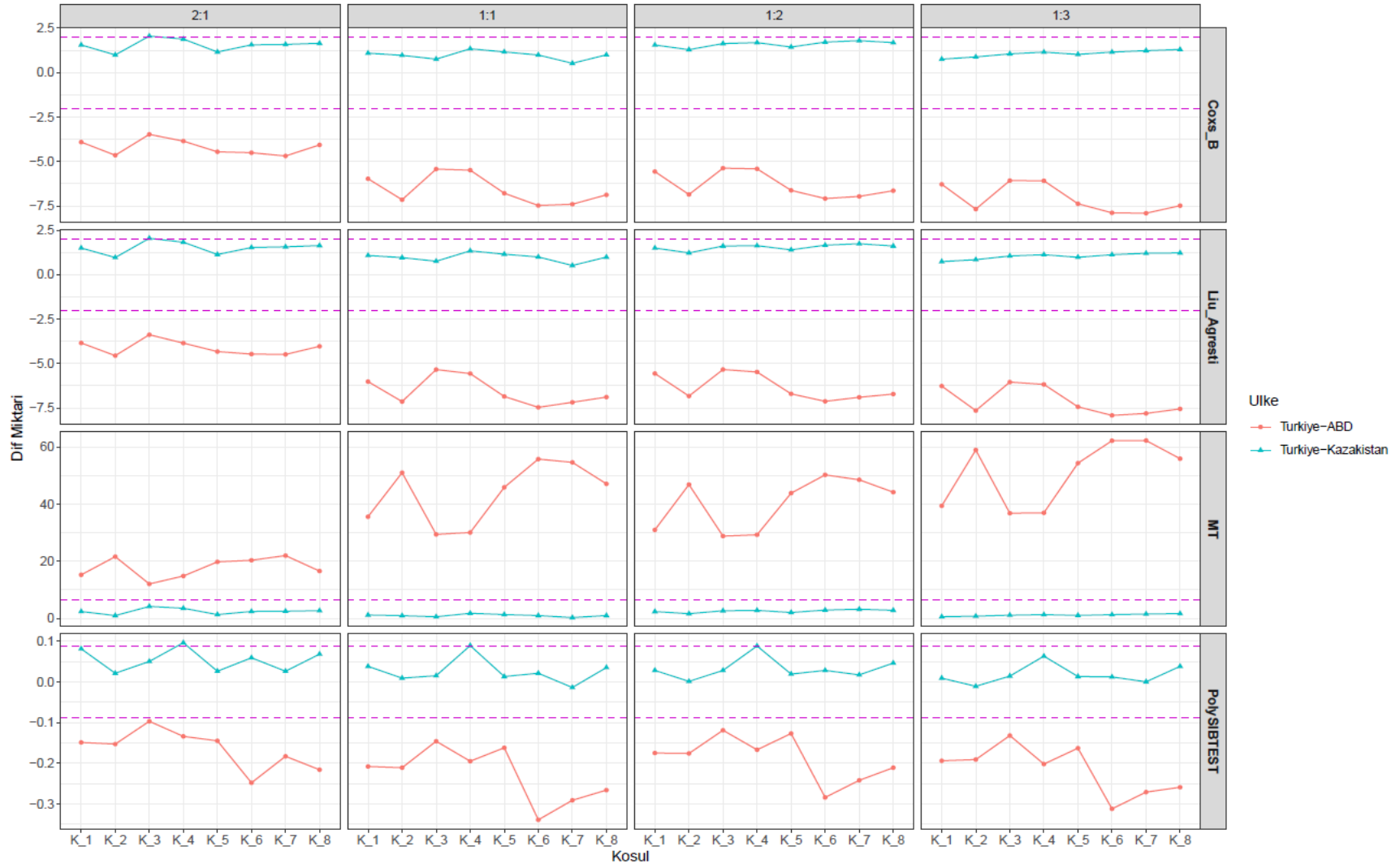
1.58 arasında; MT yönteminden elde edilen DMF miktarının 2.54 ile 23.72 arasında ve son olarak Poly-SIBTEST yönteminden elde edilen DMF miktarının ise -0.19 ile 0.11 arasında değişmektedir. Poly-SIBTEST yöntemine göre koşul 4'te diğer koşulların tersine pozitif yönde DMF tespit edilmiştir.

Madde 10'un içerdiği DMF miktarı ülkelere göre değerlendirildiğinde bütün yöntemlerde bazı koşullarda Türkiye-ABD karşılaştırmasında ortaya çıkan DMF miktarı Türkiye-Kazakistan karşılaştırmasında ortaya çıkan DMF'den yüksektir. Türkiye-ABD ve Türkiye-Kazakistan karşılaştırmalarında madde 10'un Türkiye lehine DMF sergilediği görülmektedir. Poly-SIBTEST yöntemine göre örneklem büyüklüğü oranı 1:3 iken koşul 4'te Türkiye-Kazakistan karşılaştırmasında bu madde Kazakistan lehine işlerken diğer koşullarda Türkiye lehine işlemektedir.

Alt Problem 2'ye İlişkin Bulgular

Odak grup örneklem büyüklüğü 1000 (büyük) olduğunda, ülke karşılaştırmalarında çok kategorili DMF yöntemleriyle elde edilen DMF miktarları; değişen kategori birleştirme kuralı ve odak grup:referans grup örneklem oranlarında farklılık göstermekte midir?

Odak grup örneklem büyüklüğü 1000 (büyük) olduğunda, değişen koşullarda çok kategorili DMF yöntemlerine (Mantel test, Liu- Agresti istatistiği, Cox β , poly-SIBTEST) göre Türkiye-ABD ve Türkiye-Kazakistan karşılaştırmalarından elde edilen sonuçlar M4 için Şekil 5'te, M9 için Şekil 6'da ve M10 için Şekil 7'de verilmiştir.



Şekil 5. Odak grup örneklem büyüklüğü 1000 iken M4'e ait DMF miktarının değişimi.

Şekil 5'te odak grup örneklem büyüklüğü 1000 iken M4'e ait DMF miktarının, farklı örneklem büyüklüğü oranları, DMF belirleme yöntemleri, kategori birleştirme kuralları ve ülkelere göre dağılımı yer almaktadır. M4'ün Türkiye-ABD karşılaştırmasında çeşitli koşullar altında elde edilen DMF miktarının değişimi incelendiğinde;

Örneklem büyüklüğü oranı 2:1 iken Cox's B, Liu Agresti ve MT yöntemlerinden elde edilen DMF miktarları ilk üç koşulda farklılaşmakla birlikte diğer koşullarda kendi içinde benzer düzeyde seyretmektedir. Cox's B yönteminden elde edilen DMF miktarları -4.69 ile -3.47 arasında; Liu Agresti yönteminden elde edilen DMF miktarları -4.56 ile -3.38 arasında; MT yönteminden elde edilen DMF 12.05 ile 21.98 arasında ve son olarak poly-SIBTEST yönteminden elde edilen DMF miktarları ise -0.25 ile -0.1 arasında değişmektedir. Poly-SIBTEST yönteminden elde edilen DMF miktarı bazı koşullarda yüksek bazılarında ise düşük olmakla birlikte en yüksek DMF miktarı koşul 6 da hesaplanmıştır. Bu yöntemden elde edilen DMF miktarına ait değişimin koşullar arasında sistematik bir örüntü sergilemediği söylenebilir. M4; Cox's B, Liu Agresti, MT ve poly-SIBTEST yöntemlerinin tamamında Türkiye örneklemini lehine C düzeyinde DMF sergilemektedir.

Örneklem büyüklüğü oranı 1:1 iken Cox's B, Liu Agresti ve MT yöntemlerinden elde edilen DMF miktarlarının sistematik bir örüntü sergilemediği, DMF miktarlarında koşullar bazında değişiklikler olduğu söylenebilir. Cox's B, Liu Agresti, MT ve Poly-SIBTEST yöntemlerinden elde edilen DMF miktarları sırasıyla, -7.47 ile -5.42; -7.47 ile -5.34; 29.40 ile 55.77 ve -0.34 ile -0.15 arasında değişmektedir. Yöntemlerin tamamında maddede Türkiye lehine yüksek düzey DMF olduğu görülmüş, en yüksek DMF miktarı koşul 6 için hesaplanmıştır. Ayrıca tüm yöntemler için örneklem büyüklüğü oranı 2:1'e göre DMF miktarlarında artış olduğu görülmektedir.

Örneklem büyüklüğü oranı 1:2 iken; Cox's B, Liu Agresti ve MT yöntemlerinden elde edilen DMF miktarının bir önceki örneklem büyüklüğü oranıyla kıyaslandığında benzer bir örüntü sergilediği görülmektedir. Bununla birlikte poly-SIBTEST yöntemine göre elde edilen DMF miktarları, bir önceki örneklem büyüklüğü oranına kıyasla daha düşük hesaplanmıştır. Cox's B yönteminden elde edilen DMF miktarları -7.08 ile -5.37 arasında; Liu Agresti yönteminden elde edilen DMF miktarı -7.13 ile -5.34 arasında; MT

yönteminden elde edilen DMF 28.80 ile 50.24 arasında ve son olarak poly-SIBTEST yönteminden elde edilen DMF ise -0.28 ile -0.12 arasında değişmektedir.

Son olarak örneklem büyüklüğü oranı 1:3 iken, bütün koşullarda ve bütün yöntemlerden elde edilen DMF miktarları bir önceki örneklem büyüklüğü oranından elde edilen DMF miktarına göre azalma gösterdiği belirlenmiştir. Ortaya çıkan bu değişimler koşullar bazında sistematik bir örüntü sergilememiştir. Cox's B yönteminden elde edilen DMF miktarı -7.90 ile -6.07 arasında; Liu Agresti yönteminden elde edilen DMF miktarı -7.81 ile -6.05 arasında; MT yönteminden elde edilen DMF 36.81 ile 62.27 arasında ve son olarak poly-SIBTEST yönteminden elde edilen DMF ise -0.31 ile -0.13 arasında değişmektedir.

M4'ün Türkiye-Kazakistan karşılaştırmasında;

Örneklem büyüklükleri oranı 2:1 iken Cox's B, Liu Agresti ve MT yöntemlerinden elde edilen DMF miktarlarının kendi içinde koşullar bazında benzer düzeyde olduğu, koşullar arasında ani düşüş ve artışların olmadığı söylenebilir. Poly-SIBTEST yönteminde hesaplanan DMF miktarlarında ise koşullar arasında farklılaşmalar olduğu görülmektedir. Cox's B yönteminden elde edilen DMF miktarının 0.99 ile 2.05 arasında değiştiği görülmektedir. Liu Agresti yönteminden elde edilen DMF miktarının ise koşul 2 ve koşul 5'te en düşük ve benzer olduğu; genel olarak koşullar arasında sistemli bir örüntü sergilemediği söylenebilir. Liu Agresti yönteminden elde edilen DMF miktarları 0.96 ile 2.04 arasında değişmektedir. MT yönteminden elde edilen DMF miktarlarının genel olarak benzer düzeyde seyrettiği; koşul 3 ve koşul 4'te elde edilen DMF'lerin kendi aralarında benzer ve nispeten yüksek söylenebilir. MT yönteminden elde edilen DMF miktarları 0.98 ile 4.18 arasında değişmektedir. Genel olarak değerlendirildiğinde, poly-SIBTEST yönteminden elde edilen DMF miktarının koşullar arasında düzenli bir örüntü sergilemediği ancak koşullar değiştikçe DMF miktarlarında düşüş veya artış olduğu söylenebilir. Poly-SIBTEST yönteminden elde edilen DMF miktarları 0.02 ile 0.10 arasında değişmektedir.

Örneklem büyüklüğü oranı 1:1 iken bütün yöntemlerde hemen hemen bütün koşullarda elde edilen DMF miktarlarında bir önceki örneklem büyüklüğü oranına göre azalma olduğu görülmektedir. Genel olarak Cox's B, Liu Agresti ve MT yöntemlerinden

elde edilen DMF'nin koşullar arasında ani bir deęişim göstermedięi; poly-SIBTEST yönteminde ise koşullara göre farklılaştığı söylenebilir. Cox's B yönteminden elde edilen DMF miktarı 0.52 ile 1.33 arasında, Liu Agresti yönteminden elde edilen DMF miktarı ise 0.51 ile 1.34 arasında, MT yönteminde ise 0.26 ile 1.77 arasında deęişmektedir. Poly-SIBTEST yönteminden elde edilen DMF miktarlarının koşul 4 haricinde benzer düzeylerde olup, -0.01 ile 0.09 arasında deęiştii görülmektedir.

Örneklem büyüklüğü oranı 1:2 iken yöntemlerden elde edilen DMF miktarının deęişiminin, farklı koşullarda birbirinin aynısı olacak şekilde bir örüntü sergilediği görülmektedir. Cox's B ve Liu Agresti yöntemlerinden elde edilen DMF miktarları birbirine oldukça yakın olup 1.22 ile 1.79 arasında deęişmektedir. MT yönteminden elde edilen DMF miktarı ise bir önceki örneklem büyüklüğü oranından elde edilen DMF miktarı ile oldukça benzer ve koşullar bazında birbirine yakın düzeydedir. Poly-SIBTEST yönteminden elde edilen DMF miktarının 0 ile 0.09 arasında deęiştii görülmektedir.

Son olarak örneklem büyüklüğü oranı 1:3 iken, Cox's B, Liu Agresti ve poly-SIBTEST yöntemlerinden elde edilen DMF miktarları bir önceki örneklem büyüklüğü oranından elde edilen DMF miktarına göre hemen hemen her koşulda azalma göstermiş olup düzenli bir örüntü sergilememiştir. MT yönteminden elde edilen DMF miktarlarının bir önceki örneklem oranına göre çok fazla deęişim göstermediği ve 0'a yakın düzeylerde olduğu belirlenmiştir. MT yöntemi, koşullar deęiştiiğinde de benzer sonuçlar üreterek 0.56 ile 1.66 arasında deęerler almıştır. Cox's B ve Liu Agresti yöntemlerinden elde edilen DMF miktarı 0.73 ile -1.29; poly-SIBTEST yönteminden elde edilen DMF miktarı ise -0.01 ile 0.06 arasında deęişmektedir.

M4'ün içerdii DMF miktarı ülkelere göre deęerlendirildiğinde, Türkiye-ABD karşılaştırmasında elde edilen grafiklerin kritik deęerlerden oldukça uzak olduğu dolayısıyla iki grup için hesaplanan DMF'nin yüksek düzeyde olduğu görülmektedir. Grafiklerin Cox's B, Liu Agresti ve poly-SIBTEST yöntemlerinde negatif yönde olması hesaplanan DMF'nin Türkiye örnelemi lehine olduğunu göstermektedir. Dięer yandan Türkiye-Kazakistan karşılaştırmasında elde edilen grafiklerin yöntemler için belirlenen kritik deęerleri çoğunlukla aşmadığı, dolayısıyla M4'te genel olarak yüksek düzeyde DMF olmadığı ifade edilebilir.



Şekil 6. Odak grup örneklem büyüklüğü 1000 iken Madde 9'a ait DMF miktarının değişimi.

Şekil 6'da örneklem büyüklüğü 1000 iken madde 9'a ait DMF miktarının, farklı örneklem büyüklüğü oranları, DMF belirleme yöntemleri, kategori birleştirme kuralları ve ülkelere göre dağılımı bulunmaktadır. Madde 9'un Türkiye-ABD karşılaştırmasında çeşitli koşullar altında elde edilen DMF miktarlarının değişimi incelendiğinde;

Örneklem büyüklüğü oranı 2:1 iken Cox's B, Liu Agresti ve MT yöntemlerinden elde edilen DMF'lerin koşul 6 ve 7'de nispeten yüksek, koşul 3,4 ve 5'te birbirine benzer ve nispeten düşük olduğu görülmektedir. Poly-SIBTEST yöntemlerinden elde edilen DMF ise koşullara göre düzenli bir örüntü çizmese de koşullar bazında elde edilen DMF'lerin birbirinden oldukça farklılaştığı söylenebilir. Bu örneklem büyüklüğü oranında Cox's B yönteminden elde edilen DMF miktarı 4.09 ile 5.37 arasında; Liu Agresti yönteminden elde edilen DMF miktarı 4.01 ile 5.29 arasında; MT yönteminden elde edilen DMF 16.72 ile 28.84 arasında ve son olarak poly-SIBTEST yönteminden elde edilen DMF ise 0.12 ile 0.3 arasında değişmektedir. Bu örneklem büyüklüğü oranında tüm yöntemlerden elde edilen DMF'nin C düzeyinde olduğu görülmektedir.

Örneklem büyüklüğü oranı 1:1 iken Cox's B, Liu Agresti ve MT yöntemlerinden elde edilen DMF miktarının bir önceki örneklem büyüklüğünden elde edilen DMF miktarına kıyasla büyük bir artış gösterdiği; koşullar bazında ise sistematik bir değişim çizmediği söylenebilir. Üç yöntem için de en düşük DMF miktarlarının koşul 1, 3 ve 4 için hesaplandığı görülmektedir. Bir önceki örneklem büyüklüğü oranında düşük DMF sergileyen sonuçların koşul 5'teki ani artış ve grafikte oluşturduğu değişim dikkat çekmektedir. Cox's B yönteminden elde edilen DMF miktarı 4.60 ile 7.73 arasında; Liu Agresti yönteminden elde edilen DMF miktarı 4.48 ile 7.48 arasında; MT yönteminden elde edilen DMF miktarı ise 21.03 ile 59.73 arasında değişmektedir. Poly-SIBTEST yönteminden elde edilen DMF sonuçlarının ise bir önceki örneklem büyüklüğü oranına göre her koşulda artmakla birlikte, DMF miktarlarının koşullar bazında farklı bir örüntü sergilediği görülmektedir. Ancak bu örüntü sistematik bir örüntü olmayıp Poly-SIBTEST yönteminden elde edilen DMF miktarları 0.14 ile 0.38 arasında değişmektedir. Bu örneklem büyüklüğü oranında tüm yöntemlerden elde edilen DMF C düzeyindedir.

Örneklem büyüklüğü oranı 1:2 iken Cox's B, Liu Agresti ve MT yöntemlerinden elde edilen DMF miktarında bir önceki örneklem büyüklüğü oranına kıyasla bir miktar artış

olmakla birlikte koşullar bazında benzer örüntünün sergilendiği görülmektedir. Poly-SIBTEST yönteminden elde edilen DMF sonuçları ise bir önceki örneklem büyüklüğü oranı ile benzer örüntüyü göstermekle birlikte DMF miktarlarında azalma olmuştur. Koşullar bazında ise yine sistematik bir örüntü sergilenmediği belirlenmiştir. Cox's B yönteminden elde edilen DMF miktarı 4.96 ile 8.11 arasında; Liu Agresti yönteminden elde edilen DMF miktarı 4.90 ile 8.13 arasında; MT yönteminden elde edilen DMF miktarı 24.50 ile 65.85 arasında ve son olarak poly-SIBTEST yönteminden elde edilen DMF ise 0.14 ile 0.32 arasında değişmektedir. Yine bu örneklem büyüklüğü oranında tüm yöntemlerden elde edilen DMF C düzeyindedir.

Son olarak örneklem büyüklüğü oranı 1:3 iken, bütün yöntemlerden elde edilen DMF miktarları bir önceki örneklem büyüklüğü oranından elde edilen DMF miktarıyla benzer olup koşullar bazında düzenli bir örüntü olacak şekilde değişim göstermemiştir. Cox's B yönteminden elde edilen DMF miktarı 5.29 ile 8.5 arasında; Liu Agresti yönteminden elde edilen DMF miktarı 5.29 ile 8.56 arasında; MT yönteminden elde edilen DMF 27.95 ile 73.37 arasında ve son olarak poly-SIBTEST yönteminden elde edilen DMF ise 0.13 ile 0.31 arasında değişmektedir. Bu örneklem büyüklüğü oranında tüm yöntemlerden elde edilen DMF'nin C düzeyinde olduğu görülmektedir.

Türkiye-ABD karşılaştırmasında bütün koşullar ve örneklem büyüklüğü oranları birlikte ele alındığında Cox's B, Liu Agresti ve MT yöntemlerinde örneklem büyüklüğü oranı 2:1 iken koşulların birbirine yakın sonuçlar verdiği, diğer örneklem oranlarında ise koşullar bazında sonuçların oldukça farklılaştığı görülmektedir. Tüm durumlarda hesaplanan DMF miktarları yüksek düzeyde olmakla birlikte en yüksek DMF, örneklem büyüklüğü oranı 1:3 iken hesaplanmıştır. Poly-SIBTEST yönteminde ise yine hesaplanan bütün DMF miktarları yüksek düzeyde olmakla birlikte en yüksek DMF miktarları örneklem büyüklüğü oranı 1:1 iken elde edilmiştir. Bütün koşullarda ABD lehine C düzeyinde DMF elde edilmiştir.

Madde 9'un Türkiye-Kazakistan karşılaştırmasında ortaya çıkan DMF miktarının değişimi incelendiğinde;

Örneklem büyüklüğü oranı 2:1 iken Cox's B, Liu Agresti ve MT yöntemlerinden elde edilen DMF miktarının koşullar bazında benzer düzeylerde olmakla birlikte, koşul 1,

6 ve 7'deki DMF miktarlarının diğer koşullardakinden daha yüksek olduğu görülmektedir. Her üç yöntemden elde edilen DMF miktarı örüntüsünün benzer olduğu görülmektedir. Poly-SIBTEST yönteminden elde edilen DMF'lerin koşullar bazında farklılaştığı; koşul 2 ve 3'te DMF miktarlarının kritik değerin altında olduğu dikkat çekmektedir. Cox's B yönteminden elde edilen DMF miktarının 3.14 ile 3.96 arasında; Liu Agresti yönteminden elde edilen DMF miktarının 3.01 ile 3.91 arasında; MT yönteminden elde edilen DMF miktarının 9.85 ile 15.71 arasında ve son olarak Poly-SIBTEST yönteminden elde edilen DMF miktarının ise 0.07 ile 0.16 arasında değiştiği görülmektedir.

Örneklem büyüklüğü oranı 1:1 iken özellikle koşul 3 ve 7'de Cox's B, Liu Agresti ve MT yöntemlerinden elde edilen DMF miktarının bir önceki örneklem büyüklüğü oranından elde edilen DMF miktarına kıyasla ani bir artış gösterdiği; bu artışın dağılımın örüntüsünü değiştirdiği; diğer koşullarda ise örüntüyü değiştirmeyen minimal artışların olduğu görülmektedir. Bir önceki örneklem büyüklüğü oranında tüm yöntemlerde en düşük DMF miktarının olduğu koşul 3'ün bu örneklem büyüklüğü oranında en yüksek DMF miktarları arasına girmesi oldukça dikkat çekicidir. Poly-SIBTEST yöntemlerinden elde edilen DMF sonuçlarındaki örüntünün ise bir önceki örneklem büyüklüğü oranına göre farklılaştığı görülmektedir. Daha önce koşul 2 ve 3 kritik değer altında DMF değeri verirken 1:1 örneklem büyüklüğü oranında koşul 3'e ait DMF miktarı artmış, koşul 5'e ait DMF miktarı ise azalarak kritik değerin altına düşmüştür. Cox's B yönteminden elde edilen DMF miktarının 3.35 ile 4.92 arasında; Liu Agresti yönteminden elde edilen DMF miktarının 3.14 ile 4.75 arasında; MT yönteminden elde edilen DMF miktarının 11.19 ile 24.18 arasında ve son olarak poly-SIBTEST yönteminden elde edilen DMF miktarının ise 0.06 ile 0.16 arasında değiştiği görülmektedir.

Örneklem büyüklüğü oranı 1:2 iken koşul 2, 5, 6 ve 8'de Cox's B, Liu Agresti ve MT yöntemlerinden elde edilen DMF miktarının bir önceki örneklem büyüklüğü oranına kıyasla bir miktar artış; koşul 3'te ise genel olarak azalma gösterdiği ancak bu değişimin koşullar bazında düzenli olmadığı görülmektedir. Poly-SIBTEST yöntemlerinden elde edilen DMF sonuçlarının ise bir önceki örneklem büyüklüğü oranına göre genel olarak azaldığı görülmektedir. Bu durumda kritik değerin altında kalan koşul sayısı artmıştır. Bir önceki örneklem büyüklüğü oranında DMF miktarı artan koşul 3 tekrar kritik değerin

altında kalmıştır. Cox's B yönteminden elde edilen DMF miktarının 3.81 ile 5.14 arasında; Liu Agresti yönteminden elde edilen DMF miktarının 3.65 ile 4.84 arasında; MT yönteminden elde edilen DMF miktarının 14.5 ile 26.36 arasında ve son olarak poly-SIBTEST yönteminden elde edilen DMF miktarının ise 0.09 ile 0.12 arasında değiştiği görülmektedir.

Son olarak örneklem büyüklüğü oranı 1:3 iken, Cox's B, Liu Agresti ve MT yöntemlerinden elde edilen DMF miktarları bir önceki örneklem büyüklüğü oranından elde edilen DMF miktarına göre minimal artış göstermiş olup; koşullar bazında düzenli bir örüntü olacak şekilde değişim göstermemiştir. Poly-SIBTEST yönteminde ise bu durumun aksine bütün koşullarda DMF miktarı azalmıştır. Son durumda koşul 1, 2, 3, 5 ve 7'de hesaplanan DMF değerleri kritik değer altında kalmıştır. Dolayısıyla bu koşulların DMF göstermediği söylenebilir. Cox's B yönteminden elde edilen DMF miktarının 4.30 ile 5.62 arasında; Liu Agresti yönteminden elde edilen DMF miktarının 4.08 ile 5.42 arasında; MT yönteminden elde edilen DMF miktarının 18.46 ile 31.56 arasında ve son olarak poly-SIBTEST yönteminden elde edilen DMF miktarının ise 0.06 ile 0.12 arasında değiştiği görülmektedir.

Madde 9'un içerdiği DMF miktarı ülkelere göre değerlendirildiğinde, koşulların tamamında Türkiye-ABD karşılaştırmasında elde edilen DMF'nin miktarının Türkiye-Kazakistan karşılaştırmasında ortaya çıkan DMF'den yüksek olduğu dikkat çekmektedir. Karşılaştırmalarda en yüksek DMF değerleri Cox's B, Liu Agresti ve MT yöntemlerine göre örneklem büyüklüğü oranı 1:3'te hesaplanırken; poly-SIBTEST yöntemine göre örneklem büyüklüğü oranı 2:1'te hesaplanmıştır. Madde 9, Türkiye-ABD karşılaştırmasında tüm koşullarda ABD lehine yüksek düzey DMF gösterirken; Türkiye-Kazakistan karşılaştırmasında Cox's B, Liu Agresti ve MT yöntemlerine göre tüm koşullarda Kazakistan lehine yüksek düzey DMF göstermiştir. Ancak maddenin poly-SIBTEST yöntemine göre bazı koşullarda Kazakistan lehine yüksek düzey DMF gösterdiği, bazı koşullarda ise DMF göstermediği belirlenmiştir.



Şekil 7. Odak grup örneklem büyüklüğü 1000 iken Madde 10'a ait DMF miktarının değişimi.

Şekil 7’de örneklem büyüklüğü 1000 iken madde 10’a ait DMF miktarının, farklı örneklem büyüklüğü oranları, DMF belirleme yöntemleri, kategori birleştirme kuralları ve ülkelere göre dağılımı yer almaktadır. Madde 10’un Türkiye-ABD karşılaştırmasında çeşitli koşullar altında elde edilen DMF miktarının değişimi incelendiğinde;

Örneklem büyüklüğü oranı 2:1 iken Cox’s B, Liu Agresti ve MT yöntemlerinden elde edilen DMF miktarının farklı koşullar arasındaki değişimi düzenli olmamakla birlikte son dört koşulda benzer sonuçlar verdiği görülmektedir. Bu koşuldaki DMF’lerin Cox’s B yöntemine göre -5.40 ile -4.16 arasında; Liu Agresti yöntemine göre -5.44 ile -4.06 arasında; MT yöntemine göre ise 17.27 ile 29.2 arasında değiştiği görülmektedir. Her üç yöntemde de en yüksek DMF koşul 2’de elde edilmiştir. Poly-SIBTEST yönteminden elde edilen DMF miktarlarının da koşullar bazında farklılaştığı en yüksek DMF’nin ise koşul 6’da elde edildiği söylenebilir. Bu yöntemden elde edilen DMF’ler bu örneklem büyüklüğü oranında -0.25 ile -0.11 arasında değişmektedir.

Örneklem büyüklüğü oranı 1:1 iken Cox’s B ve Liu Agresti yöntemlerinden elde edilen dağılımın bir önceki örneklem büyüklüğü oranına kıyasla kritik değerden bir miktar uzaklaştığı, dolayısıyla DMF’nin minimal düzeyde artış gösterdiği; ancak bu değişimin koşullar bazında düzenli bir örüntü sergilemediği görülmektedir. Bu örneklem büyüklüğü oranında Cox’s B yönteminden elde edilen DMF’ler -5.7 ile -4.43; Liu Agresti yönteminden elde edilen DMF’lerin ise -5.81 ile -4.46 ranjında değişmektedir. MT yöntemine göre elde edilen DMF miktarı, bir önceki örneklem büyüklüğü oranına kıyasla minimal düzeyde artmıştır. MT yönteminden elde edilen DMF miktarı 19.55 ile 32.57 arasında değişmektedir. Diğer yöntemlerin aksine Poly-SIBTEST yönteminden elde edilen DMF miktarı bir önceki örneklem büyüklüğü oranından elde edilen DMF miktarına göre bazı koşullarda azalırken bazı koşullarda artmış; bu durum dağılımın örüntüsü de değiştirmiştir. Bu örneklem büyüklüğü oranında Poly-SIBTEST yönteminden elde edilen DMF miktarı -0.21 ile 0.1 arasında değişmektedir.

Örneklem büyüklüğü oranı 1:2 iken Cox’s B, Liu Agresti ve MT yöntemlerinden elde edilen DMF miktarı bir önceki örneklem büyüklüğü oranına göre ani bir artış; Poly-SIBTEST yönteminde ise bir miktar artış göstermiştir. Genel olarak dağılımların örüntüleri

yöntemler bazında değişmemiştir. Bu örneklem büyüklüğü oranında, Cox's B yönteminden elde edilen DMF'ler -8.11 ile -6.65; Liu Agresti yönteminden elde edilen DMF'lerin ise -8.03 ile -6.48; MT yönteminden elde edilen DMF'ler 44.27 ile 65.83; son olarak Poly-SIBTEST yönteminden elde edilen DMF'ler -0.28 ile -0.15 ranjında değişmektedir.

Son olarak örneklem büyüklüğü oranı 1:3 iken, tüm yöntemlere göre hemen hemen bütün koşullarda elde edilen DMF miktarları bir önceki örneklem büyüklüğü oranından elde edilen DMF miktarlarına benzer düzeylerde olup dağılımlar benzer örüntüye sahiptir. Bu örneklem büyüklüğü oranında Cox's B yönteminden elde edilen DMF'ler -8.01 ile -6.51; Liu Agresti yönteminden elde edilen DMF'ler -7.92 ile -6.36; MT yönteminden elde edilen DMF'ler 42.43 ile 64.25; son olarak Poly-SIBTEST yönteminden elde edilen DMF'ler -0.27 ile -0.15 arasında değişmektedir.

Madde 10'un Türkiye-Kazakistan karşılaştırmasında;

Örneklem büyüklükleri oranı 2:1 iken Cox's B ve Liu Agresti yöntemlerinden elde edilen DMF miktarının koşullar bazında farklılaşmakta olup, koşul 1 ve 4'te elde edilen DMF değerleri kritik değerleri aşmamaktadır. Diğer koşullarda ise negatif yönde yüksek düzey DMF olduğu görülmektedir. Cox's B yönteminden elde edilen DMF'nin -3.3 ile -1.55; Liu Agresti yönteminden elde edilen DMF'nin ise -3.31 ile -1.57 arasında değiştiği görülmektedir. MT yönteminden elde edilen DMF miktarının koşul 1, 3 ve 4'te düşük düzeyde olmakla birlikte koşullar bazında genel olarak benzer olduğu söylenebilir. Bu örneklem büyüklüğünde MT yönteminden elde edilen DMF miktarları 2.41 ile 10.91 arasında değişmektedir. Poly-SIBTEST yönteminden elde edilen DMF miktarları incelendiğinde koşul 3'ten koşul 6'ya gidilirken negatif yönde doğrusal olarak bir artış olduğu görülmektedir. Koşul 1 ve 3'ten elde edilen DMF değerleri kritik değerleri aşmamaktadır. Poly-SIBTEST yönteminden elde edilen DMF miktarları -0.16 ile -0.07 arasında değişmektedir.

Örneklem büyüklüğü oranı 1:1 iken Cox's B ve Liu Agresti yöntemlerinden elde edilen DMF miktarları bir önceki örneklem büyüklüğü oranına göre artış göstermiştir. Daha önce kritik değeri aşmayan koşul 1 ve 4'ün bu örneklem büyüklüğü oranında kritik değerleri aşarak yüksek düzey DMF sergiledikleri görülmektedir. Bu örneklem büyüklüğü

oranında Cox's B yönteminden elde edilen DMF miktarı -2.34 ile -5.2; Liu Agresti yönteminde -2.36 ile -5.25 arasında değişmektedir. MT yönteminden elde edilen DMF miktarları incelendiğinde en yüksek DMF'nin koşul 7'de hesaplandığı görülmektedir. MT yönteminden elde edilen DMF miktarları 5.45 ile 27.07; Poly-SIBTEST yönteminden elde edilen DMF ise -0.08 ile -0.19 arasında değişmektedir. Bütün yöntemlerden elde edilen DMF miktarının koşullar arasında ve bir önceki örneklem büyüklüğü oranına kıyasla düzenli bir örüntü sergilemediği söylenebilir.

Örneklem büyüklüğü oranı 1:2 iken yöntemlerden elde edilen DMF miktarının değişiminin, farklı koşullara ve bir önceki örneklem büyüklüğü oranına göre sistematik bir örüntü sergilemediği görülmektedir. Cox's B ve Liu Agresti ve MT yöntemlerine ait örüntülerdeki değişimin büyük oranda koşul 3'ten kaynaklandığı söylenebilir. Diğer koşullarda hesaplanan DMF miktarlarında artış olurken koşul 3'te bir önceki örneklem büyüklüğü oranına göre azalma olmuştur. Cox's B yönteminden elde edilen DMF miktarı -4.85 ile -2.8 arasında; Liu Agresti yönteminden elde edilen DMF miktarı -4.84 ile -2.74 arasında; MT yönteminden elde edilen DMF miktarı 7.87 ile 23.53 arasında ve son olarak poly-SIBTEST yönteminden elde edilen DMF miktarı ise -0.18 ile -0.04 arasında değişmektedir.

Son olarak örneklem büyüklüğü oranı 1:3 iken, tüm yöntemlerde koşullar bazında elde edilen DMF miktarlarının ve dağılıma ilişkin örüntünün hemen hemen aynı olduğu söylenebilir. Cox's B, Liu Agresti ve MT yöntemlerine göre madde 10 bütün koşullarda yüksek düzey DMF sergilerken; poly-SIBTEST yöntemine göre koşul 3 ve 4'te yüksek düzey DMF'ye rastlanmamıştır. Cox's B, Liu Agresti ve MT yöntemlerine göre en yüksek DMF koşul 2'de oluşurken, poly-SIBTEST yöntemine göre en yüksek DMF koşul 6'da hesaplanmıştır. Cox's B yönteminden elde edilen DMF miktarının -5.22 ile -2.89 arasında; Liu Agresti yönteminden elde edilen DMF miktarının -5.19 ile -2.85 arasında; MT yönteminden elde edilen DMF miktarının 8.37 ile 27.22 arasında ve son olarak poly-SIBTEST yönteminden elde edilen DMF miktarının ise -0.16 ile 0.06 arasında değiştiği görülmektedir.

Madde 10'un içerdiği DMF miktarı ülkelere göre değerlendirildiğinde, koşulların çoğunluğunda Türkiye-ABD karşılaştırmasında elde edilen DMF'nin miktarının Türkiye-

Kazakistan karşılaştırmasında ortaya çıkan DMF'den yüksek olduğu dikkat çekmektedir. Karşılaştırmalarda en yüksek DMF değerleri yöntemlerin tümü için örneklem büyüklüğü oranı 1:3'te hesaplanmıştır. Madde 10, Türkiye-ABD karşılaştırmasında tüm koşullarda ve Türkiye-Kazakistan karşılaştırmasında koşulların büyük çoğunluğunda Türkiye lehine yüksek düzey DMF göstermiştir. Ancak maddenin özellikle örneklem büyüklüğü oranı 2:1 iken tüm yöntemler için bazı koşullarda yüksek düzey DMF sergilemediği de açıktır. Özellikle koşul 3 ve 4'te elde edilen DMF değerlerinin kritik değerlere oldukça yakın olduğu görülmektedir.

Alt Problem 3'e İlişkin Bulgular

Odak grup örneklem büyüklüğü 200 (küçük) olduğunda, ülke karşılaştırmalarında DAF yöntemleriyle elde edilen DAF miktarları; değişen kategori birleştirme kuralı ve odak grup:referans grup örneklem oranlarında farklılık göstermekte midir?

Odak grup örneklem büyüklüğü 200 (küçük) olduğunda, değişen koşullarda DAF yöntemlerine (AC-LOR, CU-LOR) göre Türkiye-ABD ve Türkiye-Kazakistan karşılaştırmalarından elde edilen sonuçlar M4 için Şekil 8'de, M9 için Şekil 9'da ve M10 için Şekil 10'da verilmiştir.



Şekil 8. Odak grup örneklem büyüklüğü 200 iken M4'e ait DAF miktarının değişimi.

Şekil 8'de M4'ün ülkelere ve DAF belirleme yöntemlerine göre karşılaştırıldığı DAF sonuçları yer almaktadır. Türkiye-ABD sonuçları incelendiğinde;

Örnekleme büyüklüğü oranı 2:1 iken Adım 1'de koşul 2 ve koşul 7'de birbirine çok yakın olsa da genel olarak bütün koşullarda CU-LOR yönteminden elde edilen DAF miktarının AC-LOR yönteminden elde edilen DAF miktarından daha fazla olduğu görülmektedir. Bütün koşullarda bu örneklem büyüklüğü oranında CU-LOR yönteminden elde edilen DAF miktarı benzer düzeydedir. AC-LOR yönteminden elde edilen DAF miktarı bazı koşullarda artsa da genel olarak benzer düzeyde seyrettiği söylenebilir. Bu örneklem büyüklüğünde bütün koşullarda CU-LOR yönteminden elde edilen DAF C düzeyinde iken; AC-LOR yöntemlerinden elde edilen DAF sadece koşul 2 ve koşul 7'de C düzeyindedir.

Örnekleme büyüklüğü oranı 1:1 iken, Adım 1'de CU-LOR yönteminden elde edilen DAF miktarı koşul 3 ve koşul 4 hariç bütün koşullarda azalma göstermişken; AC-LOR yönteminden elde edilen DAF ise koşul 1, koşul 3 ve koşul 4'te artış; diğer koşullarda ise azalma göstermiştir. Bu örneklem büyüklüğünde koşul 5 dışında bütün koşullarda CU-LOR yönteminden elde edilen DAF C düzeyinde iken; AC-LOR yöntemlerinden elde edilen DAF koşul 1, koşul 2, koşul 3, ve koşul 7'de C düzeyindedir.

Örnekleme büyüklüğü oranı 1:2 iken, Adım 1'de AC-LOR ve CU-LOR yöntemlerinden elde edilen DAF miktarları son dört koşulda ve özellikle koşul 2'de artış göstermiş olup; koşul 1 ve koşul 4'te benzer düzeylerde seyretmiş; koşul 3'te ise azalma göstermiştir. Bu örneklem büyüklüğü oranında CU-LOR yönteminden elde edilen DAF miktarı bütün koşullarda C düzeyinde iken; AC-LOR yönteminden elde edilen DAF miktarı ise koşul 1 ve koşul 4 dışında bütün koşullarda C düzeyindedir.

Örnekleme büyüklüğü oranı 1:3 iken, Adım 1'de AC-LOR ve CU-LOR yöntemlerinden elde edilen DAF miktarı bir önceki örneklem büyüklüğü oranına göre azalma göstermiştir. Bu örneklem büyüklüğü oranında CU-LOR yöntemiyle elde edilen DAF miktarı koşul 2, koşul 5, koşul 6, koşul 7 ve koşul 8'de C düzeyinde iken; AC-LOR yönteminden elde edilen DAF miktarı koşul 2, koşul 5, koşul 6 ve koşul 7'de C düzeyindedir.

Genel olarak Adım 1’de elde edilen sonuçlar değerlendirildiğinde, CU-LOR ve AC-LOR yöntemlerinden elde edilen sonuçların paralel olduğu; özellikle koşul 2’de ve koşul 5’te bu sonuçların çok benzer olduğu, genel olarak CU-LOR yönteminden elde edilen DAF’nin daha yüksek olduğu görülmektedir.

Adım 2’de ise DAF miktarının koşullara ve örneklem büyüklüğü oranına göre değişimi her ne kadar düzenli bir örüntü sergilemese de örneklem büyüklüğü oranı 2:1 iken bütün koşullarda CU-LOR yönteminden elde edilen DAF miktarı C düzeyinde iken; AC-LOR yönteminden elde edilen DAF miktarı koşul1, koşul2, koşul 3 ve koşul 5’te C düzeyindedir. Örneklem büyüklüğü oranı 1:1 iken yine bütün koşullarda CU-LOR yönteminden elde edilen DAF miktarı C düzeyinde iken; AC-LOR yönteminden elde edilen DAF miktarı, koşul 3, koşul 5 ve koşul 6’da C düzeyindedir. Örneklem büyüklüğü 1:2 ve 1:3 iken yine bütün koşullarda CU-LOR yönteminden elde edilen DAF miktarı C düzeyinde iken; AC-LOR yönteminden elde edilen DAF miktarı, koşul 1, koşul 2, koşul 3 ve koşul 4’te C düzeyindedir.

M4 ilk üç koşulda üç kategorili olduğu için Adım 3’te bu üç koşulda herhangi bir sonuç yer almamaktadır. Diğer koşullarda ise yine DAF miktarının düzenli bir örüntü sergilemediği görülmektedir. Örneklem büyüklüğü oranı 2:1 iken AC-LOR yönteminde koşul 6 ve koşul 7’de C düzeyinde DAF bulunurken; CU-LOR yönteminden elde edilen DAF miktarı koşul 7 haricinde bütün koşullarda C düzeyindedir. Örneklem büyüklüğü oranı 1:1 iken koşul 4, koşul 7 ve koşul 8’de CU-LOR yönteminden elde edilen DAF miktarı C düzeyinde olup; AC-LOR yönteminden elde edilen DAF miktarı hiçbir koşulda C düzeyinde değildir. Örneklem büyüklüğü oranı 1:2 iken koşul 4, koşul 6, koşul 7 ve koşul 8’de CU-LOR yönteminden elde edilen DAF miktarı; AC-LOR yönteminden elde edilen DAF miktarı ise yalnızca koşul 7’de C düzeyindedir. Son olarak örneklem büyüklüğü oranı 1:3 iken AC-LOR için koşul 7 dışında bütün koşullarda her iki yöntemden elde edilen DAF miktarı C düzeyindedir. Örneklem büyüklüğü oranları genel olarak incelendiğinde Adım 3 için hesaplanan en düşük DAF değerleri örneklem büyüklüğü oranı 1:1 olduğunda elde edilmiştir.

M4 sadece koşul 8’de dört kategorili olduğu için yalnız koşul 8’de Adım 4’e ait sonuçlar yer almaktadır. Buna göre bütün örneklem büyüklüğü oranlarında CU-LOR

yönteminden elde edilen DAF miktarının C düzeyinde olduğu, en yüksek DAF miktarının ise örneklem büyüklüğü oranı 1:2'de hesaplandığı belirlenmiştir. AC-LOR yöntemi için ise örneklem büyüklüğü oranı 2:1 olduğu durum dışında bütün örneklem büyüklüğü oranlarında C düzeyinde DAF elde edilmiştir. Örneklem büyüklüğü oranı 2:1 iken ACLOR yöntemine göre DAF değeri hesaplanamamıştır.

Türkiye-Kazakistan karşılaştırıldığında;

Örneklem büyüklüğü oranı 2:1 iken Adım 1 'de AC-LOR ve CU-LOR yöntemlerinden elde edilen DAF miktarı bazı koşullarda azalmış bazılarında ise artmıştır. Bu örneklem büyüklüğü oranında koşul 2 haricinde bütün koşullarda AC-LOR yönteminden elde edilen DAF miktarı CU-LOR yönteminden elde edilen DAF miktarından daha fazladır. Ayrıca örneklem büyüklüğü oranı 2:1 iken her iki yöntem için de koşul 2 ve koşul 4'ün zıt yönde DAF sergileme eğiliminde oldukları dikkat çekmektedir. Son üç koşulda AC-LOR yönteminden elde edilen DAF miktarı benzer düzeydedir. Bu örneklem büyüklüğünde sadece koşul 2'de CU-LOR; koşul 4, 7 ve 8'de AC-LOR yöntemlerinden elde edilen DAF C düzeyindedir. Örneklem büyüklüğü oranı 1:1 iken hesaplanan DAF miktarlarının bir önceki örneklem büyüklüğü oranına göre her iki yöntem için de koşul 5 haricinde tüm koşullarda azaldığı ve benzer düzeylerde seyrettiği görülmektedir. ACLOR yöntemine göre koşul 1,4 ve 5'te; CULOR yöntemine göre ise koşul 4 ve 5'te C düzeyinde DAF elde edilmiştir.

Son iki örneklem büyüklüğü oranında DAF miktarlarında benzer örüntüler elde edilmiş olup genel olarak koşul 1, 4 ve 5'te ACLOR yöntemine göre C düzeyinde DAF hesaplanırken; örneklem büyüklüğü oranı 1:3 iken CULOR yöntemine göre C düzeyinde DAF yoktur.

Adım 2'de örneklem büyüklüğü oranı 2:1 iken yöntemlere göre elde edilen DAF miktarları koşullar bazında farklılaşırken sistematik bir örüntü elde edilmemiştir. Koşul 1 ve 6 haricinde bütün koşullarda CU-LOR yönteminden elde edilen DAF miktarı AC-LOR yönteminden elde edilen DAF miktarından daha fazladır. M4 bu örneklem büyüklüğünde CU-LOR yöntemine göre koşul 4 ve 8 haricinde tüm koşullarda; AC-LOR yöntemine göre ise tüm koşullarda negatif yönde C düzeyinde DAF sergilemektedir.

Örneklem büyüklüğü oranı 1:1 iken hesaplanan DAF değerlerinin bir önceki örneklem büyüklüğü oranına göre azalarak kritik değerin altına düştüğü görülmektedir. Buna göre CU-LOR yöntemine göre sadece koşul 3'te, AC-LOR yöntemine göre ise 1,2, 3 ve 7'de C düzeyinde DAF elde edilmiştir. Örneklem büyüklüğü oranı 1:2 iken bazı koşullarda DAF miktarları artarken bazı koşullarda azalmıştır. Özellikle koşul 3'ten elde edilen DAF miktarı tüm koşullardan elde edilen DAF miktarlarına göre oldukça yüksektir. Bu örneklem büyüklüğü oranında AC-LOR yöntemine göre koşul 1,2 ve 3; CU-LOR yöntemine göre ise koşul 2 ve 3'te hesaplanan DAF C düzeyindedir.

Örneklem büyüklüğü oranı 1:3 olduğunda her iki yöntemle göre C düzeyinde DAF veren koşul sayısı artmıştır. Yöntemlerin koşullar bazında oluşturduğu örüntü sistematik olmamakla birlikte birbirine oldukça benzerdir. Bu durumda AC-LOR ve CU-LOR yöntemlerine göre ilk beş koşulda negatif yönde C düzeyinde DAF elde edilmiştir.

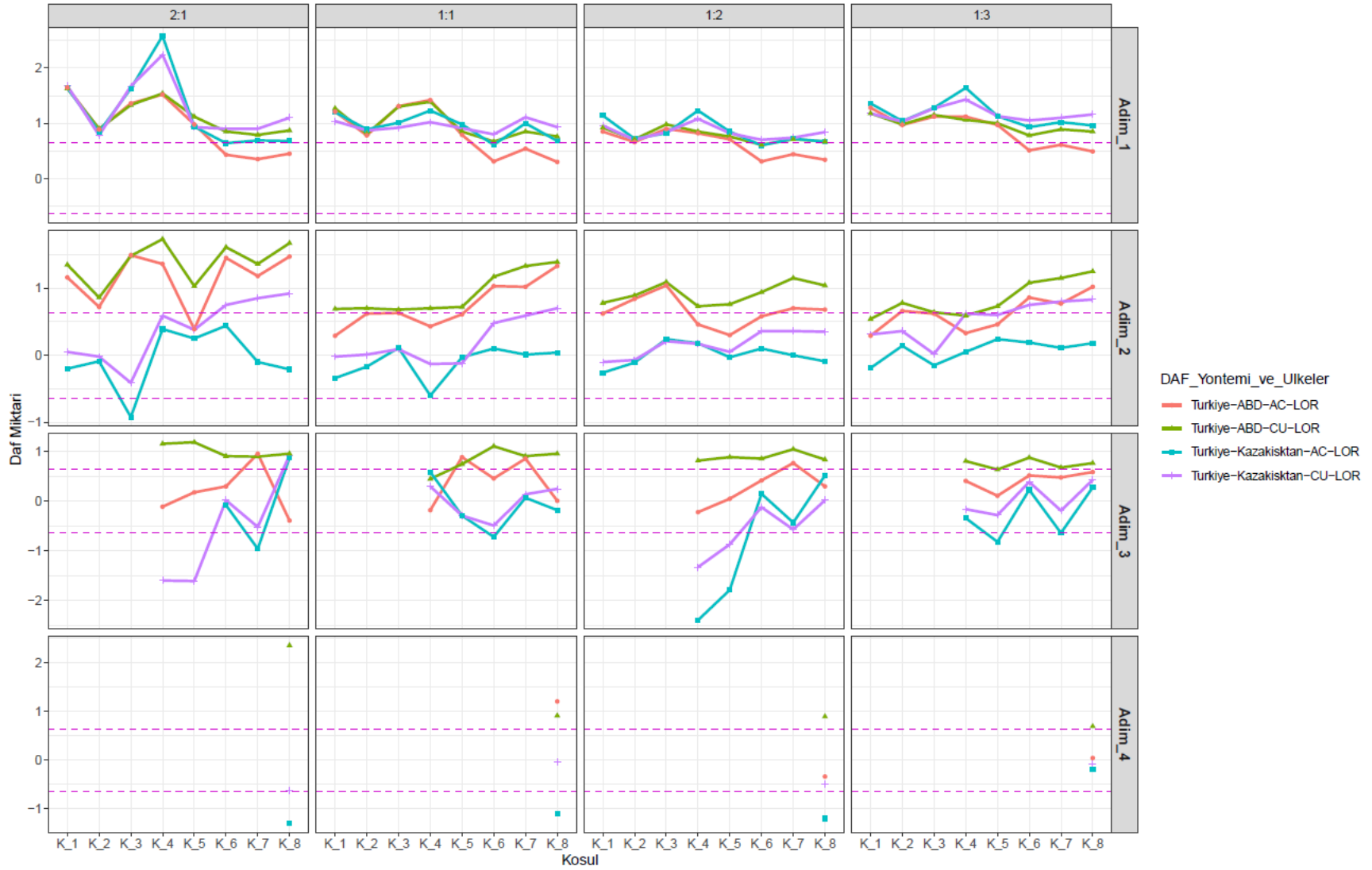
Adım 2 genel olarak yorumlandığında koşulların büyük bir kısmında Türkiye lehine yüksek düzey DAF sergilendiği söylenebilir. Yöntemler bazında en düşük DAF miktarları örneklem büyüklüğü oranı 1:1 olduğunda elde edilmiştir.

M4 ilk üç koşulda üç kategorili olduğu için Adım 3'te bu üç koşulda herhangi bir sonuç yer almamaktadır. Ayrıca örneklem büyüklüğü oranının 2:1 olduğu durumda koşul 4'te AC-LOR yöntemi ve koşul 5'te her iki yöntem de kestirim yapamadığından bu noktalar için herhangi bir değer bulunmamaktadır. Diğer koşullarda ise yine DAF miktarının düzenli bir örüntü sergilemediği görülmektedir. Örneklem büyüklüğü oranı 2:1 iken AC-LOR yönteminden C düzeyinde DAF bulunmazken; CU-LOR yönteminden elde edilen DAF miktarı koşul 4, 7 ve 8'de C düzeyindedir. Örneklem büyüklüğü oranı 1:1 iken koşul 4, koşul 5 ve koşul 7'de CU-LOR yönteminden elde edilen DAF miktarı; koşul 4, koşul 5, koşul 6 ve koşul 7'de ise AC-LOR yönteminden elde edilen DAF miktarı C düzeyindedir. Örneklem büyüklüğü oranı 1:2 iken koşul 4, koşul 5, koşul 7 ve koşul 8'de CU-LOR yönteminden elde edilen DAF miktarı; AC-LOR yönteminden elde edilen DAF miktarı koşul 4, koşul 5, koşul 6 ve koşul 7'de C düzeyindedir. Son olarak örneklem büyüklüğü oranı 1:3 iken bütün koşullarda her iki yöntemden elde edilen DAF miktarı C düzeyindedir.

M4 sadece koşul 8'de dört kategorili olduğu için yalnız koşul 8'de Adım 4'e ait sonuçlar yer almaktadır. Örneklem büyüklüğü oranının 2:1 olduğu durumda her iki yöntem

de kestirim yapamadığından bu noktalar için herhangi bir değer bulunmamaktadır. Diğer tüm örneklem büyüklüğü oranlarında her iki yöntemden elde edilen DAF miktarının C düzeyinde olduğu belirlenmiştir. Örneklem büyüklüğü oranları incelendiğinde en yüksek DAF değerlerinin örneklem büyüklüğü oranı 1:3 iken hesaplandığı görülmektedir.

M4 ülkelere göre incelendiğinde, Türkiye-Kazakistan karşılaştırmasında Adım 1’de Kazakistan lehine C düzeyinde DAF olduğu görülürken diğer adımlarda Türkiye lehine C düzeyinde DAF görülmektedir. Özellikle örneklem büyüklüğü oranı 2:1 iken koşullara göre maddenin hangi ülke lehine DAF gösterdiğinin değişmesi oldukça dikkat çekmektedir. Öyle ki koşul 2’de Türkiye lehine görülen DAF, koşul 4’te Kazakistan lehinedir. Türkiye-ABD karşılaştırmasında ise elde edilen DAF miktarları Türkiye-Kazakistan karşılaştırmasından elde edilen DAF miktarlarından daha yüksektir. Tüm adımlarda hesaplanan yüksek düzey DAF’lerin ABD lehine olduğu görülmektedir.



Şekil 9. Odak grup örneklem büyüklüğü 200 iken Madde 9'a ait DAF miktarının değişimi.

Şekil 9'da odak grup örneklem büyüklüğü 200 iken Madde 9'un ülkelere ve DAF belirleme yöntemlerine ilişkin DAF sonuçları yer almaktadır. Türkiye-ABD için DAF sonuçları incelendiğinde;

M4'ten farklı olarak madde 9'da Adım 1'de CU-LOR ve AC-LOR yöntemlerinden her iki ülke için elde edilen DAF miktarının ilk beş koşulda birbirine oldukça yakın olduğu son üç koşulda ise birbirinden farklılaştığı dikkat çekmektedir. Son üç koşulda AC-LOR yönteminden elde edilen DAF miktarı, CU-LOR yönteminden elde edilen DAF miktarından daha azdır. Genel olarak örneklem büyüklüğü oranı 2:1'den 1:1'e değişimde AC-LOR e CU-LOR yöntemlerinden elde edilen DAF miktarları genel olarak azalmıştır. Benzer şekilde örneklem büyüklüğü oranı 1:2 olduğunda yine genel olarak her iki yöntemden elde edilen DAF genel olarak azalmıştır. Son olarak örneklem büyüklüğü oranı 1:3 olduğunda her iki yöntemden elde edilen DAF miktarı artmıştır. Örneklem büyüklüğü oranı 2:1 iken bütün koşullarda CU-LOR yönteminden elde edilen DAF C düzeyinde iken; AC-LOR yöntemlerinden elde edilen DAF son üç koşul dışındaki bütün koşullarda C düzeyindedir.

Örneklem büyüklüğü oranı 2:1 iken, Adım 2'de hem CU-LOR hem de AC-LOR yöntemlerinden elde edilen DAF miktarı koşul 2 ve 5'te biraz düşük olsa da genel olarak 1.00'in üzerindedir. Koşul 3 dışında bütün koşullarda CU-LOR yönteminden elde edilen DAF miktarı AC-LOR yönteminden elde edilen DAF miktarından daha yüksektir. Bu örneklem büyüklüğünde bütün koşullarda CU-LOR yönteminden elde edilen DAF C düzeyinde iken; AC-LOR yöntemlerinden elde edilen DAF miktarı ise koşul 5 dışında C düzeyindedir.

Örneklem büyüklüğü oranı 1:1 olduğunda bir önceki örneklem büyüklüğü oranına kıyasla hem CU-LOR hem de AC-LOR yönteminden elde edilen DAF miktarı genel olarak azalmıştır. Bu örneklem büyüklüğünde bütün koşullarda CU-LOR yönteminden elde edilen DAF miktarı C düzeyinde iken, AC-LOR yöntemlerinden elde edilen DAF miktarı son üç koşulda C düzeyindedir.

Örneklem büyüklüğü oranı 1:2 olduğunda ise DAF miktarı genel olarak tekrar artış göstermiştir. Bu örneklem büyüklüğünde bütün koşullarda CU-LOR yönteminden elde edilen DAF miktarı C düzeyinde iken, AC-LOR yöntemlerinden elde edilen DAF miktarı koşul 2, koşul 3, koşul 7 ve koşul 8'de C düzeyindedir.

Son olarak örneklem büyüklüğü oranı 1:3 olduğunda, son dört koşulda artış gösterirken ilk dört koşulda azalma göstermiştir. Bu örneklem büyüklüğünde koşul 1 hariç bütün koşullarda CU-LOR yönteminden elde edilen DAF miktarı C düzeyinde iken, AC-LOR yöntemlerinden elde edilen DAF miktarı koşul 2, koşul 6, koşul 7 ve koşul 8'de C düzeyindedir.

Genel olarak Adım 2'de elde edilen sonuçlar değerlendirildiğinde, CU-LOR ve AC-LOR yöntemlerinden elde edilen sonuçların paralel olduğu; özellikle koşul 2'de ve koşul 3'te bu sonuçların çok benzer olduğu, genel olarak CU-LOR yönteminden elde edilen DAF'nin daha yüksek olduğu görülmektedir.

Adım 3'te DAF miktarının düzenli bir örüntü sergilemese de CU-LOR yönteminden elde edilen DAF'nin AC-LOR yönteminden elde edilen DAF miktarından daha yüksek olduğu görülmektedir. Örneklem büyüklüğü oranı 2:1 iken AC-LOR yönteminde yalnızca koşul 7'de C düzeyinde DAF bulunurken; CU-LOR yönteminden elde edilen DAF miktarı bütün koşullarda C düzeyindedir. Örneklem büyüklüğü oranı 1:1 iken koşul 4 dışında bütün koşullarda CU-LOR yönteminden elde edilen DAF miktarı C düzeyinde iken; AC-LOR yönteminden elde edilen DAF miktarı yalnızca koşul 5 ve koşul 7'de C düzeyindedir. Örneklem büyüklüğü oranı 1:2 iken tüm koşullarda CU-LOR yönteminden elde edilen DAF miktarı; AC-LOR yönteminden elde edilen DAF miktarı ise yalnızca koşul 7'de C düzeyindedir. Son olarak örneklem büyüklüğü oranı 1:3 iken AC-LOR yönteminden elde edilen DAF miktarı hiçbir koşulda C düzeyinde değilken; CU-LOR yönteminden elde edilen DAF bütün koşullarda C düzeyindedir. Örneklem büyüklüğü oranlarında sağa doğru gidildikçe hesaplanan DAF miktarlarında azalma gözlenmektedir.

Madde 9 sadece koşul 8'de dört kategorili olduğu için yalnız koşul 8'de Adım 4'e ait sonuçlar yer almaktadır. Örneklem büyüklüğü oranı 2:1 iken AC-LOR yöntemi kestirim yapamadığından bu nokta için herhangi bir değer bulunmamaktadır. Buna göre 1:1 örneklem büyüklüğü dışındaki bütün örneklem büyüklüğü oranlarında CU-LOR yönteminden elde edilen DAF miktarı AC-LOR yönteminden elde edilen DAF miktarından yüksektir. Bütün örneklem büyüklüğü oranlarında CU-LOR yönteminden elde edilen DAF miktarı C düzeyinde iken AC-LOR yönteminden elde edilen DAF miktarı yalnızca 1:1 örneklem büyüklüğü oranında C düzeyindedir.

Türkiye-Kazakistan karşılaştırıldığında;

Örneklem büyüklüğü oranı 2:1 iken Adım 1'de AC-LOR ve CU-LOR yöntemlerinden elde edilen DAF miktarları bazı koşullarda azalmış bazılarında ise artmıştır. Her iki yöneme göre hesaplanan DAF değerleri C düzeyinde olup koşul 4'te en yüksek değerini almıştır. Son üç koşulda elde edilen DAF miktarları benzer düzeyde olup kritik değere oldukça yakındır. Diğer örneklem büyüklüğü oranları incelendiğinde DAF miktarlarının azalma göstermekle birlikte C düzeyinde DAF sergiledikleri görülmektedir. Yöntemlere ilişkin elde edilen örüntüler oldukça benzer olup sistematik değildir.

Örneklem büyüklüğü oranı 2:1 iken, Adım 2'de hem CU-LOR hem de AC-LOR yöntemlerinden elde edilen DAF miktarlarının büyük bir kısmı kritik değerlerin altında olup 0'a oldukça yakındır. CU-LOR yöntemine göre hesaplanan DAF miktarları son üç koşulda; AC-LOR yöntemine göre hesaplanan DAF miktarları ise sadece koşul 2'de C düzeyindedir. Koşul 2'de AC-LOR yöntemine göre negatif yönde; koşul 6, 7 ve 8'de CU-LOR yöntemine göre pozitif yönde DAF elde edilmiştir.

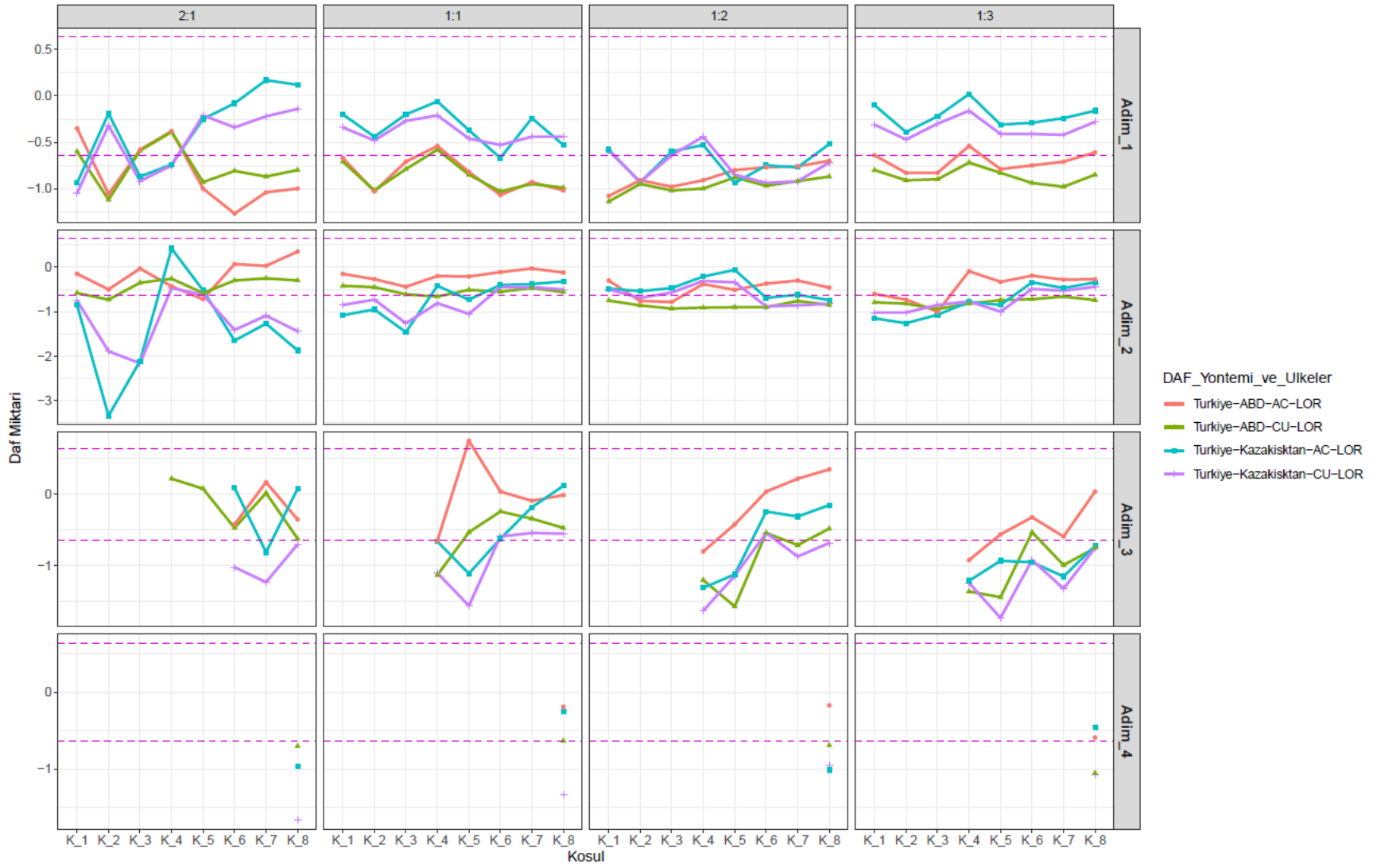
Örneklem büyüklüğü oranı 1:1 olduğu durumda bir önceki örneklem büyüklüğü oranına kıyasla her iki yöneme göre hesaplanan DAF miktarlarında azalma görülmüştür. Bu durumda AC-LOR yöntemine göre C düzeyinde DAF bulunmamaktadır. CU-LOR yöntemine göre ise sadece koşul 8'de C düzeyinde DAF elde edilmiştir. Örneklem büyüklüğü oranı 1:2 iken değerlerin daha da birbirine ve 0'a yaklaştığı görülmektedir. Bu örneklem büyüklüğü oranında yöntemlerin ve koşulların hiçbirinde C düzeyinde DAF'ye rastlanmamıştır. Örneklem büyüklüğü oranı 1:3 olduğunda ise hesaplanan DAF miktarlarında tekrar bir artış görünmektedir. AC-LOR yöntemine göre hiçbir koşulda C düzeyinde DAF'ye rastlanmazken CU-LOR yöntemine göre 4, 5, 6, 7 ve 8'de C düzeyinde DAF bulunmaktadır.

Sekil 9'da ilk üç koşulda madde 9'un üç kategorili olması dolayısıyla sadece iki adimin olması nedeniyle herhangi bir değer yer almamaktadır. Ayrıca koşul 4 ve 5'te örneklem büyüklüğü oranının 2:1 olduğu durumda AC-LOR yöntemi kestirim yapamadığından bu noktalar için herhangi bir değer bulunmamaktadır. Örneklem büyüklüğü oranı 2:1 iken koşul 4 ve 5'te CU-LOR yöntemiyle elde edilen DAF miktarının oldukça yüksek diğer üç koşulda ise daha düşük olduğu görülmektedir. Örneklem

büyüklüğü 2:1 iken CU-LOR yönteminden elde edilen DAF koşul 4,5 ve 8'de C düzeyinde iken; AC-LOR yönteminden elde edilen DAF koşul 7 ve 8'de C düzeyindedir. C düzeyindeki bu DAF miktarı koşul 4, 5 ve 7'de Türkiye lehine iken; koşul 8'de Kazakistan lehinedir. Örneklem büyüklüğü oranı 1:1 olduğunda genel olarak bir önceki örneklem büyüklüğü oranına kıyasla DAF miktarı azalma göstermiştir. Bu örneklem büyüklüğü oranında yalnızca koşul 6'da CU-LOR yönteminden elde edilen DAF C düzeyinde olup, Türkiye lehinedir. Örneklem 1:2 olduğunda özellikle koşul 4 ve koşul 5'te bir önceki örneklem büyüklüğü oranına kıyasla özellikle AC-LOR yönteminden elde edilen DAF miktarında ani bir artış meydana gelmiştir. Son üç koşulda bu ivme bu kadar keskin olmayıp, düzenli bir örüntü elde edilmemiştir. Koşul 4 ve koşul 5'te her iki yöntemden elde edilen DAF miktarı C düzeyinde olup Türkiye lehinedir. Son olarak örneklem büyüklüğü oranı 1:3 iken yine koşul 4 ve 5'te ani bir ivme ile DAF miktarında azalma meydana gelmişken son üç koşulda bu değişim bu kadar keskin değildir. Sadece AC-LOR yönteminden elde edilen DAF koşul 5 ve 7'de C düzeyinde olup Türkiye lehinedir.

Adım 4'te örneklem büyüklüğü oranının 1:3 olduğu durumlar dışındaki bütün örneklem büyüklüğü oranında özellikle AC-LOR yönteminden elde edilen DAF benzer düzeyde seyretmiştir. Son örneklem büyüklüğü oranı dışında AC-LOR yönteminden elde edilen DAF miktarı CU-LOR yöntemiyle elde edilen DAF miktarından daha yüksektir. CU-LOR yönteminden elde edilen DAF örneklem büyüklüğü oranının değişmesine bağlı olarak düzenli bir örüntü sergilememiştir. Örneklem büyüklüğü oranının 1:3 olduğu durum haricindeki bütün örneklemelerde AC-LOR yönteminden elde edilen DAF C düzeyinde olup Türkiye lehinedir.

Ülkelere göre karşılaştırıldığında, Türkiye-ABD karşılaştırmasında madde 9'a ait bütün adımlar ABD lehine C düzeyinde DAF göstermiştir. Türkiye-Kazakistan karşılaştırmasında ise Adım 1'de Kazakistan lehine; Adım 3 ve 4'te Türkiye lehine C düzeyinde DAF olduğu görülmektedir. Adım 2'de ise C düzeyinde DAF yok denecek kadar azdır. Türkiye-ABD karşılaştırmasında elde edilen DAF miktarlarının Türkiye-Kazakistan karşılaştırmasında elde edilen DAF miktarlarından daha yüksek olduğu belirlenmiştir.



Şekil 10. Odak grup örneklem büyüklüğü 200 iken Madde 10'a ait DAF miktarının değişimi.

Şekil 10'da odak grup örneklem büyüklüğü 200 iken Madde 10'un ülkelere ve DAF belirleme yöntemlerine ilişkin DAF sonuçları yer almaktadır. Türkiye-ABD karşılaştırmasında AC-LOR ve CU-LOR yöntemleri kullanılarak elde edilen DAF sonuçları incelendiğinde;

Örneklem büyüklüğü oranı 2:1 iken ilk 6 koşulda CU-LOR ve AC-LOR yöntemlerinden elde edilen DAF sonuçlarının çok benzer; son üç koşulda ise AC-LOR yönteminden elde edilen DAF sonuçlarının daha yüksek olduğu görülmektedir. Koşul 1, Kosul3,Kosul4 dışındaki bütün koşullarda her iki yöntemden elde edilen DAF C düzeyindedir.

Örneklem büyüklüğü oranı 1:1 olduğunda CU-LOR ve AC-LOR yöntemlerinden elde edilen DAF miktarları oldukça yakındır. Koşul 4 dışındaki bütün koşullarda her iki yöntemden elde edilen DAF miktarları C düzeyindedir. Her iki yöntemden elde edilen DAF düzeylerinin koşulların değişimine bağlı olarak düzenli bir örüntü çizmediği görülmektedir.

Örneklem büyüklüğü 1:3 olduğunda ise CU-LOR yönteminden elde edilen DAF miktarının AC-LOR yönteminden elde edilen DAF miktarından daha fazla olduğu görülmektedir. CU-LOR yönteminden elde edilen DAF bütün koşullarda C düzeyinde iken AC-LOR yönteminden elde edilen DAF koşul 1, 4 ve 8 dışındaki bütün koşullarda C düzeyindedir.

Genel olarak Adım 1'de elde edilen sonuçlar değerlendirildiğinde, CU-LOR ve AC-LOR yöntemlerinden elde edilen sonuçların paralel olduğu; bütün koşullarda DAF'nin Türkiye lehine olduğu görülmektedir.

Adım 2'de AC-LOR ve CU-LOR yöntemlerinden elde edilen DAF miktarı son 3 koşulda biraz farklılaşmakla birlikte genel olarak paralel seyrettiği görülmektedir. Örneklem büyüklüğü oranı 2:1 iken koşul 2 haricinde bütün koşullarda elde edilen DAF;nin -0.64, +0.64 aralığında kaldığı dolayısıyla C düzeyinde DAF göstermediği anlaşılmaktadır. Bu örneklem büyüklüğü oranında sadece koşul 2'de CU-LOR yöntemine göre C düzeyinde DAF elde edilmiştir.

Örneklem büyüklüğü oranı 1:1 olduğunda CU-LOR ve AC-LOR yöntemlerinden elde edilen DAF miktarının birbirine paralel olduğu ve bütün sonuçların -0.64,+0.64 aralığında kaldığı görülmektedir.

Örneklem büyüklüğü oranı 1:2 olduğunda CU-LOR yönteminden elde edilen DAF miktarları AC-LOR yönteminden elde edilen DAF miktarlarından daha fazladır. CU-LOR yönteminden elde edilen DAF bütün koşullarda C düzeyinde iken, AC-LOR yönteminden elde edilen DAF genel olarak C düzeyinde değildir.

Örneklem büyüklüğü oranı 1:3 olduğunda CU-LOR ve AC-LOR yöntemlerinden elde edilen sonuçların diğer örneklem büyüklüğü oranına kıyasla daha çok farklılaştığı görülmektedir. CU-LOR yönteminden elde edilen DAF miktarı bütün koşullarda C düzeyinde iken AC-LOR yönteminden elde edilen DAF miktarı ise genel olarak C düzeyinde değildir.

İlk üç koşulda madde 10, 3 kategorili olduğu için bu koşullarda Adım 3'e ilişkin herhangi bir sonuç bulunmamaktadır. Örneklem büyüklüğü oranı 2:1 olduğunda CU-LOR ve AC-LOR yöntemlerinden elde edilen DAF miktarının kritik değerlerin altında kaldığı görülmektedir. Bu örneklem büyüklüğünde koşul 4 ve 5'te AC-LOR yönteminin kestirim yapamadığı görülmektedir. Örneklem büyüklüğü oranı 1:1 iken CU-LOR ve AC-LOR yöntemlerinden elde edilen DAF miktarının farklılaştığı, CU-LOR yönteminden elde edilen DAF'nin koşul 4'te; AC-LOR yönteminden elde edilen DAF'nin ise koşul 5'te C düzeyinde olduğu belirlenmiştir. Örneklem büyüklüğü oranı 1:2 olduğunda CU-LOR yönteminden elde edilen DAF miktarının çoğunlukla belirlenen ölçüt aralığının dışında kaldığı, dolayısıyla C düzeyinde DAF gösterdiği; AC-LOR yönteminden elde edilen DAF sonuçlarının ise çoğunlukla C düzeyinde DAF göstermediği söylenebilir. Örneklem büyüklüğü oranı 1:3 olduğunda ise CU-LOR yönteminden elde edilen DAF sonuçlarının çoğunlukla C düzeyinde olduğu; AC-LOR yönteminden elde edilen sonuçların ise bir koşul dışında C düzeyinde olmadığı görülmektedir. Yalnızca koşul 4'te CU-LOR yönteminden elde edilen DAF miktarının C düzeyinde olup, son üç koşulda her iki yöntemden elde edilen sonuçların paralel olduğu söylenebilir.

Madde 10 sadece koşul 8'de dört kategorili olduğu için yalnız koşul 8'de Adım 4'e ait sonuçlar yer almaktadır. Örneklem büyüklüğü oranı 2:1 iken AC-LOR yöntemi kestirim

yapamadığından bu nokta için herhangi bir değer bulunmamaktadır. Buna göre bütün örneklem büyüklüğü oranlarında CU-LOR yönteminden elde edilen DAF miktarları AC-LOR yönteminden elde edilen DAF miktarından yüksektir. Bütün örneklem büyüklüğü oranlarında CU-LOR yönteminden elde edilen DAF miktarı C düzeyinde iken AC-LOR yöntemine göre C düzeyinde DAF görülmemektedir.

Türkiye-Kazakistan karşılaştırmasında AC-LOR ve CU-LOR yöntemleri kullanılarak elde edilen DAF sonuçları incelendiğinde;

Örneklem büyüklüğü oranı 2:1 iken Adım 1’de her iki yönetime göre ilk beş koşulda benzer olan DAF miktarları son üç koşulda farklılaşmaktadır. Genel olarak CU-LOR yönteminden elde edilen DAF miktarlarının AC-LOR yönteminden elde edilen DAF miktarlarından minimal düzeyde de olsa daha yüksek olduğu görülmektedir. Her iki yönetime göre de koşul 1, 3 ve 4’te negatif yönde C düzeyinde DAF hesaplanırken; diğer koşullarda elde edilen DAF miktarları kritik değerleri aşmamıştır.

Örneklem büyüklüğü oranı 1:1 ve 1:3 olduğu durumlarda AC-LOR ve CU-LOR yöntemlerine göre elde edilen DAF miktarlarında bazı koşullarda düşüş, bazı koşullarda ise artış görülmekle birlikte hemen hemen bütün koşullarda hesaplanan DAF değerlerinin kritik değerinin altında olduğu görülmektedir. Bu durumda örneklem büyüklüğü 1:1 iken AC-LOR yöntemine göre koşul 2’de hesaplanan DAF dışında ilgili örneklem büyüklüğü oranlarında C düzeyinde DAF yoktur.

Örneklem büyüklüğü oranı 1:2 iken Adım 1 için hesaplanan en büyük DAF değerleri elde edilmiştir. Öyle ki AC-LOR yöntemine göre koşul 2, 5, 6 ve 7’de; CU-LOR yöntemine göre ise koşul 2, 5, 6, 7 ve 8’de negatif yönde C düzeyinde DAF görülmektedir. Adım 1 genel olarak incelendiğinde her iki yöntem için de koşullar bazında benzer örüntüler elde edildiği, bu örüntülerin sistematik olmadığı ve genel olarak CU-LOR yöntemine göre daha yüksek DAF değerleri elde edildiği söylenebilir.

Örneklem büyüklüğü oranı 2:1 iken Adım 2’de her iki yönetime göre elde edilen DAF miktarları koşul 2 ve 4 haricinde tüm koşullarda oldukça birbirine benzerdir. İlgili adımda AC-LOR ve CU-LOR yöntemlerine göre koşul 4 ve 5 haricinde bütün koşullarda

negatif yönde C düzeyinde DAF görülmektedir. AC-LOR yöntemine göre en yüksek DAF değeri koşul 2'de hesaplanırken; CU-LOR yöntemine göre en yüksek DAF koşul 3'tedir.

Örneklem büyüklüğü oranı 1:1 iken koşullar bazında hesaplanan DAF miktarlarının azaldığı ve birbirine çok yaklaştığı görülmektedir. örneklem büyüklüğü oranı 1:2'de du değerler daha da azalarak kritik değerlerin altına inmiştir. Bu örneklem büyüklüğü oranında sadece son üç koşulda C düzeyinde DAF elde edilmiştir. Örneklem büyüklüğü oranı 1:3 olduğunda ise ilk koşullardaki DAF miktarları tekrar artmış, son üç koşuldaki DAF miktarları ise azalmıştır.

Sekil 10'da ilk üç koşulda madde 10'un üç kategorili olması dolayısıyla sadece iki adımın olması nedeniyle herhangi bir değer yer almamaktadır. Ayrıca koşul 4 ve 5'te örneklem büyüklüğü oranının 2:1 olduğu durumda her iki yöntem de kestirim yapamadığından bu noktalar için herhangi bir değer bulunmamaktadır. Örneklem büyüklüğü oranı 2:1 iken AC-LOR yöntemine göre koşul 7'de; CU-LOR yöntemine göre ise tüm koşullarda C düzeyinde DAF elde edilmiştir. Diğer örneklem büyüklüklerinde koşullar bazında hesaplanan DAF miktarlarında artış olduğu görülmektedir. buna göre örneklem büyüklüğü 1:3 olduğunda her iki yönteme göre tüm koşullarda C düzeyinde DAF vardır.

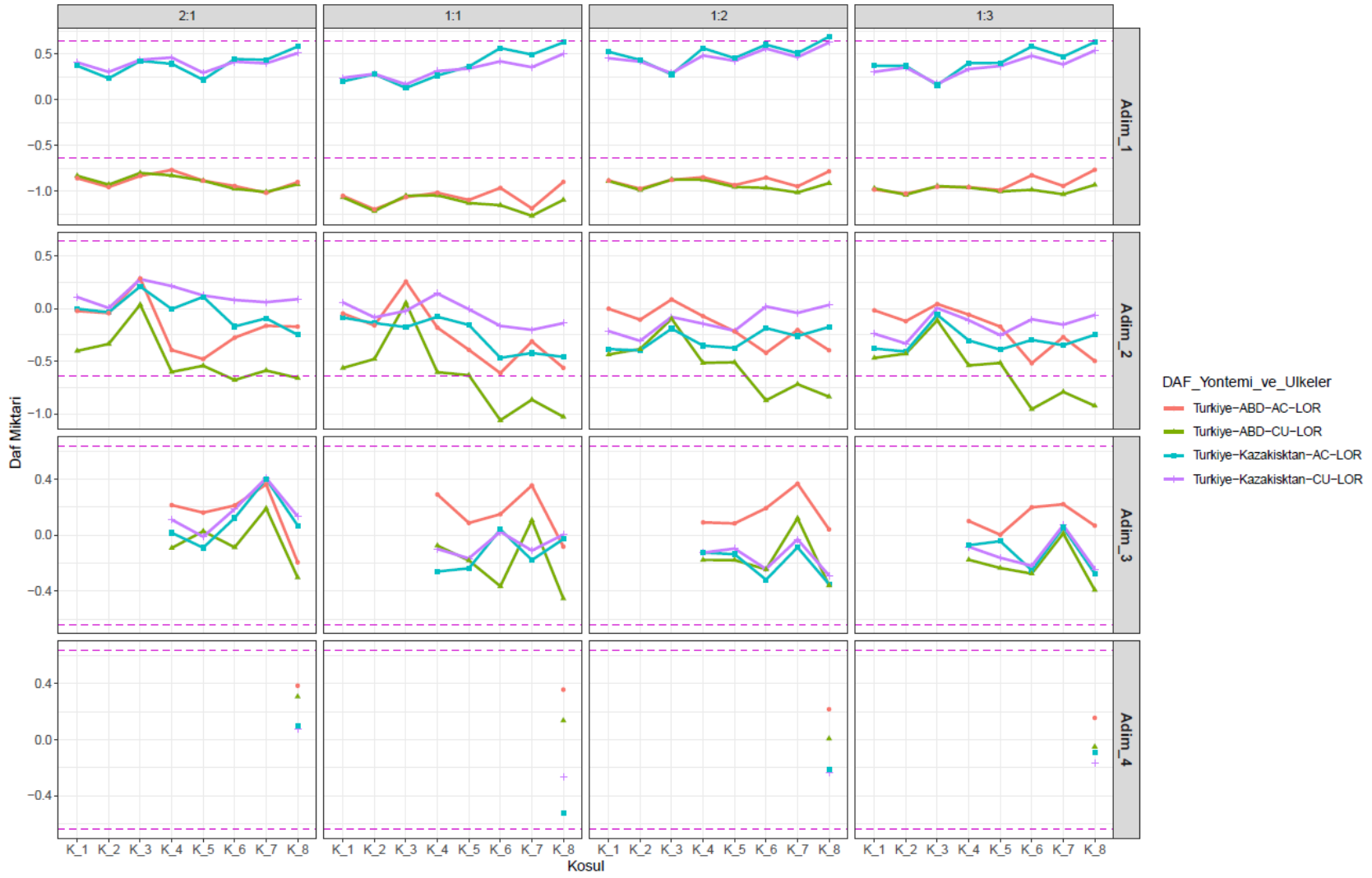
Adım 4'te CU-LOR yöntemine göre tüm örneklem büyüklüğü oranlarında hesaplanan DAF miktarı negatif yönde C düzeyindedir. AC-LOR yöntemine göre ise hesaplanan DAF miktarları değişiklik göstermektedir.

Genel olarak bakılırsa, Türkiye-ABD karşılaştırmasında madde 10'a ait adımlar incelendiğinde özellikle Adım 1'in Türkiye lehine C düzeyinde DAF içerdiği görülmektedir. Diğer adımlarda hesaplanan DAF değerleri çoğunlukla yöntemler için belirlenen kritik değerlerin altında seyretmektedir. Türkiye-Kazakistan karşılaştırmasında ise adımların çoğunlukla Türkiye lehine DAF gösterme eğiliminde olduğu söylenebilir. Türkiye-ABD karşılaştırmasında elde edilen DAF miktarlarının Adım 1 için Türkiye-Kazakistan karşılaştırmasında elde edilen DAF miktarlarından daha yüksek olduğu belirlenmiştir. Diğer adımlarda ise ülkeler bazında hesaplanan DAF miktarları benzer düzeyde ya da Türkiye-Kazakistan karşılaştırmasında elde edilen DAF miktarları daha fazladır.

Alt Problem 4'e İlişkin Bulgular

Odak grup örneklem büyüklüğü 1000 (büyük) olduğunda, ülke karşılaştırmalarında DAF yöntemleriyle elde edilen DAF miktarları; değişen kategori birleştirme kuralı ve odak grup:referans grup örneklem oranlarında farklılık göstermekte midir?

Odak grup örneklem büyüklüğü 1000 (büyük) olduğunda, değişen koşullarda DAF yöntemlerine (AC-LOR, CU-LOR) göre Türkiye-ABD ve Türkiye-Kazakistan karşılaştırmalarından elde edilen sonuçlar M4 için Şekil 11'de, M9 için Şekil 12'de ve M10 için Şekil 13'te verilmiştir.



Şekil 11. Odak grup örneklem büyüklüğü 1000 iken M4'e ait DAF miktarının değişimi.

Şekil 11’de odak grup (Türkiye) örneklem büyüklüğü 1000 iken, M4’ün ülkelere ve DAF belirleme yöntemlerine göre karşılaştırıldığı DAF sonuçları yer almaktadır. Türkiye-ABD için DAF sonuçları incelendiğinde;

Adım 1’de örneklem büyüklüğü oranı 2:1 iken CU-LOR ve AC-LOR yöntemlerinden elde edilen sonuçların çok benzer olduğu görülmektedir. Bütün koşullarda her iki yöntemden elde edilen DAF sonuçlarının C düzeyinde ve Türkiye lehine olduğu belirlenmiştir. Örneklem büyüklüğü oranı 1:1 iken, CU-LOR yönteminden elde edilen DAF miktarı koşul 6 ve koşul 8’de biraz farklılaşsa da genel olarak her iki yöntemden elde edilen sonuçların yine birbirine oldukça yakın olduğu söylenebilir. Bu örneklem büyüklüğü oranında her iki yöntemden elde edilen DAF sonuçları C düzeyinde ve Türkiye lehine olup, miktar olarak bir önceki örneklem büyüklüğü oranında elde edilen sonuçlardan biraz daha yüksektir. Örneklem büyüklüğü oranı 1:2 ve 1:3 iken, Adım 1’de CU-LOR yönteminden elde edilen DAF sonuçlarının yine oldukça örtüştüğü ve her iki yöntemden elde edilen sonuçların C düzeyinde ve Türkiye lehine olduğu belirlenmiştir.

Adım 2’de ise örneklem büyüklükleri oranı 2:1 olduğunda, CU-LOR ve AC-LOR yöntemlerinden elde edilen DAF sonuçlarının son koşullarda farklılaştığı ancak ilk koşullarda paralel seyrettiği görülmektedir. Koşullar bağlamında DAF sonuçlarının değişimi açısından herhangi bir örüntü elde edilmemiştir. Bütün koşullarda CU-LOR yönteminden elde edilen sonuçlar daha yüksek olsa da genel olarak her iki yöntemden elde edilen sonuçların C düzeyinde olmadığı belirlenmiştir. Örneklem büyüklükleri oranı 1:1 iken elde edilen DAF sonuçları, örneklem büyüklükleri oranı 2:1’de elde edilen sonuçlara bir miktar benzemektedir. Yalnız son koşullarda CU-LOR yönteminden elde edilen sonuçların C düzeyinde ve Türkiye lehine olduğu görülmektedir. Benzer durum örneklem büyüklükleri oranının 1:2 ve 1:3 olduğu durumlarda da elde edilmiştir.

Genel olarak Adım 2 sonuçları değerlendirildiğinde özellikle örneklem büyüklükleri oranı 1:1, 1:2 ve 1:3’te elde edilen sonuçların benzer olduğu söylenebilir.

Adım 3’te genel olarak bütün örneklem büyüklükleri oranında bazı koşullarda CU-LOR bazılarında AC-LOR yönteminden elde edilen DAF sonuçlar daha yüksektir. Hiçbir koşulda C düzeyinde DAF elde edilmemiştir. Son olarak Adım 4’te koşul 8’de AC-LOR

yönteminden elde edilen DAF miktarı CU-LOR yönteminden elde edilen DAF miktarından biraz fazla olsa da her iki yöntemden elde edilen sonuçlar C düzeyinde değildir.

Türkiye-Kazakistan karşılaştırmasında AC-LOR ve CU-LOR yöntemleri kullanılarak elde edilen DAF sonuçları incelendiğinde;

Adım 1'de tüm örneklem büyüklüğü oranlarında hesaplanan DAF miktarlarının koşullar bazında benzer örüntü oluşturdukları görülmektedir. Genel olarak incelendiğinde koşul 8 haricinde C düzeyinde DAF elde edilmemiştir.

Adım 2'de ise örneklem büyüklükleri oranı 2:1 ve 1:1 olduğunda, CU-LOR ve AC-LOR yöntemlerinden elde edilen DAF sonuçlarının benzer bir örüntü oluşturduğu; ancak bu örüntünün sistematik olmadığı söylenebilir. Örneklem büyüklükleri oranı 1:2 ve 1:3 olduğunda ise hesaplanan DAF miktarlarının birbirine daha yakın olduğu görülmektedir. Koşullar bağlamında DAF sonuçlarının değişimi açısından herhangi bir örüntü elde edilmemiştir. Bütün koşullarda AC-LOR yönteminden elde edilen sonuçlar daha yüksek olsa da genel olarak her iki yöntemden elde edilen sonuçların C düzeyinde olmadığı belirlenmiştir.

Adım 3 ve Adım 4'te genel olarak bütün örneklem büyüklükleri oranlarında CU-LOR ve AC-LOR yöntemlerinden elde edilen DAF sonuçları benzerlik göstermektedir. Hiçbir koşulda C düzeyinde DAF elde edilmemiştir.

M4'e ait adımların içerdiği DAF miktarları ülkelere göre değerlendirildiğinde, Türkiye-ABD karşılaştırmasında her iki yöntem için koşulların tamamında Adım1'in; AC-LOR yöntemine göre ise bazı koşullarda Türkiye lehine C düzeyinde DAF içerdiği görülmektedir. Elde edilen DAF miktarlarının Türkiye-Kazakistan karşılaştırmasında ortaya çıkan DAF'den yüksek olduğu dikkat çekmektedir. Öyle ki Türkiye-Kazakistan karşılaştırmasında Adım 1'deki Koşul 8 haricinde C düzeyinde DAF yoktur. Adım 1'de hesaplanan DAF ise Kazakistan lehinedir. Adım 3 ve 4'te ülke karşılaştırmalarının hiç birinde C düzeyinde DAF elde edilmemiştir.



Şekil 12. Odak grup örneklem büyüklüğü 1000 iken Madde 9'a ait DAF miktarının değişimi.

Şekil 12’de odak grup (Türkiye) örneklem büyüklüğü 1000 iken, madde 9’un ülkelere ve DAF belirleme yöntemlerine göre karşılaştırıldığı DAF sonuçları yer almaktadır. Türkiye-ABD için DAF sonuçları incelendiğinde;

Adım 1’de örneklem büyüklükleri oranı 2:1 olduğunda özellikle son koşullarda CU-LOR ve AC-LOR yöntemlerinden elde edilen sonuçların farklılaştığı, bunun dışında genel olarak paralel seyrettiği ve koşullar bazında sonuçların benzerlik gösterdiği görülmektedir. Bu örneklem büyüklüğü oranında genel olarak her iki yöntemden de elde edilen sonuçların C düzeyinde DAF göstermediği söylenebilir. Örneklem büyüklükleri oranı 1:1 olduğunda bütün koşullarda her iki yöntemden elde edilen sonuçların oldukça benzer olduğu görülmektedir. Koşulların değişimine göre her iki yöntemden elde edilen sonuçlar düzenli bir örüntü sergilememiştir. Genel olarak CU-LOR yönteminden elde edilen sonuçlar çok az bir miktar AC-LOR yönteminden daha yüksek olup, C düzeyinde ve ABD lehine DAF sergilemektedir. Örneklem büyüklükleri oranı 1:2 ve 1:3’te elde edilen sonuçların birbirine çok benzer olduğu görülmektedir. Son dört koşulda AC-LOR yönteminden elde edilen sonuçlar dışında her iki yöntemden elde edilen sonuçlar ABD lehine C düzeyinde DAF göstermiştir.

Adım 2’de bütün örneklem büyüklükleri oranında CU-LOR ve AC-LOR yöntemlerinden elde edilen sonuçların paralel; CU-LOR yönteminden elde edilen sonuçların AC-LOR yönteminden elde edilen sonuçlardan az bir miktar yüksek olduğu görülmektedir. AC-LOR yönteminden elde edilen DAF sonuçları genel olarak yalnızca koşul 7’de C düzeyinde ve ABD lehine DAF göstermiştir. CU-LOR yönteminden elde edilen sonuçları ise özellikle son koşullarda C düzeyinde ve ABD lehine DAF sergilemiştir.

Adım 3’te bütün örneklem büyüklükleri oranında her iki yöntemden elde edilen sonuçlar koşul 6’da ani bir artış göstermiştir. CU-LOR ve AC-LOR yöntemlerinden elde edilen DAF sonuçları paralel olup; CU-LOR yönteminden elde edilen sonuçlar bir miktar daha yüksektir. AC-LOR yönteminden elde edilen DAF sonuçları yalnızca örneklem büyüklükleri oranının 2:1 olduğu durumda koşul 6’da C düzeyinde DAF sergilemişken; CU-LOR yönteminden elde edilen sonuçlar bütün örneklem büyüklükleri oranında koşul 6, 7 ve 8’de DAF sergilemiştir. Adım 3’te C düzeyinde elde edilen bütün DAF sonuçları ABD lehinedir.

Son olarak Adım 4'te her iki yöntemden de elde edilen sonuçların C düzeyinde DAF göstermediği; CU-LOR yönteminden elde edilen sonuçların AC-LOR yönteminden elde edilen sonuçlardan bir miktar daha yüksek olduğu ve örneklem büyüklükleri oranının değişimine göre sonuçların çok fazla değişmediği görülmektedir.

Türkiye-Kazakistan karşılaştırmasında AC-LOR ve CU-LOR yöntemleri kullanılarak elde edilen DAF sonuçları incelendiğinde;

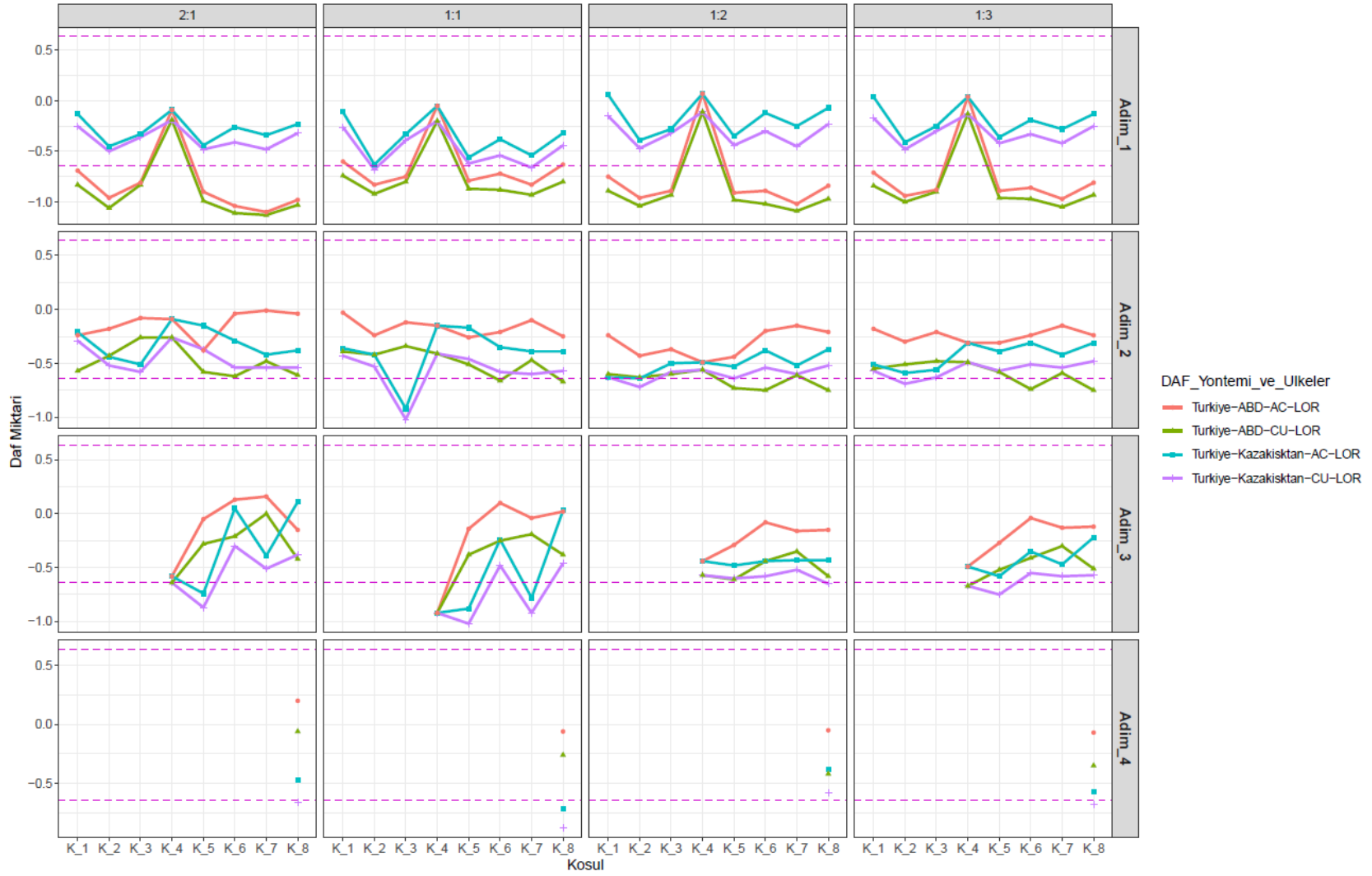
Adım 1'de örneklem büyüklükleri oranı 2:1, 1:2 ve 1:3 iken her iki yöntemde elde edilen DAF miktarlarının benzer düzeylerde seyrettiği ve tüm koşullarda C düzeyinde DAF elde edildiği görülmektedir. Örneklem büyüklüğü oranı 1:1 olduğunda ise elde edilen DAF miktarlarının diğer örneklem büyüklüğü oranlarında kıyasla ilk koşullarda artarken son üç koşulda azaldığı belirlenmiştir. Genel olarak CU-LOR yönteminden elde edilen sonuçların AC-LOR yönteminden bir miktar daha yüksek olduğu söylenebilir.

Adım 2'de bütün örneklem büyüklükleri oranında CU-LOR ve AC-LOR yöntemlerinden elde edilen sonuçlar paralel olmakla birlikte; elde edilen DAF miktarları koşullar bazında farklılaşmaktadır. Örneklem büyüklüğü oranı 2:1 iken koşul 4 ve 5'te; örneklem büyüklüğü oranı 1:1 iken koşul 6, 7 ve 8'de Kazakistan lehine C düzeyinde DAF elde edilmiştir. Diğer örneklem büyüklüğü oranlarında hesaplanan DAF miktarları kritik değerlerin altındadır.

Adım 3'te bütün örneklem büyüklüğü oranı 2:1 iken koşullardan elde edilen DAF miktarları oldukça değişkenlik göstermiştir. Öyle ki koşul 4 ve 5'te negatif yönde kritik değere yaklaşan DAF miktarı koşul 6 ve 8'de pozitif yönde kritik değere yaklaşmıştır. Bununla birlikte bu örneklem oranında C düzeyinde DAF elde edilmemiştir. Diğer örneklem büyüklüğü oranlarında CU-LOR ve AC-LOR yöntemlerinden elde edilen DAF sonuçları paralel olup; CU-LOR yönteminden elde edilen sonuçlar bir miktar daha yüksektir. Her iki yöntem için de hesaplanan DAF miktarları kritik değerlerin altındadır.

Son olarak Adım 4'te her iki yöntemden de elde edilen sonuçların C düzeyinde DAF göstermediği; CU-LOR yönteminden elde edilen sonuçların AC-LOR yönteminden elde edilen sonuçlardan bir miktar daha yüksek olduğu ve örneklem büyüklükleri oranının değişimine göre sonuçların çok fazla değişmediği görülmektedir.

Madde 9'a ait adımların içerdiği DAF miktarları ülkelere göre değerlendirildiğinde, Türkiye-ABD karşılaştırmasında her iki yöntem için de Adım1, 2 ve 3'ün ABD lehine C düzeyinde DAF içerdiği görülmektedir. Elde edilen DAF miktarlarının Adım 2 ve 3 için Türkiye-Kazakistan karşılaştırmasında ortaya çıkan DAF'den yüksek olduğu dikkat çekmektedir. Bununla birlikte Adım 1 için Türkiye-Kazakistan karşılaştırmasında ortaya çıkan DAF daha yüksek olup; Adım4 'te elde edilen DAF miktarları birbirine benzer düzeydedir. Türkiye-Kazakistan karşılaştırmasında Adım 1'deki Koşul 8 haricinde C düzeyinde DAF yoktur. Adım 1 ve 2'de hesaplanan DAF'ler ise Kazakistan lehinedir. Adım 4'te ülke karşılaştırmalarının hiç birinde C düzeyinde DAF elde edilmemiştir.



Şekil 13. Odak grup örneklem büyüklüğü 1000 iken Madde 10'a ait DAF miktarının değişimi.

Şekil 13'te odak grup (Türkiye) örneklem büyüklüğü 1000 iken, madde 10'un ülkelere ve DAF belirleme yöntemlerine göre karşılaştırıldığı DAF sonuçları yer almaktadır. Türkiye-ABD için DAF sonuçları incelendiğinde;

Adım 1'de bütün örneklem büyüklükleri oranında elde edilen sonuçların birbirine oldukça benzediği görülmektedir. Bütün örneklem büyüklükleri oranında koşul 4'te her iki yöntemden elde edilen sonuçların ani bir düşüş gösterdiği, bu koşul dışındaki koşullarda elde edilen sonuçların koşullar bazında benzer olduğu söylenebilir. Yöntemler bazında karşılaştırıldığında AC-LOR ve CU-LOR yönteminden elde edilen sonuçların çok benzer olduğu görülmektedir. Ani düşüşün yaşandığı koşul 4 dışındaki bütün koşullarda hem CU-LOR hem de AC-LOR yönteminden elde edilen DAF sonuçları C düzeyinde olup, Türkiye lehinedir.

Adım 2'de bütün örneklem büyüklükleri oranında CU-LOR yönteminden elde edilen sonuçlar AC-LOR yönteminden elde edilen sonuçlardan daha yüksektir. CU-LOR yönteminden elde edilen sonuçlar örneklem büyüklükleri oranı 1:2 ve 1:3 iken bazı koşullarda C düzeyinde DAF göstermiş olup, DAF'nin yönü Türkiye lehinedir.

Adım 3'te örneklem büyüklükleri oranı 2:1 ve 1:1 iken koşul 4'te her iki yöntemden elde edilen sonuçlar benzer olup, koşullar farklılaştıkça sonuçların da farklılaştığı ancak genel olarak paralel seyrettiği söylenebilir. Genel olarak değerlendirildiğinde, CU-LOR yönteminden elde edilen DAF miktarının AC-LOR yönteminden elde edilen sonuçlara göre daha yüksek olduğu görülmektedir. Yalnızca örneklem büyüklükleri oranı 1:1 iken koşul 4'te her iki yöntemden elde edilen sonuçların C düzeyinde DAF gösterdiği ve DAF'nin yönünün Türkiye lehine olduğu belirlenmiştir.

Son olarak Adım 4'te her iki yöntemden elde edilen DAF sonuçlarının C düzeyinde olmadığı görülmektedir.

Türkiye-Kazakistan karşılaştırmasında AC-LOR ve CU-LOR yöntemleri kullanılarak elde edilen DAF sonuçları incelendiğinde;

Adım 1'de bütün örneklem büyüklükleri oranında elde edilen sonuçların koşullar bazında farklılaşmakla birlikte; birbirine oldukça benzediği görülmektedir. Sadece örneklem büyüklüğü oranı 1:1 iken koşul 2, 5 ve 7'de C düzeyinde DAF elde edilmiş olup

diğer örneklem büyüklüğü oranlarında hiçbir koşulda C düzeyinde DAF hesaplanmamıştır.

Adım 2'de örneklem büyüklüğü oranı 1:2 iken her iki yöntem için de hesaplanan DAF değerlerinin birbirine paralel olduğu ve hiçbir koşulda C düzeyinde DAF sergilemediği görülmektedir. Örneklem büyüklüğü oranı 1:1 olduğunda koşul 3 için hesaplanan DAF miktarında ani bir artış olmuş ve Adım 2 her iki yönteme göre koşul 3'te Türkiye lehine DAF sergilemiştir. Diğer örneklem büyüklüğü oranlarında elde edilen DAF miktarları benzer düzeylerde olup ilk koşullarda C düzeyinde DAF elde edilmiştir.

Adım 3'te örneklem büyüklüğü 2:1 iken yöntemlerden elde edilen sonuçlar birbirine paralel olup koşullar bazında farklılaşmalar görülmektedir. Bu örneklem oranında sadece koşul 5'te C düzeyinde DAF elde edilmiştir. Örneklem büyüklüğü oranı 1:1 olduğunda ise DAF miktarlarının ve C düzeyinde DAF elde edilen koşulların sayısının arttığı görülmektedir. Örneklem büyüklüğü oranı 1:2 ve 1:3 iken DAF miktarları tekrar azalmakta ve kritik değerler etrafında seyretmektedir.

Son olarak Adım 4'te CU-LOR yönteminden elde edilen DAF miktarlarının çoğunlukla C düzeyinde olduğu; AC-LOR yönteminden elde edilen DAF miktarlarının ise kritik değerler altında seyrettiği görülmektedir.

Madde 10'a ait adımların içerdiği DAF miktarları ülkelere göre değerlendirildiğinde, Türkiye-ABD karşılaştırmasında her iki yöntem için de Adım1'in çoğunlukla ABD lehine C düzeyinde DAF içerdiği görülmektedir. Elde edilen DAF miktarlarının Adım 2 ve 3 için kritik değerlerin altında seyrettiği; Türkiye-Kazakistan karşılaştırmasında ise Türkiye lehine DAF sergileme eğiliminde olduğu söylenebilir. Bu durum Adım 3 için de benzer şekildedir. Adım 4'te ise Türkiye-ABD karşılaştırmasında C düzeyinde DAF elde edilmezken; Türkiye-Kazakistan karşılaştırmasında bazı örneklem büyüklüğü oranlarında Türkiye lehine C düzeyinde DAF tespit edilmiştir.

Alt Problem 5'e İlişkin Bulgular

Ülke karşılaştırmalarında; değişen kategori birleştirme kuralı ve odak grup:referans grup örneklem oranlarında çok kategorili DMF yöntemleriyle elde edilen DMF miktarları; odak grup örneklem büyüklüğüne göre farklılık göstermekte midir?

Alt problem 5'e yanıt bulmak üzere Türkiye-ABD ve Türkiye-Kazakistan karşılaştırmalarında; değişen koşullarda çok kategorili DMF yöntemleriyle (Mantel test, Liu- Agresti istatistiği, Cox β , poly-SIBTEST) elde edilen DMF miktarları odak grup örneklem büyüklüğüne göre ele alınmıştır. Bu bağlamda örneklem büyüklüğü 200 olduğu durumdaki sonuçların yer aldığı Şekil 2-4 ile örneklem büyüklüğünün 1000 olduğu durumdaki sonuçların yer aldığı Şekil 5-7'deki sonuçlar birlikte değerlendirilmiştir. Bu karşılaştırmayı görselleştirmek amacıyla grafikler oluşturulmuş ancak bu grafikler metinde hali hazırda var olan grafiklerden oluşturulduğu için Ek B'ye yerleştirilmiştir.

Ek B'de yer alan grafikler incelendiğinde, her üç madde için de odak grubun örneklem büyüklüğünün 1000 olduğu durumda elde edilen DMF miktarları örneklem büyüklüğü 200 iken elde edilen DMF miktarından daha büyük olduğu görülmektedir. Bu durum hem Türkiye-ABD hem de Türkiye-Kazakistan sonuçları için geçerlidir. Örneklem büyüklükleri oranı özellikle 1:2 ve 1:3 iken 200 ve 1000'lik örneklem için elde edilen sonuçlar arasındaki fark örneklem büyüklükleri oranı 2:1 olduğunda elde edilen farka kıyasla daha da açılmaktadır.

Ülkeler bazında incelendiğinde, Türkiye-ABD karşılaştırmasında 200 ve 1000'lik örneklemelerden elde edilen DMF farkının Türkiye-Kazakistan için 200-1000'lik örneklemeleri arasındaki farktan daha büyük olduğu görülmektedir. Genel olarak değerlendirildiğinde, 1000'lik örneklemden elde edilen DMF sonuçları 200'lük örneklemden elde edilen sonuçlarla benzer olsa da tamamen paralel olmadığı söylenebilir. Örneğin Madde 9'da Türkiye-ABD karşılaştırmasında MT yönteminden elde edilen sonuçlar incelendiğinde, örneklem büyüklüğü 200 iken koşullar değiştiğinde DMF miktarı çok fazla değişmemekte ancak örneklem büyüklüğü 1000 iken koşulların değişimine bağlı olarak elde edilen DMF sonuçları ani artma ve azalmalar göstermektedir.

Benzer bir durum Türkiye-Kazakistan karşılaştırmasında poly-SIBTEST sonuçlarında da görülmektedir.

Koşullar bazında incelendiğinde koşulların her iki örneklemden elde edilen sonuçların farkına bir etkisi olmadığı söylenebilir. Koşullar değiştikçe her iki örneklemden elde edilen sonuçlar arasındaki farkın buna bağlı olarak değişmediği görülmektedir.

Yöntem bazında karşılaştırıldığında, örneklem büyüklüğü etkisinin özellikle Türkiye-ABD sonuçlarından MT yönteminde daha fazla olduğu görülmektedir. Nitekim özellikle örneklem büyüklükleri oranı 1:2 ve 1:3 iken her iki örneklemden elde edilen sonuçlar arasındaki fark oldukça büyüktür.

Alt Problem 6'ya İlişkin Bulgular

Ülke karşılaştırmalarında; değişen kategori birleştirme kuralı ve odak grup:referans grup örneklem oranlarında DAF yöntemleriyle elde edilen DMF miktarları; odak grup örneklem büyüklüğüne göre farklılık göstermekte midir?

Alt problem 6'ya yanıt bulmak üzere Türkiye-ABD ve Türkiye-Kazakistan karşılaştırmalarında; değişen koşullarda DAF yöntemleriyle (AC-LOR, CU-LOR) elde edilen DAF miktarları odak grup örneklem büyüklüğüne göre ele alınmıştır. Bu bağlamda örneklem büyüklüğü 200 olduğu durumdaki sonuçların yer aldığı Şekil 8-10 ile örneklem büyüklüğünün 1000 olduğu durumdaki sonuçların yer aldığı Şekil 11-13'teki sonuçlar birlikte değerlendirilmiştir. Bu karşılaştırmayı görselleştirmek amacıyla grafikler oluşturulmuş ancak bu grafikler metinde hali hazırda var olan grafiklerden oluşturulduğu için Ek C'ye yerleştirilmiştir.

DAF sonuçlarının örneklem büyüklüklerine ilişkin dağılımı M4 üzerinde Türkiye-ABD karşılaştırması bağlamında incelendiğinde;

Genel olarak değerlendirildiğinde örneklem büyüklüğünün 1000 olduğu koşullarda elde edilen DAF miktarının daha büyük olduğu söylenebilir. Adım bazında incelendiğinde, bazı adımlarda AC-LOR yöntemiyle 200'lük örneklemelerde elde edilen sonuçların ani ivmelerle yükselip düştüğü görülmektedir. Örneğin Adım 2'de örneklem büyüklükleri oranı 2:1 iken koşul 6'da AC-LOR yönteminden elde edilen DAF miktarı ani

bir artış göstererek hem örneklem büyüklüğünün 1000 olduğu koşullardan hem de CU-LOR yönteminden daha fazla DAF üretmiştir. Genel olarak AC-LOR yönteminden elde edilen sonuçlarının örneklem büyüklüğü 200 iken kararlı bir dağılım sergilemediği söylenebilir. Adım 3'te örneklem büyüklüğü 200 iken bazı koşullarda AC-LOR yönteminin kestirim yapamadığı görülmektedir. CU-LOR yönteminden elde edilen sonuçlar incelendiğinde, yine örneklem büyüklüğünün 200 olduğu durumda CU-LOR yönteminden elde edilen sonuçların koşullar bazında kararlı bir dağılım göstermediği söylenebilir. Özellikle Adım 1'de örneklem büyüklüğü oranı 2:1 iken, 200 ve 1000 olduğu durumda elde edilen sonuçlar çok fazla farklılaşmazken diğer örneklem büyüklükleri oranında bu farkın bir miktar açıldığı görülmektedir. Adım 2'de ise 200 ve 1000'lik örneklemelerde elde edilen DAF sonuçlarının oldukça farklılaştığı söylenebilir.

M4 için dikkat çeken özelliklerden bir diğeri örneklem büyüklüğü 1000 iken AC-LOR yönteminden elde edilen sonuçların örneklem büyüklüğü 1000 iken CU-LOR yönteminden elde edilen sonuçlar ile paralellik göstermesidir. Buna göre örneklem büyüklüğü büyük iken CU-LOR ve AC-LOR yöntemlerinin benzer sonuçlar ürettiği söylenebilir. Son olarak Adım 1 ve Adım 2'de örneklem büyüklüklerine göre DAF sonuçları birbirinden farklı örüntüler çizse de Adım 3'te bu örüntülerin ani ivmelenmeler göstermediği belirlenmiştir.

M4 için Türkiye-Kazakistan sonuçları incelendiğinde ise;

Örneklem büyüklüğü 200 iken elde edilen sonuçların daha büyük olduğu görülmektedir. Türkiye-ABD sonuçlarına kıyasla aynı örneklem büyüklüklerinde CU-LOR ve AC-LOR yöntemleri daha paralel sonuçlar üretmiştir. Örneklem büyüklüğü 200 iken hem CU-LOR hem de AC-LOR yönteminden elde edilen sonuçların daha az kararlı olduğu, koşulların değişiminde çok fazla ani ivmemelerle artıp azaldığı görülmektedir. Örneklem büyüklüğü 1000 olduğunda her iki yöntemden elde edilen sonuçlar daha kararlı bir dağılım sergilemiştir. Örneklem büyüklüğü 1000 iken yine CU-LOR ve AC-LOR yöntemlerinin benzer sonuçlar ürettiği gözlenmiştir.

Madde 9 için Türkiye-ABD karşılaştırmasında;

Adım 1'de ve Adım 2'de örneklem büyüklüğü 200 iken CU-LOR ve AC-LOR yöntemleri benzer sonuçlar üretmiş, bu örneklem büyüklüğünde elde edilen sonuçlar, örneklem büyüklüğü 1000 iken elde edilen sonuçlardan daha yüksektir. Adım 3'te ise örneklem büyüklüklerine göre elde edilen DAF sonuçları arasındaki fark Adım 1 ve Adım 2'ye göre artmakla birlikte yine 200'lük örneklemelerde elde edilen DAF sonuçlarının daha yüksek olduğu belirlenmiştir.

Madde 9 ait Türkiye-Kazakistan sonuçları incelendiğinde;

Örneklem büyüklükleri oranı 2:1 iken Adım 1'de örneklem büyüklüğü 200 olduğunda hem CU-LOR hem AC-LOR yönteminden elde edilen sonuçlar bir koşulda ani bir artış gösterse de genel olarak bütün örneklem büyüklükleri oranında 200 ve 1000'lik örneklemelerden elde edilen sonuçlar birbirine oldukça yakındır. Adım 2'de ise farklı bir durum ortaya çıkmış olup, sonuçlar örneklem büyüklüklerine göre farklılaşmıştır. Buna göre adım 2'de bazı örneklem büyüklükleri oranında 200'lük örneklemelerde elde edilen sonuçlar bazılarında ise 1000'lik örneklemelerden elde edilen sonuçlar daha yüksektir. Buna göre Adım 2'de örneklem büyüklüklerinin etkisi belirli bir yönde değildir. Adım 3'te yine örneklem büyüklüğü 1000 olduğunda CU-LOR ve AC-LOR yöntemlerinden elde edilen sonuçların yakın olduğu görülmektedir.

Madde 10 için Türkiye-ABD karşılaştırmasına ilişkin sonuçlar incelendiğinde;

Adım 1'de koşul 4'te 1000'lik örneklemelerde ani bir ivme ile 200'lük örneklemelerden daha fazla DAF elde edildiği, bunun dışında her iki örneklem büyüklüğünde benzer sonuçlar elde edildiği gözlenmiştir. Adım 2'de ise 200'lük ve 1000'lik örneklemelerde elde edilen sonuçlar arasındaki farklılaşma artmış, genel olarak 200'lük örneklemelerde daha yüksek DAF sonuçları elde edilmiştir. Adım 3'te ise yine 200'lük örneklemelerde bazı koşullarda kestirim yapılamadığı belirlenmiştir. Genel olarak 200'lük örneklemelerde elde edilen DAF sonuçlarının daha yüksek olduğu söylenebilir.

Madde 10'un Türkiye-Kazakistan karşılaştırmasına ilişkin sonuçlar incelendiğinde;

Adım 1'de bazı koşullarda 200'lük örneklemelerden bazı koşullarda ise 1000'lik örneklemelerden elde edilen DAF sonuçları yüksek çıkmıştır. Adım 1'de örneklem büyüklükleri oranı 1:1 ve 1:3 iken her iki örneklem büyüklüğünden elde edilen sonuçlar

oldukça benzerdir. Örneklem büyüklükleri oranı 2:1 ve 1:2 iken ise 200'lük örneklemelerden elde edilen DAF sonuçları daha yüksek elde edilmiştir. Örneklem büyüklüğü 1000 olduğunda CU-LOR ve AC-LOR yöntemleri benzer sonuçlar üretmiştir.

Adım 2'de örneklem büyüklükleri oranı 2:1 iken 200'lük örneklemeler daha yüksek miktarda DAF üretmiş olup diğer örneklem büyüklükleri oranında ise her iki örneklem büyüklüğünden benzer sonuçlar elde edilmiştir.

Adım 3'te ise örneklem büyüklüğü 200 iken bazı koşullarda kestirim yapılamadığı; Adım 3'te ve Adım 4'te 200'lük örneklemelerde elde edilen sonuçlarda koşullar arasında daha çok dalgalanmalar meydana geldiği; 200'lük örneklemelerde daha fazla DAF elde edildiği belirlenmiştir.

Tartışma

Bu bölümde araştırma problemlerine ilişkin elde edilen sonuçlar birlikte ele alınarak alanyazın ile ilişkisine yönelik tartışmalara yer verilmiştir.

Çok kategorili DMF belirleme yöntemlerinin (Cox's B, Liu Agresti, MT ve poly-SIBTEST) örneklem büyüklüğü ve koşullara göre incelenmesi. DMF belirleme yöntemlerinden elde edilen sonuçlar incelendiğinde küçük örneklemde Cox's B, Liu Agresti ve MT yöntemlerinden elde edilen DMF miktarlarının birbirlerine oldukça benzer olduğu; poly-SIBTEST yönteminden elde edilen DMF miktarlarının ise diğer yöntemlerden farklılaştığı görülmektedir. Yöntemler arasından yüksek DMF sergilediği tespit edilen koşul sayısı en fazla poly-SIBTEST sonuçlarında elde edilmiştir. Diğer yöntemlerle uyumluluk göstermekle birlikte poly-SIBTEST yönteminin en fazla DMF'li madde tespit eden yöntem olduğu ve diğer yöntemlere göre daha hassas sonuçlar verdiği belirtilmektedir (Henderson, 2001; Mellor, 1995). Dört yöntemden elde edilen sonuçların birbirine en çok benzediği durumun örneklem büyüklüğü oranı 1:3 iken olduğu söylenebilir. Ancak referans grubun odak gruba oranının artmasıyla, yani odak grup örneklem büyüklüğüne karşın referans grup örneklem büyüklüğünün artmasıyla yöntemlerin DMF belirleme gücünün azalma eğiliminde olduğu; referans ve odak grup örneklem büyüklüğünün eşit olduğu durumlarda ise Tip I hatanın artma eğiliminde olduğu belirtilmiştir (Wang & Su, 2004; Zwick, 2012).

DMF belirleme yöntemlerinden elde edilen sonuçlar odak grup örneklem büyüklüğü açısından incelendiğinde ise örneklem büyüklüğü arttığında bütün yöntemlerin birbirine paralel sonuçlar verdiği görülmektedir. Yöntemlerden elde edilen DMF miktarlarında en yüksek değerler ve bu miktarlarda koşullar bazında meydana gelen değişkenlikler en fazla örneklem büyüklüğü oranları 1:2 ve 1:3 iken elde edilmiştir. Testlerin istatistiksel güç oranlarının örneklem büyüklüğünden oldukça etkilendiğini belirten çalışmalar mevcuttur (Bolt, 2002; Kristjansson ve diğerleri, 2005). Buna göre örneklem büyüklüğü arttıkça yöntemlerin daha yüksek istatistiksel güç oranına sahip olduğu belirtilmektedir (Yandı, 2017). Bu çalışmada yöntemlerden elde edilen DMF miktarları incelendiğinde yüksek düzeyde DMF miktarının büyük örneklemde daha fazla olduğu; küçük örneklemde maddelere ilişkin DMF miktarlarının çoğunlukla kritik değerlerin altında olduğu söylenebilir. Bununla birlikte yöntemlerin büyük örneklemde daha tutarlı sonuçlar verdiği görülmektedir.

DAF belirleme yöntemlerinin (AC-LOR ve CULOR) örneklem büyüklüğü ve koşullara göre incelenmesi. AC-LOR ve CULOR yöntemleri kıyaslandığında Türkiye-Kazakistan karşılaştırmasında M4 için Adım 1 ve 2’de AC-LOR yönteminden elde edilen DAF miktarlarının CU-LOR yönteminden elde edilen DAF miktarlarından daha yüksek; DAF miktarının yüksek olduğu koşul sayısının ise daha fazla olduğu görülmektedir. M4’ün diğer adımlarında ise CU-LOR yönteminden elde edilen sonuçlar daha yüksektir. Diğer yandan Madde 9 incelendiğinde ilk iki adımda bazı koşullarda AC-LOR, bazı koşullarda ise CU-LOR yönteminden elde edilen sonuçların daha yüksek olduğu, dolayısıyla koşullar bazında yöntemler arasında belirgin bir farklılaşma olmadığı söylenebilir. Bununla birlikte Madde 9’un diğer adımlarında ve Madde 10’un tüm adımlarında CU-LOR yönteminden elde edilen DAF miktarları daha yüksektir. Bu durum Türkiye-ABD karşılaştırmasında tüm maddeler için geçerlidir. Gattamorta and Penfield (2012), bu iki yöntemi kıyasladıkları çalışmalarında ardışık kategoriler yaklaşımında kullanılan AC-LOR yönteminde, sadece etki büyüklüğüne göre orta ila yüksek DAF sergileyen daha fazla adım bulunduğunu belirtmişlerdir. Hem etki büyüklüklerine hem de anlamlılık testlerine göre incelendiğinde ise kümülatif yaklaşım altında kullanılan CU-LOR yöntemine göre anlamlı DAF sergileyen adımların sayısının daha fazla olduğu görülmüştür. CU-LOR yönteminden elde edilen sonuçların daha küçük standart hatalar vermesine bağlı olarak AC-LOR yöntemine

kıyasla sonuçların istatistiksel olarak anlamlı çıkma olasılığının yüksek olduğu ifade edilmiştir. Diğer yandan kümülatif yaklaşımda tüm adımlardan gelen yanıtların kullanılması nedeniyle CU-LOR istatistiğinin; sadece ardışık kategorilerdeki yanıtları kullanan AC-LOR istatistiğine göre daha yüksek güce sahip olduğu belirtilmektedir (Ayodele, 2017).

DAF belirleme yöntemleri örneklem büyüklüklerine göre incelendiğinde Türkiye-Kazakistan ve Türkiye-ABD karşılaştırmalarında örneklemin küçük olduğu durumlarda her iki yöntemden elde edilen DAF miktarlarının büyük örnekleme kıyasla daha yüksek olduğu görülmüştür. Öyle ki Türkiye-Kazakistan karşılaştırmasında aynı madde (M4, M10) küçük örnekleme yarı yarıya yüksek DAF bulundururken; büyük örnekleme hemen hemen hiç yüksek DAF sergilememektedir. Ek'te ülke karşılaştırmaları ve örneklem büyüklüğüne göre AC-LOR ve CU-LOR yöntemlerinin yüksek DAF'li olarak işaretlediği/işaretlemediği adım sayıları verilmiştir. Türkiye-Kazakistan karşılaştırması madde ve koşulların tümü için değerlendirildiğinde 136 (%25,76) adımda her iki yönleme göre yüksek DAF olduğu; 326 (%61,74) adımda ise her iki yönleme göre yüksek DAF olmadığı tespit edilmiştir. Bununla birlikte AC-LOR yöntemine göre yüksek DAF olmayan 43 (%8,14) adımda CU-LOR yöntemine göre yüksek DAF tespit edilmiştir. CU-LOR yöntemine göre yüksek DAF olmayan 23 (%4,36) adımda ise AC-LOR yöntemine göre yüksek DAF hesaplanmıştır. Bu karşılaştırmada yöntemlerin madde adımlarını DAF açısından sınıflamadaki benzerlik oranları; odak grup örneklem büyüklüğü 200 için %77-%86 aralığında değişirken; odak grup örneklem büyüklüğü 1000 iken %89-%99 aralığında değişmektedir.

Türkiye-ABD karşılaştırmasında ise 198 (%37,5) adımda her iki yönleme göre yüksek DAF olduğu; 168 (%31,82) adımda ise her iki yönleme göre yüksek DAF olmadığı tespit edilmiştir. Bununla birlikte AC-LOR yöntemine göre yüksek DAF olmayan 159 (%30,11) adımda CU-LOR yöntemine göre yüksek DAF tespit edilmiştir. CU-LOR yöntemine göre yüksek DAF olmayan 3 (%0,57) adımda ise AC-LOR yöntemine göre yüksek DAF hesaplanmıştır. Bu karşılaştırmada yöntemlerin madde adımlarını DAF açısından sınıflamadaki benzerlik oranları; odak grup örneklem büyüklüğü 200 için %53-%72 aralığında değişirken; odak grup örneklem büyüklüğü 1000 iken %59-%89

aralığında değişmektedir. Dolayısıyla yöntemlerin madde adımlarını DAF açısından sınıflamadaki benzerlik oranlarının daha yüksek olduğu söylenebilir. Bu durum, özellikle küçük örneklerde değişmezliğin yorumlanarak nihayetinde madde revizyonu ya da kaldırılmasına ilişkin kararın verilmesinde; kullanılan yöntemlerin önemini açıkça göstermektedir.

Maddelere ait adımların yöntemler bazında yüksek düzey DAF içerip içermediğine ilişkin sınıflamalar incelendiğinde yöntemlerin benzerlik oranlarının büyük örneklerde çok daha yüksek olması oldukça dikkat çekmektedir. Özellikle örneklem büyüklüğü 1000 iken M4 ve M9'a ilişkin Türkiye-Kazakistan karşılaştırmasında CU-LOR ve AC-LOR yöntemleriyle yapılan DAF sınıflamasında yöntemlerin uyuşma yüzdeleri oldukça yüksektir (%99 ve %94). Türkiye-ABD karşılaştırmasında; küçük örneklemde yöntemlerin sınıflamaya ilişkin benzerlik oranı düşüken örneklem büyüklüğü arttığında benzerlik oranlarında da bir artış olduğu görülmektedir. Dolayısıyla yöntemlerin maddeleri DAF bakımından sınıflandırmada özellikle büyük örneklemde oldukça tutarlı sonuçlar ürettiği söylenebilir. Bu sonuca paralel olarak alanyazında, küçük DAF haricinde diğer DAF sınıflandırmalarında AC-LOR yöntemi daha yüksek DAF değeri vermekle birlikte her iki yöntemin çoğunlukla tutarlı sonuçlar verdiği belirtilmiştir (Gattamorta, 2009).

Yöntemlere ilişkin sonuçlar örneklem büyüklüğü oranları ve koşullar bazında incelendiğinde; aynı örneklem büyüklüğü oranında bazı maddelere ilişkin DAF miktarlarında artış görülürken bazı maddelerin DAF miktarlarında azalma görülmektedir. Dolayısıyla örneklem büyüklüğü oranlarının DAF sonuçlarına önemli bir etkisi olmadığı söylenebilir. Diğer yandan; incelenen koşullar sonuçları önemli ölçüde etkilememekle birlikte, koşullar değiştikçe AC-LOR yönteminden elde edilen sonuçlarda dalgalanmalar olduğu; CU-LOR yönteminde ise sonuçların koşullar bazında paralellik gösterdiği söylenebilir. Alanyazında kümülatif yaklaşım altında kestirilen DAF miktarlarının, ardışık kategoriler yaklaşımı altında tahmin edilen DAF miktarlarından daha istikrarlı olduğu belirtilmektedir (Gattamorta & Penfield, 2012; Penfield, 2008). İstikrarlı olsun ya da olmasın her iki yöntem için de adım sayısının DAF sonuçları üzerinde oluşturduğu örüntünün sistematik olmadığı görülmüştür. Ayodele (2017), çalışmasında benzer

sonuçlara ulaşarak örneklem büyüklüğü oranı ve adım sayısının DAF miktarları üzerinde istatistiksel ve pratik bir öneme sahip olmadığını belirtmiştir.

Maddelere ilişkin puan kategorilerinin işaretlenme sıklıkları incelendiğinde M4, M9 ve M10 maddelerinde bireylerin yaklaşık yarısının ilk iki seçeneğe yoğunlaştığı görülmektedir. Bununla birlikte Türkiye verisinde bireylerin yarısından fazlasının ilk seçeneği işaretlemesi kategorilerin dağılımını daha çarpık hale getirmektedir. Koşulların oluşturulması bu bağlamda incelendiğinde en çok işaretlenen 1. ve 2. seçeneklerin, üç kategorili veri için koşul 1 ve 3; dört kategorili veri için ise koşul 4'te birleştirildiği görülmektedir. En az sıklıkta işaretlenen 4. ve 5. seçenekler ise üç kategorili veri için koşul 1 ve 2; dört kategorili veri için koşul 6 da birleştirilmiştir. DMF analizleri sonucunda üç kategorili veriler için oluşturulan koşul 1, 2 ve 3 arasında en fazla DMF miktarının çoğunlukla koşul 2'de elde edildiği görülmektedir. Dört kategorili veriler (koşul 4, 5, 6 ve 7) kendi arasında değerlendirildiğinde ise sistematik bir örüntü olmamakla birlikte koşul 6'da koşul 4'e kıyasla daha fazla DMF miktarı elde edildiği söylenebilir. DAF analizi sonuçları incelendiğinde ise Adım 1'de ilk iki seçeneğin birleştirildiği koşul 1 ve 3'e ilişkin sonuçların koşul 2'ye ilişkin sonuçlardan farklılaştığı görülmektedir. Bu farklılaşma sistematik olmayıp koşul 2'ye ait sonuçlar bazı maddelerde yüksekken bazı maddelerde düşüktür. Diğer yandan dört kategorili veri için oluşturulan koşullarda Adım 1'de elde edilen DAF miktarının koşul 4 ve diğerleri arasında farklılaştığı söylenebilir. Bu farklılaşmanın yönü standart olmayıp bazı maddelerde koşul 4'te en yüksek DAF değeri elde edilirken bazı maddelerde koşul 4'te en düşük DAF miktarı elde edilmektedir.

Çok kategorili DMF ve DAF belirleme yöntemlerinden elde edilen sonuçların birlikte incelenmesi. DMF ve DAF üzerine yapılan çalışmalar incelendiğinde maddelerde aynı anda DMF/DAF analizi yapılan (Akour ve diğerleri, 2015) veya önce DMF analizi yapılarak sonrasında sadece DMF'li maddelere DAF analizi yapılan (Miller ve diğerleri, 2010) çalışmalar olduğu görülmektedir. Akour ve diğerleri, 2015); hiçbir adımında yüksek DAF sergilemeyen maddelerin aynı zamanda DMF de sergilemediğini belirtmiştir. Bununla birlikte madde adımlarında DAF olmadığı durumlarda DMF belirleyen bazı yöntemlerde Tip I hatanın yüksek olduğu görülmüştür (Ayodele, 2017). Yani nadir olmakla birlikte yöntemlerin bir kısmında DMF'li olmayan bir maddenin DMF'li olarak

işaretlendiği durumlara rastlanmıştır. Bu çalışmadan elde edilen sonuçlar incelendiğinde büyük örnekleme hiçbir adımda DAF sergilemeyen M4'ün, DMF analizleri sonuçlarının kritik değerlerin altında olduğu yani DMF sergilemediği görülmektedir. Diğer yandan küçük örnekleme örneklem büyüklüğü oranı 1:1 ve 1:2 iken M9'a ilişkin DAF sonuçları incelendiğinde Adım 1 ve 4'te elde edilen DAF miktarları yüksek ve zıt işaretlidir. İlgili maddenin DMF sonuçlarına bakıldığında ise aynı örneklem büyüklüğü oranlarında yöntemlerin çoğuna göre maddenin DMF'li olmadığı belirlenmiştir. Bu durum adımlarda görülen zıt işaretli DAF miktarlarının DMF etkisini yaklaşık olarak sıfıra indirmesinden kaynaklanabilmektedir. DMF sergilemeyen maddelere DAF analizi yapılmaması durumunda adımlara ilişkin DAF miktarları hakkında bilgi sahibi olunamamaktadır. Dolayısıyla önce DMF analizi yapıp sonrasında sadece DMF'li maddelere DAF analizi yapılması halinde adımlar hakkındaki önemli bilgilerin göz ardı edilebileceği unutulmamalıdır. Nitekim DAF miktarı adımlar boyunca işaret ve büyüklük olarak değiştiğinde birçok DMF belirleme yöntemlerinin nispeten düşük güç gösterdiği belirtilmiştir (Ankenmann ve diğerleri, 1999; Chang ve diğerleri, 1996; Penfield & Algina, 2003; Wang & Su, 2004).

DMF ve DAF analizleri birlikte incelendiğinde DMF miktarının en yüksek olduğu durumlarda (bkz. Şekil 2; TUR-ABD), ilgili maddenin adımlarından elde edilen DAF miktarları değişmekle birlikte işaretlerinin aynı olduğu görülmektedir.

Bölüm 5

Sonuç ve Öneriler

Bu bölümde araştırma problemlerine ilişkin elde edilen sonuçlar sıralanmış ve sonraki çalışmalara ilişkin öneriler sunulmuştur.

Sonuçlar

Çok kategorili DMF belirleme yöntemlerine (Cox's B, Liu-Agresti, MT ve Poly-SIBTEST) ilişkin sonuçlar.

1. Odak grup örneklem büyüklüğü küçük olduğunda Cox's B, Liu Agresti ve MT yöntemlerinden elde edilen DMF miktarlarının birbirlerine oldukça benzediği görülmüş; PolySIBTEST yönteminden elde edilen sonuçlarda ise diğer yöntemlerden farklı olarak daha fazla yüksek düzey DMF değerleri elde edilmiştir.
2. Yine odak grup örneklem büyüklüğü küçük olduğunda Cox's B, Liu Agresti, MT ve poly-SIBTEST yöntemlerinden elde edilen sonuçların koşullar bazında oluşturduğu örüntüler incelendiğinde birbirine en çok benzeyen örüntüler örneklem büyüklüğü oranı 1:3 iken elde edilmiştir.
3. DMF belirleme yöntemlerinden elde edilen sonuçlar odak grup örneklem büyüklüğü açısından incelendiğinde ise örneklem büyüklüğü arttığında bütün yöntemlere göre daha fazla yüksek DMF elde edildiği ve yöntemlerin birbirine paralel sonuçlar verdiği görülmüştür.
4. Yöntemlerden elde edilen DMF miktarlarında en yüksek değerler ve bu miktarlarda koşullar bazında meydana gelen değişkenlikler en fazla örneklem büyüklüğü oranları 1:2 ve 1:3 iken elde edilmiştir. Odak ve referans grup örneklem büyüklükleri eşit olduğunda elde edilen DMF değerleri daha düşük olup; sonuçlar koşullar bazında daha kararlıdır.

DAF belirleme yöntemlerine (AC-LOR ve CULOR) ilişkin sonuçlar.

1. AC-LOR ve CU-LOR yöntemleri kullanılarak yapılan DAF analizi sonuçlarında bazı durumlarda CU-LOR yönteminden elde edilen DAF değerlerinin; bazı durumlarda ise AC-LOR yönteminden elde edilen DAF değerlerinin daha yüksek olduğu görülmüştür.

2. DAF belirleme yöntemleri örneklem büyüklüklerine göre incelendiğinde örneklemin küçük olduğu durumlarda her iki yöntemden elde edilen DAF miktarlarının büyük örnekleme kıyasla daha yüksek olduğu görülmüştür.
3. Örneklem büyüklüğü arttıkça kümülatif yaklaşım altında kullanılan CU-LOR ve ardışık kategoriler yaklaşımında kullanılan AC-LOR yöntemlerinin, madde adımlarını DAF açısından sınıflamadaki benzerlik oranları artmaktadır.
4. DAF belirleme yöntemlerine ilişkin sonuçlar örneklem büyüklüğü oranları ve koşullar bazında incelendiğinde; aynı örneklem büyüklüğü oranında bazı maddelere ilişkin DAF miktarlarında artış görülürken bazı maddelerin DAF miktarlarında azalma görülmektedir. Dolayısıyla örneklem büyüklüğü oranlarının DAF sonuçlarına önemli bir etkisi olmadığı düşünülmektedir.
5. Kategori birleştirme kuralının DAF analizleri üzerinde sistematik bir etkisi görülmemiştir. Bununla birlikte işaretlenme sıklıkları yüksek olan kategorilerin birleştirmesiyle oluşturulan koşullardan elde edilen sonuçların diğer koşullardan elde edilen sonuçlardan farklı olduğu söylenebilir.

Çok kategorili DMF ve DAF belirleme yöntemlerinden birlikte elde edilen sonuçlar.

1. Adımlarında zıt işaretli DAF sergileyen maddelerde DMF etkisinin yaklaşık olarak sifıra indiği görülmüştür. Dolayısıyla DMF sergilemeyen maddelere DAF analizi yapılmaması durumunda, adımlara ilişkin DAF miktarlarının göz ardı edileceği unutulmamalıdır.
2. Adımlarında aynı yönde yüksek DAF sergileyen maddelerin DMF miktarlarının, adımlarında zıt yönde yüksek DAF sergileyen maddelerin DMF miktarından daha yüksek olduğu belirlenmiştir.
3. Odak grup örneklem büyüklüğü arttığında çok kategorili DMF yöntemlerinden elde edilen DMF miktarları artarken; DAF yöntemlerinden elde edilen DAF miktarlarının azaldığı görülmüştür. Bu durumda örneklem büyüklüğü arttığında madde adımlarındaki DAF değerlerinin işaret olarak zıt yönde değiştiği sonucuna ulaşılabılır.

Öneriler

Uygulayıcılar için öneriler.

1. Adımlarda görülen zıt işaretli DAF miktarlarının madde düzeyindeki DMF etkisini yaklaşık olarak sıfıra indirmesi nedeniyle; DMF sergilemeyen maddelere DAF analizi yapılmaması adımlara ilişkin DAF miktarları hakkında bilgi sahibi olunmaması anlamına gelmektedir. Dolayısıyla madde revizyonu ya da maddenin kaldırılması kararını vermek üzere sadece DMF sergileyen maddelere değil, tüm maddelere DAF analizi yapılması önerilmektedir.
2. Verinin çok kategorili olması, veride herhangi bir değişikliğe gidilmeden verileri ham haliyle kullanılması daha geçerli sonuçlar üretecektir. Ancak çeşitli nedenlerle kategori birleştirme yoluna gidilecekse hangi ardışık kategorileri birleştirileceğinin bir etkisi olmayıp, araştırmanın ve verinin doğasına uygun olarak kategori birleştirmesine gidilmesi önerilmektedir.
3. Penfield (2008); verilerin ATM'ye uygun olduğu durumlarda CU-LOR; GKPM'ye uygun olduğu durumlarda ise AC-LOR yönteminin kullanılabileceğini belirtmiştir. Bu çalışmada genel olarak AC-LOR ve CU-LOR yönteminden elde edilen sonuçların benzer olduğu, dolayısıyla modele en uygun yöntemin seçilmesi önerilmektedir.
4. Odak grup örneklem büyüklüğü arttığında çok kategorili DMF ve DAF değerlerinin zıt yönde değiştiği, ancak yöntemlerin büyük örneklemde daha benzer sonuçlar ürettiği görülmüştür. Dolayısıyla büyük örneklem üzerinde çalışılmasının daha kararlı sonuçlar üreteceği düşünülmektedir.

Sonraki araştırmalara yönelik öneriler.

1. Bu çalışmada DAF belirleme yöntemlerinin madde adımlarını DAF açısından sınıflamadaki benzerlik oranları büyük örneklemde daha yüksek çıkmıştır. Penfield, Alvarez ve diğerleri (2008); DAF sınıflandırmasında etki büyüklüğü ile birlikte istatistiksel anlamın da test edilebileceğini ancak yeterli örneklem büyüklüğü elde edildiğinde (örn; her grup için $n > 300$) tek başına etki büyüklüğünün yeterli olabileceğini belirtmiştir. Bu çalışmada küçük ve büyük örneklemde sadece etki büyüklüğüne göre sınıflandırma yapılmış olup gelecek çalışmalarda küçük örneklem için etki büyüklükleri istatistiksel olarak test edilebilir.

2. Koşullar oluşturulurken kategori birleştirme yoluna gidilerek ardışık kategoriler farklı kombinasyonlarda birleştirilmiştir. Koşullar bazında elde edilen sonuçların birleştirilen kategorilerin işaretlenme sıklıklarına göre farklılaştığı düşünülmektedir. Bu bağlamda kategorilerin işaretlenme sıklıklarının DAF sonuçlarına etkisini inceleyen çalışmalar yapılabilir.
3. Bu çalışmada analizler DAF belirleme yaklaşımlarından ardışık ve kümülatif yaklaşımlar kullanılarak yapılmıştır. Sonraki çalışmalarda diğer DAF belirleme yaklaşımları (örn; MTK yaklaşımı, lojistik regresyon yaklaşımı) kullanılarak yaklaşımlardan elde edilen sonuçların benzerlikleri ya da farklılıkları incelenebilir.
4. DMF belirleme çalışmaları incelendiğinde eşleştirme değişkeninin geçerliğini artırdığı düşüncesiyle eşleştirme değişkeni olan toplam puanın DMF'li maddelerden arındırılması gerektiğini belirten çalışmalara da rastlanmaktadır (French & Maller, 2007; Gierl ve diğerleri, 1999; Lee & Geisinger, 2016; Svetina & Rutkowski, 2014). Bu çalışmada eşleştirme değişkeni, DMF'li maddelerden arındırılmamış olup sonuçlar bu bağlamda ele alınmıştır. Sonraki çalışmalarda eşleştirme değişkeni arındırılarak analizler yapılabilir ve sonuçlar karşılaştırılabilir.
5. Simülatif veriler kullanılarak bu çalışmada manipüle edilen değişkenlerin yer aldığı benzer bir çalışma yürütülerek bu araştırmadan elde edilen sonuçların teorik bağlamda incelenmesi önerilmektedir.

Kaynaklar

- Ababneh, E., Al-Tweissi, A., & Abulibdeh, K. (2016). TIMSS and PISA impact—the case of Jordan. *Research Papers in Education, 31*(5), 542–555.
- Akour, M., Sabah, S., & Hammouri, H. (2015). Net and global differential item functioning in pisa polytomously scored science items. *Journal of Psychoeducational Assessment, 33*(2), 166–176.
- Allalouf, A., Hambleton, R. K., & Sireci, S. G. (1999). Identifying the causes of DIF in translated verbal items. *Journal of Educational Measurement, 36*(3), 185–198.
- AERA, APA & NCME. (2014). Standards for education and psychological testing. American Educational Research Association.
- Ankenmann, R. D., Witt, E. A., & Dunbar, S. B. (1999). An investigation of the power of the likelihood ratio goodness-of-fit statistic in detecting differential item functioning. *Journal of Educational Measurement, 36*(4), 277–300.
- Ayodele, A.N. (2017). *Examining power and type 1 error for step and item level tests of invariance: Investigating the effect of the number of item score levels* (Doctoral dissertation). University of Minnesota, USA.
- Baird, J. A., Johnson, S., Hopfenbeck, T. N., Isaacs, T., Sprague, T., Stobart, G., & Yu, G. (2016). On the supranational spell of PISA in policy. *Educational Research, 58*(2), 121–138.
- Benítez, I., Padilla, J.L., Hidalgo Montesinos, M. D., & Sireci, S. G. (2015). Using mixed methods to interpret differential item functioning. *Applied Measurement in Education, 29*(1), 1–16.
- Bolt, D. M. (2002). A Monte Carlo comparison of parametric and nonparametric polytomous DIF detection methods. *Applied Measurement in Education, 15*(2), 113–141.
- Büyüköztürk, Ş. (2018). *Veri analizi el kitabı*. Pegem Akademi Yayıncılık.
- Camilli, G., & Congdon, P. (1999). Application of a method of estimating DIF for polytomous test items. *Journal of Educational and Behavioral Statistics, 24*(4), 323–341.

- Camilli, G., & Shepard, L. A. (1994). *Methods for identifying biased test items*. Sage Publications.
- Chang, H. H., Mazzeo, J., & Roussos, L. (1996). Detecting DIF for polytomously scored items: An adaptation of the SIBTEST procedure. *Journal of educational measurement, 33*(3), 333–353.
- Clauser, B. E., & Mazor, K. M. (1998). Using statistical procedures to identify differentially functioning test items. An NCME instructional module. *Educational Measurement: Issues and Practice, 17*(1), 31–44.
- Cohen, A. S., Kim, S.-H., & Baker, F. B. (1993). Detection of differential item functioning in the graded response model. *Applied Psychological Measurement, 17*(4), 335–350.
- Cox, D. R. (1958). The regression analysis of binary sequences. *Journal of the Royal Statistical Society: Series B (Methodological), 20*(2), 215–232.
- Dancer, L. S., Anderson, A. J., & Derlin, R. L. (1994). Use of log-linear models for assessing differential item functioning in a measure of psychological functioning. *Journal of Consulting and Clinical Psychology, 62*(4), 710–717.
- Dorans, N. J., & Holland, P. W. (1993). DIF detection and description: Mantel Haenszel and standardization. In P. W. Holland & H. Wainer (Eds.), *Differential item functioning* (pp. 35–66). Lawrence Erlbaum Associates, Inc.
- Dorans, N. J., & Kulick, E. (1986). Demonstrating the utility of the standardization approach to assessing unexpected differential item performance on the Scholastic Aptitude Test. *Journal of Educational Measurement, 23*(4), 355–368.
- Ellis, B. B., & Raju, N. S. (2003). Test and Item Bias: What they are, what they aren't, and how to detect them. In J. E. Wall & G. R. Walz (Eds.), *Measuring up: Assessment issues for teachers, counselors, and administrators* (pp. 89–98). Greensboro, N.C.
- Elosua, P., & Wells, C. S. (2013). Detecting DIF in polytomous items using MACS, IRT and ordinal logistic regression. *Psicológica, 34*(2), 327–342.
- Embretson, S. E. (2007). Construct validity: A universal validity system or just another test evaluation procedure?. *Educational Researcher, 36*(8), 449–455.

- Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists*. Mahwah, Lawrence Erlbaum Associates, Inc.
- Emenogu, B., & Childs, R. A. (2003). *Curriculum and translation differential item functioning: A comparison of two DIF detection techniques*. <https://files.eric.ed.gov/fulltext/ED476424.pdf>
- Feng, Y. (2008). *Difference in gender differential item functioning patterns across item format and subject area on diploma examinations after change in administration procedure* (Doctoral dissertation). University of Alberta, Canada.
- French, A. W., & Miller, T. R. (1996). Logistic regression and its use in detecting differential item functioning in polytomous items. *Journal of Educational Measurement, 33*(3), 315–332.
- French, B. F., & Maller, S. J. (2007). Iterative purification and effect size use with logistic regression for differential item functioning detection. *Educational and Psychological Measurement, 67*(3), 373–393.
- Furlow, C. F., Raiford Ross, T., & Gagné, P. (2009). The impact of multidimensionality on the detection of differential bundle functioning using simultaneous item bias test. *Applied Psychological Measurement, 33*(6), 441–464.
- Gattamorta, K. A. (2009). *A comparison of adjacent categories and cumulative DSF effect estimators* (Doctoral dissertation). University of Miami, Florida.
- Gattamorta, K. A., & Penfield, R. D. (2012). A comparison of adjacent categories and cumulative differential step functioning effect estimators. *Applied Measurement in Education, 25*(2), 142–161.
- Gattamorta, K. A., Penfield, R. D., & Myers, N. D. (2012). Modeling item-level and step-level invariance effects in polytomous items using the partial credit model. *International Journal of Testing, 12*(3), 252–272.
- Gelin, M. N., & Zumbo, B. D. (2003). Differential item functioning results may change depending on how an item is scored: An illustration with the center for epidemiologic studies depression scale. *Educational and Psychological Measurement, 63*(1), 65–74.
- George, D., & Mallery, P. (2003). *SPSS for Windows step by step: A simple guide and reference. 11.0 update* (4th ed.). Allyn & Bacon.

- Gierl, M., Rogers, W. T., & Klinger, D. A. (1999). Using statistical and judgmental reviews to identify and interpret translation differential item functioning. *Alberta Journal of Educational Research, 45*(4), 1–37.
- Gonzalez-Roma, V., Hernandez, A., & Gomez-Benito, J. (2006). Power and Type I error of the mean and covariance structure analysis model for detecting differential item functioning in graded response items. *Multivariate Behavioral Research, 41*(1), 29–53.
- Göçer-Şahin, S., Gelbal, S., & Walker, C. M. (2016, October). *Impact of decreasing category number of polytomous items on DIF*. 15th International Mineral Processing Symposium (IMPS 2016), İstanbul, Turkey.
- Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of item response theory*. Sage Publications.
- Henderson, D. L. (1999). Investigation of differential item functioning in exit Examinations across item format and subject area (Doctoral dissertation). University of Alberta, Canada.
- Henderson, D. L. (2001). *Prevalence of gender DIF in mixed format high school exit examinations*. <https://files.eric.ed.gov/fulltext/ED458284.pdf>
- Holland, P. W., & Wainer, H. (1993). *Differential item functioning*. Lawrence Erlbaum Associates, Inc.
- Kabasakal, K. A. (2014). *Değişen Madde Fonksiyonunun Test Eşitlemeye Etkisi* (Doktora tezi). Hacettepe Üniversitesi, Türkiye.
- Kaiser, H. F. (1970). A second generation little jiffy. *Psychometrika, 35*(4), 401–415.
- Karami, H., & Nodoushan, M. A. S. (2011). Differential item functioning (DIF): Current problems and future directions. *Online Submission, 5*(3), 133–142.
- Kristjansson, E., Aylesworth, R., Mcdowell, I., & Zumbo, B. D. (2005). A comparison of four methods for detecting differential item functioning in ordered response items. *Educational and Psychological Measurement, 65*(6), 935–953.
- Lee, H., & Geisinger, K. F. (2016). The matching criterion purification for differential item functioning analyses in a large-scale assessment. *Educational and Psychological Measurement, 76*(1), 141–163.

- Leino, K. (2014). The relationship between ICT use and reading literacy: Focus on 15-year-old Finnish students in PISA studies. University of Jyväskylä, Finnish Institute for Educational Research.
- Liu, I.-M., & Agresti, A. (1996). Mantel-Haenszel-type inference for cumulative odds ratios with a stratified ordinal response. *Biometrics*, 1223–1234.
- Lord, F. M. (1980). *Applications of item response theory to practical testing problems*: Routledge.
- Luu, K., & Freeman, J. G. (2011). An analysis of the relationship between information and communication technology (ICT) and scientific literacy in Canada and Australia. *Computers & Education*, 56(4), 1072–1082.
- Mantel, N. (1963). Chi-square tests with one degree of freedom; extensions of the Mantel-Haenszel procedure. *Journal of the American Statistical Association*, 58(303), 690–700.
- Maydeu-Olivares, A., Drasgow, F., & Mead, A. D. (1994). Distinguishing among parametric item response models for polychotomous ordered data. *Applied Psychological Measurement*, 18(3), 245–256.
- Meade, A. W., & Lautenschlager, G. J. (2004). A comparison of item response theory and confirmatory factor analytic methodologies for establishing measurement equivalence/invariance. *Organizational Research Methods*, 7(4), 361–388.
- MEB. (2020). *PISA Nedir?*. http://pisa.meb.gov.tr/?page_id=18
- Mellor, L. T. (1995). *A comparison of four differential item functioning (DIF) methods for polytomously scored items*. The University of Texas at Austin, USA.
- Miller, T., Chahine, S., & Childs, R. A. (2010). Detecting differential item functioning and differential step functioning due to differences that should matter. *Practical Assessment, Research, and Evaluation*, 15(10), 1–13.
- Miller, T. R., & Spray, J. A. (1993). Logistic discriminant function analysis for DIF identification of polytomously scored items. *Journal of educational measurement*, 30(2), 107–122.
- Muraki, E. (1992). A generalized partial credit model: Application of an EM algorithm. *Applied Psychological Measurement*, 16(2), 159–176.

- Narayanan, P., & Swaminathan, H. (1996). Identification of items that show nonuniform DIF. *Applied Psychological Measurement, 20*(3), 257–274.
- OECD. (2019). *PISA 2018 results (volume I): What students know and can do*, PISA, OECD Publishing, Paris.
- Osterlind, S. J. (1983). *Test item bias*. Sage Publications.
- Osterlind, S. J., & Everson, H. T. (2009). *Differential item functioning* (Vol. 161). Sage Publications.
- Özberk, E. B. Ü. (2016). *Wechsler çocuklar için zeka ölçeği IV Türkçe formundaki maddelerin işlev farklılığının sosyo-ekonomik düzey ve cinsiyet açısından incelenmesi* (Doktora tezi). Ankara Üniversitesi, Ankara.
- Penfield, R. D. (2007). Assessing differential step functioning in polytomous items using a common odds ratio estimator. *Journal of educational measurement, 44*(3), 187–210.
- Penfield, R. D. (2008). Three classes of nonparametric differential step functioning effect estimators. *Applied Psychological Measurement, 32*(6), 480–501.
- Penfield, R. D. (2010). Distinguishing between net and global DIF in polytomous items. *Journal of Educational Measurement, 47*(2), 129–149.
- Penfield, R. D. (2013). DIFAS 5.0 differential item functioning analysis system user's manual. https://soe.uncg.edu/wp-content/uploads/2015/12/DIFASManual_V5.pdf
- Penfield, R. D., & Algina, J. (2003). Applying the Liu-Agresti estimator of the cumulative common odds ratio to DIF detection in polytomous items. *Journal of Educational Measurement, 40*(4), 353–370.
- Penfield, R. D., Alvarez, K., & Lee, O. (2008). Using a taxonomy of differential step functioning to improve the interpretation of DIF in polytomous items: An illustration. *Applied Measurement in Education, 22*(1), 61–78.
- Penfield, R. D., & Lam, T. C. (2000). Assessing differential item functioning in performance assessment: Review and recommendations. *Educational Measurement: Issues and Practice, 19*(3), 5–15.

- Penfield, R. D., Myers, N. D., & Wolfe, E. W. (2008). Methods for assessing item, step, and threshold invariance in polytomous items following the partial credit model. *Educational and Psychological Measurement, 68*(5), 717–733.
- Raykov, T., & Marcoulides, G. A. (2008). *An introduction to applied multivariate analysis*: Routledge.
- Roussos, L. A., & Stout, W. F. (1996). Simulation studies of the effects of small sample size and studied item parameters on SIBTEST and Mantel-Haenszel Type I error performance. *Journal of educational measurement, 33*(2), 215–230.
- Samejima, F. (1997). Graded response model. In *Handbook of modern item response theory* (pp. 85–100). Springer.
- Shealy, R., & Stout, W. (1993). A model-based standardization approach that separates true bias/DIF from group ability differences and detects test bias/DTF as well as item bias/DIF. *Psychometrika, 58*(2), 159–194.
- Simola, H. (2005). The Finnish miracle of PISA: Historical and sociological remarks on teaching and teacher education. *Comparative education, 41*(4), 455–470.
- Sireci, S. G., & Allalouf, A. (2003). Appraising item equivalence across multiple languages and cultures. *Language testing, 20*(2), 148–166.
- Sireci, S. G., & Khaliq, S. N. (2002). *An Analysis of the Psychometric Properties of Dual Language Test Forms*. <https://files.eric.ed.gov/fulltext/ED468489.pdf>
- Smith, W. F. (2009). *Language-related DIF in the WASL mathematics test*. University of Washington.
- Somes, G. W. (1986). The generalized Mantel–Haenszel statistic. *The American Statistician, 40*(2), 106–108.
- Svetina, D., & Rutkowski, L. (2014). Detecting differential item functioning using generalized logistic regression in the context of large-scale assessments. *Large-scale Assessments in Education, 2*(1), 1–17.
- Taylor, C. S., & Lee, Y. (2012). Gender DIF in reading and mathematics tests with mixed item formats. *Applied Measurement in Education, 25*(3), 246–280.
- Tezoh, T. (2015). *Exploring the relationship between students' mathematics literacy and their access to and use of information and communication technologies*

- (ICT): *Using PISA 2012 data* (Yüksek lisans tezi). Boğaziçi Üniversitesi, Türkiye.
- Tienken, C. H. (2016). PISA is coming!. *Kappa Delta Pi Record*, 52(3), 112–115.
- Tukey, J. W. (1977). *Exploratory data analysis* (Vol. 2). Reading, MA.
- Turhan, A. (2006). *Multilevel 2PL item response model vertical equating with the presence of differential item functioning*. (Doctoral dissertation). The Florida State University, USA.
- Wang, W. C., & Su, Y. H. (2004). Factors influencing the Mantel and generalized Mantel-Haenszel methods for the assessment of differential item functioning in polytomous items. *Applied Psychological Measurement*, 28(6), 450–480.
- Welch, C., & Hoover, H. (1993). Procedures for extending item bias detection techniques to polytomously scored items. *Applied Measurement in Education*, 6(1), 1–19.
- Wood, S. W. (2011). *Differential item functioning procedures for polytomous items when examinee sample sizes are small* (Doctoral dissertation). Graduate College of The University of Iowa, USA.
- Woods, C. M. (2011). DIF testing for ordinal items with poly-SIBTEST, the Mantel and GMH tests, and IRT-LR-DIF when the latent distribution is nonnormal for both groups. *Applied Psychological Measurement*, 35(2), 145–164.
- Wyse, A. E. (2013). DIF cancellation in the Rasch model. *J Appl Meas*, 14(2), 118–28.
- Yandı, A. (2017). *Ölçme eşdeğerliğini incelemede kullanılan yöntemlerin farklı koşullar altında istatistiksel güç oranları açısından karşılaştırılması*. (Doktora tezi). Ankara Üniversitesi, Türkiye.
- Zieky, M. (1993). Practical questions in the use of DIF statistics in test development. In P. W. Holland & H. Wainer (Eds.), *Differential item functioning* (pp. 337–347). Lawrence Erlbaum Associates.
- Zumbo, B. D. (1999). *A handbook on the theory and methods of differential item functioning (DIF): Logistic regression modeling as a unitary framework for binary and likert-type (ordinal) item scores*. Ottawa, ON: Directorate of Human Resources Research and Evaluation, Department of National Defense.

- Zumbo, B. D. (2007). Three generations of DIF analyses: Considering where it has been, where it is now, and where it is going. *Language assessment quarterly*, 4(2), 223–233.
- Zumbo, B. D., & Gelin, M. N. (2005). A matter of test bias in educational policy research: bringing the context into picture by investigating sociological/community moderated (or mediated) test and item bias. *Journal of Educational Research & Policy Studies*, 5(1), 1–23.
- Zwick, R. (2012). A review of ETS differential item functioning assessment procedures: Flagging rules, minimum sample size requirements, and criterion refinement. *ETS Research Report Series*, ETS RR-12-08.
- Zwick, R., Donoghue, J. R., & Grima, A. (1993). Assessment of differential item functioning for performance tasks. *Journal of educational measurement*, 30(3), 233–251.

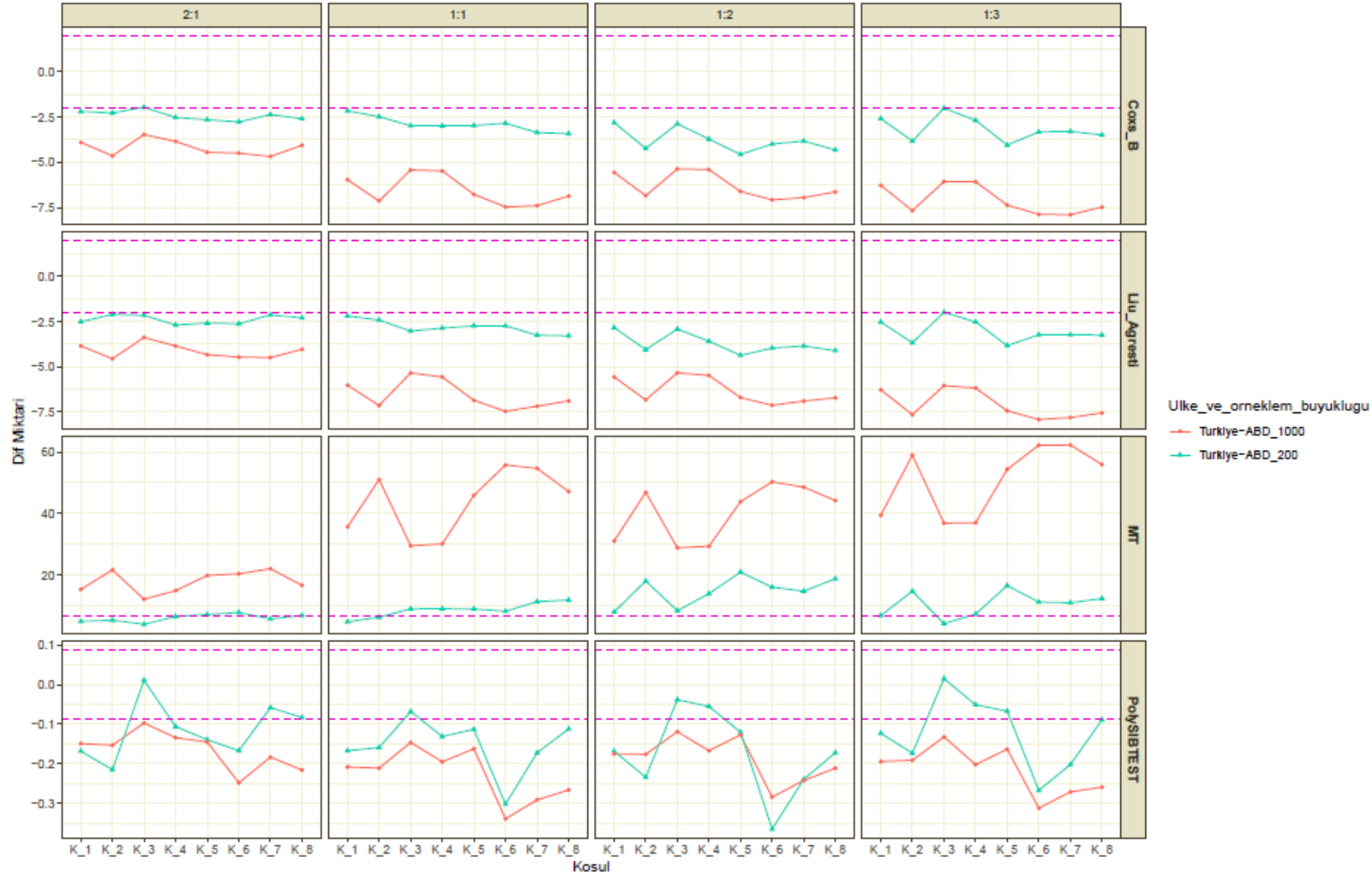
Ekler

EK-A: “Bilgi ve İletişim Teknolojilerine Aşinalık Anketi” içindeki IC011 kodlu maddeler

IC011	How often do you use digital devices for the following activities <u>at school</u>?					
		<i>Never or hardly ever</i>	<i>Once or twice a month</i>	<i>Once or twice a week</i>	<i>Almost every day</i>	<i>Every day</i>
IC011Q01TA	<Chatting online> at school.	<input type="checkbox"/> 01	<input type="checkbox"/> 02	<input type="checkbox"/> 03	<input type="checkbox"/> 04	<input type="checkbox"/> 05
IC011Q02TA	Using email at school.	<input type="checkbox"/> 01	<input type="checkbox"/> 02	<input type="checkbox"/> 03	<input type="checkbox"/> 04	<input type="checkbox"/> 05
IC011Q03TA	Browsing the Internet for schoolwork.	<input type="checkbox"/> 01	<input type="checkbox"/> 02	<input type="checkbox"/> 03	<input type="checkbox"/> 04	<input type="checkbox"/> 05
IC011Q04TA	Downloading, uploading or browsing material from the school’s website (e.g. <intranet>).	<input type="checkbox"/> 01	<input type="checkbox"/> 02	<input type="checkbox"/> 03	<input type="checkbox"/> 04	<input type="checkbox"/> 05
IC011Q05TA	Posting my work on the school’s website.	<input type="checkbox"/> 01	<input type="checkbox"/> 02	<input type="checkbox"/> 03	<input type="checkbox"/> 04	<input type="checkbox"/> 05
IC011Q06TA	Playing simulations at school.	<input type="checkbox"/> 01	<input type="checkbox"/> 02	<input type="checkbox"/> 03	<input type="checkbox"/> 04	<input type="checkbox"/> 05
IC011Q07TA	Practicing and drilling, such as for foreign language learning or mathematics.	<input type="checkbox"/> 01	<input type="checkbox"/> 02	<input type="checkbox"/> 03	<input type="checkbox"/> 04	<input type="checkbox"/> 05
IC011Q08TA	Doing homework on a school computer.	<input type="checkbox"/> 01	<input type="checkbox"/> 02	<input type="checkbox"/> 03	<input type="checkbox"/> 04	<input type="checkbox"/> 05
IC011Q09TA	Using school computers for group work and communication with other students.	<input type="checkbox"/> 01	<input type="checkbox"/> 02	<input type="checkbox"/> 03	<input type="checkbox"/> 04	<input type="checkbox"/> 05
IC011Q010HA	Using learning apps or learning websites.	<input type="checkbox"/> 01	<input type="checkbox"/> 02	<input type="checkbox"/> 03	<input type="checkbox"/> 04	<input type="checkbox"/> 05

EK-B: Maddelere İlişkin DMF Miktarlarının Örneklem Büyüklüğüne Göre Karşılaştırılması

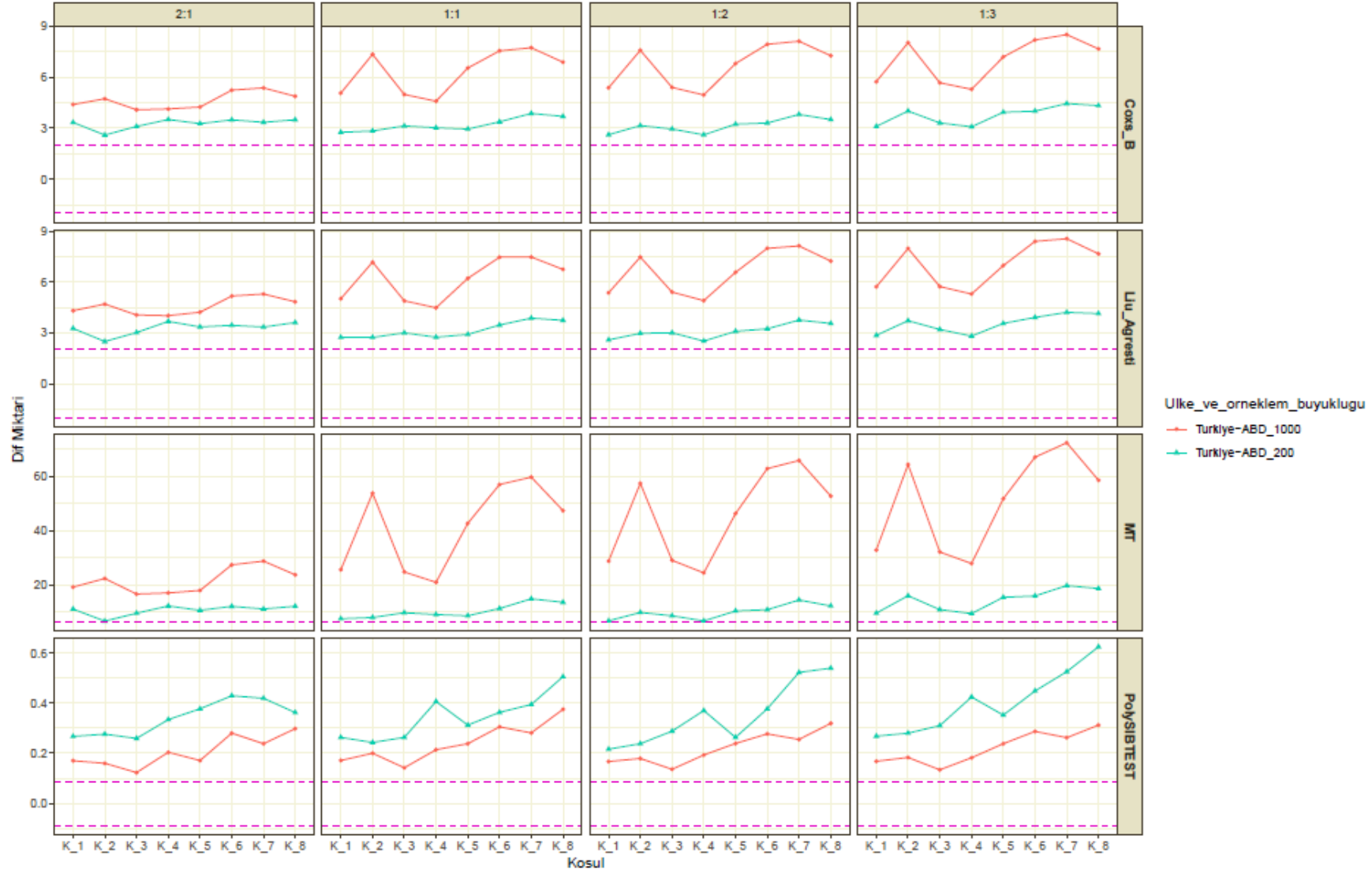
M4'e ilişkin DMF miktarlarının örneklem büyüklüğüne göre karşılaştırılması (Türkiye-ABD)



M4'e ilişkin DMF miktarlarının örneklem büyüklüğüne göre karşılaştırılması (Türkiye-Kazakistan)



M9'a ilişkin DMF miktarlarının örneklem büyüklüğüne göre karşılaştırılması (Türkiye-ABD)



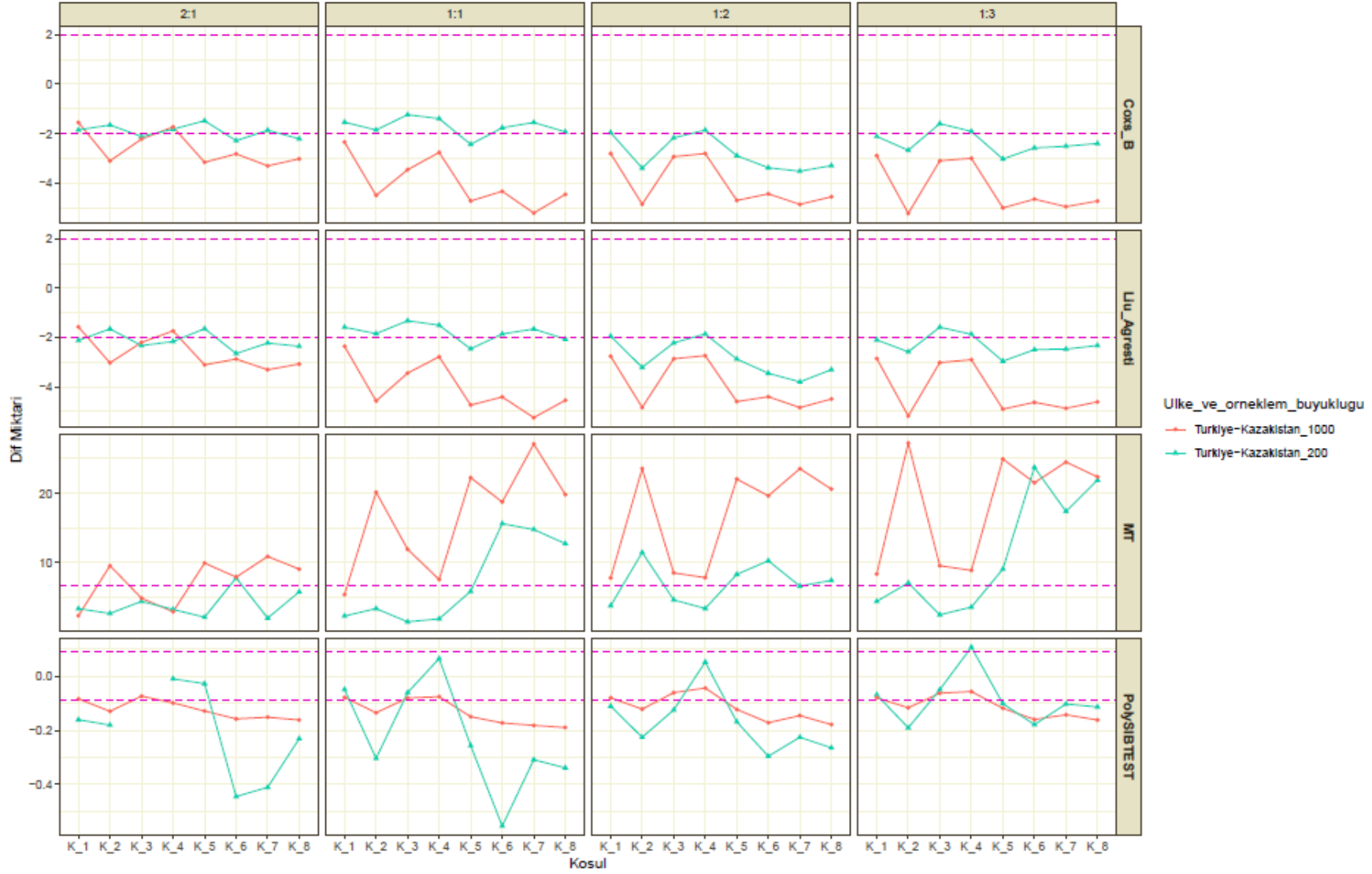
M9'a ilişkin DMF miktarlarının örneklem büyüklüğüne göre karşılaştırılması (Türkiye-Kazakistan)



M10'a ilişkin DMF miktarlarının örneklem büyüklüğüne göre karşılaştırılması (Türkiye-ABD)

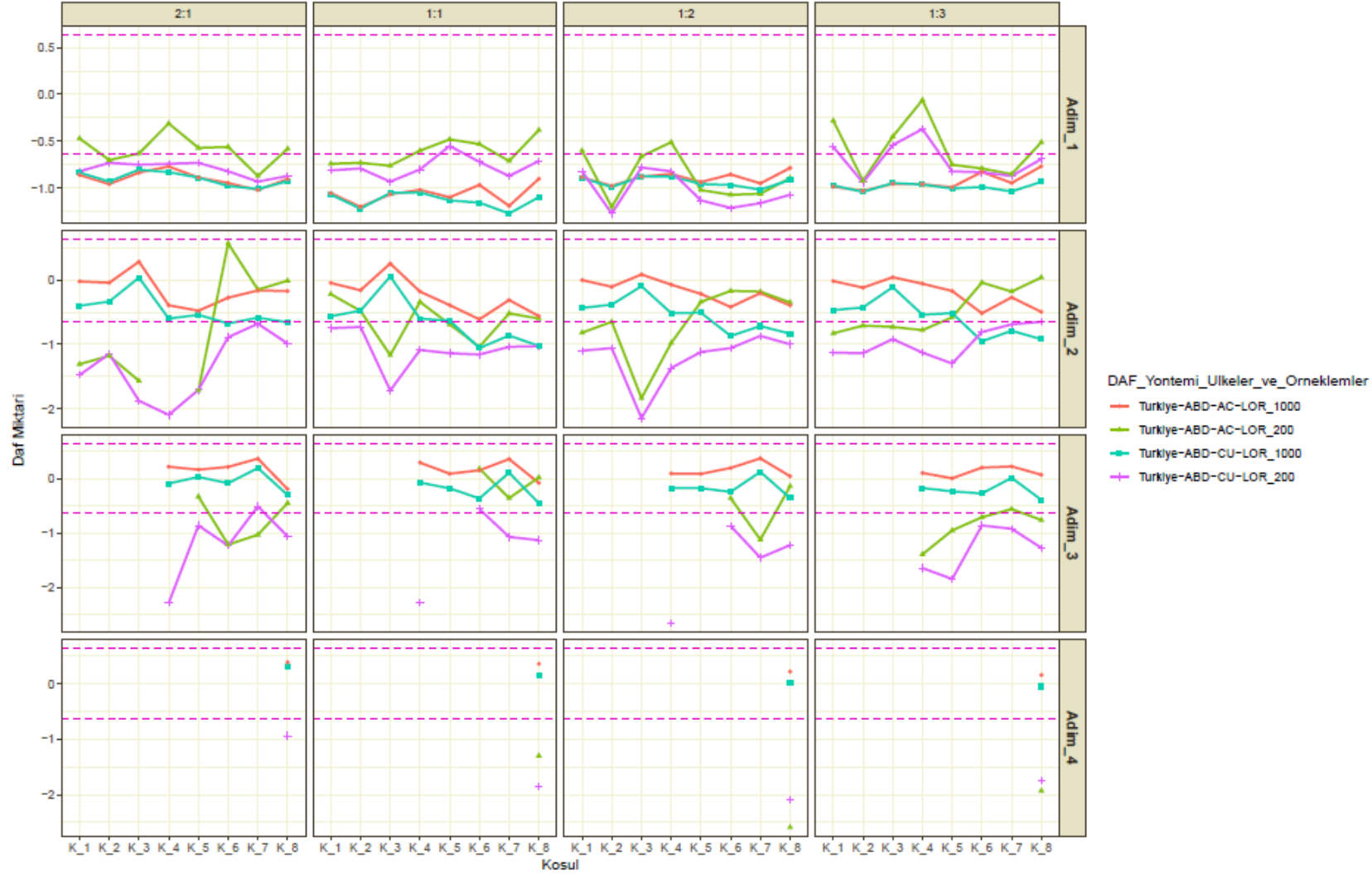


M10'a ilişkin DMF miktarlarının örneklem büyüklüğüne göre karşılaştırılması (Türkiye-Kazakistan)

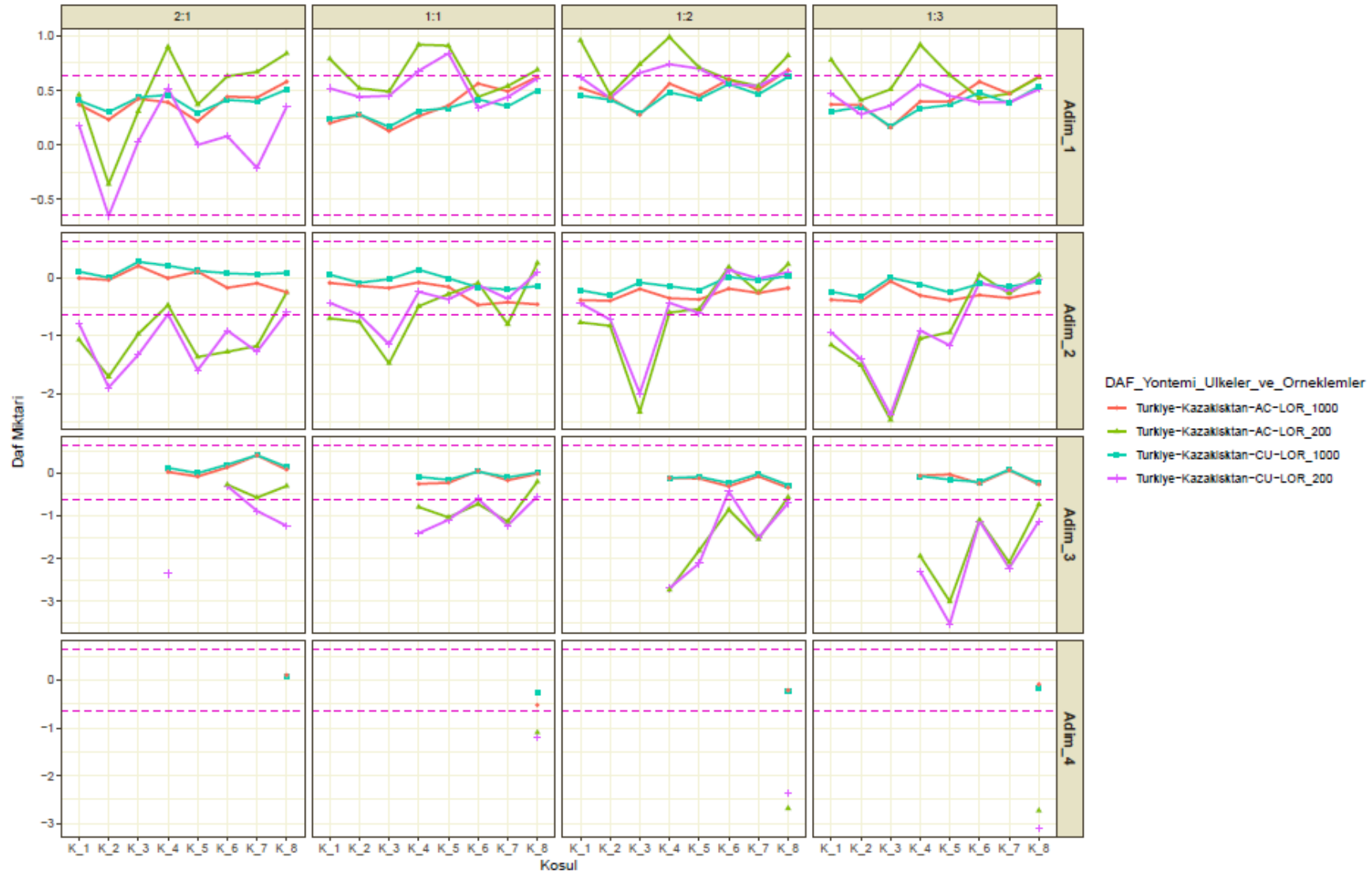


EK-C: Madde Adımlarına İlişkin DAF Miktarlarının Örneklem Büyüklüğüne Göre Karşılaştırılması

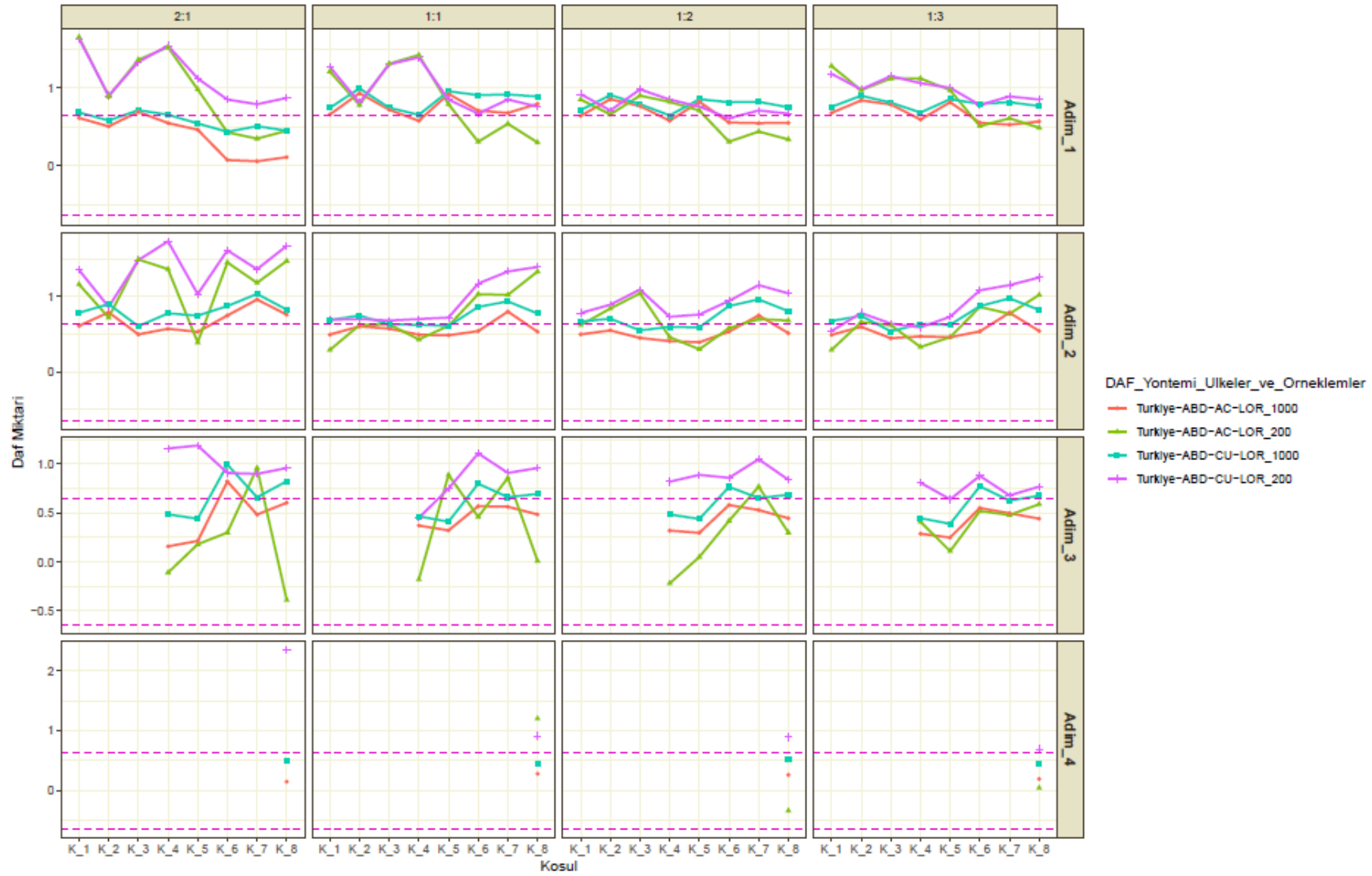
M4'e ilişkin DAF miktarlarının örneklem büyüklüğüne göre karşılaştırılması (Türkiye-ABD)



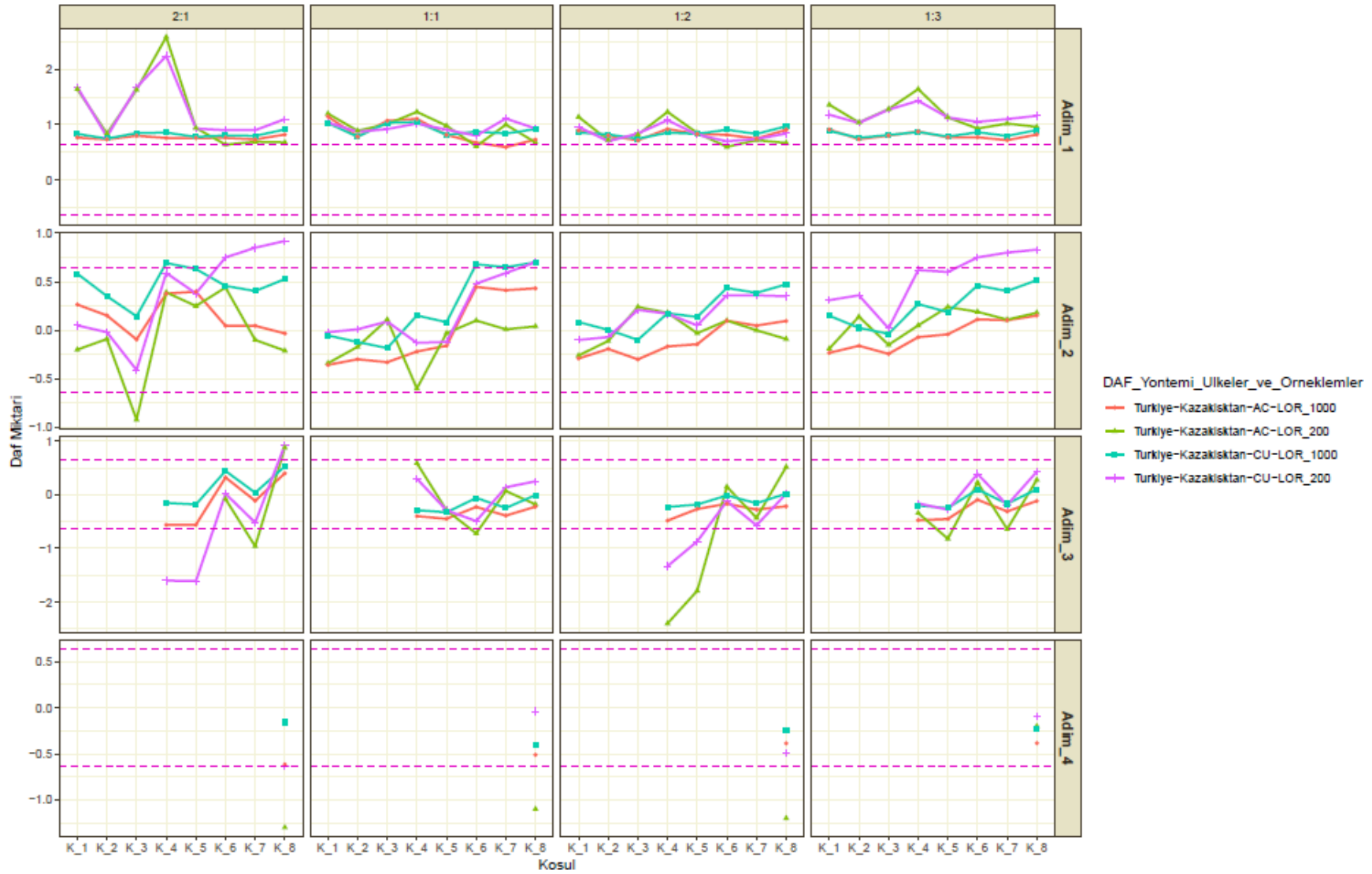
M4'e ilişkin DAF miktarlarının örneklem büyüklüğüne göre karşılaştırılması (Türkiye-Kazakistan)



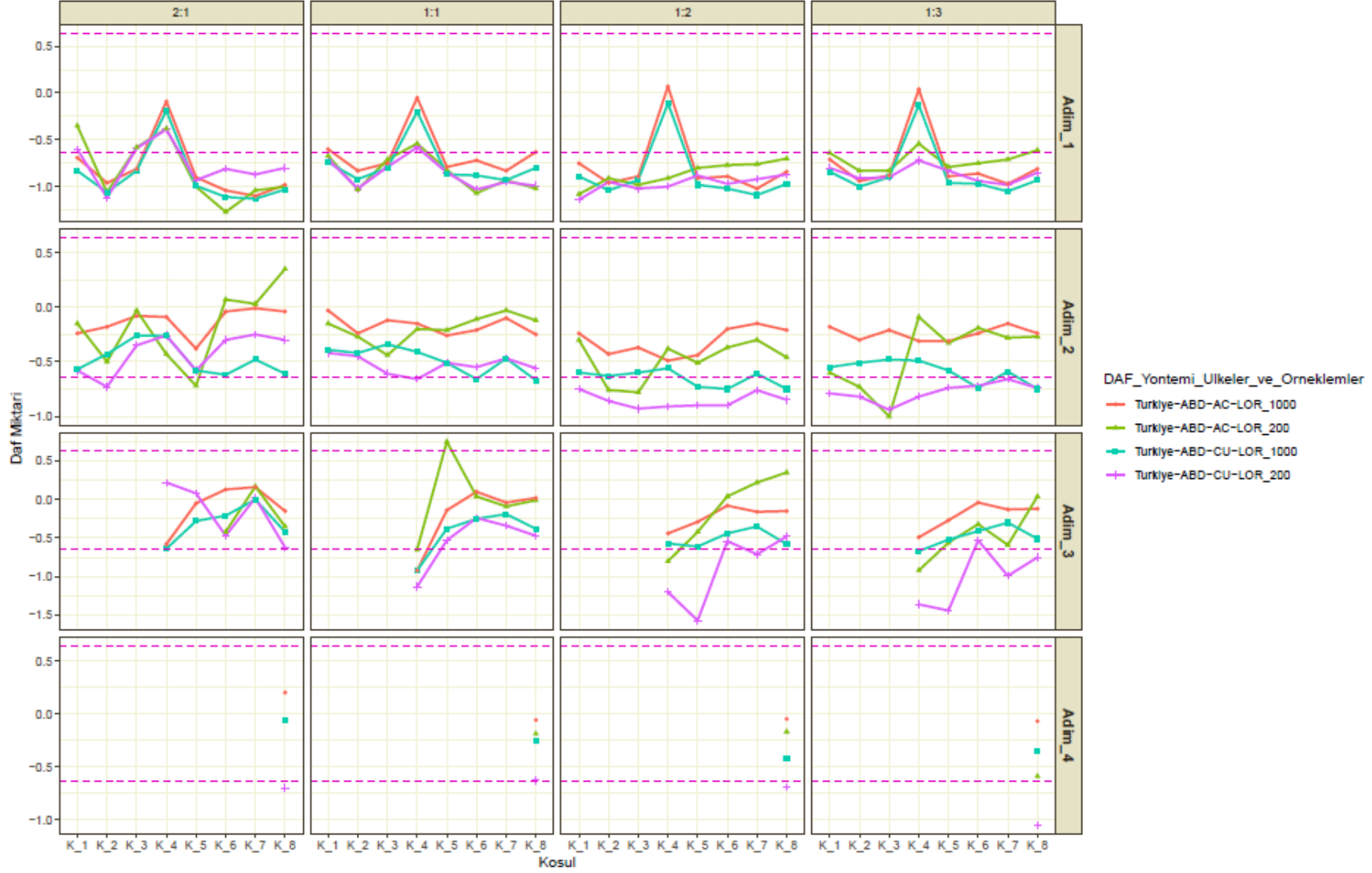
M9'a ilişkin DAF miktarlarının örneklem büyüklüğüne göre karşılaştırılması (Türkiye-ABD)



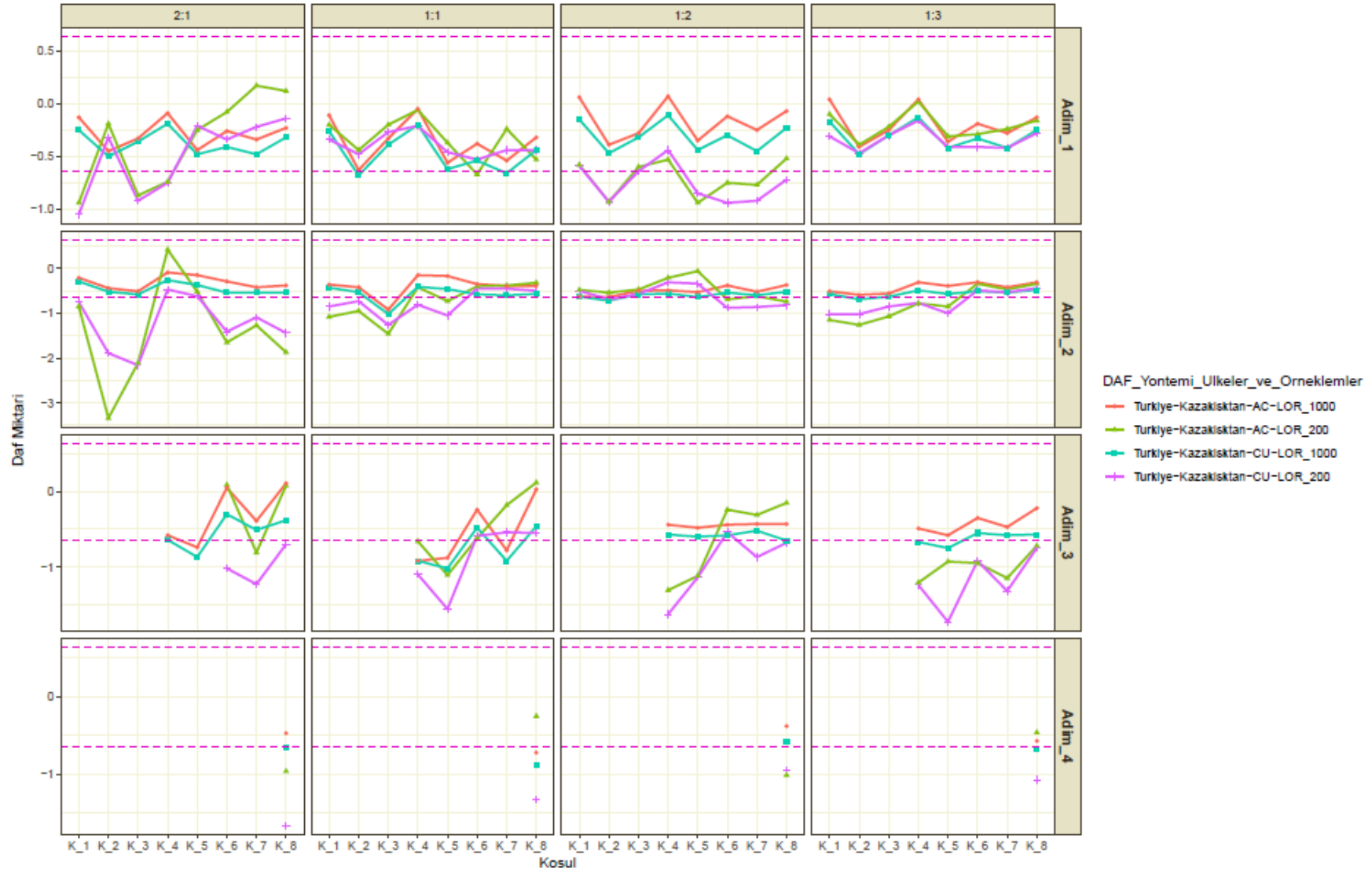
M9'a ilişkin DAF miktarlarının örneklem büyüklüğüne göre karşılaştırılması (Türkiye-Kazakistan)



M10'a ilişkin DAF miktarlarının örneklem büyüklüğüne göre karşılaştırılması (Türkiye-ABD)



M10'a ilişkin DAF miktarlarının örneklem büyüklüğüne göre karşılaştırılması (Türkiye-Kazakistan)



EK-Ç: AC-LOR ve CU-LOR Yöntemlerinin Maddeleri Sınıflamasına (Yüksek DAF'lı/DAF'sız) İlişkin Benzerlik Oranları

AC-LOR DAF Miktarı	Madde		CU-LOR DAF Miktarı								
			Örneklem büyüklüğü 200			Örneklem büyüklüğü 1000					
			Yüksek DAF	Diğer	Toplam	Benzerlik (%)	Yüksek DAF	Diğer	Toplam	Benzerlik (%)	
	M4	Türkiye-Kazakistan	Yüksek DAF	32	13	45	79,55	-	1	1	98,86
			Diğer	5	38	43	-	87	87		
			Toplam	37	51	88	-	88	88		
		Türkiye-ABD	Yüksek DAF	40	1	41	53,41	32	-	32	87,5
				Diğer	40	7	47	11	45	56	
				Toplam	80	8	88	43	45	88	
	M9	Türkiye-Kazakistan	Yüksek DAF	32	8	40	77,27	31	-	31	94,32
				Diğer	12	36	48	5	52	57	
				Toplam	44	44	88	36	52	88	
		Türkiye-ABD	Yüksek DAF	43	-	43	55,68	24	-	24	59,10
				Diğer	39	6	45	36	28	64	
				Toplam	82	6	88	60	28	88	
M10	Türkiye-Kazakistan	Yüksek DAF	36	1	37	86,36	5	-	5	88,64	
			Diğer	11	40	51	10	73	83		
			Toplam	47	41	88	15	73	88		
	Türkiye-ABD	Yüksek DAF	33	2	35	71,59	26	-	26	88,64	
			Diğer	23	30	53	10	52	62		
			Toplam	56	32	88	36	52	88		

EK-D: Etik Komisyonu Onay Bildirimi



Hacettepe Üniversitesi
Eğitim Bilimleri Enstitüsü
Tez Çalışması/Araştırma Etik Komisyon İzin Muafiyeti Formu

F46

17 / 06 / 2021

Hacettepe Üniversitesi
Eğitim Bilimleri Enstitüsü
Eğitim Bilimleri Ana Bilim Dalı Başkanlığına

Tez/Araştırma Başlığı	Çok Kategorili Puanlanan Maddelerde Değişen Madde ve Adım Fonksiyonu Belirleme Yöntemlerinin İncelenmesi
------------------------------	--

Yukarıda başlığı/konusu verilen tez/araştırma çalışmam,

1. İnsan ve hayvan üzerinde deney niteliği taşımamaktadır.
2. Biyolojik materyal (kan, idrar vb. biyolojik sıvılar ve numuneler) kullanılmasını gerektirmemektedir.
3. Beden bütünlüğüne veya ruh sağlığına müdahale içermemektedir.
4. Anket, ölçek (test), mülakat, odak grup çalışması, gözlem, deney, görüşme gibi teknikler kullanılarak katılımcılardan veri toplanmasını gerektiren nitel ya da nicel yaklaşımlarla yürütülen araştırmalar niteliğinde değildir.
5. Diğer kişi ve kurumlardan temin edilen veri kullanımını (kitap, belge vs.) gerektirmektedir. Ancak bu kullanım, diğer kişi ve kurumların izin verdiği ölçüde Kişisel Bilgilerin Korunması Kanuna riayet edilerek gerçekleştirilecektir.

Çalışmada kullanacağım veriler:

- (x) Kamusal erişime açık (buraya yazınız): PISA 2018 öğrenci anketi verisi
() Özel izin ve onaya tabi (buraya yazınız):
() Üretilmiş veri (buraya yazınız):
() Diğer (buraya yazınız):

Yükseköğretim Kurumları Etik Kurulları ve Komisyonlarının Yönergelerini inceledim ve bunlara göre çalışmamın yürütülebilmesi için herhangi bir Etik Komisyondan/Kuruldan izin alınmasına gerek olmadığını; aksi durumda doğabilecek her türlü hukuki sorumluluğu kabul ettiğimi ve yukarıda vermiş olduğum bilgilerin doğru olduğunu beyan ederim.

Gereğini saygılarımla arz ederim.

Yasemin KUZU

Araştırmacı Bilgileri

Adı Soyadı	Yasemin KUZU
Öğrenci ise No	N15145861
Ana Bilim Dalı	Eğitim Bilimleri
Programı	Eğitimde Ölçme ve Değerlendirme
Statüsü	<input type="checkbox"/> Yüksek Lisans <input checked="" type="checkbox"/> Doktora <input type="checkbox"/> Bütünleşik Dr. <input type="checkbox"/> Diğer

Danışman Görüşü ve Onayı*

Tezde 2018 PISA verisi kullanılmıştır. Veriler OECD'nin resmi internet sitesinden indirilmiştir.

Prof. Dr. Selahattin GELBAL
(İmza)

*Tez ve tezden üretilen yayınlarda gerekli

EK-E: Etik Beyanı

Hacettepe Üniversitesi Eğitim Bilimleri Enstitüsü, tez yazım kurallarına uygun olarak hazırladığım bu tez çalışmada,

- tez içindeki bütün bilgi ve belgeleri akademik kurallar çerçevesinde elde ettiğimi,
- görsel, işitsel ve yazılı bütün bilgi ve sonuçları bilimsel ahlak kurallarına uygun olarak sunduğumu,
- başkalarının eserlerinden yararlanılması durumunda ilgili eserlere bilimsel normlara uygun olarak atıfta bulunduğumu,
- atıfta bulunduğum eserlerin bütününe kaynak olarak gösterdiğimi,
- kullanılan verilerde herhangi bir tahrifat yapmadığımı,
- bu tezin herhangi bir bölümünü bu üniversitede veya başka bir üniversitede başka bir tez çalışması olarak sunmadığımı

beyan ederim.

17/06/2021

Yasemin KUZU

EK-F: Doktora Tez Çalışması Orijinallik Raporu

17/06/2021

HACETTEPE ÜNİVERSİTESİ

Eğitim Bilimleri Enstitüsü

Eğitim Bilimleri. Ana Bilim Dalı Başkanlığına,

Tez Başlığı: Çok Kategorili Puanlanan Maddelerde Değişen Madde ve Adım Fonksiyonu Belirleme Yöntemlerinin İncelenmesi

Yukarıda başlığı verilen tez çalışmamın tamamı (kapak sayfası, özetler, ana bölümler, kaynakça) aşağıdaki filtreler kullanılarak **Turnitin** adlı intihal programı aracılığı ile kontrol edilmiştir. Kontrol sonucunda aşağıdaki veriler elde edilmiştir:

Rapor Tarihi	Sayfa Sayısı	Karakter Sayısı	Savunma Tarihi	Benzerlik Oranı	Gönderim Numarası
15/05/2021	103	166653	28/05/2021	%2	1586554486

Uygulanan filtreler:

- Kaynaklar hariç
- Alıntılar dâhil
- 5 kelimedenden daha az örtüşme içeren metin kısımları hariç

Hacettepe Üniversitesi Eğitim Bilimleri Enstitüsü Tez Çalışması Orijinallik Raporu Alınması ve Kullanılması Uygulama Esasları'nı inceledim ve çalışmamın herhangi bir intihal içermediğini; aksinin tespit edileceği muhtemel durumda doğabilecek her türlü hukuki sorumluluğu kabul ettiğimi ve yukarıda vermiş olduğum bilgilerin doğru olduğunu beyan eder, gereğini saygılarımla arz ederim.

Ad Soyadı: Yasemin KUZU

Öğrenci No.: N15145861

Ana Bilim Dalı: Eğitim Bilimleri

Programı: Eğitimde Ölçme ve Değerlendirme

Statüsü: Y.Lisans Doktora Bütünleşik Dr.

İmza

DANIŞMAN ONAYI

UYGUNDUR.

Prof. Dr. Selahattin GELBAL

EK-G: Dissertation Originality Report

17/06/2021

HACETTEPE UNIVERSITY

Graduate School of Educational Sciences

To The Department of Educational Sciences

Thesis Title: Investigation of Differential Item and Differential Step Functioning Procedures in Polytomous Items

The whole thesis that includes the *title page, introduction, main chapters, conclusions and bibliography section* is checked by using **Turnitin** plagiarism detection software take into the consideration requested filtering options. According to the originality report obtained data are as below.

Time Submitted	Page Count	Character Count	Date of Thesis Defense	Similarity Index	Submission ID
15/05/2021	103	166653	28/05/2021	2%	1586554486

Filtering options applied:

1. Bibliography excluded
2. Quotes included
3. Match size up to 5 words excluded

I declare that I have carefully read Hacettepe University Graduate School of Educational Sciences Guidelines for Obtaining and Using Thesis Originality Reports; that according to the maximum similarity index values specified in the Guidelines, my thesis does not include any form of plagiarism; that in any future detection of possible infringement of the regulations I accept all legal responsibility; and that all the information I have provided is correct to the best of my knowledge.

I respectfully submit this for approval.

Name Lastname: Yasemin KUZU
Student No.: N15145861
Department: Educational Sciences
Program: Educational Measurement and Evaluation
Status: Masters Ph.D. Integrated Ph.D.

Signature

ADVISOR APPROVAL

APPROVED

Prof. Selahattin GELBAL

EK-H: Yayımlama ve Fikrî Mülkiyet Hakları Beyanı

Enstitü tarafından onaylanan lisansüstü tezimin/raporumun tamamını veya herhangi bir kısmını, basılı (kâğıt) ve elektronik formatta arşivleme ve aşağıda verilen koşullarla kullanıma açma iznini Hacettepe Üniversitesine verdiğimi bildiririm. Bu izinle Üniversiteye verilen kullanım hakları dışındaki tüm fikri mülkiyet haklarım bende kalacak, tezimin tamamının ya da bir bölümünün gelecekteki çalışmalarda (makale, kitap, lisans ve patent vb.) kullanım hakları bana ait olacaktır.

Tezin kendi orijinal çalışmam olduğunu, başkalarının haklarını ihlal etmediğimi ve tezimin tek yetkili sahibi olduğumu beyan ve taahhüt ederim. Tezimde yer alan telif hakkı bulunan ve sahiplerinden yazılı izin alınarak kullanılması zorunlu metinlerin yazılı izin alınarak kullandığımı ve istenildiğinde suretlerini Üniversiteye teslim etmeyi taahhüt ederim.

Yükseköğretim Kurulu tarafından yayınlanan "**Lisansüstü Tezlerin Elektronik Ortamda Toplanması, Düzenlenmesi ve Erişime Açılmasına İlişkin Yönerge**" kapsamında tezim aşağıda belirtilen koşullar haricince YÖK Ulusal Tez Merkezi / H.Ü. Kütüphaneleri Açık Erişim Sisteminde erişime açılır.

- o Enstitü/Fakülte yönetim kurulu kararı ile tezimin erişime açılması mezuniyet tarihinden itibaren 2 yıl ertelenmiştir. ⁽¹⁾
- o Enstitü/Fakülte yönetim kurulunun gerekçeli kararı ile tezimin erişime açılması mezuniyet tarihimden itibaren ... ay ertelenmiştir. ⁽²⁾
- o Tezimle ilgili gizlilik kararı verilmiştir. ⁽³⁾

17/06/2021

Yasemin KUZU

"*Lisansüstü Tezlerin Elektronik Ortamda Toplanması, Düzenlenmesi ve Erişime Açılmasına İlişkin Yönerge*"

(1) Madde 6. 1. Lisansüstü teze ilgili patent başvurusu yapılması veya patent alma sürecinin devam etmesi durumunda, tez danışmanının önerisi ve enstitü anabilim dalının uygun görüşü üzerine enstitü veya fakülte yönetim kurulu iki yıl süre ile tezin erişime açılmasının ertelenmesine karar verebilir.

(2) Madde 6. 2. Yeni teknik, materyal ve metotların kullanıldığı, henüz makaleye dönüşmemiş veya patent gibi yöntemlerle korunmamış ve internette paylaşılması durumunda 3. şahıslara veya kurumlara haksız kazanç; imkânı oluşturabilecek bilgi ve bulguları içeren tezler hakkında tez danışmanının önerisi ve enstitü anabilim dalının uygun görüşü üzerine enstitü veya fakülte yönetim kurulunun gerekçeli kararı ile altı ayı aşmamak üzere tezin erişime açılması engellenebilir.

(3) Madde 7. 1. Ulusal çıkarları veya güvenliği ilgilendiren, emniyet, istihbarat, savunma ve güvenlik, sağlık vb. konulara ilişkin lisansüstü tezlerle ilgili gizlilik kararı, tezin yapıldığı kurum tarafından verilir. Kurum ve kuruluşlarla yapılan işbirliği protokolü çerçevesinde hazırlanan lisansüstü tezlere ilişkin gizlilik kararı ise, ilgili kurum ve kuruluşun önerisi ile enstitü veya fakültenin uygun görüşü üzerine üniversite yönetim kurulu tarafından verilir. Gizlilik kararı verilen tezler Yükseköğretim Kuruluna bildirilir.

Madde 7.2. Gizlilik kararı verilen tezler gizlilik süresince enstitü veya fakülte tarafından gizlilik kuralları çerçevesinde muhafaza edilir, gizlilik kararının kaldırılması halinde Tez Otomasyon Sistemine yüklenir

* Tez danışmanının önerisi ve enstitü anabilim dalının uygun görüşü üzerine enstitü veya fakülte yönetim kurulu tarafından karar verilir.