

Çok Düzeyli Yapısal Eşitlik Modelleri Üzerine Örnek Bir Uygulama

Seda CAN*

İzmir Ekonomi Üniversitesi

Oya SOMER**, Mediha KORKMAZ*** ve Seda DURAL****

Ege Üniversitesi

Özet

Sosyal bilimlerde birçok araştırma probleminin test edilebilmesi amacıyla toplanan veri setleri sıklıkla çok düzeyli bir yapıya sahiptir. Tek düzeyli analizlerin sınırlılıkları nedeniyle, uzun yıllar yapılan görgül çalışmalarda gözlemlerin bağımsızlığı varsayımının ihlalini göz ardı edilmiş, modellerde analiz birimleri olarak ya bireyler ya da gruplar kullanılmıştır. Hiyerarşik/kümelenmiş verilerle çeşitli alanlarda yürütülen araştırmalardaki bu tür sorunlar, çok düzeyli modelleme tekniklerinin gelişmesine yol açmıştır. Çok Düzeyli Yapısal Eşitlik Modellemeleri ile hiyerarşik verinin her bir düzeyi (grup-içi ve gruplar arası) için eş zamanlı test edilebilecek modellerin geliştirilebilmesine olanak sağlanmaktadır. Bu çalışmada, analiz Mplus 5.1 (Muthén ve Muthén, 2008) programının Monte Carlo simülasyonu kapsamındaki veri üretimi (*data generation*) özelliği kullanılarak elde edilen veriler üzerinden iki düzeyli yapısal bir modelin incelenmesi ve araştırmacılara örnek bir uygulama sunulması amaçlanmıştır. Bu çerçevede, Mplus programı kullanılarak yapılan analiz üzerinden modelin tanıtımı yapılmış, Mplus sentaksı açıklanmış ve bulguların yorumlanması üzerinde durulmuştur.

Anahtar sözcükler: Yapısal eşitlik modellemeleri, çok düzeyli yapısal eşitlik modelleri

Abstract

The data collected in order to test the research questions in social sciences usually have multilevel data structures. Because of the limitations of single level analysis, the violation of independence of observations assumption in empirical studies was ignored; either groups or individuals were used as units in the analysis for a long period of time. Such kinds of problems in the studies with hierarchical/clustered data brought about the development of multilevel modeling techniques. Multilevel Structural Equation Modeling allows developing different models which can be tested for each level of the hierarchical data (within group and between groups) simultaneously. In current study, it was aimed to investigate two-level structural model by using data generated from the Monte Carlo data generation facility of Mplus program and to present a demonstration for researchers. Within this context, the model was introduced using the analysis performed by Mplus 5.1, the Mplus syntax used in the analysis was explained and the interpretation of the model parameters was discussed.

Keywords: Structural equation modeling, multilevel structural equation models

Sosyal bilimlerde birçok araştırma probleminin test edilebilmesi amacıyla toplanan veri setleri sıklıkla çok düzeyli bir yapıya sahiptir. Sınıflardaki öğrenciler, mahallelerdeki bireyler, hastanelerdeki hastalar, şirketlerdeki çalışanlar, ailelerdeki kardeşlerden toplanan veriler hiyerarşik yapıda ya da kümelenmiş verilere örnek oluşturmaktadır. Kümelenmiş veriler ayrıca bazı araştırma desenlerinin sonucu olarak da elde edilebilmektedir. Örneğin, araştırmaya katılacak bireyler tabakalama

* İzmir Ekonomi Üniversitesi, Ölçme ve Değerlendirme Uzmanı, seda.canet@ieu.edu.tr

** Prof. Dr., Ege Üniversitesi, Psikoloji Bölümü, oya.somer@ege.edu.tr

*** Yrd. Doç. Dr., Ege Üniversitesi, Psikoloji Bölümü, mediha.korkmaz@ege.edu.tr

**** Arş. Gör., Ege Üniversitesi, Psikoloji Bölümü, dseda@mu.edu.tr

örneklemeye seçilmiş olabilir (belli şehir ve coğrafi bölgelerdeki mahallelerden seçilen bireyler gibi). Ayrıca tekrarlı ölçümlerin, araştırmaya katılan bireylere yuvalandığı (*nest*) boylamsal çalışmalardan elde edilen veriler de araştırma deseni sonucu elde edilen hiyerarşik verilere örnek oluşturmaktadır (Hox, 1995).

Tek düzeyli analizlerin sınırlılıkları sebebiyle, uzun yıllar yapılan görgül çalışmalarda, ya bireyler ya da gruplar analiz birimleri olarak modellerde yer almışlardır. Araştırmacılar ölçülen değişkenlerin doğasına ilişkin yaptıkları gözlemlerin bağımsızlığı varsayımının ihlalini göz ardı ederek tek düzeyli bir analiz yapabilmek için değişkenleri bir düzeyden diğerine taşımışlardır.

Hiyerarşik/kümelenmiş verilerle çeşitli alanlarda yürütülen araştırmalardaki bu tür sorunlar, çok düzeyli modelleme tekniklerinin gelişmesine yol açmıştır (Hox, 1995). Hiyerarşik yapıdaki verinin analiz edilmesine işaret eden yöntemler literatürde çeşitli isimlerle kullanılmaktadır; çok düzeyli regresyon modelleri (*multilevel regression models*), hiyerarşik lineer modeller (*hierarchical linear models*), karmaşık ve seçkisiz etkiler modelleri (*mixed and random effects models*), seçkisiz katsayılar modelleri (*random coefficients model*) ve çok düzeyli kovaryans yapıları modelleri (*multilevel covariance structure models*) (Heck, 2001).

Çok düzeyli modeller araştırmacılara veri hiyerarşisinin her bir düzeyindeki değişkenler arasındaki ilişkileri araştırmak için matematiksel bir modelleme sunmaktadır. Çok düzeyli modelleme, düzeylerin eş zamanlı olarak ele alınmasına imkan vermesi bakımından, sadece birey ya da sadece grup düzeyinde analiz etme probleminin ortadan kalkmasını sağlamaktadır (Heck, 2001).

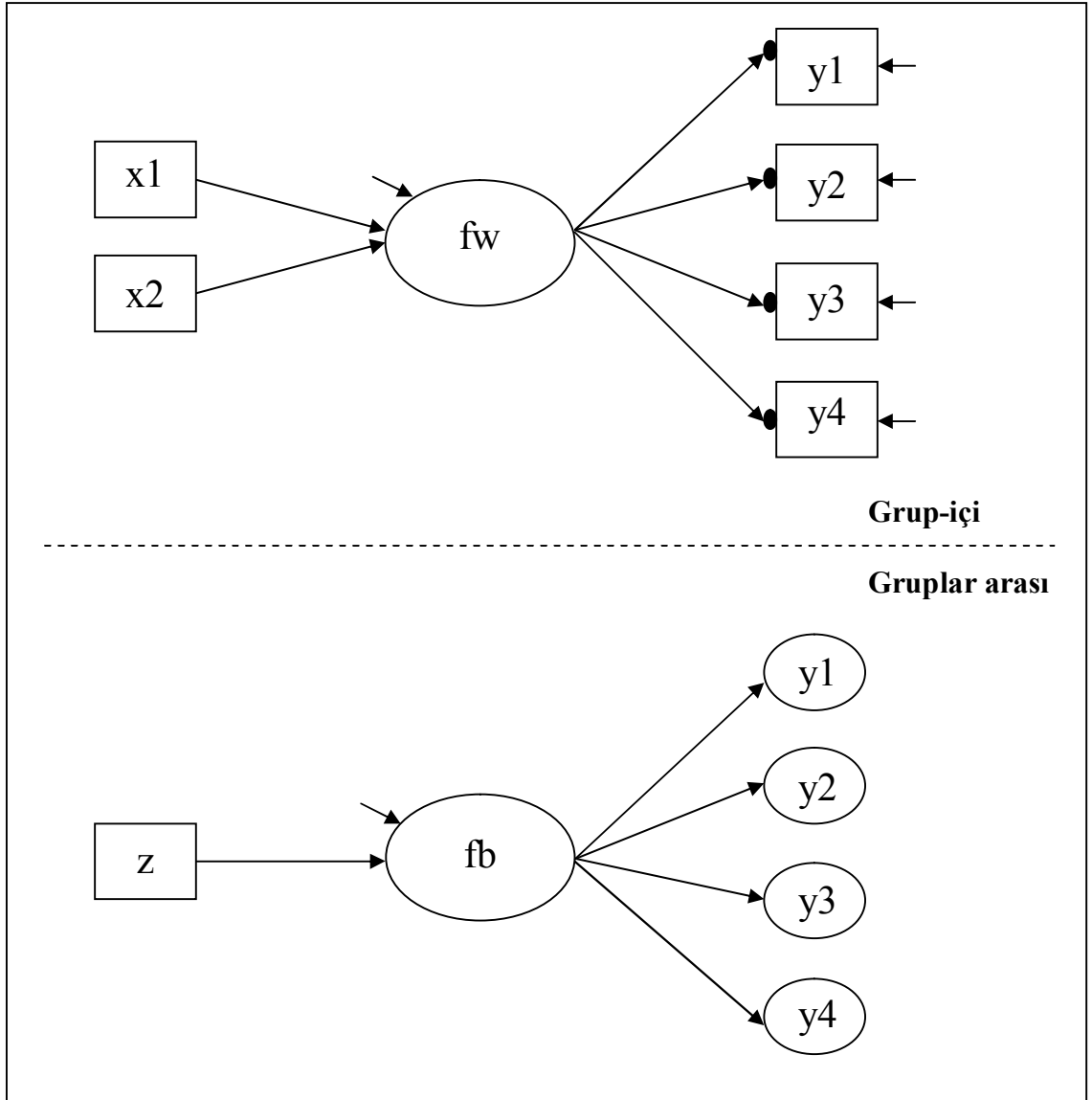
Çok düzeyli modelleme ve Yapısal Eşitlik Modellemesi - YEM (*Structural Equation Modeling - SEM*) farklı kavramsal ve yöntemsel kaynaklardan gelişmiştir (Mehta ve Neale, 2005). Çok düzeyli modellemede gözlenen varyans gruplar arası ve grup-içi bileşenlerine ayrılmaktadır. YEM ise verinin ortalama ve kovaryans yapılarını modellemek üzere geliştirilmiştir. Sosyal bilimlerde hipotezlerin test edilmesinde hem hiyerarşik hem de çok değişkenli verilerin birlikte kullanılması sıklıkla görülen bir durum olduğundan, her iki modellemenin birleştirilmesine ihtiyaç duyulmuştur (Bauer, 2003; Bentler ve Liang, 2003; Kaplan ve Elliot, 1997; Muthén, 1991, 1994). Böylelikle, her iki yaklaşımı birleştiren ve hiyerarşik verinin her bir düzeyi (birey ve grup düzeyi) için eş zamanlı test edilebilecek yapısal eşitlik modellerinin geliştirilebilmesine olanak sağlayan Çok Düzeyli Yapısal Eşitlik Modellemeleri (*Multilevel Structural Equation Modeling*) çalışmaya başlanmıştır.

Hiyerarşik verilerin yapısal eşitlik modelleriyle analizinde; (1) gözlenen değişkenler yoluyla örtük değişkenlerin tanımlanmasına odaklanan iki-düzeyli ölçme modelleri, (2) gözlenen değişkenler arası iki düzeyli ilişkileri araştıran yol (*path*) analizleri ve (3) örtük ve gözlenen değişkenler arası ilişkilere odaklanan iki düzeyli yapısal (*structural*) modeller araştırılabilmektedir (Heck, 2001).

Bu çalışmada, simüle edilen veriler üzerinden iki düzeyli yapısal bir modelin incelenmesi ve araştırmacılara örnek bir uygulama sunulması amaçlanmıştır. Bu amaçla, analizin her bir düzeyinde farklı modellerin test edilebilmesine örnek oluşturacak şekilde, grup-içi düzeyde iki ve gruplar arası düzeyde bir yordayıcı değişkenin yer aldığı modele (Şekil 1) uygun veri seti üretilmiş, ilgili model sentaksı açıklanmış ve elde edilen bulgular yorumlanmıştır.

Yöntem

Çalışmada yer alan analiz Mplus 5.1 (Muthén ve Muthén, 2008) programının Monte Carlo simülasyonu kapsamındaki veri üretimi (*data generation*) özelliği kullanılarak elde edilen veriler üzerinden yapılmıştır. Çalışmanın amacı doğrultusunda, Şekil 1’de sunulan çok düzeyli YEM’e uygun her birinde 15 kişinin yer aldığı 40 gruptan oluşan, örneklem büyüklüğü 600 olan normal dağılıma sahip bir veri seti üretilmiştir.



Şekil 1. Çok Düzeyli YEM.

Analizler

Çalışmanın bu aşamasında 15'er kişilik 40 grubun yer aldığı 600 gözlemlik üretilen veri setine Mplus programı kullanılarak çok düzeyli YEM analizi yapılmıştır.

Test edilen modelin daha anlaşılır olması bakımından Şekil 1'deki model bir örnek üzerinden düşünüldüğünde;

- y1 – y4 akademik başarı örtük değişkeninin gözlenen değişkenleri
- fw grup-İçi düzeyde, fb gruplar arası düzeyde akademik başarı örtük değişkenleri
- x1 yordayıcı değişkeni, grup-İçi düzeyde akademik başarı örtük değişkenindeki varyasyonu açıklayan “öğrencinin ailesinin aylık geliri”

- x2 yordayıcı değişkeni, grup-içi düzeyde akademik başarı örtük değişkenindeki varyasyonu açıklayan “öğrencinin kardeş sayısı”
- z yordayıcı değişkeni, gruplar arası düzeyde akademik başarı örtük değişkenindeki varyasyonu açıklayan “sınıf öğretmeninin deneyimi”

olarak tanımlanabilir. Söz konusu modele ilişkin Mplus sentaksı aşağıda verilmiştir:

```

TITLE:      Çok düzeyli YEM için Mplus sentaksı
DATA:      FILE IS sim6.dat;
VARIABLE:  NAMES ARE y1 - y4 x1 x2 z clus;
           CLUSTER = clus;
           WITHIN = x1 x2;
           BETWEEN = z;
ANALYSIS:  TYPE = TWOLEVEL;
MODEL:     %WITHIN%
           fw BY y1-y4;
           fw ON x1 x2;
           %BETWEEN%
           fb BY y1-y4;
           fb ON z ;
OUTPUT:    STANDARDIZED;

```

Program komutları incelendiğinde, “VARIABLE” başlığı altında modelde yer alan gözlenen değişkenlerin isimleri yazılmıştır. Ayrıca bu değişkenlerin sonuna, hiyerarşik veri setindeki küme/grubu tanımlamak üzere bir değişken eklenmiş ve eklenen bu değişken yine “VARIABLE” başlığı altında “CLUSTER” komutunda tanımlanmıştır. Son olarak yine bu başlık altında birey düzeyindeki (x1, x2) ve grup düzeyindeki (z) gözlenen yordayıcı değişkenler tanımlanmaktadır.

“ANALYSIS” başlığı altında ise gözlemlerin bağımsız olmama durumunu verinin her bir düzeyini tanımlayarak modelleyen ve iki düzeyli modelleme olduğu durumda kullanılan analiz tipi “TWOLEVEL” olarak tanımlanmaktadır.

“MODEL” başlığı altında analiz için iki düzeyine ilişkin tanımlamalar ayrı ayrı yapılmaktadır. Birey düzeyindeki tanımlamalar “%WITHIN%”, grup düzeyindeki tanımlamalar ise “%BETWEEN%” alt başlıkları altında verilmektedir. Her iki düzeyde de “BY” komutu kullanılarak y1-y4 gözlenen değişkenleri fw, ve fb örtük değişkenlerine bağlanmıştır. fw ve fb’nin metriği, program tarafından otomatik olarak “BY” komutundan sonra gelen ilk gözlenen değişkenin faktör yükü 1’e sabitlenerek sağlanmaktadır.

“MODEL” başlığı altında her iki düzeyde de en alt satırda yer alan “ON” komutuyla değişkenler arasındaki regresyon katsayılarının tahminlenmesi sağlanır. Bu örnekte grup- içi düzey için örtük değişkenin (fw) yordayıcı değişkenler üzerindeki regresyon katsayısını elde edebilmek için “ON” komutundan sonra bu düzeyin yordayıcı değişkenleri olan x1 ve x2, gruplar arası düzey için örtük değişkenin (fb) yordayıcı değişkenler üzerindeki regresyon katsayısını elde edebilmek için grup düzeyinin yordayıcı değişkeni olan z “ON” komutundan sonra yazılmıştır.

Bulgular

Gözlenen dört değişken için hesaplanan grup içi (*intra*class) korelasyon katsayıları 0.20 ile 0.22 arasında değişmektedir ve medyan değeri 0.21’dir. Grup içi korelasyon katsayısının 0.05’den yüksek olması gözlemlerin bağımsızlığı varsayımının ihlaline ve buna bağlı olarak da çok düzeyli analiz gereksinim olduğuna işaret etmektedir (Muthén ve Satorra, 1995).

Tanımlanan çok düzeyli yapısal model için yapılan analiz sonucunda elde edilen uyum indekslerinden bazıları; $\chi^2 (13) = 25.3$, $p = .02$, $NC = 1.95$, $CFI = .99$, $RMSEA = 0.04$ olarak bulunmuştur ve bu değerler modelin iyi bir uyuma sahip olduğunu göstermektedir. Ayrıca hem grup- içi hem de gruplar arası düzeyler için ayrı ayrı gösterilen SRMR (grup- içi = .016, gruplar arası = .013) indeksi de model veri uyumunun gruplar arası düzeyde biraz daha iyi olmakla birlikte her iki düzeyde de oldukça iyi olduğuna işaret etmektedir.

Test edilen modelin standardize olmayan parametre tahminleri Ek 1’de yer almaktadır. Grup- içi düzeyindeki standardize faktör yükleri .80 ile .83 arasında değişmektedir. Gruplar arası düzeyindeki standardize faktör yükleri ise 1’e çok yakındır. Tahminlenen faktör yükleri grup düzeyinde daha yüksek olmakla birlikte her iki düzeyde de Ek 1’in son sütununda görülebileceği gibi istatistiksel olarak anlamlıdır ($p < .001$).

Grup- içi düzeydeki yordayıcıların regresyon katsayılarına ilişkin değerler incelendiğinde ise x1 değişkenine ilişkin standardize katsayı .73 ($p < .001$), x2 değişkenine ilişkin standardize katsayı ise .005 ($p > .05$) olarak bulunmuştur. Grup düzeyindeki z yordayıcı değişkeni için regresyon katsayısı ise .67 ($p < .01$) olarak bulunmuştur. Bu sonuçlara göre; öğrencilerin ailelerinin aylık gelir miktarı arttıkça öğrencilerin akademik başarıları artmakta, sınıf öğretmeninin deneyimi arttıkça sınıfların akademik başarıları artış göstermektedir. x1 yordayıcı değişkeni grup- içi düzeyde akademik başarı örtük değişkenindeki varyansın %53’ünü, z yordayıcı değişkeni ise gruplar arası düzeyde varyansın %43’ünü açıklamaktadır.

Tartışma

Bu çalışmada, çok düzeyli YEM kapsamında Mplus 5.1 (Muthén ve Muthén, 2008) programının Monte Carlo simülasyonu kapsamındaki veri üretimi (*data generation*) özelliği kullanılarak elde edilen veriler üzerinden analiz yapılmış ve araştırmacılara örnek bir uygulama sunulmuştur. Bu amaçla, analizin her bir düzeyinde farklı modellerin test edilebilmesine örnek oluşturacak şekilde, grup- içi ve gruplar arası düzeyde farklı değişkenlerin yer aldığı modele uygun veri seti üretilmiş, ilgili model sentaksı açıklanmış ve elde edilen bulguların yorumlanması üzerinde durulmuştur.

Söz konusu modeller, bulgularda da ifade edildiği gibi hem grup- içi hem de gruplar arası düzeyde parametrelerin tahminlenmesine dolayısıyla sadece birey ya da grup düzeyinde değil hem birey hem de grup düzeyinde bulguların yorumlanmasına olanak sağlamaktadır.

Çok düzeyli YEM için genel istatistiksel model karmaşık ve zor olarak görünse de özellikle son yıllarda bu modellerin test edilmesinde kullanılabilecek bilgisayar programlarının da gelişmesiyle yapısal eşitlik modellemeleri kapsamında çok düzeyli modeller sıklıkla çalışılmaya başlanmıştır (Julian, 2001). Yapısal eşitlik modellemelerinin analizlerinde yaklaşık 30 yıl öncesine kadar LISREL kullanılan tek program iken son yıllarda farklı birçok program, modellerin test edilmesinde yaygın bir şekilde kullanılmaya başlanmıştır. Bu programların arasında son yıllarda kullanımı giderek yaygınlaşan Mplus programı, özellikle çok düzeyli modellerin analizinde çok daha anlaşılabilir ve sade bir program diline sahip olması bakımından söz konusu programın bu alanda çalışan araştırmacılar açısından daha pratik olacağı düşünülmektedir.

Kaynaklar

- Bauer, D. J. (2003). Estimating multilevel linear models as structural equation models. *Journal of Educational and Behavioral Statistics*, 28, 135-167.
- Bentler, P. M., & Liang, J. (2003). Two-level mean and covariance structures: Maximum likelihood via and EM algorithm. In S. P. Reise & N. Duan (Eds.), *Multilevel modeling: Methodological advances, issues, and applications* (pp. 53-70). Hillsdale, NJ: Erlbaum.
- Heck, R. H. (2001). Multilevel modeling with SEM. In J. A. Marcoulides & R. E. Schumacker (Eds.), *New developments and techniques in structural equation modeling* (pp. 89-127). Lawrence Erlbaum Associates.
- Hox, J. J. (1995). *Applied multilevel analysis*. Amsterdam: TT-Publikaties.

-
- Julian, M. W. (2001). The consequences of ignoring multilevel data structures in nonhierarchical covariance modeling. *Structural Equation Modeling, 8*(3), 325-352.
- Kaplan, D., & Elliot, P. R. (1997). A didactic example of multilevel structural equation modeling applicable to the study of organizations. *Structural Equation Modeling, 4*, 1-24.
- Kline, R. B. (2005). *Principles and practice of structural equation modeling*. New York: The Guilford Press.
- Mehta P. R., & Neale M. C. (2005). People are variables too: multilevel structural equations modeling. *Psychological Methods, 10*, 259-284.
- Muthén, B. (1991). Multilevel factor analysis of class and student achievement components. *Journal of Educational Measurement, 28*, 338-354.
- Muthén, B. (1994). Multilevel covariance structure analysis. *Sociological Methods and Research, 22*, 376-398.
- Muthén, B. & Satorra, A. (1995). Complex sample data in structural equation modeling. In P. Marsden (Eds.), *Sociological Methodology*, (pp. 216-316).
- Muthén, L. K. ve Muthén, B. O. (2008). *Mplus (Version 5.1)* [Computer software]. Los Angeles: Muthén, ve Muthén.

Ek.1. Çok düzeyli YEM için parametre tahminlerine ilişkin standardize edilmemiş sonuçlar.

MODEL RESULTS					
		Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
Within Level					
FW	BY				
	Y1	1.000	0.000	999.000	999.000
	Y2	1.026	0.051	19.917	0.000
	Y3	1.004	0.050	19.970	0.000
	Y4	1.026	0.057	17.875	0.000
FW	ON				
	X1	1.017	0.056	18.247	0.000
	X2	0.742	4.153	0.179	0.858
Residual Variances					
	Y1	0.956	0.072	13.264	0.000
	Y2	1.088	0.084	13.019	0.000
	Y3	0.883	0.066	13.288	0.000
	Y4	1.007	0.083	12.156	0.000
	FW	0.874	0.085	10.251	0.000
Between Level					
FB	BY				
	Y1	1.000	0.000	999.000	999.000
	Y2	1.058	0.101	10.521	0.000
	Y3	1.020	0.088	11.617	0.000
	Y4	0.992	0.072	13.865	0.000
FB	ON				
	Z	0.646	0.109	5.932	0.000
Intercepts					
	Y1	0.142	0.113	1.251	0.211
	Y2	0.129	0.129	1.001	0.317
	Y3	0.152	0.110	1.381	0.167
	Y4	0.069	0.116	0.600	0.549
Residual Variances					
	Y1	0.003	0.023	0.140	0.888
	Y2	0.028	0.027	1.037	0.300
	Y3	0.004	0.023	0.194	0.846
	Y4	0.020	0.031	0.646	0.518
	FB	0.419	0.168	2.498	0.012