Multivariate Dependence Modeling and Some Financial and Economic Consequences/

Menpekotasun Anizkoitzaren Modelizazioa eta Ondorio Finantzario eta Ekonomiko Batzuk

Jone Ascorbebeitia Bilbatua

2021

eman ta zabal zazu

Universidad del País Vasco

Euskal Herriko Unibertsitatea

Cover designed using images from Freepik.com

Multivariate Dependence Modeling and Some Financial and Economic Consequences/ Menpekotasun Anizkoitzaren Modelizazioa eta Ondorio Finantzario eta Ekonomiko Batzuk

Jone Ascorbebeitia Bilbatua

Supervisors/Zuzendariak:

Eva Ferreira García

Susan Orbe Mandaluniz

eman ta zabal zazu

Universidad del País Vasco / Euskal Herriko Unibertsitatea

Bilbao, July 2021

"Life isn't about waiting for the storm to pass, it's about learning to dance in the rain"

Sherrilyn Kenyon

"Bizitzan egin daitekeen akatsik larriena akats bat egiteko beldurra izatea da"

Elbert Hubbard

Acknowledgements/Eskerrak

Lehenik eta behin nire eskerrik beroenak eman nahi dizkiet nire zuzendariak diren Eva Ferreira García eta Susan Orbe Mandalunizi. Urte hauetan laguntza eta babes izugarria eman didazue biok eta asko ikasi dut zuei esker, bai esparru akademikoan eta bereziki baloratzen dudan garapen pertsonalean. Nirekin eramango ditut gonbidatutako kafe-uneak eta samurtasunez gordetzen ditut baita Alemaniako abesti eta beste oroitzapen guztiak.

Eskertu nahiko nituzke baita ere Ekonomia Aplikatua III (Estatistika eta Ekonometria) eta gaur egungo Metodo Kuantitatiboak departamentuetako kide guztiak jasotako harrera eta arretagatik.

I would like to thank the Department of Mathematical Stochastics in the University of Freiburg (Germany) for all the support before, during, and after my stay in Germany, it was a pleasure to be there. In particular, I would like to thank Thorsten Schmidt and Monika Hattenbach, thank you very much. I would also like to thank all the participants that attended my presentations in different conferences as well as the anonymous reviewers of the accepted papers for their useful comments that helped to improve this thesis chapters.

Bestalde, mila esker urte hauetan zehar bidelagun izan zareten guztioi. Alde batetik, nirekin prozesu honetako momentu onak eta txarrak partekatu dituzuen doktoregai guztiei. Bestetik, unibertsitatetik kanpo ondoan izan ditudan lagunei esandako "zuk eingozu" bakoitzagatik.

Eskerrik asko nire familiari beti eskutik helduta edukitzeagatik. Lehenik eta behin, amama eta aititeri, lasterketa honetako helmugan lehen lerroan egon beharko zinateketelako. Izekitori, izeko baino askoz gehiago izateagatik eta aitari, nigan duzun konfiantza infinituagatik. Azkenik, esker bereziak eman nahi dizkiot amari:

"Bere ondoan bizi izan naiz erregina bat bezela,
ezin zitekeen hobeto bete ama on baten papera,
segundu batez pentsatuz gero ni epaile bat naizela,
bere buruan jarriko nuke Euskal Herriko txapela."
(Miren Amuriza eta Maddalen Arzallus)

Euskal Herriko Unibertsitateko ikertzaileak prestatzeko Doktorego Aurreko Programaren bidez finantzatutako tesia.

Contents

List of Tables

List of Figures

Sarrera

Tesi honetan zehar menpekotasun egiturak aztertuko ditugu. Izan ere, 2007-08 urteko mundu mailako krisiak bultzatuta, finantza aldagaiek osatzen duten egituren azterketak gero eta indar handiagoa hartu du. Nahiko erraza izan daiteke bi aldagai ditugunean egitura hoiek aztertzea, baina aldagai gehiago ditugunean azterketa asko konplikatu daiteke. Hain zuzen ere, praktikan gero eta aldagai gehiago izan behar ditugu kontuan eta, haien baterako egiturak aztertzeko binaka sortzen dituzten egituretan oinarritzea aukera bat den arren, aldagaien binakako efektuez gain orotara dituzten efektuak ere kontuan hartu behar ditugula irakatsi du esperientziak.

Zentzu honetan, baterako banaketa funtzioek informazio aberatsa ematen dute aldagaiek sortzen dituzten egituren inguruan. Baterako banaketak aztertzeko kopulak erabiltzen ditugu. Aldagaiek banaketa normala ez dutenean, kopulek menpekotasun egiturak banaketa eliptikoek baino malgutasun handiagoz modelizatzeko aukera ematen dute. Finantzen arloan bereziki, kopulak oso erabiliak dira arrisku mota ezberdinak kudeatzeko tresnak ematen dituztelako (Embrechts et al., 2003; Palaro eta Hotta, 2006; Jorion, 2007; Hull, 2018). Beste esparru batzuetan ere kopulak oso erabiliak izan ohi dira, hala nola, medikuntzan (Panagiotelis et al., 2012; Rabhi eta Bouezmarni, 2019), airearen kutsadurari buruzko ikerketetan (Noh et al., 2013), ikerketa ozeanografikoetan (Huser eta Wadsworth, 2019) eta makroekonomian (Granger et al., 2006; Dowd, 2008; Loaiza-Maya et al., 2018).

Lehenengo kapituluan kopulen oinarrizko definizioak aurkezten dira. Kapitulu honen lehenengo zatiak bi aldagaietako kopulak deskribatzen ditu, definizio nagusiak eta dituzten ezaugarri edo berezitasunak aipatuz. Gainera, bi aldagaietako analisian dauden kopula parametrikoen sailkapena eta oinarrizkoak diren kopula parametrikoak aurkezten dira. Bestalde, baterako egiturek aldagai gehigarrien eragina ere jaso dezakete. Ildo honetatik, hirugarren aldagai batek aurreko bien baterako banaketan izan dezakeen eragina neurtzeko tresnak ematen dira. Aldagai anitzeko baterako egiturek duten garrantzia dela eta, bigarren zatian lehenengoan aurkeztutako emaitzen orokorpena ematen dugu.

Gaur egun, aldagaiek gero eta egitura konplexuagoak dituztenez, parametrizazioak dituen abantailak labur geratu ohi dira. Konponbide bat estimazio metodo ez-parametrikoen erabilera da. Metodo parametrikoekin alderatuz, malgutasun handiagoa dute eta forma funtzional zehatz bati lotuta ez dauden egiturak modelizatzea baimentzen dute. Horregatik, kapitulu honetatik aurrera modelizazio ez-parametrikoa erabiliko dugu.

Aldagaien baterako egiturak aztertzerakoan haien menpekotasun maila neurtzea bilatzen dugu. Aldagai anitzen menpekotasuna neurtzeko prozedurarik erabiliena binakako korrelazio linealen batazbestekoa kalkulatzea da (ikusi adibidez Joe, 1990; Longin eta Solnik, 1995; Moskowitz, 2003; Capiello et al., 2006; Pollet eta Wilson, 2010, beste batzuen artean). Hasiera batean, neurri erakargarria da interpretatzeko duen erraztasuna dela eta, baina kontuan izan behar da bi desabantaila nagusi dituela. Alde batetik, korrelazio linealetan oinarritutako neurri bat da eta beraz, aproposa izan daiteke banaketa simetrikoa duten aldagaientzat baina ez da bat ere egokia aldagaiek mutur pisutsuak eta banaketa asimetrikoak dituztenean, ekonomia eta finantzen arloko aldagaietan ohikoak diren ezaugarriak. Pearson-en korrelazioa eta banaketa normaletatik haratago, Kendall-en tauak linealtasuna beharrean monotonia hautematen du eta ez du banaketa simetrikoekiko murrizketarik. Testuinguru honetan, Kendall eta Smith (1940)-ek binakako Kendall-en tauan oinarritutako aldagai anitzetarako proposamena egiten dute. Binakako eraginetan oinarritutako neurriek ordea, aldagai multzoak bere osotasunean sortzen dituen informazio topologikoaren zati bat galtzen dute, arinago aipatu dugun bezala. Hori dela eta, Joe (1990)-k baterako egitura aintzat hartzen duen eta Kendall-en tauan oinarritzen diren neurrien familia bat proposatzen du.

Bigarren kapituluan dugun helburuetako bat aldagai anitzeko baldintzatutako Kendall-en taua estimatzea da, menpekotasun egituran aldagai gehigarri batek duen eragina aztertzeko. Horretarako, Joe (1990)-k proposatutako neurria menpekotasun baldintzatuaren kasura hedatzen dugu aldagaiak independenteak eta berdinki banatuta dauden kasurako. Zentzu honetan, menpekotasunaren estimaziorako gure proposamena Gijbels et al. (2011)-k baldintzatutako Kendall-en taurako proposatutako neurriaren aldagai anitzetarako hedapen bat da.

Estimazio ez-parametrikoan leuntasun parametroaren aukeraketa garrantzitsua da eta dentsitatearen edo banaketa funtzioen testuinguruan jadanik prozedura ezberdinak proposatu dira (Silverman, 1986; Altman eta Leger, 1995; Sarda, 1993; Bowman et al., 1998). Kapitulu honetan, Kendall-en tauari dagokion parametroaren *plug-in* aukeraketa algoritmo bat proposatzen dugu Derumigny eta Fermanian (2019)-ren proposamenaren ordezko aukera bezala.

Kapitulu honetako bigarren helburua baldintzatutako Kendall-en tauaren gaineko murrizketa linealen kontrasteak egitea da. Gijbels et al. (2017) ikerlariek Kendall-en tauan oinarritutako estatistikoa proposatzen dute baldintzatutako kopula kopula partzialarekin bat datorren aztertzeko. Gure kasuan, Wald motako estatistiko bat proposatzen dugu edozein motako murrizketa linealen kontrasteak egiteko aukera ematen duena, esaterako, baldintzatutako menpekotasun konstantea, tau ezberdinen arteko murrizketa linealak edo populazio ezberdinen arteko baldintzatutako tauen murrizketak. Gainera, proposatutako estatistikoaren portaera simulazio ikerketa baten bidez aztertzen dugu.

Aplikazio enpirikoari dagokionez, RCI Europar Batasuneko eskualdeen konpetitibitate indizearekin lan egiten dugu. Zehatz mehatz, eraginkortasun eta berrikuntza taldeetako oinarriek dituzten erlazioak eta eskualdeetako erakundeen kalitateak erlazio horietan duen eragina aztertzen dira. Horretarako, proposatutako Kendall-en taua eta Wald motako estatistikoa erabiltzen ditugu. Estatistikoak, denbora tarte ezberdinetan erakundeen kalitatearen eragina aldatuz joan den ikustea ere onartzen du.

Hirugarren kapituluan finantzetan ematen diren menpekotasun egituretan jartzen dugu fokua. Finantza aldagaien ezaugarri nabarmen bat haien denborarekiko aldakortasuna da (Krishnan et al., 2009; Ferreira et al., 2011; Ferreira eta Orbe, 2018). Longin eta Solnik (1995) eta Engle (2002) ikerlanetan ere, finantzazko denbora serieak eta kopulen menpekotasun parametroak denborarekin aldatu egiten direla ondorioztatzen da. Hori dela eta, menpekotasun egituren denborarekiko aldakortasuna eredu parametriko ezberdinen bidez aztertu ohi da (Palaro eta Hotta, 2006; Jondeau eta Rockinger, 2006; Patton, 2006; Hafner eta Reznikova, 2010; Hafner eta Manner, 2012; Oh eta Patton, 2013). Guk, kapitulu honetan, aldakortasun hau barneratzen duen kopularen eta Kendall-en tauaren estimatzaile ez-parametrikoak proposatzen ditugu tokiko egonkorrak diren aldagai eta α-mixing koefizienteetarako (Ascorbebeitia et al., 2021), Oh eta Patton (2018) lanean azpimarratutako hutsune bat betetzen delarik. Gainera, proposatutako bi estimatzaileen sendotasun eta normalitate asintotikoaren emaitza nagusiak ematen ditugu.

Kendall-en tauak praktikan izan dezakeen erabileraren erakargarritasuna erakusteko, bere portaera simulazio ikerketa baten bitartez aztertzen dugu eta hain erabilia den DCC korrelaziotik eratorritako hein korrelazioarekin konparatzen dugu. Emaitzek Kendall-en tauaren nagusitasuna erakusten dute, batez ere konplexuak diren egituretan. Ohartu proposatutako estimatzaileak, binakako tauen batazbestekoak ez bezala, baterako banaketak ematen dituen aldibereko erlazioak kontuan hartzen dituela. Ezaugarri hau bereziki garrantzitsua da menpekotasun anizkoitzaren neurri egoki gisa hartzeko.

Proposatutako Kendall-en tauaren estimatzailea Europako bost indize nagusien menpekotasunak 2005etik 2020ra arte duen bilakaera aztertzeko erabiltzen dugu. Honela, DAX, FTSE 100, IBEX 35, CAC 40 eta SMI indizeek osatzen duten egitura topologikoa hartzen dugu aztergai. Emaitzen arabera, menpekotasuna handiagoa da Euro Stoxx-ak prezio baxua duenean eta txikiagoa prezioa altua denean, eta honek menpekotasuna merkatuaren beheranzko joerekin erlazionatuta dagoen ikustera bultzatzen gaitu.

Horretarako, Adrian eta Brunnermeier (2016)-ek arrisku sistemikoa neurtzeko proposatutako $\Delta CoVaR$ neurria indize-zorro baten menpekotasunak Euro Stoxx-aren kuantiletan duen eragina ebaluatzeko erabiltzen dugu. Kontuan izan, menpekotasun maila altuek duten eragina dela bereziki interesatzen zaiguna. Orain arte egindakoarekiko koherentzia mantentzeko $CoVaR$ eta $\Delta CoVaR$ neurriak kuantil ez-parametrikoen bidez estimatzen dira.

Emaitzen arabera, estimatutako aldagai anitzeko Kendall-en taua egunean-eguneko menpekotasun mailaren adierazle ona dela ondorioztatu daiteke. Honek, menpekotasunak Euro Stoxx-aren kuantilak eragiten dituen aztertzera eramaten gaitu zuzen zuzenean. Honela, Jeong et al. (2012)-k kuantil baldintzatuen murrizketetarako proposatutako Granger-en kausalitate testaren bertsio eraldatu batek kausalitatearen aldeko ebidentzia ematen du.

Finantzetan aldagaien baterako erlazioek duten garrantzia dela eta, interesgarria iruditzen zaigu, herrialde ezberdinetako indizeez gain, finantza merkatuetan akzioek dituzten erlazioak nolakoak diren aztertzea. Horretarako, sare teoria erabiliko dugu.

Laugarren kapituluan sare teorian oinarrizkoak diren kontzeptuak azaltzen dira. Lehenik eta behin, sareen definizioa eta sareetan parte hartzen duten elementuak aurkezten dira. Sare moten barnean *klike* izenekoak azpimarratu daitezke dituzten ezaugarri bereziak direla eta. Sareen azterketa egin ahal izateko algebra matrizialak ematen dituen tresnak erabiltzen dira. Izan ere, sareen egitura topologikoak matrizeen bitartez adierazi ditzakegu, non matrizeko elementu bakoitzaren posizioak sareko bi akzio adierazten dituen eta elementuaren balioak akzio horiei dagokien lotura. Matrizeen propietate aljebraikoak erabiliz sareen oreka kontzeptua aurkezten da. Praktikan propietate hau sistemen egonkortasunarekin erlazionatzen da eta beraz, bereziki interesgarria izan daiteke akzio sareeak aztertzeko erabiltzea. Ohartu egonkortasun eza ezaugarri negatibo bezala hautematen dela sistemari ematen dion ziurgabetasuna dela eta. Kontuan izan behar da baita ere edozein sistema kanpo faktoreen eraginpean dagoela. Finantzetan esaterako, merkatuak testuinguru ekonomiko eta politiko batean daude murgilduta eta hauek merkatuen gorabeherak baldintzatu ditzakete. Beraz, sareak aztertzerakoan eragin hoiek kontuan nola hartu daitezkeen azaltzen da.

Bosgarren kapituluan finantza merkatuen azterketa sakonago bat egiten dugu eta akzioek sortzen dituzten egiturak aztertzen ditugu. Izan ere, azken finantza krisiak finantza aldagaiek dituzten erlazioekin lotutako arrisku mota berri bat utzi zuen agerian, arrisku sistemikoa, eta zentzu honetan, finantza merkatuen egiturak eta barne erlazioek arrisku sistemikoari buruz asko dute esateko (Scott, 2012).

Kutsadura nola ematen den aztertzeko *credit default swap (CDS)* bezala ezagutzen diren finantza deribatu subiranu eta korporatiboak erabiltzea ohikoa da (Acharya et al., 2014; Alter eta Beyer, 2014; De Bruyckere et al., 2013). Testuinguru honetan, Allen et al. (2010) ikerlariek grafoek arrisku sistemikoaren sorrera eta hedapena aztertzeko finantzetan izan dezaketen erabilgarritasuna azpimarratzen dute. Honela, Billio et al. (2012)-k, beste batzuen artean, finantza erakunde ezberdinak aintzat hartzen dituzten sareak sortzen dituzte eta bankuek ezusteko gertaerak zabaltzen beste edozein finantza erakundek baino eginkizun handiagoa dutela aurkitzen dute. Honi erlazionatuta, Gross eta Siklos (2020)-en lana aurkitu daiteke, zeinetan CDS deribatuetan oinarritutako sareak eraikitzen diren. Lortzen dituzten emaitzen artean, banku arriskua eta arrisku subiranua azpimarratzen dituzte korporazioen kreditu arriskuaren eragile bezala. Gainera, krisi garaietan finantzazkoak ez diren sektoreetara ematen den arrisku transmisioa denboran zehar mantendu egin dela ondorioztatzen dute.

Hori dela eta, herrialde ezberdinetako akzio sareetan ematen diren aldaketak eta aldaketa hauek finantza eta finantzazkoak ez diren sektoreekin duten erlazioa ikertu nahi dugu. Horretarako, bederatzi herrialderen akzio merkatuetako sareak eraikiko ditugu. Alde batetik, GIIPS bezala ezagutzen diren herrialdeak kontsideratzen dira, eta bestetik, Europa eta mundu mailan garrantzitsuak diren ekonomiak, hala nola, Alemania, Frantzia, Estatu Batuak eta Japonia.

Sareak eraikitzeko akzio bikoteak zein irizpideren arabera erlazionatuko diren zehaztatu behar da. Literaturan korrelazio linealetan oinarritutako neurriak erabiltzen dira (ikusi adibidez Mantegna, 1999; Birch et al., 2016; Brida et al., 2016; Wang et al., 2018; Zhao et al., 2018; Guo et al., 2018), beti ere irizpide horiek erlazio positiboak adierazten dituztelarik. Hala ere, berehalakoa da erlazioen zeinua deuseztatzeak informazio galera inplikatzen duela (Heiberger, 2014; Stavroglou et al., 2019). Ondorioz, aldi berean erlazio positibo eta negatiboen erabilera onartzen duen zeinudun grafo teoria erabiltzen dugu (Harary et al., 1953) bederatzi herrialdeetako sareak eraikitzeko. Zehatz mehatz, akzioen arteko loturak ezartzeko hein korrelazioak kontsideratuko ditugu.

Finantza aldagaiek duten denborarekiko aldakortasuna dela eta, eraikitzen ditugun sareen egiturek urteetan zehar izan duten bilakaera aztertzen dugu Bardoscia et al. (2021)-k nabarmentzen duten hutsunea betez. Horretarako, 2005ean hasi eta 2020ko irailera arteko bederatzi herrialdeetako akzioen errentagarritasunak erabiltzen ditugu. Sareetan denborarekiko aldakortasuna aintzat hartzeko akzioen arteko loturak proposatutako Kendall-en tauaren bitartez zehazten dira. Jadanik arinago aipatu dugun bezala, sareen oreka propietate interesgarria izan daiteke. Zentzu honetan, Estrada eta Benzi (2014) ikertzaileek proposatutako sareen oreka neurria erabiltzen dugu sareen egonkortasunak denboran zehar izan duen bilakaera aztertzeko. Azken honek finantza merkatua bere guztiz orekatutako bertsiotik zenbat aldentzen den neurtzen du.

Sareen oreka koefizienteak 2011. urteko irailean egonkortasun galera bortitz bat erakusten du aztertutako bost herrialdetarako, urte horretako abuztuaren 8-ko *astelehen beltzaren* ondoren ematen dena. Nabarmentzekoa da egonkortasun galera finantza sektorekoak ez diren eta kapitalizazio baxua duten akzio multzo txiki batek sortzen duela. Izan ere, multzo horietako akzioak oso negatiboki erlazionatuta daude haien artean eta baita sareko gainontzeko akzioekin ere. Beraz, talde txiki bateko akzioen elkarrekiko konektibitateak garrantzia handia duela erakusten digute emaitzek, bereziki akzio merkatu osoaren egonkortasuna eta arrisku sistemikoari begira. Zentzu honetan, hainbat ikerlanetan azpimarratutako huts egiteko elkar konektatuegi dauden akzioen garrantzia bermatzen da (Markose et al., 2012; Acemoglu et al., 2015). Ohartu akzio merkatuetan honelako efektu bat hautematen den lehen aldia dela eta hortaz, ikerketa honek erregulazio, estaldura, zein aktibo-zorroen esleipen estrategiak hobetzeko beste abiapuntu bat ematen duela.

1

Kopulen Teoria

Kopula Parametrikoak

1 Sarrera

Kapitulu honetan kopulei buruzko oinarrizko teoria aurkeztuko dugu (informazio gehiagorako ikusi Nelsen (2007) oinarrizko erreferentzia bezala). Lehenik eta behin, bi dimentsiotako kopuletan zentratuko gara. Kopulak zer diren eta nondik datozen ikusiko dugu. Gero, literaturan aurkitu daitezkeen kopula parametrikoen sailkapena eta mota bakoitzean gailentzen diren kopulak errepasatuko ditugu. Kapitulu honetan ikusiko ditugun kopula parametrikoak oinarrizkoak dira kopula teorian. Bestalde, bi aldagaietako erlazioari hirugarren aldagai gehigarri baten eragina aztertzeko tresnak aurkeztuko ditugu. Praktikan askotan bi aldagai baino gehiagorekin lan egin behar izaten dugu eta ondorioz, bi dimentsiotan ematen diren propietateen orokorpenak behar ditugu. Kapitulu honetako bigarren zatian hain zuzen, lehenengo zatiko emaitzen orokorpenak ematen dira.

Notazioari dagokionez, aurrerantzean $\mathbb{R}$ erabiliko dugu $(-\infty, +\infty)$ zuzen erreala adierazteko eta $\overline{\mathbb{R}}$ aldiz $[-\infty, +\infty]$ zuzen errealaren hedapenak adierazteko.

2 Bi aldagaietako kopulak

Definizioa 1. Izan bitez S_1 eta S_2 $\overline{\mathbb{R}}$-ren azpimultzoak eta F, $D(F) = S_1 \times S_2$ eremuan definitutako funtzioa. Izan bedi baita erpin guztiak $D(F)$ eremuan dituen $J = [x_1, x_2] \times [y_1, y_2]$ laukizuzena. Orduan, *J-ren F-bolumena*

$$V_F(J) = F(x_2, y_2) - F(x_2, y_1) - F(x_1, y_2) + F(x_1, y_1)$$

adierazpenaren bidez emanda dago. Ohartu $V_F(J)$, J laukizuzenaren F-masa dela baita: J laukizuzenean F-ren lehen ordenako deribatuak $\Delta_{x_1}^{x_2} F(x,y) = F(x_2, y) - F(x_1, y)$ eta $\Delta_{y_1}^{y_2} F(x,y) = F(x, y_2) - F(x, y_1)$ bezala definitu badaitezke, orduan J laukizuzenaren F-bolumena J-ren gaineko F-ren bigarren ordenako deribatua da,

$$V_F(J) = \Delta_{x_1}^{x_2} \Delta_{y_1}^{y_2} F(x,y).$$

Definizioa 2. F bi aldagaietako funtzio erreala *2-gorakorra* dela esaten da $V_F(J) \geq 0$ betetzen bada erpinak $D(F)$ eremuan dituen J laukizuzen guztietarako. Bestela, F *2-beherakorra* dela esaten da.

F funtzioa 2-gorakorra denean, J laukizuzenaren F-bolumenari askotan *J-ren F-neurria* esaten zaio. Esan beharra dago 2-gorakorrak diren funtzioei *sasi monotonoak* ere esaten zaiela. Ohartu baita F funtzioa 2-gorakorra izateak ez duela inplikatzen, ezta inplizituki azaltzen ere, F argumentu bakoitzean ez-beherakorra denik.

Adibidea 1.

- Izan bedi $I^2 = [0,1]^2$ eremuan definitutako $F(x,y) = \max(x,y)$ funtzioa. F funtzioa ez-beherakorra da x-en eta baita y-n ere, baina $V_F([0,1]^2) = \Delta_{y_1}^{y_2} \Delta_{x_1}^{x_2} \max(x,y) = \max(x_2, y_2) - \max(x_1, y_2) - \max(x_2, y_1) + \max(x_1, y_1) = -1$ da. Hortaz, F ez da 2-gorakorra.

- Izan bedi $I^2 = [0,1]^2$ eremuan definitutako $F(x,y) = (2x-1)(2y-1)$ funtzioa. Orduan, $V_F([0,1]^2) = \Delta_{y_1}^{y_2} \Delta_{x_1}^{x_2} \left((2x-1)(2y-1)\right) = \big((2x_2-1)-(2x_1-1)\big)\left((2y_2-1)-(2y_1-1)\right) = 4 \geq 0$ dela kontuan izanik, F funtzioa 2-gorakorra da. Hala ere, x-en funtzio beherakorra da $y \in (0, \frac{1}{2})$ bakoitzerako eta y-n funtzio beherakorra da $x \in (0, \frac{1}{2})$ bakoitzerako.

Suposatu S_1 eta S_2 azpimultzoek a_1 eta a_2 elementu minimoak dituztela, hurrenez hurren. $F : S_1 \times S_2 \to \mathbb{R}$ funtzioa $(x_1, x_2) \in S_1 \times S_2$ guztietarako *oinarrituta* dagoela esaten da baldin eta $F(x, a_2) = 0 = F(a_1, y)$ betetzen bada.

Lema 1. *Izan bitez S_1 eta S_2 $\overline{\mathbb{R}}$-ren bi azpimultzo eta F, $S_1 \times S_2$ eremuan definitutako 2-gorakorra den funtzio oinarritua. Orduan, F argumentu bakoitzean ez-beherakorra da.*

Demagun orain S_1 eta S_2 azpimultzoetako bakoitzak b_1 eta b_2 izendaturiko elementu handien bat duela. Orduan $F : S_1 \times S_2 \to \mathbb{R}$ funtzioak $F_1(x) = F(x, b_2)$ eta $F_2(y) = F(b_1, y)$ funtzio marjinalak ditu $x \in S_1$ eta $y \in S_2$ guztietarako. Ohartu F_1 eta F_2 funtzioak $D(F_1) = S_1$ eta $D(F_2) = S_2$ eremuetan definituta daudela, hurrenez hurren.

Adibidea 2. Izan bedi $[-1, 1] \times [0, \infty]$ eremuan definitutako $F(x, y) = \frac{(x+1)(e^y-1)}{x+2e^y-1}$ funtzioa. $F(x, 0) = 0$ eta $F(-1, y) = 0$ betetzen direnez, F oinarrituta dago eta dagozkion marjinalak $F_1(x) = F(x, \infty) = \frac{(x+1)(e^\infty-1)}{x+2e^\infty-1} \to \frac{x+1}{2}$ eta $F_2(y) = F(1, y) = \frac{e^y-1}{e^y} = 1 - e^{-y}$ dira.

Definizioa 3. *2 dimentsioko azpikopula edo 2-azpikopula* C' funtzio bat da hurrengo propietateak betetzen dituena:

i) C'-ren eremua $D(C') = S_1 \times S_2$ da, non S_1 eta S_2 $[0, 1]$-ren bi azpimultzo diren.

ii) C' oinarritua eta 2-gorakorra da.

iii) $C'(u_1, 1) = u_1$ eta $C'(1, u_2) = u_2$, $u_1 \in S_1$ eta $u_2 \in S_2$ guztietarako.

Ohartu $0 \leq C'(u_1, u_2) \leq 1$ dela $(u_1, u_2) \in D(C')$ guztietarako, beraz $rang(C') \subseteq I = [0, 1]$.

Definizioa 4. *Bi dimentsioko kopula*, *2-kopula* edo beste gabe *kopula*, $[0, 1]^2$ eremuan definitutako 2-azpikopula bat da. Beraz, *kopula* bat $C : [0, 1]^2 \to [0, 1]$ eremuan definitutako $C(u_1, u_2) = P\,(U_1 \leq u_1, U_2 \leq u_2)$ funtzio bat dela esango dugu, ondoko propietateak betetzen dituena:

i) $C(u_1, 0) = C(0, u_2) = 0$.

ii) $C(u_1, 1) = u_1$ eta $C(1, u_2) = u_2$.

iii) $V_C([u_1, v_1] \times [u_2, v_2]) = C(v_1, v_2) - C(u_1, v_2) - C(v_1, u_2) + C(u_1, u_2) \geq 0$, non $u_1, u_2, v_1, v_2 \in [0, 1]$ guztietarako, $u_1 \leq u_2$ eta $v_1 \leq v_2$.

Kopula eta azpikopula kontzeptuen arteko ezberdintasuna oso txikia dela ematen duen arren (ezberdintasuna definizio eremuan dago), Sklar-en teoreman garrantzitsua izango da.

Kopulen artean ondoren aurkeztuko diranak oinarrizkoak dira. Aldagaien independentziarekin konsistentea den kopulari $\Pi(u_1, u_2) = u_1 u_2$ *biderkadura kopula* esaten zaio. Kasu honetan baterako banaketa funtzioa banaketa funtzio marjinalen biderkadura izango da. Beste kopula garrantzitsu bat $M(u_1, u_2) = \min(u_1, u_2)$ kopula da, *Fréchet-en goi bornea* izenaz ezagutzen dena. Ez dago Fréchet-en goi kopulak baino balio handiagoak hartu ditzakeen beste kopularik. Zorizko aldagaiak kopula honen bidez erlazionatuta daudenean, menpekotasun zehatz positiboa dutela esaten da. Azken honekin erlazionatuta *Fréchet-en behe bornea* ere aurki daiteke, $W(u_1, u_2) = \max(u_1 + u_2 - 1, 0)$. Ez dago balio txikiagoa hartu dezakeen beste kopularik eta kopula honen bitartez erlazionatutako zorizko aldagaiek menpekotasun zehatz negatiboa dute. Edozein kopulatan parametroren bat aldatu ahala kopula horrek Fréchet-en kopuletako batera konbergitzen badu, menpekotasun positiboa edo negatiboa inplikatzen duela esaten da. Gainera, $(u_1, u_2) \in [0, 1]^2$ eta C kopula guztietarako betetzen den $W(u_1, u_2) \leq C(u_1, u_2) \leq M(u_1, u_2)$ erlazioari *Fréchet-Hoeffding borneen desberdintza* esaten zaio.

Sklar-en teorema kopulen teoriako teorema nagusia da. Teorema honek, kopulek banaketa marjinalen eta baterako banaketa funtzioen arteko erlazioan duten eginkizuna zehazten du. Sklar-en teorema aurkeztu aurretik beharrezkoa da zenbait kontzeptu ezagutzea.

Definizioa 5. Banaketa funtzio bat $\overline{\mathbb{R}}$-n definitutako G funtzio ez-beherakorra da, non $G(-\infty) = 0$ eta $G(\infty) = 1$.

Definizioa 6. Bi aldagaietako *baterako banaketa funtzioa* $\overline{\mathbb{R}}^2$-n definitutako F funtzio bat da eta hurrengo bi propietateak betetzen ditu:

i) F, 2-gorakorra da.

ii) $F(x, -\infty) = F(-\infty, y) = 0$ eta $F(\infty, \infty) = 1$.

Bigarren baldintzatik F oinarritua dela ondorioztatu dezakegu. Gainera, $D(F) = \overline{\mathbb{R}}$ denez, F baterako banaketa funtzioak $F_1(x) = F(x, \infty)$ eta $F_2(y) = F(\infty, y)$ marjinalak ditu. Funtzio marjinal hauek ere banaketa funtzioak dira era berean.

Teorema 2. ***(Sklar-en teorema)*** *Izan bedi F baterako banaketa-funtzio bat eta F_1 eta F_2, F-ren banaketa funtzio marjinalak. Orduan, C kopula bat existitzen da, non $x, y \in \overline{\mathbb{R}}$ guztietarako*

$$F(x, y) = C(F_1(x), F_2(y)). \tag{1.1}$$

F_1 eta F_2 banaketa funtzio marjinalak jarraituak badira C kopula bakarra da eta hortaz, kopula modu unibokoan zehaztuta dago $rang(F_1) \times rang(F_2)$ eremuan. Alderantziz, C kopula bat bada eta F_1 eta F_2 banaketa funtzioak badira, orduan (1.1) *adierazpenean definitutako F funtzioa banaketa funtzio bateratua da F_1 eta F_2 marjinalak dituena.*

Sklar-en teoremak baterako banaketa funtziorako adierazpen bat ematen du kopularen eta banaketa funtzio marginalen menpean. Adierazpen horri buelta ematen bazaio, kopula baterako banaketa funtzioaren eta funtzio marjinalen alderantzizkoen menpe jarri daiteke. Hala ere, banaketa funtzio marjinalen bat hertsiki gorakorra ez bada, orduan alderantzizkoak ez du ohiko zentzua izango, eta hori dala eta, banaketa funtzioen "sasi-alderantzizkoak" definitzea beharrezkoa da.

Definizioa 7. F banaketa funtzio bat bada, *F-ren sasi-alderantzizkoa* hurrengo bi propietateak betetzen dituen eta $[0,1]$ tartean definituta dagoen edozein F^{-1} funtzio da:

i) baldin eta $\ell \in Im(F)$ eta $x \in \mathbb{R}$ badira, $F^{-1}(\ell) = x$ eta $F(x) = \ell$. Hortaz, $\ell \in Im(F)$ guztietarako, $F(F^{-1}(\ell)) = \ell$ betetzen da.

ii) baldin eta $\ell \notin Im(F)$ bada, $F^{-1}(\ell) = inf\{x : F(x) \geq \ell\} = sup\{x : F(x) \leq \ell\}$.

Ohartu F hertsiki gorakorra denean sasi-alderantzizkoa ohiko alderantzizko funtzioa da.

Sklar-en alderantzizko teoremaren arabera, edozein bi banaketa funtzio (ez dute familia berdinekoak izan behar) kopularen batekin bateratu daitezke; eta ondorioz bi aldagaietako banaketa funtzio baliagarri bat definituta dago. Ekonomia eta estatistikan honelako emaitzak bereziki interesgarriak dira, aldagai bakarreko banaketa funtzio parametriko ugari aurkitu arren aldagai anitzeko banaketa funtzioen multzoa oso txikia baita. Izan ere, baliteke aldagai anitzeko banaketa funtzioen konplexutasunak edo funtzio hauen iraganeko erabilera eskasak arazo hau sorrarazi izana.

Korolarioa 3. *Izan bitez H bi aldagaietako banaketa funtzioa eta F_1 eta F_2 bere banaketa funtzio marjinal jarraituak. Izan bitez baita F_1^{-1} eta F_2^{-1} banaketa funtzio marjinalen sasi-alderantzizkoak. Orduan, $C : [0,1]^2 \to [0,1]$ kopula bakarra existitzen da, non $C(u_1, u_2) = H(F^{-1}(u_1), G^{-1}(u_2))$, $(u_1, u_2) \in [0,1]^2$ guztietarako.*

Korolario honen arabera, aldagai anitzeko edozein banaketa funtzioren kopula lortzea posiblea da jatorrizko banaketa funtzioaren marjinalak edozein izanik. Beste era batera esanda, edozein bi banaketa funtzio marjinal eta hauei lotutako aldagaiak erlazionatzen dituen kopula ezagutzen baldin badugu, baterako banaketa funtzioa lortu daiteke. Propietate honetan ikusi daiteke kopulen bitartez aldagai anitzeko banaketa funtzioen multzoa are gehiago handitzen dela.

Oharra 1. Izan bitez X eta Y bi zorizko aldagai eta demagun $F(x,y) = C(F_X(x), F_Y(y))$ betetzen dela. Sklar-en teoremaren arabera, baterako probabilitatea kopula eta marjinalen artean banatu daiteke, kopulak X eta Y aldagaien arteko lotura bakarrik adierazten duelarik. Hortaz, banaketa funtzioaren bitartez adierazitako probabilitate bateratuan ez bezala, kopulek aldagaien portaera marjinala multzotik banatzen dute eta hori dala eta, kopulei menpekotasun funtzio esaten zaie. Kasu honetan, C_{XY} kopulari X eta Y zorizko aldagaien arteko kopula esaten zaio.

Proposizioa 4. *Kopula funtzioek ondorengo bi propietateak betetzen dituzte:*

i) Izan bitez C_{XY} X eta Y-ren kopula eta α eta β hertsiki gorakorrak diren bi funtzio. Orduan, $C_{\alpha(X)\beta(Y)} = C_{XY}$. Beraz, C_{XY} inbariantea dela esan dezakegu hertsiki gorakorra den X eta Y aldagaien edozein transformaziorako.

ii) Izan bitez C_1 eta C_2 bi kopula. Orduan, C_1 kopula C_2 baino txikiagoa da ($C_1 \prec C_2$) baldin eta $u_1, u_2 \in [0,1]$ guztietarako $C_1(u_1,u_2) \leq C_2(u_1,u_2)$ betetzen bada.

Hurrengo teoremaz baliatuz posible dugu orain kopula baten dentsitatea definitzea.

Teorema 5. *Izan bedi C kopula. Orduan, $u_1, u_2 \in [0,1]$ guztietarako $\frac{\partial C}{\partial u_1}$ eta $\frac{\partial C}{\partial u_2}$ deribatu partzialak existitzen dira, non $0 \leq \frac{\partial C(u_1,u_2)}{\partial u_1} \leq 1$ eta $0 \leq \frac{\partial C(u_1,u_2)}{\partial u_2} \leq 1$ betetzen diren, hurrenez hurren. Gainera, $u_1 \mapsto \frac{\partial C(u_1,u_2)}{\partial u_2}$ eta $u_2 \mapsto \frac{\partial C(u_1,u_2)}{\partial u_1}$ funtzioak ez beherakorrak dira eta $[0,1]$ eremuan definituta daude.*

Definizioa 8. Izan bedi C bi aldagaietako kopula bat. Kopula baten *dentsitatea*

$$c(u_1,u_2) = \frac{\partial C(u_1,u_2)}{\partial u_1 \partial u_2}$$

bezala definitzen da. Honela, $c(u_1,u_2)$ kopularen dentsitatea ezagutzen badugu, F banaketa funtzioaren f dentsitate funtzioa kopula dentsitatearen menpean

$$f(x,y) = c(F_X(x), F_Y(y)) f_X(x) f_Y(y)$$

bezala definituta geratzen da. Azken emaitza hau Sklar-en kopulen dentsitateetarako teorema moduan ere ezagutzen da.

2.1 Kopula motak

Kopula parametriko mota ezberdin asko daude (Nelsen, 2007), baina sailkapenak egiteko irizpide ezberdinak daudenez, zaila da mota guztien sailkapen zehatz bat aurkitzea. Esaterako, kopulak menpekotasunaren arabera edo euskarriaren arabera (jarraitua edo diskretua) sailkatu daitezke.

- Kopula inplizitua edo esplizitua:
    - Kopula inplizituak: Kopularen forma funtzionalak banaketa funtzio ezagun batekin bat egiten du (Gauss-en kopula edo Student-en t kopula).
    - Kopula esplizituak: Kopula bakoitzak forma funtzional erraz bat du. Talde honetakoak dira Gumbel, Clayton eta Frank bezalako kopulak.
- Menpekotasunaren arabera sailkatutako kopulak: Kopula eliptikoak, kopula normalak, mutur balioetako kopulak, kopula arkimedearrak eta HRT kopulak.

2.1.1 Kopula eliptikoak

Definizioa 9. *Kopula eliptikoak banaketa funtzio eliptikoei lotutako kopulak dira. Kopula eliptikoek menpekotasun erlazio simetrikoak adierazten dituzte, banaketa funtzioaren muturrei erreparatu gabe. Kopula eliptiko guztiek itxura bera dute:*

$$\begin{aligned} C_\rho(u_1,u_2) &= (1-\rho^2)^{-1/2}\int_{-\infty}^{\Phi_{g,1}^{-1}(u_1)}\int_{-\infty}^{\Phi_{g,2}^{-1}(u_2)} g\left(\frac{x^2-\rho xy+y^2}{\sqrt{1-\rho^2}}\right)dxdy \\ &= H_\rho(\Phi_{g,1}^{-1}(u_1),\Phi_{g,2}^{-1}(u_2)). \end{aligned}$$

Gauss-en eta Student-en t kopulak dira kopula eliptiko ezagunenak, simetrikoak eta erabilterrazak dagozkien banaketak ezagunak direlako.

Gauss-en kopula

Gauss-en kopula banaketa funtzio normalari lotutako menpekotasun funtzioa da eta kopula inplizituen taldean sartzen da. ρ korrelazio koefizientea bada, orduan Gauss-en kopula $C(u_1,u_2;\rho) = \Phi_\rho(\Phi^{-1}(u_1),\Phi^{-1}(u_2))$ bezala definitzen da, hau da,

$$C(u_1,u_2;\rho) = \int_0^{\Phi^{-1}(u_1)}\int_0^{\Phi^{-1}(u_2)} \frac{1}{2\pi\sqrt{1-\rho^2}} exp\left(-\frac{x^2+y^2-2\rho xy}{2(1-\rho^2)}\right)dxdy,$$

bere dentsitatea

$$c(u_1,u_2;\rho) = (1-\rho^2)^{-1/2} exp\left(\frac{x^2+y^2-2\rho xy}{2(1-\rho^2)}+\frac{x^2+y^2}{2}\right)$$

delarik. Gauss-en kopula kopula simetrikoa da, hau da, $C(u_1,u_2) = C(u_2,u_1)$. Kopula honek muturretan menpekotasun nulua edo oso baxua du aldagaien korrelazioa 1-en ezberdina denean. Hori dela eta, finantza errentagarritasunen menpekotasuna modelizatzeko ez da kopula oso egokia. Izan ere, badirudi finantza errentagarritasunak handiagoak eta negatiboak direnean korrelazio handiagoa dutela handiak eta positiboak direnean baino, muturretako menpekotasuna asimetrikoa dela aditzera ematen duena.

Student-en t kopula

Student-en t kopula Student-en t banaketa funtzioari lotutako menpekotasun funtzioa da eta kopula inplizituen taldean sartzen da. Bi aldagairen kasuan $C(u_1, u_2; \rho, \nu) = \mathbf{t}_{\rho,\nu}(t^{-1}(u_1), t^{-1}(u_2))$ bezala definitzen da. Student-en t banaketaren dentsitatea ezagutzea beharrezkoa da kopula definitu ahal izateko. Beraz, $f(x, y) = (2\pi(1-\rho^2)^{-1/2})^{-1}\left(1 + \frac{x^2-2\rho xy+y^2}{\nu(1-\rho^2)}\right)^{-\frac{\nu+2}{2}}$ da, non ν askatasun graduak eta ρ korrelazio koefizientea diren. Gogoratu dezagun kopulak aldagaitzak direla transformazio gorakor monotonoetarako. Beraz, $C(t(\nu, \mu, \Sigma)) = C(t(\nu, 0, \mathbf{P}))$, non $\mathbf{P}$ korrelazio matrizea den.

Student-en t banaketak orokorrean $\mathbf{X} \stackrel{d}{=} \mu + \sqrt{\frac{\nu}{S}}\mathbf{Z}$ adierazpena badu $\mathbf{Z} \sim N(\mathbf{0}, \Sigma)$ izanik, orduan $\mathbf{X}$-ek $\mathbf{t}_\nu$ banaketa bivariantea du batazbestekoa μ eta kobariantza matrizea $\frac{\nu}{\nu-2}\Sigma$ dituena, $\nu > 2$. Kontuan izan $\nu \leq 2$ bada $\mathbf{X}$-en kobariantza matrizea ez dagoela definituta. Hortaz, Studen-en t kopula eta dagokion kopula dentsitatea

$$\begin{aligned} C(u_1, u_2; \rho) &= \mathbf{t}_\nu(t_{\nu_1}^{-1}(u_1), t_{\nu_2}^{-1}(u_2)) = \int_{-\infty}^{t_\nu^{-1}(u_1)} \int_{-\infty}^{t_\nu^{-1}(u_2)} f(x,y)dxdy \\ &= \int_{-\infty}^{t_\nu^{-1}(u_1)}\int_{-\infty}^{t_\nu^{-1}(u_2)} \frac{1}{2\pi\sqrt{1-\rho^2}}\left(1+\frac{x^2-2\rho xy+y^2}{\nu(1-\rho^2)}\right)^{-\frac{\nu+2}{2}} dxdy \end{aligned}$$

eta

$$c(u_1, u_2; \rho) = \frac{K(\nu)}{\sqrt{1-\rho^2}}\left(1 + \frac{x^2 - 2\rho xy + y^2}{\nu(1-\rho^2)}\right)^{-\frac{\nu+2}{2}} \left(\left(1 + \frac{x^2}{\nu}\right)\left(1 + \frac{y^2}{\nu}\right)\right)^{\frac{\nu+1}{2}}$$

dira hurrenez hurren, non $K(\nu) = \frac{\nu}{2}\Gamma(\frac{\nu}{2})^2\Gamma(\frac{\nu+1}{2})^{-2}$ eta ν kopularen askatasun gradu kopurua diren.

2.1.2 Mutur balioetako kopulak

Mutur balioetako kopulak oso erabilgarriak dira banaketa funtzio marginaletan muturretako gertaerei garrantzia handiagoa ematen dieten erlazioak adierazteko. Kopula hauek independenteak eta berdinki banatutako aldagaien maximoei lotutako kopulen limiteak dira, beti ere limitea existitzen den kasuetan. Izan bedi bi dimentsiotako zorizko aldagaiez sortutako $(X_1, Y_1), ..., (X_n, Y_n)$ lagina, non aldagaiak F_X eta G_Y marginalekiko eta H_{XY} baterako banaketarekiko askeak eta berdinki banatuak diren. Bestalde, izan bitez baita $M_n = \max\{X_1, ..., X_n\}$

eta $N_n = \max\{Y_1, ..., Y_n\}$ bi aldagai, non $F^n(x) = P(M_n \leq x)$ eta $G^n(y) = P(N_n \leq y)$ dagozkien banakako banaketa funtzioak eta $H_n(x, y) = P(M_n \leq x, N_n \leq y)$ haien baterako banaketa funtzioa diren. (M_n, N_n) bikoteari eta bere limiteari lotutako C kopula *mutur balioetako kopula* bat dela esaten da. Mutur balioetako kopulek menpekotasun positiboa modelizatzen dute.

Deheuvels-en teoremaren arabera (Deheuvels, 1991), mutur balioko C kopula batek $C^{\ell}(u_1^{\ell^{-1}}, u_2^{\ell^{-1}}) = C(u, v)$ betetzen du $\ell > 0$ guztietarako eta ondorioz, mota hauetako kopulek, hurrengo teoreman azaltzen den bezala, adierazpen konkretu bat dute (Pickands, 1981).

Teorema 6. ***(Mutur balioko kopulen adierazpena)*** *C kopula bat mutur balioko kopula bat da baldin eta soilik baldin hurrengo erlazioa betetzen duen eta* $[0, 1]$ *tartean definituta dagoen F funtzio erreal bat existitzen bada:*

$$C(u_1, u_2) = exp\left\{-log(u_1u_2)F\left(\frac{log(u_2)}{log(u_1u_2)}\right)\right\}$$

edo baliokideki,

$$C(e^{-u_1}, e^{-u_2}) = exp\left\{-(u_1 + u_2)F\left(\frac{u_2}{log(u_1 + u_2)}\right)\right\}.$$

F funtzioari ***Pickands-en menpekotasun funtzioa*** *esaten zaio eta baldintza hauek betetzen ditu:*

i) *Ganbila da* $[0, 1]$ *tartean.*

ii) $\max(\ell, 1 - \ell) \leq F(\ell) \leq 1$, $\ell \in [0, 1]$ *guztietarako.*

2.1.3 Kopula arkimedearrak

Kopula arkimedearrak, finantzetako eta geoestatistikako arloetan erabiliak izan aurretik, beste arlo batzuetan aplikatuak izan dira, hots Clayton (1978), Oakes (1982, 1986) eta Cook eta Johnson (1981, 1986) lanetan. Arrazoi asko dira kopula arkimedearrak hainbeste erabiltzea bultzatu dutenak. Izan ere, inplementatzeko, bornatzeko eta simulazio desberdinak egiteko ematen duten erraztasunak, malgutasunak eta dituzten propietate matematikoek erakargarriak egiten dituzte. Hala ere, kopula arkimedearren desabantaila bat argumentuekiko simetria da, $C(u_1, u_2) = C(u_2, u_1)$.

Definizioa 10. Izan bitez $\psi : [0, 1] \to [0, \infty)$ motako funtzio jarraituak, hertsiki beherakorrak eta ganbilak, eta demagun $\psi(0) = \infty$ eta $\psi(1) = 0$ baldintzak betetzen dituztela. Izan bedi baita Ψ funtzio horien multzoa. Orduan, Ψ multzoko ψ funtzio bakoitzak

$$C(u_1, u_2) = \psi^{-1}(\psi(u_1) + \psi(u_2))$$

motako kopula bat sorrarazten du, non $u_1, u_2 \in [0,1]$. ψ funtzioari *kopularen sortzaile arkimedearra* esaten zaio eta bere sasi-alderantzizkoa ψ^{-1} hurrengo eran dago definituta:

$$\psi^{-1}(\ell) = \begin{cases} \psi^{-1}(\ell) & 0 \leq \ell < \psi(0) \\ \\ 0 & \psi(0) \leq \ell < \infty. \end{cases}$$

$\psi(0) = \infty$ bada, C kopula hertsiki arkimedearra dala esaten da.

Datozenak kopula arkimedearren familian parte hartzen duten oinarrizko kopulak dira.

Independentzia kopula

Independentzia kopula kasuan funtzio sortzailea $\psi(\ell) = -ln(\ell)$ da eta hortaz, bere alderantzizkoa $\psi^{-1}(x) = e^{-x}$. Orduan, eta aurrekoa kontuan hartuz, independentzia kopula

$$C(u_1, u_2) \;=\; \psi^{-1}(-ln(u_1) - ln(u_2)) = \psi^{-1}(-ln\ (u_1 u_2)) = e^{ln(u_1 u_2)} = u_1 u_2$$

izango da $c(u_1, u_2) = 1$ dentsitatearekin. Ohartu independentzia kopula arinago definitutako $\Pi(u_1, u_2) = u_1 u_2$ biderkadura kopula dela.

Frank kopula

Frank kopulak $\psi(\ell) = -ln\left(\frac{e^{-\delta\ell}-1}{e^{-\delta}-1}\right)$ du funtzio sortzaile bezala, $\delta \in \mathbb{R}$. Orduan, ψ-ren alderantzizko funtzioa $\psi^{-1}(x) = -\frac{1}{\delta} ln\left(1 + \frac{e^{-\delta}-1}{e^x}\right)$ dela ondorioztatu daiteke eta hortaz, Frank kopularen adierazpena

$$\begin{aligned} C(u_1, u_2) &= \psi^{-1}\left(-ln\frac{e^{-\delta u_1}-1}{e^{-\delta}-1} - ln\frac{e^{-\delta u_2}-1}{e^{-\delta}-1}\right) = \psi^{-1}\left(ln\frac{(e^{-\delta}-1)^2}{(e^{-\delta u_1}-1)(e^{-\delta u_2}-1)}\right) \\ &= -\frac{1}{\delta} ln\left(1 + \frac{(e^{-\delta u_1}-1)(e^{-\delta u_2}-1)}{e^{-\delta}-1}\right) \end{aligned}$$

da, zeinen dentsitatea

$$c(u_1, u_2) = \frac{-\delta e^{-\delta u_1} e^{-\delta u_2}(e^{-\delta}-1)}{\left((e^{-\delta}-1) + (e^{-\delta u_1}-1)(e^{-\delta u_2}-1)\right)^2}$$

den.

Gumbel kopula

Gumbel kopula kopularen kasuan funtzio sortzailea $\psi(\ell) = (-ln(\ell))^\delta$ da, non $\delta \geq 1$. ψ funtzioaren alderantzizkoa $\psi^{-1}(x) = exp(-x^{\frac{1}{\delta}})$ bada, kopula

$$C(u_1,u_2) = \psi^{-1}\big((-ln(u_1))^{\delta}+(-ln(u_2))^{\delta}\big)=exp\left(-\Big([-ln(u_1)]^{\delta}+[-ln(u_2)]^{\delta}\Big)^{\frac{1}{\delta}}\right)$$

eran definitzen da eta bere dentsitatea

$$c(u_1,u_2) = C(u_1,u_2)\big((-ln\ (u_1))^{\delta}+(-ln\ (u_2))^{\delta}\big)^{\frac{2}{\delta}+2}(ln(u_1)ln(u_2))^{\delta-1}\Big\{1+(\delta-1)[(-ln\ (u_1))^{\delta}+ (-ln\ u_2)^{\delta}]^{-\frac{1}{\delta}}\Big\}$$

adierazpenak ematen du. $\delta > 1$ bada, Gumbel kopulak menpekotasun positiboa du goiko muturrean. Hala ere, δ-k hartzen dituen balio ezberdinen arabera Gumbel independentzia kopula ($\delta = 1$) eta menpekotasun zehatz positiboaren artean mugitzen da, $\delta \to \infty$ denean Fréchetén goi-borneko kopulara konbergitzen baitu.

<u>Clayton kopula</u>

Clayton kopularen sortzailea $\psi(\ell) = \frac{1}{\delta}(\ell^{-\delta} - 1)$ eta bere alderantzizkoa $\psi^{-1}(x) = (\delta x + 1)^{-\frac{1}{\delta}}$ funtzioak dira. Orduan,

$$C(u_1,u_2) = \psi^{-1}\left(\frac{1}{\delta}\left(u_1^{-\delta} + u_2^{-\delta} - 2\right)\right) = (u_1^{-\delta} + u_2^{-\delta} - 1)^{-\frac{1}{\delta}}$$

eta

$$c(u_1,u_2) = \frac{\partial C(u_1,u_2)}{\partial u_1 \partial u_2} = (u_1 u_2)^{-\delta-1}(\delta+1)(u_1^{-\delta} + u_2^{-\delta} - 1)^{-2-\frac{1}{\delta}}$$

adierazpenek Clayton kopula eta haren dentsitatea zehazten dute. Clayton kopulak menpekotasun asimetrikoa du muturretan. $\delta > 0$ guztietarako, kopulak goiko muturrean menpekotasuna nulua eta beheko muturrean menpekotasuna positiboa ditu. Gainera, δ-ren balio desberdinetarako, Clayton kopulak menpekotasun gradu desberdinak hartu ditzake. Gumbel kopularekin bezala, Clayton kopularen menpekotasuna zehatz positiboa izango da $\delta \to \infty$ denean, Fréchetén goi-borneko kopulara konbergitzen baitu.

2.2 Kopula baldintzatuak

Zorizko bi aldagairen erlazioa aztertzeaz gain, praktikan interesgarria da hirugarren aldagai batek jatorrizko bi aldagaien arteko erlazioan nola eragiten duen aztertzea. Beraz, X, Y eta Z zorizko aldagaiak badira, helburua Z aldagaiaren $z \in Z$ edozein baliorako X eta Y aldagaien erlazioa nola aldatzen den aztertzea da. Orduan, X eta Y aldagaien baterako banaketa funtzioa eta funtzio marjinalak, $Z = z$ baldintzapean, $F_z(x, y) = F(x, y|Z = z) = P(X \leq x, Y \leq y|Z = z)$, $F_{1z}(x) = P(X \leq x|Z = z)$ eta $F_{2z}(y) = P(Y \leq y|Z = z)$ dira, hurrenez hurren. Aurrerago ikusiko den bezala, $Z = z$ balio bat emanda, X eta Y-ren menpekotasun egitura *kopula baldintzatuaren* bidez deskribatzen da. Honela, datozen emaitzek kopula baldintzatuak definituko dituzte.

Definizioa 11. Bi aldagaietako banaketa funtzio baldintzatua eskuinetik jarraitua den $F_z : \overline{\mathbb{R}}^2 \to [0, 1]$ funtzioa da eta ondoko propietateak betetzen ditu:

i) $F_z(y_1, -\infty) = F_z(-\infty, y_2) = 0$ eta $F_z(\infty, \infty) = 1$.

ii) $V_{F_z}([x_1, x_2] \times [y_1, y_2]) \equiv F_z(x_2, y_2) - F_z(x_1, y_2) - F_z(x_2, y_1) + F_z(x_1, y_1) \geq 0$, $x_1, x_2, y_1, y_2 \in \overline{\mathbb{R}}$ guztietarako, non $x_1 \leq x_2$, $y_1 \leq y_2$.

Lehenengo baldintzak banaketa funtzioaren goi eta behe borneak ematen ditu eta bigarrenak $[x_1, x_2] \times [y_1, y_2]$ eremuko edozein punturen probabilitatea ez-negatiboa dela ziurtatzen du.

Definizioa 12. Kopula baldintzatu bat $C_z : [0, 1]^2 \to [0, 1]$ bezalako funtzio bat da eta $u_1, u_2, v_1, v_2 \in [0, 1]$ guztietarako hurrengo baldintzak betetzen ditu:

i) $C_z(u_1, 0) = C_z(0, u_2) = 0$.

ii) $C_z(u_1, 1) = u_1$ eta $C_z(1, u_2) = u_2$.

iii) $V_{C_z}([u_1, v_1] \times [u_2, v_2]) \equiv C_z(v_1, v_2) - C_z(u_1, v_2) - C_z(v_1, u_2) + C_z(u_1, u_2) \geq 0$, $u_1 \leq u_2$ eta $v_1 \leq v_2$ izanik.

Banaketa funtzioekin bezala, lehenengo baldintzak kopularen goi eta behe borneak ematen ditu. Gainera, *ii*) baldintzak banaketa marjinalak uniformeak izatea ziurtatzen du. Azkenik, *iii*) baldintzak puntu baten probabilitatearen ez-negatibotasuna adierazten du.

Aurreko hiru baldintzak $\overline{\mathbb{R}}^2$ eremura hedatuz, kopula baldintzatua U_1 eta U_2 zorizko aldagaien baterako banaketa funtzio baldintzatua moduan definitu daiteke, non marjinalek $[0, 1]$ tartean banaketa uniformea jarraitzen duten:

$$C_z^*(u_1,u_2) = \begin{cases} 0 & u_1 < 0 \textit{ edo } u_2 < 0, \\ C_z(u_1,u_2) & (u_1,u_2) \in [0,1] \times [0,1], \\ u_1 & u_1 \in [0,1],\ u_2 > 1, \\ u_2 & u_1 > 1,\ u_2 \in [0,1], \\ 1 & u_1 > 1,\ u_2 > 1. \end{cases}$$

Arinago ikusitako Sklar-en bi teoremei lotuta badago kopula baldintzatueri egokitutako Sklar-en teorema ere (Patton, 2006).

Teorema 7. *Izan bitez F_z bi aldagairen baterako banaketa funtzioa eta F_{1z} eta F_{2z} bere marjinal jarraituak. Orduan, $C_z : [0,1]^2 \to [0,1]$ kopula baldintzatu bakarra existitzen da, non $x, y \in \overline{\mathbb{R}} = [-\infty, \infty]$ guztietarako,*

$$F_z(x,y) = C_z(F_{1z}(x), F_{2z}(y)). \tag{1.2}$$

Alderantziz, C_z kopula baldintzatua eta F_{1z} eta F_{2z}, X eta Y bi zorizko aldagaien marjinal baldintzatuak badira, orduan (1.2) ekuazioarekin definitutako F_z funtzioa F_{1z} eta F_{2z} marjinalak dituen banaketa funtzio baldintzatua da.

Kopula baldintzatuetarako Sklar-en teoremaren bitartez, kopula baldintzatua baterako eta banakako banaketa funtzioen bitartez adierazi daiteke baldintzatugabeko kopuletarako egin daitekeen antzerako era batean.

Korolarioa 8. *Izan bitez F_z baterako banaketa funtzioa, C_z 2-kopula, F_{1z} eta F_{2z} marjinalak eta F_{1z}^{-1} eta F_{2z}^{-1} azken hauen sasi-alderantzizkoak, hurrenez hurren. Orduan, $\mathbf{u} \in [0,1]^2$ guztietarako, $C_z(u_1,u_2) = F_z(F_{1z}^{-1}(u_1), F_{2z}^{-1}(u_p))$.*

3 Aldagai anitzeko kopulak

Atal honetan aurreko ataletako emaitzak aldai anitzeko kasura orokortuko ditugu. Hala ere, baliokidetasuna ez da kasu guztietan ematen eta, hori dala eta, kontu handiz jokatu behar da. Lehenik eta behin, definitu ditzagun aurrerago baliagarriak izango diran zenbait kontzeptu.

Definizioa 13. Izan bedi p zenbaki oso positibo bat. *p-kutxa* bat $\overline{\mathbb{R}}$-ko p tarte itxiren arteko edozein biderketa kartesiar da. Edozein $J = [a_1, b_1] \times ... \times [a_p, b_p]$ p-kutxa $J = [\mathbf{a}, \mathbf{b}] =$

$[(a_1, ..., a_p), (b_1, ..., b_p)]$ ere idazten da. Bereziki, $a_i = a$ eta $b_i = b$ bada $i = 1, ..., p$ guztietarako, $J = [a, b]^p$ idatzi daiteke.

Gainera, *J p-kutxa endekatua* dela esaten da baldin eta $1 \leq i \leq p$ existitzen bada non $a_i = b_i$. Bestela, J *p-kutxa ez endekatua* edo *endekatu gabea* dela esaten da. Beste alde batetik, J-ren *erpinak* $(c_1, ..., c_p)$ puntuak dira, non $i = 1, ..., p$ bakoitzerako $c_i = a_i$ edo $c_i = b_i$ betetzen da. Endekatu gabeko J kutxa batek 2^p erpin ditu.

Definizioa 14. Izan bitez $S_1, S_2, ..., S_p$, $\overline{\mathbb{R}}$-ren azpimultzoak eta F p-dimentsioko funtzio errala, non $D(F) = S_1 \times S_2 \times ... \times S_p$ bere izate eremua den. Izan bedi baita erpinak $D(F)$ eremuan dituen J p-kutxa endekatu gabea. Orduan, *J p-kutxaren F-bolumena* $V_F(J) = \sum_i sgn(\mathbf{c})F(\mathbf{c})$ adierazpenaren bidez emanda dago, non batukaria J-ren $\mathbf{c}$ erpin guztien gainean hartzen den eta

$$sgn(\mathbf{c}) = \begin{cases} 1 & c_k = a_k \text{ bada } k \text{ indize kopurua bikoitia denean} \\ \\ -1 & c_k = a_k \text{ bada } k \text{ indize kopurua bakoitia denean.} \end{cases}$$

$J = [\mathbf{a}, \mathbf{b}]$-ren *F-bolumena* beste modu baliokide batean ere definitu daiteke. Hain zuzen ere, J-ren F-bolumena F-ren J gaineko p ordenako aldakuntza da, hau da,

$$V_F(J) = \Delta_{\mathbf{a}}^{\mathbf{b}} F(\boldsymbol{\ell}) = \Delta_{a_p}^{b_p} \Delta_{a_{p-1}}^{b_{p-1}} ... \Delta_{a_2}^{b_2} \Delta_{a_1}^{b_1} F(\boldsymbol{\ell}),$$

non lehenengo ordenako p aldakuntzak $\Delta_{a_k}^{b_k} F(\boldsymbol{\ell}) = F(\ell_1, ..., \ell_{k-1}, b_k, \ell_{k+1}, ..., \ell_p) - F(\ell_1, ..., \ell_{k-1}, a_k, \ell_{k+1}, ..., \ell_p)$ diren.

Adibidea 3. Izan bitez $F \in \overline{\mathbb{R}}^3$ funtzioa eta J, $[x_1, x_2] \times [y_1, y_2] \times [z_1, z_2]$ 3-kutxa. Orduan, J-ren F-bolumena

$$\begin{aligned} V_F(J) &= \Delta_{\mathbf{a}}^{\mathbf{b}} F(\boldsymbol{\ell}) = \Delta_{a_3}^{b_3} \Delta_{a_2}^{b_2} \Delta_{a_1}^{b_1} F(\boldsymbol{\ell}) = \Delta_{a_3}^{b_3} \Delta_{a_2}^{b_2} [F(b_1, \ell_2, \ell_3) - F(a_1, \ell_2, \ell_3)] \\ &= \Delta_{a_3}^{b_3} \{F(b_1, b_2, \ell_3) - F(b_1, a_2, \ell_3) - F(a_1, b_2, \ell_3) + F(a_1, a_2, \ell_3)\} \\ &= F(b_1, b_2, b_3) - F(b_1, b_2, a_3) - F(b_1, a_2, b_3) + F(b_1, a_2, a_3) \\ &\quad -F(a_1, b_2, b_3) + F(a_1, b_2, a_3) + F(a_1, a_2, b_3) - F(a_1, a_2, a_3) \\ &= F(x_2, y_2, z_2) - F(x_2, y_2, z_1) - F(x_2, y_1, z_2) + F(x_2, y_1, z_1) \\ &\quad -F(x_1, y_2, z_2) + F(x_1, y_2, z_1) + F(x_1, y_1, z_2) - F(x_1, y_1, z_1) \end{aligned}$$

da, kasu honetan $(a_1, a_2, a_3) = (x_1, y_1, z_1)$ eta $(b_1, b_2, b_3) = (x_2, y_2, z_2)$ betetzen dela kontuan izanda.

Definizioa 15. $F \in \overline{\mathbb{R}}^p$ funtzio erreala *p-gorakorra* dela esaten da $V_F(J) \geq 0$ bada erpinak $D(F)$ eremuan dituen J p-kutxa guztietarako. Bestela, F *p-beherakorra* dela esaten da.

Gainera, F funtzioa p-beherakorra da baldin eta soilik baldin $-F$ funtzioa p-gorakorra bada. Bestalde, J kutxa endekatu gabea bada, orduan $V_F(J) = 0$ da.

Demagun $F \in \overline{\mathbb{R}}^p$ funtzio erreal baten eremua $D(F) = S_1 \times ... \times S_p$ dela, non S_k bakoitzak a_k elementu minimo bat duen. F *oinarrituta* dagoela esaten da gutxienez k baterako $\ell_k = a_k$ diren $\boldsymbol{\ell} \in D(F)$ guztietarako $F(\boldsymbol{\ell}) = 0$ betetzen bada. S_k bakoitza ez-hutsa bada eta b_k elementu handien bat badu, orduan F-k marjinalak izango ditu. *F-ren dimentsio bateko marjinalak* $D(F_k) = S_k$ eremuan $x \in S_k$ guztietarako definituta dauden

$$F_k(x) = F(b_1, ..., b_{k-1}, x, b_{k+1}, ..., b_p) \tag{1.3}$$

funtzioak dira.

Adibidea 4. Izan bedi $[-1, 1] \times [0, \infty] \times [0, \frac{\pi}{2}]$ eremuan definitutako hurrengo funtzioa:

$$F(x, y, z) = \frac{(x+1)(e^y - 1)sin(z)}{x + 2e^y - 1}.$$

$F(0, y, z) = 0$, $F(x, 0, z) = 0$ eta $F(x, y, 0) = 0$ betetzen diranez, F oinarrituta dago eta bere 1-dimentsioko marjinalak $F_1(x) = F(x, \infty, \frac{\pi}{2}) = \frac{(x+1)(e^\infty-1)}{x+2e^\infty-1} \to \frac{x+1}{2}$, $F_2(y) = F(1, y, \frac{\pi}{2}) = \frac{e^y-1}{e^y} = 1 - e^{-y}$ eta $F_3(z) = F(1, \infty, z) = \frac{e^\infty-1}{e^\infty} sin(z) = sin(z)$ dira.

F-ren bi dimentsiotako marjinalak aldiz, $F_{1,2}(x, y) = F(x, y, \frac{\pi}{2}) = \frac{(x+1)(e^y-1)}{x+2e^y-1}$, $F_{2,3}(y, z) = F(1, y, z) = \frac{e^y-1}{e^y} sin(z) = (1 - e^{-y}) sin(z)$ eta $F_{1,3}(x, z) = F(x, \infty, z) = \frac{(x+1)(e^\infty-1)sin(z)}{x+2e^\infty-1} \to \frac{(x+1)sin(z)}{2}$ izango dira.

Aurrerantzean dimentsio bateko marjinalei *marjinala* bakarrik esango diegu eta $k \geq 2$ denean, k-dimentsioko marjinalak *k-marjinalak* izango dira.

Lema 9. *Izan bitez* $S_1, ..., S_p$ $\overline{\mathbb{R}}$*-ren azpimultzo ez hutsak eta* F, $S_1 \times ... \times S_p$ *eremuan definitutako funtzio n-gorakor oinarritua. Orduan,* F *ez-beherakorra da argumentu bakoitzean; hau da,* $(\ell_1, ..., \ell_{k-1}, x, \ell_{k+1}, ..., \ell_p)$ *eta* $(\ell_1, ..., \ell_{k-1}, y, \ell_{k+1}, ..., \ell_p)$ F*-ren eremuan badaude eta* $x < y$ *bada, orduan*

$$F(\ell_1, ..., \ell_{k-1}, x, \ell_{k+1}, ..., \ell_p) \leq F(\ell_1, ..., \ell_{k-1}, y, \ell_{k+1}, ..., \ell_p).$$

Definizioa 16. p-dimentsioko banaketa funtzio bat $\overline{\mathbb{R}}^p$-n definitutako F funtzio bat da eta propietate bi betetzen ditu:

i) F p-gorakorra da.

ii) $\boldsymbol{\ell} \in \overline{\mathbb{R}}^p$ guztietarako $F(\infty, ..., \infty) = 1$ eta $F(\boldsymbol{\ell}) = 0$, non gutxienez k baterako $\ell_k = -\infty$ den.

Hortaz, F oinarritua da bigarren baldintza dela eta. Gainera, $D(F) = \overline{\mathbb{R}}$ denez, (1.3)-ren bitartez adierazitako p-dimentsioko banaketa funtzioaren marjinalak banaketa funtzioak dira era berean. Banaketa funtzio marjinal hauei $F_1, ..., F_p$ esango zaie $p \geq 3$ denean.

Definizioa 17. *p dimentsioko azpikopula edo p-azpikopula* C' funtzio bat da hurrengo propietateak betetzen dituena:

i) C'-ren eremua $D(C') = S_1 \times ... \times S_p$ da, non S_k bakoitza 0 eta 1 barne dituen $[0, 1]$-ren azpimultzo bat den.

ii) C' oinarritua eta p-gorakorra da.

iii) C'-k (dimentsio bateko) C'_k marjinalak ditu, $k = 1, ..., p$, eta $u \in S_k$ guztietarako $C'_k(u) = u$ betetzen dute.

Bi dimentsioko kopuletan bezala, $\mathbf{u} \in D(C')$ guztietarako $0 \leq C'(\mathbf{u}) \leq 1$, beraz $rang(C')$ ere $[0, 1]$-ren azpimultzo bat da. Ondorioz, p-kopulak hurrengo definizioaren bidez definitu ditzakegu.

Definizioa 18. *p dimentsioko kopula bat edo p-kopula bat* $[0, 1]^p$ eremuan definitutako C azpikopula bat da. Baliokideki, *p-kopula* $C : [0, 1]^p \to [0, 1]$ eran definitutako funtzio bat da eta ondoko propietateak betetzen ditu:

i) $\mathbf{u} \in [0, 1]^p$ guztietarako, gutxienez $\mathbf{u}$-ren koordenatu bat 0 bada $C(\mathbf{u}) = 0$ da eta u_k izan ezik $\mathbf{u}$-ren beste koordenatu guztiak 1 badira, orduan $C(\mathbf{u}) = u_k$.

ii) $V_C([\mathbf{a}, \mathbf{b}]) \geq 0$ $\mathbf{a}, \mathbf{b} \in [0, 1]^p$ guztietarako non $\mathbf{a} \leq \mathbf{b}$.

Adibidea 5. Izan bedi $C(u_1, u_2, u_3) = u_3 \min(u_1, u_2)$. Orduan, $u_3 = 0$ bada, $C(u_1, u_2, 0) = 0$ betetzen da, eta $u_1 = u_2 = 1$ bada, $C(1, 1, u_3) = u_3$.

Beste alde batetik, izan bedi $J = [a_1, b_1] \times [a_2, b_2] \times [a_3, b_3]$, non $a_k \leq b_k$. Orduan, J-ren C-bolumena

$$\begin{aligned} V_C(J) &= \Delta_{a_2}^{b_2} \Delta_{a_1}^{b_1} (C(u_1, u_2, b_3) - C(u_1, u_2, a_3)) = \Delta_{a_2}^{b_2} \Delta_{a_1}^{b_1} ((b_3 - a_3) \min(u_1, u_2)) \\ &= (b_3 - a_3) \Delta_{a_2}^{b_2} \Delta_{a_1}^{b_1} \min(u_1, u_2) \geq 0. \end{aligned}$$

Ohartu C-ren 2-marjinalak 2-kopulak direla lehenago ikusi bezala:

$$C_{1,2}(u_1,u_2) = C(u_1,u_2,1) = \min(u_1,u_2) = M(u_1,u_2)$$

$$C_{1,3}(u_1,u_3) = C(u_1,1,u_3) = u_3 \min(u_1,1) = \Pi(u_1,u_3)$$

$$C_{2,3}(u_2,u_3) = C(1,u_2,u_3) = u_3 \min(1,u_2) = \Pi(u_2,u_3)$$

Zorizko aldagaietarako definitutako Sklar-en teormaren p-dimentsiorako orokorpena ere modu zuzenean egin daiteke.

Teorema 10. ***(Sklar-en teorema p dimentsiotan)*** *Izan bitez F, $X_1, ..., X_p$ zorizko aldagaien baterako banaketa funtzioa eta $F_1, ..., F_p$ dagozkien banaketa funtzio marjinalak. Orduan, p-dimentsioko $C : [0,1]^p \to [0,1]$ kopula bat existitzen da non $(x_1, ..., x_p) \in \overline{\mathbb{R}}^p$ guztietarako,*

$$F(x_1, x_2, ..., x_p) = C(F_1(x_1), F_2(x_2), ..., F_p(x_p)). \tag{1.4}$$

$F_1, ..., F_p$ banaketa funtzio marjinal guztiak jarraituak badira, orduan C kopula bakarra da eta hortaz, modu unibokoan zehaztuta dago $rang(F_1) \times ... \times rang(F_p)$ eremuan. Alderantziz, C kopula bat bada eta $F_1, ..., F_p$ banaketa funtzioak badira, orduan (1.4) *adierazpenean definitutako F funtzioa p-dimentsioko baterako banaketa funtzioa da $F_1, ..., F_p$ banaketa funtzio marjinalak dituena.*

Korolarioa 11. *Izan bitez F baterako banaketa funtzioa, C p-kopula eta $F_1, ..., F_p$ marjinalak. Izan bitez baita $F_1^{-1}, ..., F_p^{-1}$ aurreko $F_1, ..., F_p$ banaketa funtzioen sasi-alderantzizkoak, hurrenez hurren. Orduan, $\mathbf{u} \in [0,1]^p$ guztietarako, $C(u_1, ..., u_p) = F(F_1^{-1}(u_1), ..., F_p^{-1}(u_p))$.*

3.1 Kopula motak

Atal honetan aurretik bi aldagaietarako aztertutako kopula parametrikoak aldagai anitzetara orokortuko ditugu.

3.1.1 Kopula eliptikoak

Gauss-en kopula

Bi aldagaietako kopula $p > 2$ dimentsiotara modu zuzen batean hedatu daiteke. Bi aldagairen kasuan baterako banaketa eta marjinal normaletik ondorioztatzen den bezala, p-dimentsioko kasuan ere aldagai anitzeko banaketa funtziotik eta marjinal bakunetatik ondorioztatzen da, hots,

$$C(u_1, ..., u_p) = P(\Phi(X_1) \le u_1, ..., \Phi(X_p) \le u_p) = \Phi_p(\Phi^{-1}(u_1), ..., \Phi^{-1}(u_p)).$$

Kontuan izan kopula honek ez duela adierazpen espliziturik onartzen eta integral baten bitartez baino ezin daitekela adierazi. Gauss-en kopulan parametroak $\mathbf{P}$ korrelazio matrizearen osagaiak dira. Kopula independentea ($\mathbf{P} = \mathbf{I}_p$ identitate matrizea denean) eta goi borne kopula ($\mathbf{P} = \mathbf{1}_{p\times p}$ matrizea denean) Gauss-en kopularen kasu partikularrak dira. Gainera, Gauss-en aldagai anitzeko kopularen dentsitatea

$$c(u_1, ..., u_p; \mathbf{P}) = |\mathbf{P}|^{-\frac{1}{2}} exp(-\frac{1}{2}\boldsymbol{\xi}'(\mathbf{P}^{-1} - I_p)\boldsymbol{\xi})$$

adierazpenaren bidez emanda dago, non $\boldsymbol{\xi} = (\xi_1, ..., \xi_p)'$ eta ξ_i banaketa normal estandarraren koantilak diren, $i = 1, ..., p$. Aurreko Gauss-en kopula, $(u_1, ..., u_p)$ puntuan ebaluatutako funtzio moduan kontsideratu behar da, ez $\boldsymbol{\xi} = (\xi_1, ..., \xi_p)'$-ren funtzio bat bezala.

Student-en t kopula

Student-en t kopula Student-en aldagai anitzeko t banaketa funtzioari lotutako menpekotasun funtzioa da. Honela, $X_1, ..., X_p$ zorizko aldagaietarako, Student-en $t \sim \mathbf{t}_p(\nu, \mu, \Sigma)$ aldagai anitzeko banaketaren dentsitate funtzioa

$$f(\mathbf{x}) = \frac{\Gamma(\frac{\nu+p}{2})}{\Gamma(\frac{\nu}{2})\sqrt{(\pi\nu)^p|\Sigma|}}\left(1 + \frac{(\mathbf{x}-\boldsymbol{\mu})'\Sigma^{-1}(\mathbf{x}-\boldsymbol{\mu})}{\nu}\right)^{-\frac{\nu+p}{2}}$$

da, non ν askatasun graduak, $\boldsymbol{\mu}$ batazbestekoen bektorea eta Σ positiboki definitutako dispertsio matrizea diren. Aldagai anitzeko Student-en t banaketak $\mathbf{X} \stackrel{d}{=} \boldsymbol{\mu} + \sqrt{\frac{\nu}{S}}\mathbf{Z}$ adierazpena du, non $\mathbf{Z} \sim N(0, \Sigma)$ eta $S \sim \chi^2_\nu$ diren. Beraz, Studen-en t kopula ondoko eran definitu daiteke:

$$\begin{aligned} C(u_1, ..., u_p; \mathbf{P}) &= \mathbf{t}_\nu(t^{-1}_{\nu_1}(u_1), ..., t^{-1}_{\nu_p}(u_p)) = \int_{-\infty}^{t^{-1}_{\nu_1}(u_1)} \cdots \int_{-\infty}^{t^{-1}_{\nu_p}(u_p)} f(\mathbf{x})d\mathbf{x} \\ &= \int_{-\infty}^{t^{-1}_{\nu_1}(u_1)} \cdots \int_{-\infty}^{t^{-1}_{\nu_p}(u_p)} \frac{\Gamma(\frac{\nu+p}{2})}{\Gamma(\frac{\nu}{2})\sqrt{(\pi\nu)^p|\Sigma|}}\left(1+\frac{(\mathbf{x}-\boldsymbol{\mu})'\Sigma^{-1}(\mathbf{x}-\boldsymbol{\mu})}{\nu}\right)^{-\frac{\nu+p}{2}} d\mathbf{x}. \end{aligned}$$

Ikusi daitekeen moduan, Student-en t baterako banaketaren eta marjinalen askatasun graduek ez dute berdinak izan behar. Gainera, Student-en t kopularen dentsitatea

$$c(u_1, ..., u_p; \mathbf{P}) = K(\nu)\frac{1}{|P|^{1/2}}\left(1 + \frac{\mathbf{x}'\mathbf{P}^{-1}\mathbf{x}}{\nu}\right)^{-\frac{\nu+p}{2}} \prod_{i=1}^{p}\left(1 + \frac{x_i^2}{\nu}\right)^{\frac{\nu+1}{2}}$$

adierazpenaren bidez emonda dago, non $\mathbf{x}$=$(T^{-1}_{\nu_1}(u_1), ..., T^{-1}_{\nu_p}(u_p))$ eta $K(\nu) = \Gamma(\frac{\nu}{2})^{p-1}\Gamma(\frac{\nu+1}{2})^{-p} \times \Gamma(\frac{\nu+p}{2})$ diren. Kasu honetan $\mathbf{P} = \mathbf{I}_p$ denean ez da kopula independentea lortzen, aldagai anitzeko t banaketan korrelazio hutsak ez baitu independentzia inplikatzen. Aldiz, $\mathbf{P} = \mathbf{1}_{p\times p}$ denean goi borne kopula lortzen da oraingoan ere.

3.1.2 Mutur balioetako kopulak

M, Π eta W oinarrizko 2-kopulen p-dimentsiotarako hedapenak M^p, Π^p eta W^p bezala izendatzen dira,

$$\begin{aligned} M^p(\mathbf{u}) &= \min(u_1, ..., u_p), \\ \Pi^p(\mathbf{u}) &= u_1 \cdots u_p, \\ W^p(\mathbf{u}) &= \max(u_1 + ... + u_p - p + 1, 0). \end{aligned}$$

M^p eta Π^p funtzioak p-kopulak dira $p \geq 2$ guztietarako. W^p funtzioak aldiz, ez ditu p-kopula izateko baldintzak betetzen $p > 2$ danean. Hala ere, posible da Fréchet-Hoeffding borneen inekuazioaren orokorpena ezartzea.

Teorema 12. *C' edozein p-azpikopularako eta $\mathbf{u} \in D(C')$ guztietarako, $W^p(\mathbf{u}) \leq C'(\mathbf{u}) \leq M^p(\mathbf{u})$.*

Azkenik, mutur balioetako kopulekin erlazionatutako ondoko emaitza aurkezten dugu.

Teorema 13. *Izan bitez $X_1, ..., X_p$ zorizko aldagai jarraituak, non $p \geq 2$. Orduan,*

i) *$X_1, ..., X_p$ askeak edo independenteak dira baldin eta soilik baldin $X_1, ..., X_p$-ren p-kopula Π^p bada.*

ii) *$X_1, ..., X_p$ zorizko aldagai bakoitza multzoko beste edozein aldagairen funtzio hertsiki gorakorra da baldin eta soilik baldin $X_1, ..., X_p$ aldagaien p-kopula M^p bada.*

3.1.3 Kopula arkimedearrak

Mutur balioetako kopulekin egiten den antzera, kopula arkimedearren definizioa ere zuzenean orokortu daiteke:

$$C(\boldsymbol{u}) = \psi^{-1}(\psi(u_1) + \ldots + \psi(u_p)), \tag{1.5}$$

non $u_i \in [0, 1], i = 1, \ldots, p$. C funtzioak 2-kopula arkimedearrak iteratuz lortu daitezke, hau da, $C(\boldsymbol{u}) = C(C(u_1, \ldots, u_{p-1}), u_p)$. Hala ere, kopulak konposatzeko teknika honek ez du beti balio. Izan ere, $\psi(\ell) = 1 - \ell$ funtzio sortzailea (1.5) adierazpenean ordezkatuz W^p funtzioa lortzen da eta arinago esan bezala, W^p ez da kopula bat $p \geq 3$ denean. Zentzu honetan ψ funtzioaren propietateak zeintzuk izan behar diren aztertzea beharrezkoa da.

Definizioa 19. $g(\ell)$ funtzio bat J tarte batean *guztiz monotonoa* da tarte horretan jarraitua bada eta ordena guztietako deribatu positiboak baldin baditu, hau da, $\ell \in J$ eta $k > 0$ guztietarako $(-1)^k \frac{\partial^k g(\ell)}{\partial \ell^k} \geq 0$ betetzen bada.

Ondorioz, $c>0$ batzuetarako $g(c)=0$ betetzen bada eta $[0,\infty)$ tartean g funtzioa guztiz monotonoa bada, orduan g zeroren bardina da $[0,\infty)$ tartean. Hori dela eta, ψ^{-1} sasi-alderantzizkoa guztiz monotonoa bada, orduan funtzio positiboa izango da $[0,\infty)$ tartean. Hurrengo teoremak baldintza beharrezkoak eta nahikoak ematen ditu ψ funtzio bat p-kopula arkimedear baten funtzio sortzailea izateko.

Teorema 14. *Izan bitez $\psi:[0,1]\to[0,\infty]$ hertsiki beherakorra eta jarraitua den funtzio bat, non $\psi(0)=\infty$, $\psi(1)=0$ eta ψ^{-1} bere alderantzizkoa den. $C:[0,1]^p\to[0,1]$ funtzioa (1.5) adierazpenaren bidez definitutako funtzioa bada, orduan $p>2$ guztietarako C p-kopula bat da baldin eta soilik baldin $[0,\infty)$ tartean ψ^{-1} guztiz monotonoa bada.*

Orokorrean kopula arkimedearren aldagai anitzetarako dentsitate funtzioa

$$c(\mathbf{u})=\psi^{-1}_{(p)}(\psi(u_1)+\ldots+\psi(u_p))\prod_{i=1}^{p}\psi'(u_i)$$

adierazpenaren bidez emanda dago. p-kopulak eraikitzeko errazak izan arren, muga batzuk ere badituzte. Kontuan izan behar dugu p-kopula arkimedear baten k-marjinal guztiak orokorrean berdinak direla. Gainera, parametro bat edo bi egoteak menpekotasun egituraren izaera mugatu egiten du.

Independentzia kopula

Kontuan izanik independentzia kopularen funtzio sortzailea $\psi(\ell)=-ln(\ell)$ eta bere alderantzizkoa $\psi^{-1}(x)=e^{-x}$ direla, aldagai anitzeko independentzia kopula eta dagokion kopula dentsitatea $C(\mathbf{u})=u_1\cdots u_p=\Pi^p$ eta $c(\mathbf{u})=1$ dira, hurrenez hurren.

Frank kopula

Frank kopularen kasuan, bere aldagai anitzeko kopula $\psi(\ell)=-ln\left(\frac{e^{-\delta\ell}-1}{e^{-\delta}-1}\right)$ funtzio sortzailerako

$$C(\mathbf{u})=\psi^{-1}\left(-ln\left(\frac{e^{-\delta u_1}-1}{e^{-\delta}-1}\right)-\ldots-ln\left(\frac{e^{-\delta u_p}-1}{e^{-\delta}-1}\right)\right)=-\frac{1}{\delta}ln\left(1+\frac{\Pi_{i=1}^{p}(e^{-\delta u_i}-1)}{(e^{-\delta}-1)^{p-1}}\right)$$

izango da.

Gumbel kopula

Funtzio sortzailea $\psi(\ell)=(-ln(\ell))^\delta$ eran definituta duela kontuan izanik, non $\delta\geq 1$ den, Gumbel-en p-kopula hurrengoa da:

$$C(\mathbf{u})=\psi^{-1}\big((-ln\ u_1)^\delta+\ldots+(-ln\ u_p)^\delta\big)=exp\left(-\left[(-ln\ u_1)^\delta+\ldots+(-ln\ u_p)^\delta\right]^{\frac{1}{\delta}}\right).$$

Clayton kopula

Clayton kopularen kasuan funtzio sortzailea $\psi(\ell) = \frac{1}{\delta}(\ell^{-\delta} - 1)$ denez, Clayton-en p-kopula hurrengo eran definitzen da:

$$C(\mathbf{u}) = \psi^{-1}\left(\frac{1}{\delta}\sum_{i=1}^{p}(u_i^{-\delta}-1)\right) = \psi^{-1}\left(\frac{1}{\delta}\left(\sum_{i=1}^{p}u_i^{-\delta}-p\right)\right) = \left(\sum_{i=1}^{p}u_i^{-\delta}-p+1\right)^{-\frac{1}{\delta}}.$$

Kasu honetan erraza da Clayton-en p-kopularen dentsitatea kalkulatzea:

$$c(\mathbf{u}) = \left(\sum_{i=1}^{p}u_i^{-\delta} - p + 1\right)^{-\frac{1}{\delta}-p} \prod_{i=1}^{p}(1 + (i-1)\delta)\, u_i^{-\delta-1}.$$

3.2 Kopula baldintzatuak

Aurreko ataletan kopula baldintzatuak bi aldagairen kasuan aztertu ditugu. Orain, emaitza horiek aldagai anitzen kasura hedatuko ditugu. Izan bitez $Y_1, \ldots, Y_p$ eta Z zorizko aldagaiak. Orduan, $(Y_1, \ldots, Y_p)$ aldagaien banaketa marjinalak eta baterako banaketa funtzioa, $Z = z$ baldintzapean, $F_{1z}(y_1) = P(Y_1 \leq y_1 | Z = z), \ldots, F_{pz}(y_p) = P(Y \leq y | Z = z)$ eta $F_z(y_1, \ldots, y_p) = P(Y_1 \leq y_1, \ldots Y_p \leq y_p | Z = z)$ dira, hurrenez hurren.

Definizioa 20. Kopula baldintzatu bat $C_z : [0,1]^p \to [0,1]$ funtzio bat da hurrengo propietateak betetzen dituena:

i) $\mathbf{u} \in I^p$ guztietarako eta $z \in Z$ bakoitzerako gutxienez $\mathbf{u}$-ren koordenatu bat 0 baldin bada, $C_z(\mathbf{u}) = 0$ betetzen da, eta u_k izan ezik $\mathbf{u}$-ren beste koordenatu guztiak 1 badira, orduan $C_z(\mathbf{u}) = u_k$.

ii) $V_{C_z}([\mathbf{a}, \mathbf{b}]) \geq 0$ betetzen da $z \in Z$ eta $\mathbf{a}, \mathbf{b} \in I^p$ guztietarako, non $\mathbf{a} \leq \mathbf{b}$ den.

Bi adagaien kasuan egin den antzerako era batean, kopula baldintzatuetarako Sklar-en teorema aldagai anitzeko analisira egokitu daiteke.

Teorema 15. *Izan bitez F_z $Y_1, ..., Y_p$ aldagaien baterako banaketa funtzioa eta $F_{1z}, \ldots, F_{pz}$ aldagai hoiei dagozkien banaketa funtzio marjinal jarraiak. Orduan, $C_z : [0,1]^p \to [0,1]$ kopula baldintzatu bakarra existitzen da non $(y_1, \ldots, y_p) \in \overline{\mathbb{R}}^p$ guztietarako,*

$$F_z(y_1, \ldots, y_p) = C_z(F_{1z}(y_1), \ldots, F_{pz}(y_p)) \tag{1.6}$$

den. Alderantziz, C_z kopula baldintzatua eta $Y_1, \ldots, Y_p$ zorizko aldagaien marjinal baldintzatuak $F_{1z}, \ldots, F_{pz}$ badira, orduan aurreko (1.6) adierazpenarekin definitutako F_z funtzioa $F_{1z}, \ldots, F_{pz}$ marjinalak dituen banaketa funtzio baldintzatua da.

Arinago egin den bezela, kopula baldintzatua baterako eta banakako banaketa funtzioen menpean adierazi daiteke aurreko teoreman oinarrituz. Beraz, F_z aldagai anitzeko baterako banaketa funtzioa, C_z p-kopula eta $F_{1z}, ..., F_{pz}$ banaketa marjinalak baldin badira, $(u_1, ..., u_p) \in [0,1]^p$ guztietarako $C_z(u_1, ..., u_p) = F_z(F_{1z}^{-1}(u_1), ..., F_{pz}^{-1}(u_p))$.

Aldagai anitzeko kopula baldintzatuak orokortzeko era bat kopula ez-parametrikoak erabiltzea da. Hauek ez dute banaketaren forma funtzional zehatzik jarraitzen eta ondorioz, kopula parametrikoekiko ematen duten malgutasuna abantaila bat da banaketa asimetrikoak eta buztan pisutsuak dituzten aldagai multzoak ditugunean. Horrela, aldagai anitzeko kopula baldintzatua estimatzeko hurrengo estimatzaile ez-parametrikoa erabiliko dugu:

$$\hat{C}_z(\boldsymbol{u}) \;=\; \sum_{i=1}^{n} w_i(z,h) I\{Y_{1i} \leq \hat{F}_{1z}^{-1}(u_1), ..., Y_{pi} \leq \hat{F}_{pz}^{-1}(u_p)\},$$

non $\{w_i(z,h)\}$, $(z-Z_i)/h$ puntuaren menpekoa den pisuen segida bat eta h leuntasun parametroa diren. $I\{\cdot\}$ funtzio indikatzailea da eta $\hat{F}_{jz}(y) = \sum_{i=1}^{n} w_i(z,h) I\{Y_{ji} \leq y\}$, Y_j aldagaiaren $F_{jz}^{-1}(u) = inf\{y : F_{jz}(y) \geq u\}$ baldintzatutako kuantil funtzioaren estimatzaile ez-parametrikoa. Nadaraya-Watson pisuak erabiliz gero, pisuen segida $\{w_i(z,h)\} = k((z - Z_i)/h)/\sum_j k((z - Z_j)/h)$ bezala definituta geratzen da, non k kernel funtzio bat den.

4 Ondorioak

Kapitulu honetan kopulei buruzko oinarrizko definizio eta emaitza teorikoak azaltzen dira. Lehenengo azpiatalean bi aldagaien kasuari dagozkion emaitzak erakusten ditugu eta bigarrenean emaitza hauen aldagai anitzen kasurako orokorpenak. Emaitza garrantzitsuenen artean Sklar-en teorema aurkitzen dugu, baterako banaketa funtzioa eta banaketa marjinalen arteko lotura kopula funtzioaren bidez ezartzen duena. Kopula ezagunenak kopula parametrikoak dira, adierazpen itxi baten bitartez emanda daudenak. Kopulak irizpide ezberdinen arabera sailkatu daitezke. Hiru kopula mota nagusi daude: eliptikoak, mutur balioetako kopulak eta kopula arkimedearrak. Kapituluan zehar mota bakoitzean aurkitu daitezkeen kopulen zehaztapenak ematen ditugu. Azkenik kopula baldintzatuak ere azaltzen ditugu, aldagai osagarri batek baterako banaketan duen eragina aztertzen dutenak. Ildo honetatik, kopula baldintzatuen orokorpena kopula ez-parametrikoen bidez egin daiteke. Hain zuzen ere, azken hauek erabiliko ditugu hemendik aurrera.

Bibliografia

Clayton, D. G. (1978), "A Model for Association in Bivariate Life Tables and its Application in Epidemiological Studies of Familial Tendency in Chronic Disease Incidence," *Biometrika*, 65(1), 141–151.

Cook, R. D. eta Johnson, M. E. (1981), "A Family of Distributions for Modelling Non-Elliptically Symmetric Multivariate Data," *Journal of the Royal Statistical Society: Series B (Methodological)*, 43(2), 210–218.

——— (1986), "Generalized Burr-Pareto-Logistic Distributions with Applications to a Uranium Exploration Data Set," *Technometrics*, 28(2), 123–131.

Deheuvels, P. (1991), "On the Limiting Behavior of the Pickands Estimator for Bivariate Extreme-Value Distributions," *Statistics & Probability Letters*, 12(5), 429–439.

Nelsen, R. B. (2007), *An Introduction to Copulas,* New York: Springer Science & Business Media.

Oakes, D. (1982), "A Model for Association in Bivariate Survival Data," *Journal of the Royal Statistical Society: Series B (Methodological)*, 44(3), 414–422.

——— (1986), "Semiparametric Inference in a Model for Association in Bivariate Survival Data," *Biometrika*, 73(2), 353–361.

Patton, A. J. (2006), "Modelling Asymmetric Exchange Rate Dependence," *International Economic Review*, 47(2), 527–556.

Pickands, J. (1981), "Multivariate Extreme Value Distribution," in *Proceedings 43th, Session of the I.S.I*, The Hague, International Statistical Institute, pp. 859–878.

2

Testing Conditional Multivariate Rank Correlations

The Effect of Institutional Quality on Factors Influencing Competitiveness

1 Introduction

The joint distribution of two or more variables is a topic of clear interest in statistics. Actually, it provides a wealth information to analyze, such as the degree of dependence and the effects of conditional variables on comovements, among others. A classic measure of the degree of dependence is pairwise linear correlation. For a multivariate setting, the average of the pairwise linear correlations is widely used as a measure of multivariate dependence (see e.g. Joe, 1990; Longin and Solnik, 1995; Moskowitz, 2003; Capiello et al., 2006; Pollet and Wilson, 2010).

Beyond Pearson's correlation and normal distributions, Kendall's tau provides a descriptive statistic that detects monotonicity rather than linearity and has the benefit of not being restricted to symmetric distributions. For a multivariate context, Kendall and Smith (1940) suggest the average of pairwise taus as a coefficient of agreement. Another approach is followed by Simon (1977), who suggests a sign function based multivariate extension. In a related paper, Joe (1989) finds that the latter proposal does not meet the properties for consideration as a concordance measure. Therefore, Joe (1990) suggests a family of unconditional measures based on Kendall's tau as multivariate concordance measures beyond the average of pairwise taus. In this line, Nelsen (1992) and Nelsen (1996) analyze Kendall's tau based bivariate and multivariate measures.

Our first aim is to estimate a conditional multivariate Kendall's tau to analyze the effect of a variable of interest in the strengh of dependence. To that end, we extend the proposal made by Joe (1990) to the conditional case and define a multivariate conditional Kendall's tau estimator as an extension of Gijbels et al. (2011), with the corresponding asymptotic results.

Bandwidth selection is an important practical issue in nonparametric estimation. In this sense, Silverman (1986) is a well known reference for nonparametric densities as well as Altman and Leger (1995), Sarda (1993), and Bowman et al. (1998) for nonparametric distributions. In order to select the bandwidth for the nonparametric Kendall's tau estimator, we propose to minimize its mean squared error based on plug-in steps as in Gijbels et al. (2011). As a data driven method, we provide a jackknife method for estimating bias and variance. We also derive a simulation study to show the good performance of the multivariate estimator and the bandwidth selection procedure in practice.

The second aim of the methodological part of our study is to provide a statistic to test for the structure of dependence. For unconditional multivariate independence there are well known tests in the literature (see e.g. Genest and Rémillard, 2004). When variables are normally distributed, a test based on linear correlation is suitable. For more general cases, Leung et al. (2018) propose a rank correlation based test for independence between variables in high dimensions. In a related paper, Mao (2018) proposes a new test with better performance for large and small samples that deals with size distortion problems of the proposal in Leung et al. (2018) in small sample sizes. Strzalkowska-Kominiak and Stute (2013) also propose rank correlation based statistics to test for independence with survival data.

Here, we are interested in testing for conditional dependences, and in particular, for general restrictions among conditional Kendall's taus. In a related paper, Gijbels et al. (2017) analyze the so-called simplifying assumption, which assumes that the conditional copula coincides with the partial copula. To do so, they propose tests for this assumption expressed as a linear restriction between conditional Kendall's taus and compare their performance with tests based on conditional copulas such as the proposed by Acar et al. (2013).

Our proposal is to derive a Wald type statistic to test for general linear restrictions. The statistic enables tests to be run for specific, interesting cases such as constant conditional dependence, linear restrictions between different Kendall's taus, and equality of conditional Kendall's tau across different populations.

To study its performance in practice, we run a simulation study and compare the results with the proposal in Gijbels et al. (2017) for those situations where their test can be applied. Results support the idea that the Wald type test performs better for many different models, with a noteworthy advantage in computational cost.

For an empirical application we consider the European Union Regional Competitiveness Index (EU-RCI) data for 2019 for 268 regions. The index is formed by eleven pillars that help to identify the strengths and weaknesses in competitiveness of each region. All pillars provide indicators for the prosperity of a region and they are linked to one another. For instance, the more prosperous a region is, the better the indicators for all groups are expected to be. Obviously, this relationship is not perfect, and that is one reason to draw up a combined index. The overall index will reflect this combination of pillars. A classic question of interest is whether changes in one pillar are related to changes in another. Here another question arises: How the value of one pillar can influence or is related to the intensity of cross relations between other pillars. The test proposed is used to test for any significant effect of one pillar on the relationship between others.

In particular, we focus on the quality of regional institutions, which plays a very important role in the prosperity of a region. Dimant and Tosato (2018) provide a very helpful review of empirical results from the past few decades. One of the pillars in the Basic group is the Institutions index, which reports the quality of governments at regional and national levels, measured through perception and experience as collected via a survey. The interest lies in detecting the contribution of institutional quality to the relationship between pillars and in testing for changes in those relationships conditional on institutional quality levels. Moreover, tests for any other type of restriction can be run, such as changes in effects across different waves.

We find that tests of this type have interesting applications in practice in many different fields, particularly in medicine. For instance, Echouffo-Tcheugui et al. (2018) find that higher cortisol levels are associated with worse memory and visual perception. Moreover, morning rises in cortisol levels have been found to increase with body mass index. In this context, it would be interesting to test whether the relationship between cortisol as a "stress hormone" and the cognitive performance of patients changes with body mass index. Related to the current COVID-19 pandemic, Toyoshima et al. (2020) analyzes individuals mutations in SARS-CoV-2 genome sequences and their relationship with fatality rates, concluding that some virus variants are significantly correlated with them. Since host differences contribute to variations in response to pathogens, this test enables the effects of factors such as the genetics, age, and obesity of patients to be assessed in the relationships measured by the rank correlation to better understand the spread of COVID-19 and improve vaccine efficacy.

The rest of the chapter is structured as follows. Section 2 states the main results for the nonparametric estimators and sets out the practical estimation and testing procedures. Section 3 provides the simulation studies that show the performance of the smoothing parameter selection and of the test proposed. Section 4 applies the methodology to the link between regional efficiency and innovation pillars conditional on the quality of institutions. Section 5 concludes.

2 Estimation and tests

Let $\boldsymbol{Y} = \{Y_j\}_{j=1}^p$ be a set of p variables, and $F_1, ..., F_p$ and F their continuous marginal and joint distributions, respectively. In this context, Sklar (1959) states that there is a unique copula function $C\colon [0,1]^p \to [0,1]$ such that $F(\boldsymbol{y}) = F(y_1, ..., y_p) = C(F_1(y_1), ..., F_p(y_p))$. That is, copulas are joint distribution functions whose marginals are standard uniform variables. Patton (2006) extends this result to conditional copulas and states that, given a covariate Z, there is a unique copula $C_z : [0,1]^p \to [0,1]$ such that $F_z(\boldsymbol{y}) = C_z(F_{1z}(y_1), ..., F_{pz}(y_p))$, where $F_{jz}(y) = P(Y_j \leq y|Z = z)$, for any $y \in Y_j$, $j = 1, ..., p$. Inverting Sklar's theorem, the C_z function can be expressed as $C_z(\boldsymbol{u}) = F_z(F_{1z}^{-1}(u_1), ..., F_{pz}^{-1}(u_p))$ in terms of the joint and marginal distribution functions, where $\boldsymbol{u} = (u_1, ..., u_p) \in [0,1]^p$ and $F_{jz}^{-1}(u) = inf\{y : F_{jz}(y) \geq u\}$ is the z-conditional quantile function of Y_j.

To estimate conditional copulas, Gijbels et al. (2011) propose a nonparametric estimator in a bivariate context. We use the natural extension to the multivariate conditional copula estimator,

$$\hat{C}_z(\boldsymbol{u}) \quad = \quad \sum_{i=1}^{n} w_i(z,h) I\{Y_{1i} \leq \hat{F}_{1z}^{-1}(u_1), ..., Y_{pi} \leq \hat{F}_{pz}^{-1}(u_p)\}, \tag{2.1}$$

where $\{w_i(z,h)\}$ is a sequence of weights depending on $(z - Z_i)/h$ and h is the bandwidth. Considering Nadaraya-Watson weights, $\{w_i(z,h)\} = k((z - Z_i)/h)/\sum_j k((z - Z_j)/h)$, where k is a kernel function. $I\{\cdot\}$ is the indicator function and $\hat{F}_{jz}(y) = \sum_{i=1}^n w_i(z,h) I\{Y_{ji} \leq y\}$ is the nonparametric conditional j-marginal estimator. It is noteworthy that the bandwidth in this case does not have the usual smoothing effect as in regression. In fact, when the bandwidth h increases, the copula estimator $\hat{C}_z$ approaches the empirical distribution function.

To quantify the degree of dependence we estimate the Kendall's tau coefficient as a measure of the ordinal association between two measured quantiles. The multivariate Kendall's tau is defined as in Joe (1990), $\tau = (2^{p-1} - 1)^{-1} \left(2^p \int_{I^p} C(\boldsymbol{u}) dC(\boldsymbol{u}) - 1\right)$, where $I^p = [0,1]^p$. The multivariate Kendall's tau accounts for common comovements beyond pairwise effects and quantifies simultaneous concordance. Thus, more variables imply more constraints to be met at the same time and so, fewer concordances are expected. The distortion that the number of variables can produce is mitigated by the p-dependent correction factor included in the definition of the tau. Note that Kendall's tau is a measure of dependence that depends only on the copula and not on the marginals. We also note the advantage of the multivariate Kendall's tau over the pairwise average as an overall dependence measure, since it accounts for multivariate distribution and not only for bivariate effects. As the multivariate nonparametric estimator of τ we consider

$$\hat{\tau} = \frac{1}{2^{p-1}-1}\Big(\frac{2^p}{n(n-1)}\sum_{i=1}^{n}\sum_{j=1}^{n} I\{\mathbf{Y}_i < \mathbf{Y}_j\} - 1\Big),$$

where $\boldsymbol{Y}_i = (Y_{1i}, ..., Y_{pi})$ and $I\{\mathbf{Y}_i < \mathbf{Y}_j\} = I\{Y_{1i} < Y_{1j}, ..., Y_{pi} < Y_{pj}\}$.

An extended version of the multivariate Kendall's tau proposed by Joe (1990) to conditional copulas can be defined as

$$\tau_z = \frac{1}{(2^{p-1}-1)}\left(2^p \int_{I^p} C_z(\boldsymbol{u})dC_z(\boldsymbol{u}) - 1\right). \tag{2.2}$$

The nonparametric estimator proposed is

$$\hat{\tau}_z = \frac{1}{2^{p-1}-1}\Big(\frac{2^p}{1-\sum_{i=1}^{n} w_i(z,h)^2}\sum_{i,j=1}^{n} w_i(z,h)w_j(z,h)I\{\mathbf{Y}_i < \mathbf{Y}_j\} - 1\Big). \tag{2.3}$$

Note that, as for the copula estimator, the conditional Kendall's tau (2.3) tends to the unconditional empirical Kendall's tau as the bandwidth increases. This estimator generalizes the bivariate estimator in Gijbels et al. (2011). The asymptotic normality of the conditional Kendall's tau estimator is established by Veraverbeke et al. (2011) for the bivariate case. The next proposition generalizes the consistency and asymptotic normality of the multivariate conditional Kendall's tau estimator in (2.3) under the usual set of assumptions:

A1. $(\boldsymbol{Y}_i, Z_i)$, $i = 1, ..., n$ are $i.i.d.$ tuples.

A2. The conditional joint distribution $F_z(\cdot)$ and the density of the covariable Z, $f(z)$, have continuous first and second order derivatives with respect to z, all denoted with the respective primes.

A3. The kernel is a bounded symmetric second order kernel with compact support $\Omega = [-1, 1]$. Moreover, $c_k = \int_\Omega k(\eta)\eta^2 d\eta$ and $d_k = \int_\Omega k(\eta)^2 d\eta$ are non-zero quantities.

A4. $h \to 0$ and $nh \to \infty$ as $n \to \infty$.

Proposition 1. *Under assumptions A1 to A4, the conditional Kendall's tau estimator $\hat{\tau}_z$ defined in (2.3) is a consistent estimator of τ_z defined in (2.2), where the asymptotic bias is* $Bias(\hat{\tau}_z) = 2^{p-1}h^2 c_k((2^{p-1}-1)f(z))^{-1}\int_{\mathbb{R}^p} g_z(\boldsymbol{y})d\boldsymbol{y} + o(h^2)$ *with* $g_z(\boldsymbol{y}) = F_z(\boldsymbol{y})\Big(f_z(\boldsymbol{y})f''(z) + 2f'_z(\boldsymbol{y})f'(z) + f''_z(\boldsymbol{y})f(z)\Big) + f_z(\boldsymbol{y})\Big(F_z(\boldsymbol{y}) \times f''(z) + 2F'_z(\boldsymbol{y})f'(z) + F''_z(\boldsymbol{y})f(z)\Big)$.

Moreover, assuming that $h = o(n^{-1/5})$ *and* $\int_\Omega k(\eta)^\zeta d\eta \neq 0$ *for* $\zeta > 2$, *if* $G_z = 2^p(2^{p-1}-1)^{-1}\Big(\int_{I^p} C_z(\boldsymbol{u})dC_z^L(\boldsymbol{u}) + \int_{I^p} C_z^L(\boldsymbol{u})dC_z(\boldsymbol{u})\Big)$ *is a Gaussian variable, the asymptotic variance of* $\hat{\tau}_z$ *is given by the asymptotic variance of* G_z, $\sigma^2(G_z)$, *and*

$$(nh)^{1/2}\left(\hat{\tau}_z - \tau_z\right) \xrightarrow{d} G_z.$$

Provided that Kendall's tau can be written as a functional of the copula, the derivation of the limiting distribution for the multivariate estimator (2.3) is straightforward from the asymptotic normality of the multivariate conditional copula estimator $(nh)^{1/2}\Big(\hat{C}_z(\boldsymbol{u}) - C_z(\boldsymbol{u})\Big) \xrightarrow{d} C_z^L$, where C_z^L is a Gaussian variable with zero mean and asymptotic variance $\sigma^2(C_z^L) = d_k f(z)^{-1}C_z(\boldsymbol{u})\big(1 - C_z(\boldsymbol{u})\big)$, and the Hadamard differentiability of that functional (tangentially to the set of continuous functions on $[0,1]^p$).

2.1 Bandwidth selection

Classic proposals for selecting the smoothing parameter are based on the rule of thumb, cross-validation or plug-in methods. Smoothing parameter selection for distribution functions have been proposed by Altman and Leger (1995), Sarda (1993) and Bowman et al. (1998). Derumigny and Fermanian (2019) propose a cross-validation bandwidth selection procedure for the conditional Kendall's tau. Here, we propose a plug-in pointwise bandwidth selection method for the nonparametric conditional Kendall's tau by minimizing the overall mean squared error of the conditional tau.

The bias and variance for computing the MSE are estimated via the jackknife method based on Quenouille (1956) for bias and Tukey (1958) for variance. The procedure is an iterative process strongly related to the bootstrap resampling method proposed by Efron (1979). Actually, the jackknife is a linear approximation of the bootstrap (Abdi and Williams, 2010) that entails lower computational costs and is more suitable for small data samples (Oyeyemi, 2008; Efron, 1982). The main steps for selecting the bandwidth for the conditional Kendall's tau are summarized in Algorithm 1. We consider $h_0 = 0.9An^{-1/5}$ (Silverman, 1986) as the initial bandwidth for variance estimation, where $A = \min(\gamma/1.34,\ \sigma(Z))$ and γ is the interquartile range of the covariable Z. The initial bandwidth for the bias is taken as proposed in Gijbels et al. (2011).

Algorithm 1 Bandwidth selection for Kendall's tau

Require: A fixed conditioning value z and an initial bandwidth $h_0 = 0.9An^{-1/5}$
  Set initial bandwidth $h_v^{(r)} = h_0$ for variance
  STEP 1
  Set initial bandwidth $h_b^{(r)} = 2\gamma h_v^{(r)} n^{1/10}$ for bias
  for $l = 1$ to n **do**
    Define $Y_j^{-l} = \begin{cases} (Y_{j2}, \ldots, Y_{jn}) & l=1, \quad j=1,...,p \\ (Y_{j1}, \ldots, Y_{j(l-1)}, Y_{j(l+1)} \ldots, Y_{jn}) & l \geq 2, \quad j=1,...,p \end{cases}$
    Define $Z^{-l} = \begin{cases} (Z_2, \ldots, Z_n) & l=1 \\ (Z_1, \ldots, Z_{(l-1)}, Z_{(l+1)} \ldots, Z_n) & l \geq 2 \end{cases}$
    Calculate $\hat{\tau}_{v,z}^{-l}$ and $\hat{\tau}_{b,z}^{-l}$, the estimator (2.3) for $(\boldsymbol{Y}^{-l}, Z^{-l})$ with $h_v^{(r)}$ and $h_b^{(r)}$ respectively
  end for
  $\overline{\hat{\tau}}_{v,z} = mean(\hat{\tau}_{v,z}^{-l})$
  $\hat{\sigma}_J^2(\hat{\tau}_z) = \frac{n-1}{n} \sum_l (\hat{\tau}_{v,z}^{-l} - \overline{\hat{\tau}}_{v,z})^2$
  $\overline{\hat{\tau}}_{b,z} = mean(\hat{\tau}_{b,z}^{-l})$
  STEP 2
  Calculate $\hat{\tau}_{b,z}$, the estimator (2.3) for $(\boldsymbol{Y}, Z)$ using $h_b^{(r)}$
  $\widehat{Bias}_J(\hat{\tau}_z) = (n-1)(\overline{\hat{\tau}}_{b,z} - \hat{\tau}_{b,z})$
  STEP 3
  Obtain $\hat{h}_v^{(r+1)} = argmin_{h \in \mathbb{R}^+} (nh)^{-1} \hat{\sigma}_J^2(\hat{\tau}_z) + h^4 \widehat{Bias}_J(\hat{\tau}_z)^2$
  STEP 4
  Put $h_v^{(r)} = \hat{h}_v^{(r+1)}$ and go to STEP 1
  Repeat steps above for $\ell > 0$ times until $|\hat{h}_v^{(r+\ell)} - \hat{h}_v^{(r+1+\ell)}| < \varepsilon,$ for a prefixed $|\varepsilon| < 1$

2.2 Testing for restrictions in conditional dependence

In this section we propose a test for linear restrictions for all null hypothesis that can be expressed as

$$H_0 : \boldsymbol{R} \boldsymbol{\tau_z} = \boldsymbol{r}, \tag{2.4}$$

where $\boldsymbol{\tau_z} = (\tau_{z_1}, \ldots, \tau_{z_m})$ is an m-dimensional column vector of Kendall's taus, $z_1, ..., z_m$ are m fixed conditioning values that constitute a uniform subsample of the covariable Z, $\boldsymbol{R}$ is a $q \times m$ matrix of rank $q \leq m$, and $\boldsymbol{r}$ is a q-dimensional column vector where q is the number of restrictions to be tested. Both $\boldsymbol{R}$ and $\boldsymbol{r}$ are deterministic. The alternative is $H_a : \boldsymbol{R} \boldsymbol{\tau_z} \neq \boldsymbol{r}$. The test statistic under H_0 is

$$\mathcal{J}_n = nh(\boldsymbol{R} \hat{\boldsymbol{\tau}}_{\boldsymbol{z}} - \boldsymbol{r})'(\boldsymbol{R} \boldsymbol{V}_{\hat{\tau}_z} \boldsymbol{R}')^{-1}(\boldsymbol{R} \hat{\boldsymbol{\tau}}_{\boldsymbol{z}} - \boldsymbol{r}), \tag{2.5}$$

where $\boldsymbol{V}_{\hat{\tau}_z}$ is a consistent estimator of the covariance matrix $\boldsymbol{V}_{\tau_z}$, and h is the bandwidth. The following proposition establishes the asymptotic distribution of the test statistic $\mathcal{J}_n$ in (2.5) and the asymptotic local power for local alternatives of type $H_a(\xi_n) : \boldsymbol{R}\boldsymbol{\tau}_{\boldsymbol{z}} = \boldsymbol{r} + \xi_n \boldsymbol{\varsigma}$, where $\boldsymbol{\varsigma}$ is a $q \times 1$ non-zero deterministic column vector and $\xi_n \to 0$ as $n \to \infty$.

Proposition 2. *Consider the same assumptions as in Proposition 1. Under the null hypothesis, the $\mathcal{J}_n$ statistic asymptotically has a χ^2 distribution with q degrees of freedom, χ^2_q.*

Under the local alternative $H_a(\xi_n)$ with $\xi_n = (nh)^{-1/4}$, the $\mathcal{J}_n$ statistic is also asymptotically distributed as $\chi^2_q(\delta_n)$, a non-centered χ^2 distribution with q degrees of freedom and the noncentrality parameter $\delta_n = (nh)^{-1/2}\boldsymbol{\varsigma}'(\boldsymbol{R}\boldsymbol{V}_{\tau_z}\boldsymbol{R}')^{-1}\boldsymbol{\varsigma}$.

The limiting distributions can be obtained directly from the joint asymptotic normality of the conditional Kendall's tau when conditioned to different covariate points derived in Derumigny and Fermanian (2019)(Proposition 9) and Slutsky's theorem. Note that for local alternatives $(nh)^{1/2}(\boldsymbol{R}\hat{\boldsymbol{\tau}}_{\boldsymbol{z}} - \boldsymbol{r}) \xrightarrow{d} N(\xi_n \boldsymbol{\varsigma}, \boldsymbol{R}\boldsymbol{V}_{\tau_z}\boldsymbol{R}')$, so the statistic distribution is a non-centered χ^2.

The main steps for the practical implementation of the above test are presented in Algorithm 2. Without loss of generality, we consider a single bandwidth value, although it can be generalized to local bandwidth values.

Algorithm 2 Hypothesis test procedure

Require: Determine a set of points $\boldsymbol{z} = (z_1, \ldots, z_m)$ for $m < n$ such that $|z_\ell - z_{\ell'}| > h$ $\ell, \ell' = 1, \ldots, m, \quad \forall \ell \neq \ell'$
Define $\boldsymbol{R}$ and $\boldsymbol{r}$ according to H_0
Consider $h \sim 2n^{-1/5}h_{opt}$ where h_{opt} is obtained using Algorithm 1
STEP 1
for each $z_\ell \in \boldsymbol{z}$ **do**
  Estimate $\hat{\tau}_{z_\ell}$ as in (2.3)
end for
STEP 2
Define $\hat{\boldsymbol{\tau}}_{\boldsymbol{z}} = (\hat{\tau}_{z_1}, \ldots, \hat{\tau}_{z_m})'$
Estimate the elements in $\boldsymbol{V}_{\hat{\tau}_z}$ under H_0, as the sample variance of $\hat{\boldsymbol{\tau}}_{\boldsymbol{z}}$
STEP 3
Compute $\mathcal{J}_n = (\boldsymbol{R}\hat{\boldsymbol{\tau}}_{\boldsymbol{z}} - \boldsymbol{r})'(\boldsymbol{R}\boldsymbol{V}_{\hat{\tau}_z}\boldsymbol{R}')^{-1}(\boldsymbol{R}\hat{\boldsymbol{\tau}}_{\boldsymbol{z}} - \boldsymbol{r})$
Reject H_0 if $\mathcal{J}_n > \chi^2_{q|1-\alpha}$ for a prefixed α

The null hypothesis in expression (2.4) accounts for many possible situations. In particular, it enables to test for conditionally constant dependence. Alternative tests to determine whether there are covariate effects in conditional copulas can be found in Gijbels et al. (2017) and Derumigny and Fermanian (2017). Specifically, Gijbels et al. (2017) review some existing procedures based purely on conditional copula structures and introduce some nonparametric proposals using conditional Kendall's tau. In this particular case, $q = m - 1$, $\boldsymbol{R}$ is a $(m-1) \times m$ matrix with ones in the main diagonal and -1 values in the upper diagonal, and $\boldsymbol{r} = \mathbf{0}_{(m-1)\times 1}$. Due to the complexity of the asymptotic variance-covariance matrix, we use a permutation procedure to estimate $\boldsymbol{V}_{\hat{\tau}_z}$ under the null hypothesis: Keep Z fixed and obtain permuted $\{(Y_{1i}^b, ..., Y_{pi}^b)\}_{i=1}^n$ $p-$tuples from $\{(Y_{1i}, ..., Y_{pi})\}_{i=1}^n$ for a large number of permutations B. Then, with the permuted samples estimate $\{\hat{\tau}_{z_\ell}^b\}_{b=1}^B$ for each $\ell = 1, ..., m$ and compute the sample variance of the set of estimated conditional Kendall's tau.

The statistic $\mathcal{J}_n$ can be also used to test linear restrictions across different waves. Let s_1 and s_2 be two independent populations. Then $\hat{\boldsymbol{\tau}}_{\boldsymbol{z}}$ is a $2m \times 1$ stacked vector accounting for the conditional dependence in the two populations, $\hat{\boldsymbol{\tau}}_{\boldsymbol{z}} = (\hat{\tau}_{z_1}^{s_1}, ..., \hat{\tau}_{z_m}^{s_1}, \hat{\tau}_{z_1}^{s_2}, ..., \hat{\tau}_{z_m}^{s_2}) = (\hat{\boldsymbol{\tau}}_{\boldsymbol{z}}^{s_1\,\prime}, \hat{\boldsymbol{\tau}}_{\boldsymbol{z}}^{s_2\,\prime})'$. $\boldsymbol{R} = (\boldsymbol{I}_m, -\boldsymbol{I}_m)$, where $\boldsymbol{I}$ is the identity matrix and $\boldsymbol{r} = \mathbf{0}_{2m\times 1}$. The variance-covariance matrix $\boldsymbol{V}_{\hat{\tau}_z}$ is now a block diagonal matrix with $(\boldsymbol{V}_{\hat{\tau}_z}^{s_1}, \boldsymbol{V}_{\hat{\tau}_z}^{s_2})$ in the diagonal and $\boldsymbol{Cov}(\hat{\boldsymbol{\tau}}_{\boldsymbol{z}}^{s_1}, \hat{\boldsymbol{\tau}}_{\boldsymbol{z}}^{s_2})$ in the nondiagonal. The permutation procedure in this context is quite different since it has to be adapted into an appropriate resampling procedure. For each sample $s = s_1, s_2$ a bootstrap procedure can be implemented: Bootstrap $\{(Y_{s,1i}^b, ..., Y_{s,pi}^b, Z_{s,i}^b)\}_{i=1}^n$ $(p+1)-$tuples from $\{(Y_{s,1i}, ..., Y_{s,pi}, Z_{s,i})\}_{i=1}^n$ for a sufficiently large number of times B. Estimate $\hat{\tau}_{s,z_\ell}^b$ for each bootstrapped sample and calculate the variances $\hat{\sigma}^2(\hat{\tau}_{z_\ell}) = (2B)^{-1}\sum_{s,b}(\hat{\tau}_{s,z_\ell}^b - \overline{\hat{\tau}_{z_\ell}})^2$. Then, set $\boldsymbol{V}_{\hat{\tau}_z}$ to be a diagonal matrix with size $2m \times 2m$ and the estimated values $\{\hat{\sigma}^2(\hat{\tau}_{z_\ell})\}_{\ell=1}^m$.

These two applications of the $\mathcal{J}_n$ statistic are implemented in the simulation study presented in the next section.

3 Simulation study

3.1 Bandwidth robustness

We consider two models to study the performance of the conditional Kendall's tau estimator and the robustness of the bandwidth selection from Algorithm 1. For the sake of simplicity and ease of comparison, the data is generated under constant conditional dependence. In the first model (*Model L*) variables Y_1 and Y_2 depend linearly on a third variable Z: $Y_{1i} = 7Z_i + \varepsilon_{1i}$ and $Y_{2i} = 9Z_i + \varepsilon_{2i}$. In the second model (*Model NL*) the dependence of Y_1 and Y_2 on Z is nonlinear: $Y_{1i} = 4e^{Z_i} + \varepsilon_{1i}$ and $Y_{2i} = 5e^{Z_i} + \varepsilon_{2i}$. In both models $Z_i \sim U(0,1)$ is uniformly distributed, and the error terms are such that $\varepsilon_{1i} \sim N(0,1)$ and $\varepsilon_{2i} = \rho\varepsilon_{1i} + \sqrt{1-\rho^2}\epsilon_i$, where $\epsilon_i \sim N(0,1)$ and $\rho = 0, 0.75$. Note that normality implies a direct link between Pearson's linear correlation coefficient and Kendall's tau. Therefore, $\rho = 0$ means that the dependence of Y_1 and Y_2 is fully explained by the relationship with Z, while $\rho = 0.75$ indicates that there is an added dependency not related to Z. For the two models considered $S = 1000$ samples of sizes $n = 250, 500$ and 1000

are generated. The Epanechnikov kernel is used and the smoothing parameter for the Kendall's tau estimator is selected via Algorithm 1.

The results (reported in Appendix A.1) show that the conditional Kendall's tau estimator is highly sensitive to the selection of the correct smoothing parameter. As expected, for high values of the smoothing parameter the conditional Kendall's tau tends to an unconditional value. The smoothing parameter selection proposed in Section 2.1 performs quite well and the estimated conditional Kendall's tau figures are quite close to the real conditional Kendall's tau, regardless of the dependence structure between the variables.

3.2 Testing for constant conditional Kendall's tau

This section studies the size and power performance of the test proposed in Section 2.2. Given that the linear restriction accounts for the constant conditional Kendall's tau among others, the proposed statistic is an alternative to the statistics proposed by Gijbels et al. (2017). We analyze the behavior of the test statistic for $H_0 : \tau_{z_1} = \cdots = \tau_{z_m}$, where $z_1, \ldots, z_m$ are m conditioning values of a covariable Z. For the purpose of comparison, the V_{n1} statistic proposed in Gijbels et al. (2017) is considered. Beyond bivariate models, we also study the behavior of the two test statistics when data is generated from multivariate dependence structures.

We consider the independence case and two different settings: Models based on single copulas and models based on mixture copulas. For all models, $S = 1000$ samples are generated with sample sizes $n = 250, 500$, and 1000.

Independence setting.

Model 1: Data is generated assuming a nonlinear specification for the marginals: $Y_{1i} = 4e^{Z_i} + \varepsilon_{1i}$, $Y_{2i} = 5e^{Z_i} + \varepsilon_{2i}$, where $\epsilon_i \sim N(0, 1)$, $\rho = 0$, and the conditioning variable $Z_i \sim U(0, 1)$.

Single copula setting.

Five cases are generated from single copula models $C(F_1(Y_{1i}), \cdots, F_p(Y_{pi}); \theta)$, where θ is the dependence parameter of the copula in each model.

Model 2: The data generating process comes from a bivariate Clayton copula where the marginals $\{Y_{1i}\}$ and $\{Y_{2i}\}$ are two sequences of independent and normally distributed variables with zero mean and unit variance and the dependence parameter is $\theta_2(Z_i) = Z_i^2/(Z_i^2 + 1)$ with $Z_i \sim U(0, 6)$.

Model 3: Data is generated from a bivariate Gumbel copula with marginals $Y_{1i} = 2\sin(2\pi/3 \times(Z_i - 2) - 1) + \varepsilon_{1i}$ and $Y_{2i} = \varepsilon_{2i}$, where $\{\varepsilon_{1i}\}$ and $\{\varepsilon_{2i}\}$ are two independent sequences that have density $1 - |x|$ on the support $[-1, 1]$, the dependence parameter is given by $\theta_3(Z_i) = e^{0.5} + 1$, and $Z_i \sim U(2, 5)$ (This model is defined as 'Model 1' in Gijbels et al. (2017), Section 5).

Model 4: Data is generated as in *Model 3* above, but for the dependence parameter of the Gumbel copula the function $\theta_4(Z_i) = e^{1.5-0.4Z_i}+1$ is taken (This model is defined as 'Model 2' in Gijbels et al. (2017), Section 5).

Model 5: Data comes from a multivariate Clayton copula with $p = 3$. The sequences $\{Y_{1i}\}$ and $\{Y_{2i}\}$ are generated as in *Model 3* and $Y_{3i} = 5+3\cos(14\pi/3\,(Z_i-2)-7)+\varepsilon_{3i}$, where $\varepsilon_{3i} \sim N(0, 1)$. The copula dependence parameter is $\theta_5(Z_i) = \theta_3(Z_i)$, where $Z_i \sim U(2, 5)$.

Model 6: The data generating process is similar to that in *Model 5*. In this model the marginal $\{Y_{3i}\}$ is a sequence of an independent random variable with standard normal distribution and the Clayton copula dependence parameter functional has changed to $\theta_6(Z_i) = e^{1.5+0.4Z_i} + 1$.

Mixture copula setting.

The second set, *Models 7* to *10*, considers data generated from mixture copulas

$$C^{MIX}(u_1, \cdots, u_p)=wC_a(u_1, \cdots, u_p; \theta_a)+(1-w)C_b(u_1, \cdots, u_p; \theta_b),$$

where θ_ℓ is the dependence parameter for $\ell = a, b$ and w is the weight function.

Model 7: In this case, bivariate Clayton and Gumbel copulas are considered for C_a and C_b with $w = 0.3$ whose marginals $\{Y_{1i}\}$ and $\{Y_{2i}\}$ are two sequences of independent random variables with a standard normal distribution, and where the dependence parameters are $\theta_a(Z_i) = e^{0.5}$ and $\theta_b(Z_i) = 1.2$, and $Z_i \sim U(0, 6)$.

Model 8: Data comes from a mixture between two bivariate Frank copulas with $w = 0.3$. The marginals Y_1 and Y_2 are generated as in *Model 3*, $\theta_a(Z_i) = Z_i^3/(Z_i^3 + 1)$, $\theta_b(Z_i) = \theta_6(Z_i)$, and $Z_i \sim U(2, 5)$.

Model 9: Multivariate Clayton and Gumbel copulas (with $p = 3$) are taken as C_a and C_b with $w = 0.7$ respectively. $Y_{1i} = 2\sin(\pi\,(Z_i/3 - 1)) + \varepsilon_{1i}$, $Y_{2i} = \varepsilon_{2i}$, and $Y_{3i} = \varepsilon_{3i}$, where $\{\varepsilon_{1i}\}$, $\{\varepsilon_{2i}\}$, and $\{\varepsilon_{3i}\}$ are three sequences of independent, normally distributed random variables with zero mean and unit variance. The dependence parameters are defined as $\theta_a(Z_i) = sin(4\pi/7)+1$ and $\theta_b(Z_i) = e^{2.5}+1$, and $Z_i \sim U(0, 6)$.

Model 10: Data is generated as in *Model 8*, considering multivariate ($p = 3$) versions for the Frank copulas. The third marginal $\{Y_{3i}\}$ is a sequence of an independently distributed random variable with density $1 - |x|$ on the support $[-1, 1]$.

The rejection frequencies for each model are presented in Table 2.1 for levels $\alpha = 1\%, 5\%$, and 10% and for the different sample sizes. *Models 1, 3, 5, 7*, and *9* consider a constant conditional dependence, so the null hypothesis holds. Thus, these results report the size. By contrast, for *Models 2, 4, 6, 8*, and *10* the conditional dependence is Z-dependent so the results report the power of the test. The 3-column first block contains the rejection frequencies for the proposed $\mathcal{J}_n$ test statistic when the conditional points are fixed on the 5% equally spaced points of the sample, $\mathcal{J}_n^{5\%}$. In practice, we find that using 5% equally spaced points provides optimal results with a low computational cost. The 3-column last block shows the results for the V_{n1} statistic proposed by Gijbels et al. (2017). Additional rejection frequency results for the test statistic when the conditional points are fixed on the 2.5% and 10% equally spaced points of the sample ($\mathcal{J}_n^{2.5\%}$ and $\mathcal{J}_n^{10\%}$) are given in Appendix A.2, Table A.1. The performance of the two tests is studied in the case of unknown marginals.

The results support the idea that the test statistic is appropriate for testing constant conditional dependence in bivariate and multivariate contexts. *Models 3* and *4* are also analyzed by Gijbels et al. (2017) for $n = 100$ and $\alpha = 0.05$. The results obtained for those models seem to be in line with the ones obtained in Gijbels et al. (2017) at least for the small sample size. Indeed, $\mathcal{J}_n^{5\%}$ appears to be more powerful than the V_{n1} statistic in most cases, with a clear improvement in the case of mixture copulas. Note that in this case, compared to the structures studied in Gijbels et al. (2017), the competing models are mixtures between two Archimedean copulas where the dependence parameters are Z-dependent and the marginals are assumed to be unknown. Moreover, the performance of the test statistics is studied in larger sample sizes than in the analysis in Gijbels et al. (2017). Note also that while the $\mathcal{J}_n^{5\%}$ statistic uses only some equally spaced points of the variables, the V_{n1} operates with the estimates of Kendall's tau conditional on all values of Z. In small sample sizes this may not be a big drawback, but it entails high computational costs for large samples, which is a practical issue that must be taken into account. Moreover, $\mathcal{J}_n^{5\%}$ is computationally less expensive than V_{n1}. Therefore, the results support the conclusion that the $\mathcal{J}_n^{5\%}$ statistic is a good alternative for testing for constant conditional dependence.

3.3 Testing for equal conditional Kendall's tau across samples

This section uses the $\mathcal{J}_n$ statistic to test for equality of the conditional Kendall's tau across two samples s_1 and s_2 in a bivariate case. In this case, the null is $H_0 : \boldsymbol{\tau}_{\boldsymbol{z}}^{s_1} = \boldsymbol{\tau}_{\boldsymbol{z}}^{s_2}$, where

Table 2.1: Rejection frequencies for simulated models (S=1000 replications).

		$\mathcal{J}_n^{5\%}$			V_{n1}		
			n			n	
	α	250	500	1000	250	500	1000
Model 1	1%	1.8	0.7	0.8	1.4	0.2	1.1
	5%	5.8	4.5	4.9	6.9	4.8	3.9
	10%	10.6	8.9	10.2	13.1	8.5	8.9
Model 2	1%	97.2	98.7	99.6	3.7	19.1	57.0
	5%	98.3	99.8	100.0	12.4	39.1	80.5
	10%	99	99.8	100.0	20.1	51.3	87.5
Model 3	1%	1.0	2.2	1.0	0.5	1.5	2.6
	5%	4.1	6.3	4.8	4.3	6.7	10.9
	10%	8.8	11.1	10.3	10.1	11.5	17.6
Model 4	1%	83.7	96.4	99.4	7.3	10.6	25.5
	5%	91.6	99.1	99.8	21	28.4	47.2
	10%	94.5	99.5	99.9	33.3	37.7	57.7
Model 5	1%	3.0	1.7	1.9	1.0	1.2	1.0
	5%	6.6	4.8	5.2	7.1	5.7	5.1
	10%	10.0	7.2	9.5	13.6	12.9	14.4
Model 6	1%	17.6	69.3	99.3	2.9	9.2	60.9
	5%	31.2	84.2	99.7	8.1	25.9	80.4
	10%	40.6	89.9	99.8	14.0	36.9	86.5
Model 7	1%	0.9	1.7	1.7	2.3	0.6	0.8
	5%	4.4	6.9	8.1	8.1	3.9	5.6
	10%	8.7	12.3	14.2	12.5	6.2	11.5
Model 8	1%	6.2	22.9	56.5	11.3	16.7	34.4
	5%	19.3	43.0	71.8	24.6	30.4	54.3
	10%	30.1	55.7	78.6	32.7	40.8	62.5
Model 9	1%	2.4	1.5	2.4	0.2	1.8	2.8
	5%	7.4	6.5	5.6	1.4	10.0	8.1
	10%	12.5	11.2	9.0	2.8	15.3	14.8
Model 10	1%	22.1	75.5	93.5	12.7	30.4	53.7
	5%	43.9	87.2	97.0	31.2	50.7	76.6
	10%	56.1	91.6	97.8	44.4	62.6	84.0

$\boldsymbol{\tau}_{\boldsymbol{z}}^s = (\tau_{z_1}^s, \ldots, \tau_{z_m}^s)$ for m conditioning values of the covariable Z in the sample s. Note that the conditional Kendall's tau may change with Z.

In order to analyze the performance of the statistic in line with the sample size in the empirical part, we simulate two samples with $n = 250$ observations under four scenarios:

Scenario 1: The two samples are generated as in *Model 2*.

Scenario 2: First and second samples are generated as in *Models 2* and *4*, respectively.

Scenario 3: The samples are generated as in *Model 8*.

Scenario 4: One sample is generated as in *Model 8* and another sample as in *Model 4*.

The results from *Scenarios 1* and *3* report the size of the test and those from *Scenarios 2* and *4* report its power. In all cases 4000 replications are taken, considering the 5% equally spaced values of the covariate Z as conditioning points. Table 2.2 presents the rejection frequencies for *Scenarios 1* to *4* for different significance levels. The statistic provides adequate results in this aplication even for complicated structures such as mixtures of copulas.

Table 2.2: Rejection frequencies of $\mathcal{J}_n^{5\%}$ statistic for simulated scenarios (sample size $n = 250$ and $S = 4000$ replications).

	$\alpha = 1\%$	$\alpha = 5\%$	$\alpha = 10\%$
Scenario 1	1.35	4.97	9.87
Scenario 2	43.32	61.32	70.92
Scenario 3	0.62	3.97	7.72
Scenario 4	73.87	87.07	91.92

4 Empirical application

The European Regional Competitiveness Index (RCI) has been drawn up by the European Commission every three years since 2010[1]. It comprises more than 70 indicators for measuring the ability of regions to offer an attractive, sustainable environment for firms and residents to live and work (Annoni and Dijkstra, 2019). The final index is formed by eleven pillars grouped into three general categories: Basic, Efficiency, and Innovation. Basic comprises five pillars: Institutions ($INST$), Macroeconomic Stability, Infrastructure, Health, and Basic Education. Efficiency comprises three pillars: Higher Education (HE), Labor market efficiency (L), and Market size (M). Innovation also comprises three: Technological readiness (TR), Business sophistication (BS), and Innovation (I). The measures of the pillars are provided as z-scores and the values are such that higher means better. The data set contains the z-scores for the eleven pillars, the three categories, and the RCI index at regional level with 268 European regions for 2019.

It is well known that there are links between some individual pillars and the macroeconomy. For instance, the role of higher education in economic growth has been widely analyzed in economic models (Lucas, 1988). There is also a large body of literature that links innovation

[1]Data set and more details available on the European Commission website: https://ec.europa.eu/regional_policy/en/information/maps/regional_competitiveness/

to economic growth (see e.g. Furman et al., 2002; Hasan and Tucci, 2010; Pradhan et al., 2016; Maradana et al., 2017, among others). On that basis, it is helpful for political regulation and economic purposes to detect how higher education levels can affect a firm's innovation. Haiyan et al. (2020) find evidence for China that a highly educated stock of human capital plays an important role in both the probability and the quantity of innovation at firms.

As mentioned in the Introduction, the quality of institutions is a key determinant in the prosperity of regions, and the RCI index also takes this into account. The Institutions (*INST*) pillar in the Basic category covers regional indicators for corruption, quality, and impartiality among others. It is based on the European Quality of Government Index (EQI), a survey on corruption and governance at a regional level within the EU conducted by the Quality Government Institute at the University of Gothenburg.[2] As suggested in the studies collected in Dimant and Tosato (2018), there is a link between institutional quality and growth.

Based on the above ideas, we seek to detect whether low institutional quality hinders transfers between higher education and innovation results. Thus, the objective of this section is to study the link between higher education and innovation, conditional on institutional quality. Specifically, we study whether institutional quality helps to increase the bivariate and multivariate relationships between other pillars. To that end, we compute the conditional dependence coefficients using estimator (2.3) assessed on five quantiles $\boldsymbol{z} = (z_1 = 0.05,\ z_2 = 0.15,\ z_3 = 0.5,\ z_4 = 0.85,\ z_5 = 0.95)$ of *INST*. In particular, we are interested in four different hypotheses formulated as follows:

1. There is no concordance between higher education and innovation, measured by Kendall's tau: $H_0^{(1)} : \tau = 0$. A positive tau is expected, supporting the idea that more competitive regions are linked to higher indicator scores. This hypothesis is tested using the normal asymptotic distribution of the empirical Kendall's tau (Prokhorov, 1994).

2. The concordance between higher education and innovation is fully explained by institutional quality: $H_0^{(2)} : \tau_z = 0$, where z is the measure for institutional quality, the *INST* pillar. This is true if the pillars are z-conditional independent, which is a sufficient but not necessary condition for $\tau_z = 0$.

3. Institutional quality does not explain any of the concordance between higher education and innovation: $H_0^{(3)} : \tau_z = \tau$. This means that the unconditional and conditional degrees of concordance between pillars are equal, whatever the level z of institutional quality.

4. The concordance between higher education and innovation might not depend on the level of

[2] See https://qog.pol.gu.se/data/datadownloads/qog-eqi-data for details.

quality of institutions: $H_0^{(4)} : \tau_{z_k} = \tau_{z_l}$ is tested to check whether the degree of concordance between pillars depends or not on the standard of the quality of governance, where z_k and z_l $(z_k \neq z_l)$ denote two levels of institutional quality. Rejection would be evidence of a link between pillars that varies according to institutional quality.

The fourth hypothesis is especially interesting, since its rejection provides a starting point for studying how quality of institutions makes the links between pillars stronger or weaker. Analyzing such conditional comovements is very useful for policy makers with a view to controlling the impact of their interventions.

For the sake of illustration and to provide further empirical results, we also consider the dependence between the Efficiency and Innovation groups and between higher education and the other pillars in these two groups (Labor Market Efficiency, Market Size, Business Sophistication, and Technological Readiness). First we test for $H_0^{(1)}$ and, as expected, find that all the unconditional Kendall's taus become significant. For the other three hypotheses, we adapt the test statistic $\mathcal{J}_n$ described in Subsection 2.2 to each hypothesis. To analyze $H_0^{(2)}$, set $\boldsymbol{R}$ as the identity matrix $\boldsymbol{I}_m$ and $\boldsymbol{r} = \boldsymbol{0}_{m\times 1}$. To test for $H_0^{(3)}$, $\boldsymbol{R} = \boldsymbol{I}_m$ and $\boldsymbol{r} = \tau \cdot \boldsymbol{1}_{m\times 1}$. Finally, for $H_0^{(4)}$, (2.4) as detailed in Subsection 2.2 is considered.

Table 2.3 contains the estimated unconditional Kendall's tau in the second column and the estimated conditional coefficients in columns 3 to 7. The last block of columns summarizes the tests results.

Table 2.3: Kendall's tau coefficients based on 2019 RCI data.

Relations	$\hat{\tau}$	$\hat{\tau}_{q_{0.05}}$	$\hat{\tau}_{q_{0.15}}$	$\hat{\tau}_{q_{0.5}}$	$\hat{\tau}_{q_{0.85}}$	$\hat{\tau}_{q_{0.95}}$	$H_0^{(2)}$	$H_0^{(3)}$	$H_0^{(4)}$
Eff.-Inn.	0.678	0.097	0.623	0.484	0.204	0.733	***	***	***
HE-L	0.447	0.036	0.248	0.448	0.054	0.357	***	***	*
HE-M	0.197	-0.039	-0.053	-0.294	-0.009	-0.108		***	
HE-TR	0.405	0.077	0.215	0.382	0.052	-0.093	**	***	*
HE-BS	0.308	0.253	0.073	-0.056	0.154	-0.154		***	
HE-I	0.528	0.091	0.307	0.406	0.456	0.530	***	***	**
HE-L-M	0.363	0.025	0.155	0.020	0.151	0.029		***	
TR-BS-I	0.518	0.366	0.517	0.245	0.255	0.069	***	***	**
HE-L-M-TR-BS-I	0.387	0.112	0.225	0.171	0.025	0.118	***	***	

NOTE: The values in column 2 are the unconditional coefficients, while columns 3 to 7 show the conditional ones. Asterisks indicate level of significance: $^{***}p < 0.01$, $^{**}p < 0.05$ and $^{*}p < 0.1$.

There is a noteworthy dependence between higher education and innovation, which grows stronger as institutional quality becomes higher. For the other pillars the results are different. An analysis of the link between *HE* and *BS* reveals that the idea that they are conditionally independent cannot be rejected. That is, for a given level of institutional quality, no dependence is found between *HE* and *BS* through the conditional Kendall's tau. For the relationship between *HE* and *TR*, *INST* provides a further contribution for regions with medium institutional quality, which might be linked to specific regions.

For multivariate relationships, the dependence is always positive. There is a low dependence between the pillars in the Efficiency group and the results reveal a constant effect of institutional quality among them. Pillars in the Innovation group are unconditionally more closely linked than those in the Efficiency group. Moreover, the Innovation group pillars are significantly affected by the quality of institutions and show a higher dependence for lower quality levels. The results show in fact that governance quality has a clear impact on the comovements of variables related to innovation. Nevertheless, the last row of Table 2.3 shows that there is no significant effect for the multivariate relationship between these six pillars.

Regardless of whether or not conditional dependence is constant, the quality perception effect can vary from one period to another. To test whether the effect of institutional quality on the link between indicators is consistent over a three-year period, we consider RCI index data for 2016 and compare similarities in the behavior patterns between the different stress periods. We find that in general there are no significant changes in the dependency between pillars from 2016 to 2019, conditional on the quality of institutions. The test reveals that at the 5% level *HE-M* is the only link in which the institutional quality effect changes significantly over a three-year period.

5 Conclusions

In this paper we consider a nonparametric conditional copula to estimate conditional joint dependence in a multivariate context. As an overall measure of dependence, we compute a multivariate version of the rank correlation through a nonparametric conditional Kendall's tau estimator.

Selecting the bandwidth for nonparametric estimators is an important task for densities, distributions and regression. We derive a smoothing parameter selection procedure for the conditional Kendall's tau and provide a simulation study to show its performance. The proposed procedure is based on the minimization of the global mean squared error, where the bias and variance terms are obtained using a jackknife approach.

A Wald type statistic is used to test whether there is any significant linear restriction in Kendall's tau conditional to some values of the covariate Z. As in Gijbels et al. (2017), the procedure enables conditional independence and constant conditional dependence to be tested for. The asymptotic distributions and the procedure for practical implementation are provided. We conduct a simulation study to analyze size and power of the proposed test with bivariate and multivariate competing models for different sample sizes. The results show that the statistic performs well for different types of restriction, even when quite complex joint distributions are considered. Its ease of implementation, low computational cost, and wide range of applications make it a useful procedure for testing many different specifications of conditional rank correlations.

The methodology is applied to analyze the multivariate dependence between pillars in the 2019 RCI index. Specifically, we focus on the effect of the quality of institutions on the relationship between the Efficiency and Innovation pillars. The results are quite interesting and encourage further study.

First, there is a clear positive joint relationship between pillars, as expected. Second, the joint relationship between pairs such as innovation and higher education is only partially explained by the quality of institutions. Moreover, there is evidence in favor of a joint dependence that increases with the quality of institutions. In other words, the lower the quality of institutions, the weaker the link between innovation and higher education. This may be an interesting starting point for studying whether there is a causal link between the quality of institutions and the ability of regions to transfer human capital to innovation results. This goes beyond the scope of this work, but we believe that it opens up a promising research area.

References

Abdi, H. and Williams, L. (2010), "Jackknife," in *Encyclopedia of Research Design* (Vol. 2), ed. N.J. Salkind, Thousand Oaks, CA: SAGE Publications, pp. 655–660.

Acar, E. F., Craiu, R. V., and Yao, F. (2013), "Statistical Testing of Covariate Effects in Conditional Copula Models," *Electronic Journal of Statistics*, 7, 2822–2850.

Altman, N. and Leger, C. (1995), "Bandwidth Selection for Kernel Distribution Function Estimation," *Journal of Statistical Planning and Inference*, 46(2), 195–214.

Annoni, P. and Dijkstra, L. (2019), "The EU Regional Competitiveness Index 2019," *Luxemburg: Publications Office of the European Union.*

Bowman, A., Hall, P., and Prvan, T. (1998), "Bandwidth Selection for the Smoothing of Distribution Functions," *Biometrika*, 85(4), 799–808.

Capiello, L., Engle, R., and Sheppard, K. (2006),"Asymmetric Dynamics in the Correlations of Global Equity and Bond Returns," *Journal of Financial Econometrics*, 4, 537–572.

Derumigny, A. and Fermanian, J.-D. (2017),"About Tests of the "Simplifying" Assumption for Conditional Copulas," *Dependence Modeling*, 5(1), 154–197.

——— (2019),"On Kernel-Based Estitmation of Conditional Kendall's Tau: Finite-Distance Bounds and Asymptotic Behavior," *Dependence Modeling*, 7(1), 292–321.

Dimant, E. and Tosato, G. (2018),"Causes and Effects of Corruption: What has Past Decade's Empirical Research Taught us? A Survey," *Journal of Economic Surveys*, 32, 335–356.

Echouffo-Tcheugui, J. B., Conner, S. C., Himali, J. J., Maillard, P., DeCarli, C. S., Beiser, A. S., Vasan, R. S., and Seshadri, S. (2018),"Circulating Cortisol and Cognitive and Structural Brain Measures," *Neurology*, 91(21), 1961–1970.

Efron, B. (1979),"Bootstrap Methods: Another Look at the Jackknife," *The Annals of Statistics*, pp. 1–26.

——— (1982), *The Jackknife, the Bootstrap, and Other Resampling Plans*, Philadelphia, PA: Society for Industrial and Applied Mathematics.

Furman, J. L., Porter, M. E., and Stern, S. (2002),"The Determinants of National Innovative Capacity," *Research Policy*, 31(6), 899–933.

Genest, C. and Rémillard, B. (2004),"Test of Independence and Randomness Based on the Empirical Copula Process," *Test*, 13(2), 335–369.

Gijbels, I., Veraverbeke, N., and Omelka, M. (2011),"Conditional Copulas, Association Measures and Their Applications," *Computational Statistics and Data Analysis*, 55, 1919–1932.

Gijbels, I., Omelka, M., and Veraverbeke, N. (2017),"Nonparametric Testing for no Covariate Effects in Conditional Copulas," *Statistics: A Journal of Theoretical and Applied Statistics*, 51, 475–509.

Haiyan, D., Ahmed, K., and Nanere, M. (2020),"Life Cycle, Competitive Strategy, Continuous Innovation and Firm Performance," *International Journal of Innovation Management*, 2150004, 1–23.

Hasan, I. and Tucci, C. L. (2010),"The Innovation-Economic Growth Nexus: Global Evidence," *Research Policy*, 39(10), 1264–1276.

Joe, H. (1989), "Relative Entropy Measures of Multivariate Dependence," *Journal of the American Statistical Association*, 84, 157–164.

——— (1990), "Multivariate Concordance," *Journal of Multivariate Analysis*, 35, 12–30.

Kendall, M. G. and Smith, B. B. (1940), "On the Method of Paired Comparisons," *Biometrika*, 31, 324–345.

Leung, D., Drton, M., et al. (2018), "Testing Independence in High Dimensions with Sums of Rank Correlations," *The Annals of Statistics*, 46(1), 280–307.

Longin, F. and Solnik, B. (1995), "Is the Correlation in International Equity Returns Constant: 1960-1990?," *Journal of International Money and Finance*, 14, 3–26.

Lucas, R. E. (1988), "On the Mechanics of Economic Development," *Journal of Monetary Economics*, 22, 3–42.

Mao, G. (2018), "Testing Independence in High Dimensions Using Kendall's Tau," *Computational Statistics and Data Analysis*, 117, 128–137.

Maradana, R. P., Pradhan, R. P., Dash, S., Gaurav, K., Jayakumar, M., and Chatterjee, D. (2017), "Does Innovation Promote Economic Growth? Evidence from European Countries," *Journal of Innovation and Entrepreneurship*, 6, 1–23.

Moskowitz, T. J. (2003), "An Analysis of Covariance Risk and Pricing Anomalies," *The Review of Financial Studies*, 16, 417–457.

Nelsen, R. B. (1992), "On Measures of Association as Measures of Positive Dependence," *Statistics & Probability Letters*, 14, 269–274.

——— (1996), "Nonparametric Measures of Multivariate Association," in *Distributions with Fixed Marginals and Related Topics, IMS Lecture Notes-Monograph Series* (Vol. 28), eds. L. Rüschendorf, B. Schweizer and M. D. Taylor, Hayward, CA: Institute of Mathematical Statistics, pp. 223–232.

Oyeyemi, G. (2008), "Comparison of Bootstrap and Jackknife Methods of Re-sampling in Estimating Population Parameters," *Global Journal of Pure and Applied Sciences*, 14(2), 217–220.

Patton, A. J. (2006), "Modelling Asymmetric Exchange Rate Dependence," *International Economic Review*, 47, 527–556.

Pollet, J. M. and Wilson, M. (2010), "Average Correlation and Stock Market Returns," *Journal of Financial Economics*, 96, 364–380.

Pradhan, R. P., Arvin, M. B., Hall, J. H., and Nair, M. (2016),"Innovation, Financial Development and Economic Growth in Eurozone Countries," *Applied Economics Letters*, 23(16), 1141–1144.

Prokhorov, A. (1994),"Kendall Coefficient of Rank Correlation," in *Encyclopedia of Mathematics*, ed. M. Hazewinkel, NL: Springer.

Quenouille, M. H. (1956),"Notes on Bias in Estimation," *Biometrika*, 43, 353–360.

Sarda, P. (1993),"Smoothing Parameter Selection for Smooth Distribution Functions," *Journal of Statistical Planning and Inference*, 35(1), 65–75.

Silverman, B. W. (1986), *Density Estimation for Statistics and Data Analysis*, London: Chapman and Hall.

Simon, G. (1977),"Multivariate Generalization of Kendall's Tau with Application to Data Reduction," *Journal of the American Statistical Association*, 72, 367–376.

Sklar, M. (1959),"Fonctions de Répartition à N Dimensions et Leurs Marges," *Publications de L'Institut Statistique de L'Université de Paris*, 8, 229–231.

Strzalkowska-Kominiak, E. and Stute, W. (2013),"Empirical Copulas for Consecutive Survival Data," *Test*, 22(4), 688–714.

Toyoshima, Y., Nemoto, K., Matsumoto, S., Nakamura, Y., and Kiyotani, K. (2020),"SARS-CoV-2 Genomic Variations Associated with Mortality Rate of COVID-19," *Journal of Human Genetics*, 65(12), 1075–1082.

Tukey, J. (1958),"Bias and Confidence in not Quite Large Samples," *The Annals of Mathematical Statistics*, 29, 614.

Veraverbeke, N., Omelka, M., and Gijbels, I. (2011),"Estimation of a Conditional Copula and Association Measures," *Scandinavian Journal of Statistics*, 38, 766–780.

Appendix A: Additional results.

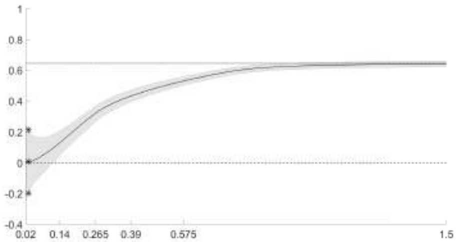
A.1 Bandwidth robustness

Figures A.1 and A.2 show the estimated conditional Kendall's tau trend (solid black line) and its %95 pointwise confidence interval (grey shading) for each model simulated and a range of bandwidth values. Results are shown for sample size $n = 1000$ (results for $n = 250, 500$ show similar features). The asterisk indicates the estimated Kendall's tau value for the optimal bandwidth selected by the proposed algorithm. The figures in the first row are for *Model L*, where the dependence between variables relies on the conditional variable Z, and those in the second row are for *Model NL*, where the dependence goes beyond the conditional variable. For the sake of illustration, each row shows the results obtained conditional on percentiles $z_{0.25}$, $z_{0.50}$, and $z_{0.75}$ of variable Z. Horizontal dashed and dotted-dashed lines represent the real conditional and unconditional Kendall's tau respectively.

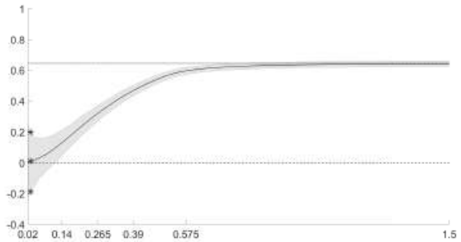
A.2 Testing for conditional Kendall's tau

Table A.1 shows the rejection frequencies of the $\mathcal{J}_n$ test for *Models 1-10.* The results for levels $\alpha = 1\%, 5\%$, and 10% by rows and different sample sizes by columns are given for all the models. The 3-column first block contains the rejection frequencies when the conditional points are fixed on the 2.5% equally spaced points of the sample, $\mathcal{J}_n^{2.5\%}$, while the 3-column second block shows those for conditional points fixed on the 10% points of the sample, $\mathcal{J}_n^{10\%}$.

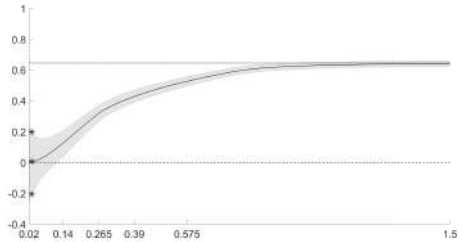
Conditional Kendall's tau and 95% confidence interval for *Model L* and $\rho = 0$.



(a) $Z = z_{0.25}$

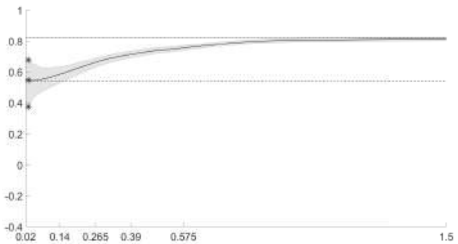


(b) $Z = z_{0.50}$

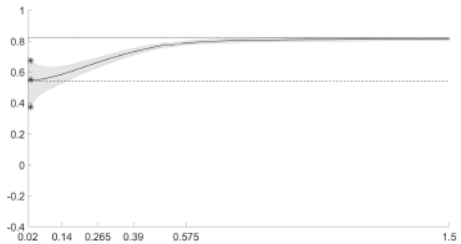


(c) $Z = z_{0.75}$

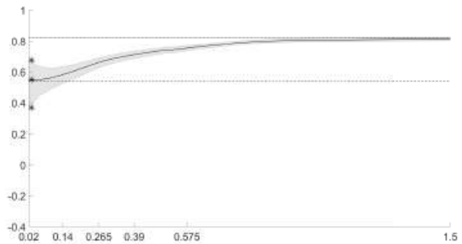
Conditional Kendall's tau and 95% confidence interval for *Model L* and $\rho = 0.75$.



(d) $Z = z_{0.25}$



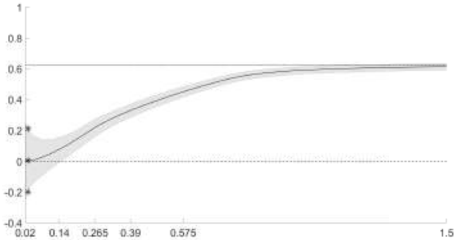
(e) $Z = z_{0.50}$



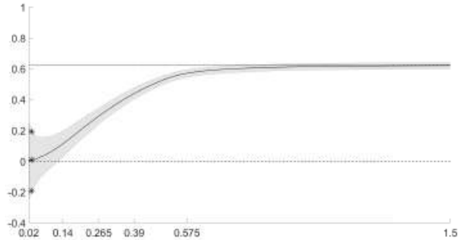
(f) $Z = z_{0.75}$

Figure A.1: Estimated Kendall's tau (solid black line) conditional on percentiles (a) 25, (b) 50, and (c) 75 of Z with 95% pointwise confidence interval (grey shaded) for variables generated for *Model L*, where dependence is linear and $\rho = 0, 0.75$. The asterisk represents the estimated Kendall's tau for the optimal bandwidth. Horizontal dashed and dotted-dashed lines represent the real conditional and unconditional Kendall's tau respectively. The X-axis shows the range of bandwidths.

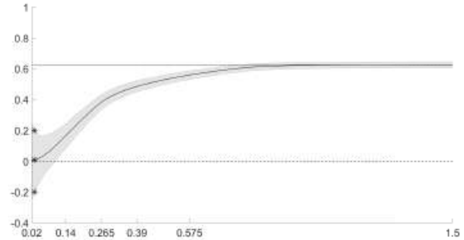
Conditional Kendall's tau and 95% confidence interval for *Model NL* and $\rho = 0$.



(a) $Z = z_{0.25}$

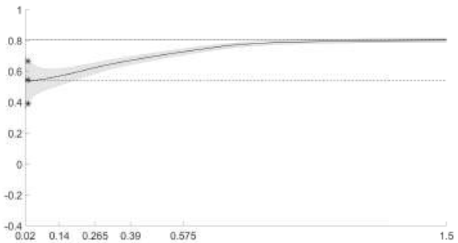


(b) $Z = z_{0.50}$

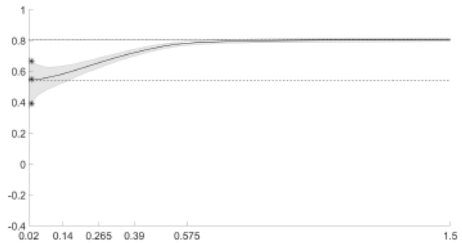


(c) $Z = z_{0.75}$

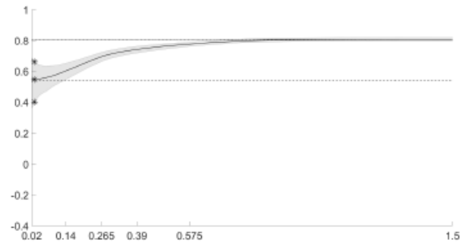
Conditional Kendall's tau and 95% confidence interval for *Model NL* and $\rho = 0.75$.



(d) $Z = z_{0.25}$



(e) $Z = z_{0.50}$



(f) $Z = z_{0.75}$

Figure A.2: Estimated Kendall's tau (solid black line) conditional on percentiles (a) 25, (b) 50, and (c) 75 of Z with 95% pointwise confidence interval (grey shaded) for variables generated for *Model NL*, where dependence is nonlinear and $\rho = 0, 0.75$. The asterisk represents the estimated Kendall's tau for the optimal bandwidth. Horizontal dashed and dotted-dashed lines represent the real conditional and unconditional Kendall's tau respectively. The X-axis shows the range of bandwidths.

Table A.1: Rejection frequencies for simulated models ($S = 1000$ replications).

		$\mathcal{J}_n^{2.5\%}$			$\mathcal{J}_n^{10\%}$		
			n			n	
	α	250	500	1000	250	500	1000
Model 1	1%	1.1	1.0	1.3	1.6	2.3	3.0
	5%	3.7	5.2	4.3	6.2	7.2	10.3
	10%	7.2	10.1	9.9	11.4	13.1	15.6
Model 2	1%	69.0	83.5	98.4	99.8	99.7	100.0
	5%	81.5	91.8	99.7	99.8	99.9	100.0
	10%	87.0	95.4	99.8	99.9	100.0	100.0
Model 3	1%	1.4	1.5	0.5	1.2	5.2	3.9
	5%	6.0	6.2	4.6	4.6	10.6	8.9
	10%	11.2	10.3	10.7	9.2	14.9	14.9
Model 4	1%	45.6	73.7	97.2	96.9	99.3	99.7
	5%	62.8	85.6	99.0	98.4	99.8	99.9
	10%	69.9	90.1	99.4	98.7	99.9	100.0
Model 5	1%	1.7	1.5	2.1	1.8	1.8	1.6
	5%	5.8	4.7	6.9	5.4	4.0	5.1
	10%	10.2	9.2	10.7	9.5	6.9	8.6
Model 6	1%	9.7	38.8	97.2	31.5	83.4	99.7
	5%	19.0	59.1	99.4	46.8	90.5	99.8
	10%	26.8	70.5	99.8	57.9	92.9	99.9
Model 7	1%	2.5	1.7	1.8	2.0	5.6	6.2
	5%	7.4	5.4	7.5	6.6	12.9	13.9
	10%	13.2	10.8	14.6	13.0	17.5	19.5
Model 8	1%	4.3	11.7	39.0	11.7	39.6	70.0
	5%	13.3	26.3	59.1	27.1	55.7	78.9
	10%	20.7	39.3	71.5	36.4	64.6	82.6
Model 9	1%	4.8	1.0	1.1	3.6	4.3	5.2
	5%	11.6	4.7	3.6	11.3	10.3	9.2
	10%	17.3	9.8	6.9	16.4	15.1	12.7
Model 10	1%	11.6	45.3	85.9	45.6	86.4	96.8
	5%	26.4	66.4	93.1	67.4	93.0	98.1
	10%	37.1	76.4	95.5	77.2	95.5	98.5

3

The Effect of Dependence on European Market Risk

A Nonparametric Time Varying Approach

Published in Journal of Business & Economic Statistics, DOI: 10.1080/07350015.2021.1883439

1 Introduction

Financial theory associates low or even negative dependence between assets with a stronger risk-return relationship. However, as pointed out in Chu et al. (2019), *"... diversification may increase systemic risk as it makes banks more similar to each other by holding similar portfolios, exposing them to the same risks"*. It is therefore more important than ever for risk management to estimate properly the dependence between different variables, especially if the degree of multivariate dependence can itself be considered a risk factor for the financial system. Unfortunately, variables in finance tend to have heavy tails and non-Gaussian distributions. The parametric multivariate distributions widely used in other empirical analyses are not appropriate and nor is the measure of correlation given by Pearson's linear correlation.

In this framework, copulas provide a more flexible structure for modeling multivariate dependence and their use has increased substantially in finance. They provide a useful tool, especially in risk management, due largely to the regulation requirements of the Basel Committee and the recent financial crisis. Embrechts et al. (2003b), Palaro and Hotta (2006), Jorion (2007), and Hull (2018), among others, use copulas to calculate risk measures. Embrechts et al. (2003a) and McNeil et al. (2015) are also major references for the analysis of parametric copulas and their classification into different types. To allow for more flexibility, Deheuvels (1979) and Fermanian et al. (2004) analyze empirical unconditional copula processes and their weak convergence in a nonparametric context. Fermanian and Scaillet (2003) study the nonparametric estimation of multivariate unconditional copulas and their derivatives under stationary and α-mixing variables using kernel-based estimators. Gijbels et al. (2011) and Veraverbeke et al. (2011) use nonparametric estimation methods to study conditional copulas under stationary i.i.d. variables. Indeed, Gijbels et al. (2011) define conditional copulas as smoothed empirical distributions to account for the influence of a covariate, while Veraverbeke et al. (2011) analyze their weak convergence and derive the asymptotic distributions for them and some conditional dependence measures. Gijbels et al. (2012) also extend those results from a univariate covariate context to include multivariate and functional covariates. This paper proposes a time varying estimator which generalizes the empirical copula estimator to account for the time variation of the distribution function under local stationary and α-mixing variables.

The time varying behavior of distributions is another characteristic in finance. Studies such as Krishnan et al. (2009), Ferreira et al. (2011), and Ferreira and Orbe (2018) show that the comovements between assets and indexes change over time. Related to this point, Patton (2012) surveys multivariate copula-based models with many applications in economic and financial time series and discusses parametric and semiparametric estimation methods of the models existing in the literature and the inference methods for these cases.

A widely used measure of the degree of dependence between several variables is the average of pairwise linear correlations (see e.g. Joe, 1990; Longin and Solnik, 1995; Moskowitz, 2003; Capiello et al., 2006; Pollet and Wilson, 2010). Beyond Pearson's correlation, Taylor (2007) extends a set of axioms given by Scarsini (1984) for measures of concordance to p-tuples. Schmid and Schmidt (2007) suggest different extensions of the unconditional Spearman's rho and show their asymptotic normality as well as a nonparametric estimator based on empirical copulas. Joe (1990) proposes a family of measures of multivariate concordance and Nelsen (1992) and Nelsen (1996) analyze bivariate and multivariate measures. The former proposes Kendall's tau as an aggregate measure that goes beyond the average for the pairwise Kendall's taus. This is a descriptive statistic that detects monotonicity rather than linearity and which is not restricted to symmetric distributions. We propose an estimator for a multivariate time varying Kendall's tau which has the added advantage of measuring dependence based only on the copula, without using specific marginal distributions. This is especially important when distributions are far from being Gaussian.

Time varying dependence structures have been analyzed using several parametric methodologies. Longin and Solnik (1995) and Engle (2002) show that financial time series correlations and copula dependence parameters often vary over time. To account for this, some authors have considered different approaches to introduce a time varying component into the copula. Palaro and Hotta (2006) and Jondeau and Rockinger (2006) consider time varying marginal distributions while Patton (2006) proposes a GARCH-based estimator for the copula parameter. The approach taken by Hafner and Reznikova (2010), known as the Semiparametric Dynamic Copula (SDC) model, relies on a two-step estimation method based on GARCH and Dynamic Conditional Correlation (DCC) structures where the copula dependence parameter is semiparametrically estimated. A model such as the Stochastic Autoregressive Copula (SCAR) proposed by Hafner and Manner (2012) suggests an estimator with a time varying copula parameter driven by an independent stochastic process. In their application to stock indexes, these authors show that SCAR outperforms the DCC copula model and the model of Patton (2006) in most cases. A different approach is taken by Oh and Patton (2013), who consider estimating copula parameters via a simulated method of moments type approach.

In this paper, the time varying dependence distribution is modeled through a nonparametric multivariate copula with nonparametric marginals that allows for local stationary variables. The main results for consistency and asymptotic normality of the proposed estimator are provided. This result enables nonparametric margins to be used together with a time varying copula, thus filling a gap pointed out by Oh and Patton (2018).

A nonparametric multivariate time varying Kendall's tau estimator is also provided and its consistency and asymptotic normality are shown. To motivate its use in empirical analysis, we conduct a simulation study and compare the estimated rank correlations computed from the proposed estimator and the DCC method developed by Engle (2002). Our findings show that the nonparametric proposal produces a better estimation of the dependence, especially for complex situations. In fact, we find that the nonparametric time varying Kendall's tau performs better in capturing dependence trends for individual and mixture copulas. Unlike the average of bivariate Kendall's taus, this estimator also takes advantage of the multivariate distribution since it accounts for simultaneous association, as remarked in Joe (1990). This is relevant for it to be considered an adequate measure of multivariate dependence.

For the empirical analysis we use multivariate measures to analyze the trend in the dependence of the five main European indexes over more than twenty years. We focus on the multivariate relationship between the DAX (Germany), the FTSE 100 (UK), the IBEX 35 (Spain), the CAC 40 (France), and the SMI (Switzerland). We use daily data from these stock indexes with a sample period from January 1994 to March 2020.

To provide a quantitative measure of the link between multivariate dependence and Euro Stoxx returns, the nonparametric multivariate time varying Kendall's tau estimator is used as a daily measure of dependence in the European financial system. The results show the trend in the indexes' common dependence structure. The plot of the proposed Kendall's tau estimates and the Euro Stoxx index shows their trends over time, with higher dependence when the Euro Stoxx is low and vice versa. On that basis, we use this estimator of dependence to check whether stronger asset dependence is related to bear market scenarios, particularly with worse value-at-risk. To that end, the systemic risk measure $\Delta CoVaR$ proposed by Adrian and Brunnermeier (2016) is adapted for index analysis. They use the difference between conditional quantiles to asses firms' contributions to systemic risk, while we use it to asses the influence of portfolio dependence on the Euro Stoxx quantile. That is, the $CoVaR$ is computed for the Euro Stoxx, conditioned to different quantiles of the nonparametric multivariate time varying Kendall's tau. Thus, if the $CoVaR$ moves to the left when dependence is very high, the degree of dependence between indexes (or portfolios) must be taken into account as a risk factor. In line with our proposal, we estimate the $CoVaR$ through a nonparametric conditional quantile estimator. The sample period analyzed means that the resulting $\Delta CoVaR$ includes the recession periods in Europe and the US and the first few weeks of the COVID-19 pandemic.

The results reveal that the Kendall's tau proposed provides a good indicator for measuring the daily degree of dependence. It is therefore of interest to test for causality on the Euro Stoxx quantile. In a related study, Bouezmarni et al. (2012) set out to analyze Granger causality through a conditional independence test based on nonparametric copulas and the Hellinger distance and apply it to study the nonlinear effects of the asymmetric volatility phenomenon on stock index futures. They find evidence of stronger nonlinear effects from stock index returns to trading volume than from volume to stock returns. Hence, a key issue in this paper is to test whether the quantile of Euro Stoxx at a given time is influenced only by its own past or if there is an additional effect when past information on dependence is also considered. To that end, we use a modified version of the Granger causality test driven by Jeong et al. (2012) which is a nonparametric test for conditional quantile restrictions. The test provides evidence in favor of causality.

The rest of the paper is organized as follows. Section 2 examines the nonparametric multivariate time varying estimators for the copula and Kendall's tau and provides the main asymptotic results. Section 3 presents a simulation study. Section 4 estimates Kendall's tau for the five main European indexes, using different estimators. Section 5 reports the contribution of the dependence level to changes in Euro Stoxx quantiles, and presents the causality test. Section 6 concludes. Details of proofs and simulations are available as Appendix.

2 Multivariate time varying copulas

Sklar (1959) ensures that for any set of p variables $\boldsymbol{Y} = \{Y_j\}_{j=1}^p$ with continuous marginals $F_1, ..., F_p$ and joint distribution F, there is a unique multivariate copula function $C\colon [0,1]^p \to [0,1]$ such that $F(\boldsymbol{y}) = F(y_1, ..., y_p) = C(F_1(y_1), ..., F_p(y_p))$. In a time varying context, $C_t : [0,1]^p \to [0,1]$ such that $F_t(\boldsymbol{y}) = F_t(y_1, ..., y_p) = C_t(F_{1t}(y_1), ..., F_{pt}(y_p))$. We define the following nonparametric estimator, which is a natural smoother considering local stationary variables as defined by Dahlhaus (1997),

$$\hat{C}_t(\boldsymbol{u}) = \frac{1}{Th}\sum_{r=1}^{T} K_{h,tr} I\{Y_{1r} \leq \hat{F}_{1t}^{-1}(u_1), ..., Y_{pr} \leq \hat{F}_{pt}^{-1}(u_p)\} \tag{1}$$

for any $\boldsymbol{u} \in [0,1]^p$, where $K_{h,tr} = k\Big((t-r)/(Th)\Big)$, $k(\cdot)$ denotes the kernel weight used to introduce smoothness into the distribution of the variables, $h > 0$ is the bandwidth that regulates the degree of smoothness, $\hat{F}_{jt}(y) = (Th)^{-1}\sum_{r=1}^{T} K_{h,tr} I\{Y_{jr} \leq y\}$ denotes the nonparametric time varying estimator of the j-marginal distribution, and $I\{\cdot\}$ is the indicator function. Therefore, in this context the smoother accounts for the local distribution structure. Gijbels et al. (2011) propose a similar estimator to compute conditional copulas using i.i.d. samples and smoothing around the conditioned value of a covariate. In our case, a time varying context is considered and the i.i.d. context is generalized to allow for α-mixing local stationary variables.

Some authors, such as Casas et al. (2020), Ferreira et al. (2011), and Kristensen (2012), derive statistical properties for time varying kernel estimators in the regression context for local stationary variables. Under a similar setting in terms of assumptions for the variables involved and the kernels used, we derive the main results for the time varying copula estimator, which are stated in the theorems below. An outline of the proofs is presented and the details are provided in Appendix B.1.

Theorem 1. *(Copula consistency) Assume that: i)* $\mathbf{Y}$ *is a p-dimensional local stationary variable, the sequences* $\{Y_{jt}\}_{t=1}^T$, $j = 1, ..., p$ *are strong* α*-mixing, and it holds for the* α *coefficients that* $T^{-1}\sum_{l=1}^{T-1}\alpha(l) = o\left(T^{-6/5}\right)$*; ii) the joint distribution of* $\mathbf{Y}$ *is time varying, such that the marginal distributions and the copula are functions of* t/T *and they are twice differentiable with respect to the time index; and iii) the kernel is a bounded symmetric second order kernel with* *compact support* $\Omega = [-1, 1]$.

Thus, if h *tends to zero and* Th *tends to infinity as* T *tends to infinity, the mean squared error (MSE) of the copula estimator defined in (1) is*

$$MSE\left(\hat{C}_t(\boldsymbol{u})\right) = Bias^2\left(\hat{C}_t(\boldsymbol{u})\right) + Var\left(\hat{C}_t(\boldsymbol{u})\right),$$

where $c_k = \int_\Omega k(\eta)\eta^2 d\eta$, $d_k = \int_\Omega k(\eta)^2 d\eta$, $b_k = \int_\Omega k(\eta)^4 d\eta \neq 0$ *and*

$$Bias\left(\hat{C}_t(\boldsymbol{u})\right) = \frac{h^2 c_k}{2} C_t''(\boldsymbol{u}) + o\left(h^2\right), \quad Var\left(\hat{C}_t(\boldsymbol{u})\right) = \frac{d_k}{Th} C_t(\boldsymbol{u})\left(1 - C_t(\boldsymbol{u})\right) + o\left(\frac{1}{Th}\right).$$

This result provides the consistency and the leading terms for bias and variance. The value of h that minimizes the $MSE(\hat{C}_t(\boldsymbol{u}))$ gives the expected order of $MSE = O((Th)^{-4/5})$. The result considers a time series context with serial dependence and time varying distributions which generalizes the usual *i.i.d.* context.

Theorem 2. *(Asymptotic normality of the copula estimator) Consider the same assumptions as in Theorem 1, plus* $h = o(T^{-1/5})$ *and* $\int_\Omega k(\eta)^{12} d\eta < \infty$*. Thus, the bias becomes negligible with respect to the variance and*

$$(Th)^{1/2}\left(\hat{C}_t(\boldsymbol{u}) - C_t(\boldsymbol{u})\right) \xrightarrow{d} C_t^L,$$

where C_t^L *is a Gaussian variable with zero mean and asymptotic variance* $\sigma^2(C_t^L) = d_k C_t(\boldsymbol{u})(1 - C_t(\boldsymbol{u}))$.

Outline of proofs (for details see Appendix B.1): Theorem 1 is proved in three steps, which are also of interest for practical purposes in the estimation procedure. First, we consider that the marginals are known and there is no serial dependence. The result arises with the use of standard tools in nonparametric estimation. Second, to allow for estimated marginals and obtain the same leading terms, we use different smoothing parameters for the marginals than for the copula, with those for the marginals being optimal and that for the copula suboptimal. The use of the same smoothing parameter results in a more cumbersome leading term, but still we end up with the same final order. For the sake of simplicity and a clearer understanding of the results, we present the first case. Finally, the serial dependence under α-mixing variables is analyzed through the difference in MSE between the serial dependence and the independent case, and it is proved that the difference tends to zero at a faster rate than the order of MSE for the independent case. For asymptotic normality, the proof follows similar steps to those in Ferreira et al. (2011).

It is noteworthy that the unknown copula affects both asymptotic terms, i.e. bias and variance. Moreover, for the copula estimator the leading term in the variance has an expected structure similar to a Bernoulli distribution. Note that in regression there is a clear separation between mean and variance and the unknown regression function only appears in the bias term (see Table 1 in Fan, 1992, e.g.).

Next, we seek to measure the degree of multivariate dependence through Kendall's tau. A multivariate version of Kendall's tau is defined as in Joe (1990), $\tau = (2^p\int_{I^p} C(\boldsymbol{u})dC(\boldsymbol{u}) - 1)/(2^{p-1}-1)$. This multivariate Kendall's tau measures simultaneous concordance and quantifies the degree of dependence, accounting for common comovements and not for average pairwise dependence only. Therefore, the more variables there are, the more constraints must be satisfied and the fewer concordances are expected. To mitigate this effect, the multivariate Kendall's tau value includes a correction factor. As mentioned above, it provides a marginal distribution free measure of joint dependence that relies only on the copula. Beyond pairwise measures, this estimator takes advantage of the multivariate distribution. Its overall nature and its link to systemic effects make it a candidate for an overall risk measure.

The multivariate time varying Kendall's tau is defined as an extension of the bivariate case. Following Joe (1990), for a time varying context we define

$$\tau_t = \frac{1}{2^{p-1}-1}\left(2^p \int_{I^p} C_t(\boldsymbol{u})dC_t(\boldsymbol{u}) - 1\right). \tag{2}$$

As a particular case, note that when the copula is constant and $p = 2$, the usual bivariate Kendall's tau is obtained. Thus, our proposal is to estimate the multivariate time varying Kendall's tau through the estimator

$$\hat{\tau}_t = \frac{1}{2^{p-1}-1}\left(\frac{2^p}{T^2h^2 - \sum_{r=1}^T K_{h,tr}^2}\sum_{r,s=1}^T K_{h,tr}K_{h,ts}I\{Y_{1r}<Y_{1s},\dots,Y_{pr}<Y_{ps}\} - 1\right), \tag{3}$$

where the weights are based on the recommendations given by Gijbels et al. (2011).

The theorem below states the consistency and the asymptotic distribution of the multivariate Kendall's tau estimator (3) (the proof is given in Appendix B.1).

Theorem 3. *(Kendall's tau consistency and asymptotic normality) Under the same assumptions as in Theorem 1, the time varying Kendall's tau estimator* $\hat{\tau}_t$ *defined in (3) is a consistent estimator of* τ_t *defined in (2), where the asymptotic bias is* $Bias(\hat{\tau}_t) = (2^{p-1}h^2c_k/(2^{p-1}-1))\int_{\mathbb{R}^p}\left(F_t(\boldsymbol{y})f_t''(\boldsymbol{y}) + F_t''(\boldsymbol{y})f_t(\boldsymbol{y})\right)d\boldsymbol{y} + o\left(h^2\right)$. *Moreover, if* Z_t *is a Gaussian variable given by*

$$Z_t = \frac{2^p}{2^{p-1}-1}\left(\int_{I^p} C_t(\boldsymbol{u})dC_t^L(\boldsymbol{u}) + \int_{I^p} C_t^L(\boldsymbol{u})dC_t(\boldsymbol{u})\right),$$

where $C_t^L(\boldsymbol{u})$ *is the limiting distribution of* $(Th)^{1/2}\left(\hat{C}_t(\boldsymbol{u}) - C_t(\boldsymbol{u})\right)$, *then the asymptotic variance of* $\hat{\tau}_t$ *is given by the asymptotic variance of* Z_t, $\sigma^2(Z_t)$. *Additionally, considering the* *assumptions in Theorem 2,*

$$(Th)^{1/2}(\hat{\tau}_t - \tau_t) \xrightarrow{d} Z_t.$$

These theoretical results enable us to further study the performance of the estimator in finite samples and to obtain reliable confidence intervals. To that end, we next conduct a simulation study.

3 Simulation study

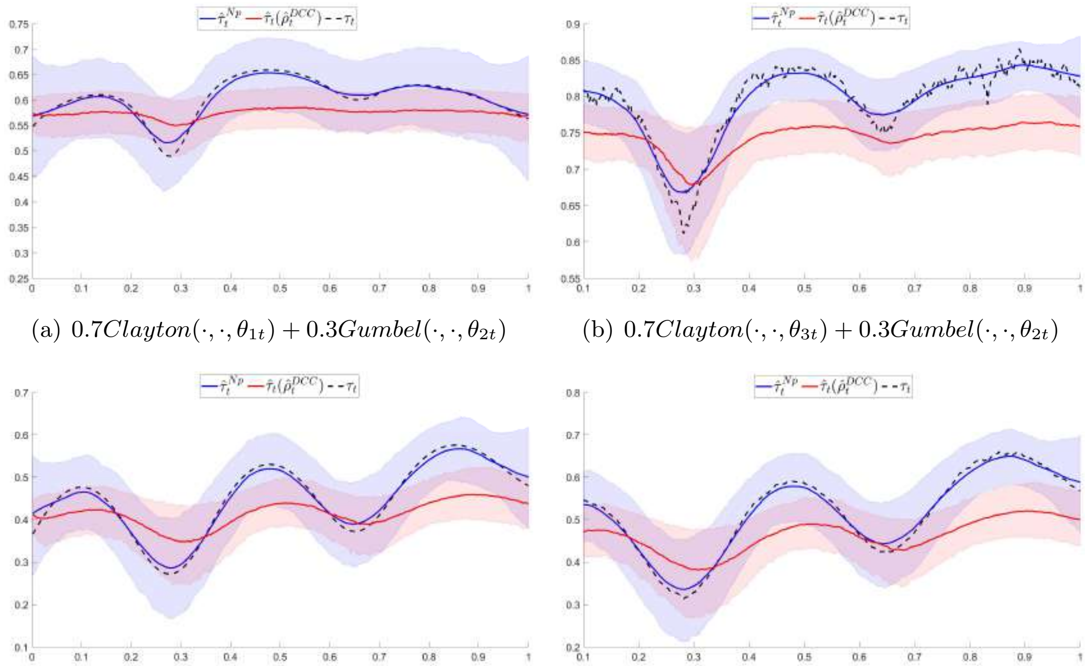
This section looks at the performance of the proposed Kendall's tau estimator. A bivariate context is simulated to enable a comparison to be run with the Kendall's tau obtained from the parametric DCC. We also implement the SDC two-step method in spite of its computational cost, which is extremely high even in the first step. Details of computational costs can be found in Appendix B.2.

For the simulation we consider the mixture copulas

$$C^{MIX}(F(z_{1,t}),F(z_{2,t}),t)=wC_a(F(z_{1,t}),F(z_{2,t}),\theta_{at})+(1-w)C_b(F(z_{1,t}),F(z_{2,t}),\theta_{bt}), \quad 0\leq w\leq 1,$$

where $\theta_{\ell t} = \theta_\ell(t/T)$ is the dependence parameter for $\ell = a, b$. Mixture copulas for modeling dependence structures in finance can be found in Fortin and Kuzmics (2002), Dias and Embrechts (2010), and Loaiza-Maya et al. (2018) among others. For C_a and C_b we consider the Clayton, Gumbel, and Frank copulas. Clayton copulas exhibit strong left tail dependence and relatively weak right tail dependence. Gumbel copulas exhibit strong right tail dependence and relatively weak left tail dependence, and Frank copulas can show strong negative or positive dependence between the marginals.

All the scenarios simulated are based on GARCH marginals where $z_{j,t} = \sigma_{j,t}\varepsilon_{j,t}$, such that $\sigma^2_{j,t} = 5 \cdot 10^{-6} + 0.05z^2_{j,t-1} + 0.9\sigma^2_{j,t-1}$, $j = 1, 2$, and $\varepsilon_{j,t}$ are independent $N(0, 1)$. For the dependence parameter in the copula structure three functional forms are considered: A sinusoidal form $\theta_{1t} = 2 + sin(10t/(3T))$, a trimodal form $\theta_{2t} = 3 + 1.9sin(50t/(3T)) + 4(t/T)^2$, and the autoregressive structure $\theta_{3t} = 0.9\theta_{3,t-1}+\epsilon_t$, where $\epsilon_t \sim N(1, 0.5)$ with an initial value of $\theta_{3,0} = 2$. These scenarios are inspired by the models in Hafner and Reznikova (2010). For the first two the processes are locally stationary whereas the last one remains stationary. $S = 1000$ samples of $T = 500$ observations are considered. The sample size for the empirical part is much larger than 500 but we consider this figure to analyze the performance of the estimator with smaller sizes. This is also a usual size in nonparametric estimation and is suitable for copula estimation (Gijbels et al., 2011) and for quantile testing (Jeong et al., 2012), which is used below. The results also support sizes large enough to ensure the accuracy of the proposed estimator. The Epanechnikov kernel is used and the smoothing parameters are selected minimizing the corresponding mean squared error.



(a) $0.7Clayton(\cdot,\cdot,\theta_{1t}) + 0.3Gumbel(\cdot,\cdot,\theta_{2t})$ (b) $0.7Clayton(\cdot,\cdot,\theta_{3t}) + 0.3Gumbel(\cdot,\cdot,\theta_{2t})$

(c) $0.3Clayton(\cdot,\cdot,\theta_{1t}) + 0.7Frank(\cdot,\cdot,\theta_{2t})$ (d) $0.3Clayton(\cdot,\cdot,\theta_{3t}) + 0.7Frank(\cdot,\cdot,\theta_{2t})$

Figure 3.1: Real and estimated time varying Kendall's tau with confidence intervals. Figures (a) and (b) show the mixture of Clayton and Gumbel copulas, while Figures (c) and (d) show the mixture of Clayton and Frank copulas. The figures on the left use smooth functions for the dependence parameters and those on the right use an autoregressive dependence for the Clayton copula.

Figure 3.1 presents true values of Kendall's tau (dashed black lines) together with the estimated time varying Kendall's tau coefficients and the pointwise confidence interval obtained from both the nonparametric estimator and the DCC procedure (blue and red lines respectively) when different weights and dependence parameter functions are considered. Figures 3.1(a) and 3.1(b) show the results for a mixture of asymmetric copulas: Clayton (more weight) and Gumbel. The former is based on sinusoidal and trimodal dependence structures and the latter uses an autoregressive dependence for the Clayton copula. In Figures 3.1(c) and 3.1(d) the Frank symmetric copula replaces the Gumbel copula but the dependence structures stay the same. The nonparametric estimator gives satisfactory results in capturing the time varying dependence path. By contrast, the rank correlation obtained from the estimated DCC correlations is unable to capture the trend. Actually, it provides a very flat estimation with a very narrow interval, so it fails to give a good estimation of the real time varying Kendall's tau.

Table 3.1: Kendall's tau squared bias, variance, and MSE (1000 replications).

	$\hat{\tau}_t^{Np}$			$\hat{\tau}_t(\hat{\rho}_t^{DCC})$		
Copulas	$Bias^2$	Var.	MSE	$Bias^2$	Var.	MSE
$0.7Clayton(\cdot,\cdot,\theta_{1t})+0.3Gumbel(\cdot,\cdot,\theta_{2t})$	$0.6\cdot 10^{-4}$	$4.0\cdot 10^{-3}$	$4.0\cdot 10^{-3}$	$2.1\cdot 10^{-3}$	$1.2\cdot 10^{-3}$	$3.3\cdot 10^{-3}$
$0.7Clayton(\cdot,\cdot,\theta_{3t})+0.3Gumbel(\cdot,\cdot,\theta_{2t})$	$3.4\cdot 10^{-4}$	$1.5\cdot 10^{-3}$	$1.8\cdot 10^{-3}$	$3.9\cdot 10^{-3}$	$1.4\cdot 10^{-3}$	$5.3\cdot 10^{-3}$
$0.3Clayton(\cdot,\cdot,\theta_{1t})+0.7Frank(\cdot,\cdot,\theta_{2t})$	$1.4\cdot 10^{-4}$	$5.3\cdot 10^{-3}$	$5.4\cdot 10^{-3}$	$5.1\cdot 10^{-3}$	$2.6\cdot 10^{-3}$	$7.7\cdot 10^{-3}$
$0.3Clayton(\cdot,\cdot,\theta_{3t})+0.7Frank(\cdot,\cdot,\theta_{2t})$	$2.0\cdot 10^{-4}$	$5.0\cdot 10^{-3}$	$5.2\cdot 10^{-3}$	$6.9\cdot 10^{-3}$	$3.1\cdot 10^{-3}$	$10\cdot 10^{-3}$

NOTE: The table shows squared bias, variance, and MSE values for the mixtures of copulas considered. Results for the nonparametric time varying Kendall's tau estimator are given in columns two to four. Columns five to seven show the results obtained from the DCC correlation based Kendall's tau. These results were obtained by running 1000 replications.

Table 3.1 shows the squared bias, variance, and MSE obtained for the estimated dependence structures for the cases shown in Figure 3.1. The first column indicates the copula used, the weights, and the dependence parameters. The remaining columns show the squared bias, variance, and MSE for S=1000 replications. The results in the three-column second block are for the nonparametric estimator and those in the three-column third block are for the DCC-based estimator. Observe that although the variance is of the same order for all cases, the nonparametric estimator shows less bias. In summary, the time varying nonparametric Kendall's tau estimator is an interesting alternative estimator for capturing dependence trends.

4 Time varying Kendall's tau for European indexes

This section uses the methodology developed before to analyze the multivariate time varying dependence between five European national indexes (DAX, FTSE 100, IBEX 35, CAC 40, and SMI) and the Euro Stoxx, with data obtained from the Yahoo Finance website `https://es.finance.yahoo.com`. The data set contains daily closing index prices from 01-03-1994 to 03-09-2020.

For the sake of illustration, Table 3.2 presents the static linear correlation coefficients between domestic indexes and Euro Stoxx returns. As expected, linear correlation is high for all pairs.

To account for the time varying dependence, we compute different alternatives and analyze the empirical similarities and differences between them. We estimate the multivariate Kendall's tau proposed in (3) and the time varying version of the average of the pairwise Kendall's taus

Table 3.2: Unconditional linear correlation coefficients matrix for index returns.

Index returns	DAX	FTSE 100	IBEX 35	CAC 40	SMI	Euro Stoxx
DAX	1	0.788	0.772	0.859	0.765	0.921
FTSE 100		1	0.761	0.857	0.779	0.860
IBEX 35			1	0.845	0.722	0.881
CAC 40				1	0.794	0.956
SMI					1	0.814

NOTE: The table contains the static linear correlation coefficients for the European index returns. The first block of columns shows the linear correlation matrix for the national indexes returns while the last column presents the unconditional dependence between each national index returns and the Euro Stoxx returns.

proposed by Kendall and Smith (1940). We also compare the results with the Kendall's tau computed from the DCC correlation. There is no analogous multivariate DCC estimator, so we compute the average of DCC correlations to provide another measure of multivariate dependence (see Moskowitz, 2003; Capiello et al., 2006; Pollet and Wilson, 2010, among others). The results are summarized in Figure 3.2.

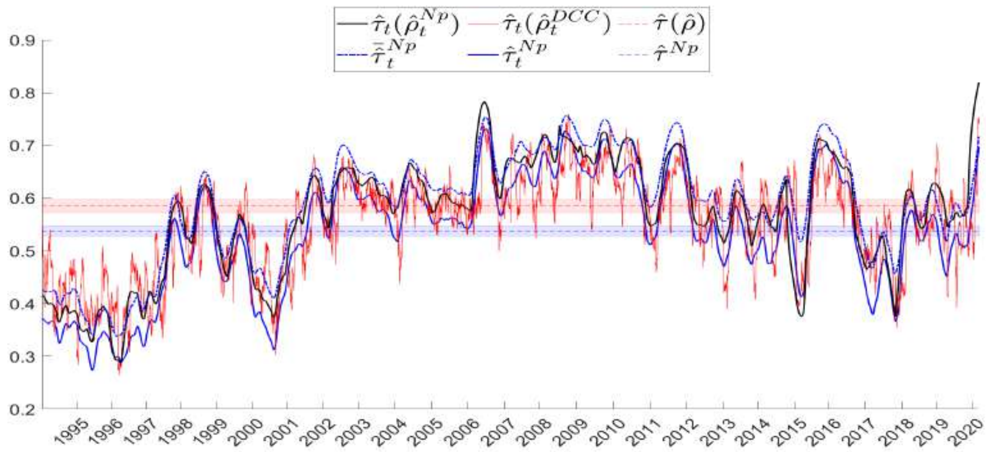


Figure 3.2: Multivariate dependence estimates. The solid and dash-dotted blue lines are the multivariate and pairwise average nonparametric time varying Kendall's taus respectively. The solid red line is the Kendall's tau computed from the pairwise average DCC correlations and the solid black line is the tau computed from the pairwise average nonparametric correlations. The constant nonparametric multivariate Kendall's tau and the rank correlation computed from the constant pairwise average linear correlation are represented with the dashed blue and red lines respectively, together with their corresponding 95% confidence intervals.

The solid blue line in Figure 3.2 is the nonparametric multivariate time varying Kendall's tau estimation, which is accompanied by the pairwise average nonparametric time varying tau (dash-dotted blue line), and the nonparametric multivariate constant Kendall's tau (dashed blue line) and its confidence interval at the 95% level. The estimates of Kendall's tau obtained with the pairwise average DCC correlations are represented by the solid red line and the solid black line is the rank correlation computed from the pairwise average nonparametric correlation. We also plot the tau computed from the unconditional linear correlation average (dashed red line) and its confidence interval. Intervals are constructed using the local block bootstrap method in Paparoditis and Politis (2002), since it is appropriate for inference under a locally stationary setting.

As expected, the nonparametric multivariate time varying tau is lower than the average of the nonparametric pairwise time varying Kendall's tau. The former appears more intuitive due to the requirement of common comovements in the multivariate definition. However, in this context all returns have strong positive dependence and all estimations show similar patterns on different scales (a more in-depth study is provided in Appendix B.2). As can be observed, the nonparametric multivariate time varying Kendall's tau provides a smoother estimation than the one computed from the DCC correlations, making the results easier to interpret (Comparisons with time varying Kendall's tau based on parametric copulas, obtained using the SCAR estimation method, have also been studied. The results show a very wiggly performance, in line with those obtained from the DCC correlations (see Appendix B.2.4)). Finally, the confidence intervals for the unconditional estimators provide evidence that support the idea that there is time varying multivariate dependence.

Figure 3.3 shows the trend for the nonparametric multivariate time varying Kendall's tau estimates taking into account all five indexes, together with the 95% confidence interval, obtained by the local block bootstrap method mentioned previously. The results for multivariate dependence obtained with the Kendall's tau estimator are in line with those obtained in the literature for dynamic correlations.

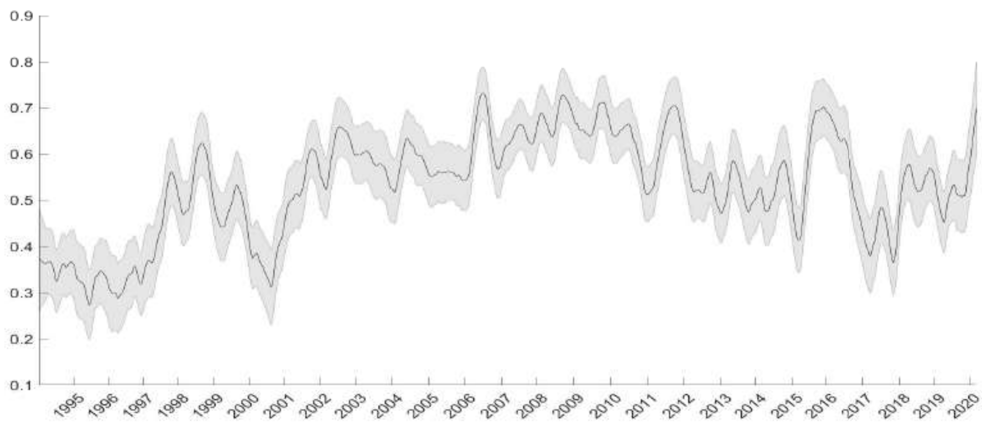


Figure 3.3: Nonparametric multivariate time varying Kendall's tau for the five index returns together with the corresponding 95% confidence interval.

A look at the estimated multivariate dependence trend reveals that from 1994 to mid 1997 the dependence was quite stable. Between 1997 and 1999 (a period that includes the dot.com crisis) dependence first increased and then decreased in January 1999 (when the Euro was introduced as the single official currency). Instabilities and a high dependence downslide continued in 2000, coinciding with a collapse in the value of the Euro in August 2000 when the European Central Bank raised interest rates. These results are in line with the evidence found with dynamic correlations (see Capiello et al., 2006). In 2001, there was a down peak that coincided with the terrorist attack on the World Trade Center in New York City. As time went by, the trend in Kendall's tau became very unstable between 2003 and 2007, but on general lines the dependence level continued to increase. Between 2007 and 2012 there were also high dependence peaks during the US subprime crisis and when the global financial crisis hit Europe. Similar results have been obtained by Wang and Moore (2012), who investigate the effect of the US subprime crisis on the links for sovereign bond CDS spreads between the US and other markets. They conclude that the failure of Lehman Brothers increased dynamic correlations significantly in European countries. Later in the recovery period, events such as the Greek crisis around 2016 can be seen to have resulted in increased peaks in dependence. These results are also echoed by Samitas and Tsakalos (2013), who conclude that correlation between European indexes actually increased significantly during the subprime crisis. During the Greek crisis correlations remained high, with the most closely correlated indexes being those of PIIGS. Additionally, the results we have obtained for bivariate dependence in this period show more abrupt increases in those coefficients related to FTSE returns during Brexit (see Appendix B.2.4). Since late 2019 common dependence has increased again, coinciding with the health alert in Europe as a result of the COVID-19 pandemic.

Figure 3.4 shows the nonparametric multivariate time varying Kendall's tau together with the Euro Stoxx prices. The dash-dotted line is the multivariate Kendall's tau, which is plotted on the left Y-axis, and the solid line is the Euro Stoxx prices series, plotted on the right Y-axis. The dark gray shaded bars represent recession periods in Europe and the light gray shaded bars show recession periods in the US. The figure shows an inversely proportional relationship between multivariate dependence and overall European index prices during recessions, which supports the evidence found in the literature that in recession periods prices decrease and correlation increases while in bull market periods prices increase but there is lower correlation.

5 Euro Stoxx quantiles conditional on Kendall's tau

Adrian and Brunnermeier (2016) propose $\Delta CoVaR$ as a systemic risk measure. This measure is based on the value-at-risk of an institution j (usually a portfolio or market) conditional on

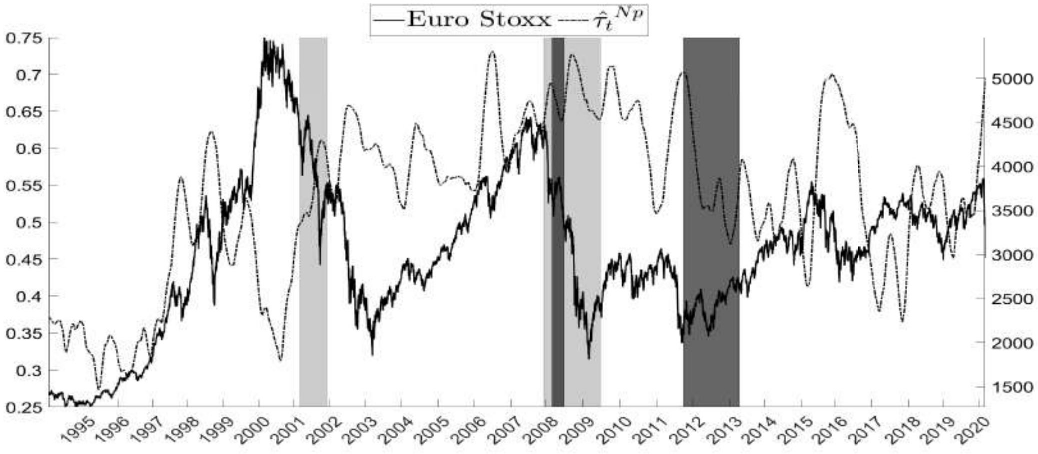


Figure 3.4: Multivariate time varying Kendall's tau and Euro Stoxx. The figure presents the nonparametric multivariate time varying Kendall's tau estimates with the dash-dotted line and the corresponding scale on the left Y-axis. The solid line is the Euro Stoxx prices series, plotted on the right Y-axis. The light and dark gray shaded bars denote recession periods in the US and Europe respectively.

another institution i (a bank or another firm in the portfolio/market). They define $\Delta CoVaR_q^{j|i}$ for a quantile q as $CoVaR_q^{j|VaR_q^i} - CoVaR_q^{j|VaR_{0.5}^i}$, and measure how much the VaR of a global portfolio j changes when institution i moves from the median to quantile q. The distributions of i and j are linked (i is part of j), so a movement is expected. In practice, for the global portfolio they compute weighted average market equity-valued returns for all financial institutions.

In this section the nonparametric multivariate time varying Kendall's tau is interpreted as a risk factor, so measuring it can help to detect the contribution of the degree of dependence within the market to the VaR of market returns. We therefore measure changes in the Euro Stoxx VaR when the estimated Kendall's tau moves from the median to a very high value. Recall that this measure is not distorted by marginal distributions, so it provides a suitable measure of dependence in the financial context not linked to the distribution of returns on each index. Given that the market returns distribution is linked to the joint distribution of the index returns, a movement is expected in a way similar to that of institutions in Adrian and Brunnermeier (2016). We are interested in measuring changes in the $CoVaR$ of the Euro Stoxx at risky quantiles ($q = 0.05, 0.01$) when dependence moves from the median to $1-q = 0.95, 0.99$, which become the risky quantiles for dependence.

The $\Delta CoVaR$ in this context provides a measure of the contribution of stronger index dependence on the risky tail of the market distribution. This is linked to evidence of stronger dependence in bear periods. As an illustration, we first estimate the static nonparametric $\Delta CoVaR$ conditional on the national indexes and the estimated Kendall's tau. Table 3.3 shows the values for quantile $q = 0.05$ in the second column and quantile $q = 0.01$ in the third column. The results show that the contribution of dependence is almost as great as those of each individual index return series.

Table 3.3: Constant $\Delta CoVaR$ estimates for Euro Stoxx returns.

Conditioning variable	$\Delta\widehat{CoVaR}_{q=0.05}$	$\Delta\widehat{CoVaR}_{q=0.01}$
$\widehat{VaR}_q^{IBEX35}$	−0.0264	−0.0465
$\widehat{VaR}_q^{DAX}$	−0.0262	−0.0448
$\widehat{VaR}_q^{FTSE\ 100}$	−0.0236	−0.0738
$\widehat{VaR}_q^{CAC\ 40}$	−0.0235	−0.0472
$\widehat{VaR}_q^{SMI}$	−0.0237	−0.0714
$\widehat{VaR}_{1-q}^{\hat{\tau}_t^{Np}}$	−0.0201	−0.0382

NOTE: This table presents the static $\Delta CoVaR$ estimates for Euro Stoxx returns conditional on national indexes and multivariate dependence risky quantiles.

To account for time varying behavior, we consider $\Delta CoVaR_{t,q}^{X|Y} = CoVaR_{t,q}^{X|VaR_{t,q}^Y} - CoVaR_{t,q}^{X|VaR_{t,0.5}^Y}$. In the systemic risk context, Adrian and Brunnermeier (2016) estimate time varying $\Delta CoVaR$ using quantile regressions conditional on a vector of lagged state variables. Oh and Patton (2018), by contrast, introduce this behavior by considering time varying factor copula models. Here, we estimate $\widehat{CoVaR}_{t,q}^{X|Y=y}$ through the time varying conditional marginal estimator given by $\hat{F}_t^{X|Y=y}(x) = \sum_{r=1}^T w_r(t,y) I\{X_r \leq x\}$, where $w_r(t,y) = K_{h,tr}K_{h,ry} / \Big(\sum_{s=1}^T K_{h,ts}K_{h,sy} \Big)$ and $K_{h,ry} = k\Big((Y_r - y)/h \Big)$ to account for conditional behavior. We calculate the quantile q to obtain the estimated $CoVaR$. Hence, to obtain $\Delta\widehat{CoVaR}_{t,q}^{X|Y}$ we estimate $\widehat{CoVaR}_{t,q}^{X|Y=y}$ for $y = \widehat{VaR}_{t,q}$ and $y = \widehat{VaR}_{t,0.5}$.

Figure 3.5 shows the scatter-plot of quarterly $\Delta CoVaR_{t,q}$ estimates of Euro Stoxx returns against the $VaR_{t,1-q}$ estimates of $\hat{\tau}_t^{Np}$ with blue points for $q = 0.05$ and with red points for $q = 0.01$. The results are given in terms of losses, so the $\Delta CoVaR$ dynamics can be related to Kendall's tau dynamics. The results show a positive, increasing relationship between high dependence quantiles and risky tails of market returns.

The right Y-axis in Figure 3.6 shows the time variation of the estimated $\Delta CoVaR$ of the Euro Stoxx when dependence, measured by Kendall's tau, moves from the median to the high 0.95 and 0.99 quantiles (represented by the dashed and solid red lines respectively). We plot the static $\Delta CoVaR$ estimates with the corresponding confidence interval and also include the static $\Delta CoVaR$ estimates for the individual national indexes. The results in this figure are given in terms of losses to make them easier to compare with the Kendall's tau dynamics. The left Y-

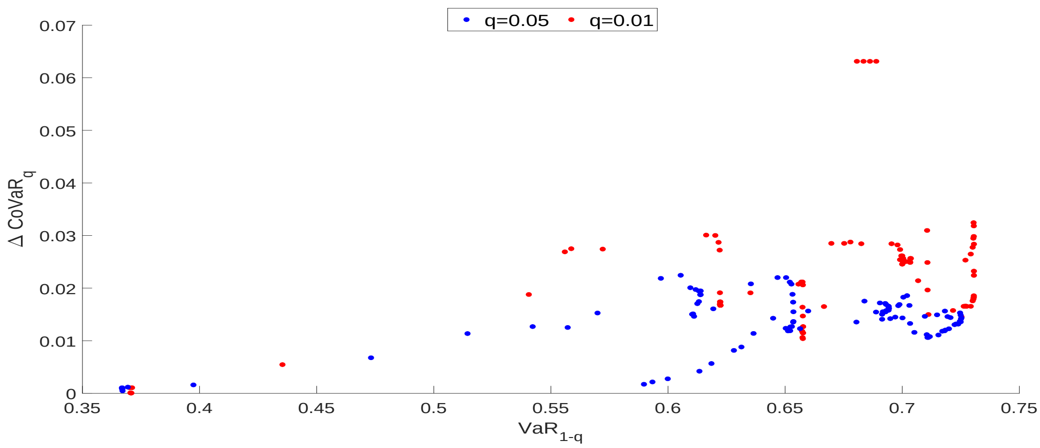


Figure 3.5: Euro Stoxx returns $\Delta CoVaR_{t,q}$ versus high quantiles of $\hat{\tau}_t^{Np}$. Scatter-plot of quarterly $\Delta CoVaR_{t,q}$ estimates of Euro Stoxx returns (Y-axis) against the estimated high quantiles of $\hat{\tau}_t^{Np}$ (X-axis). Results for $q = 0.05$ and $q = 0.01$ are represented by blue and red dots respectively. The results for $\Delta\widehat{CoVaR}_{t,q}$ are given in terms of losses.

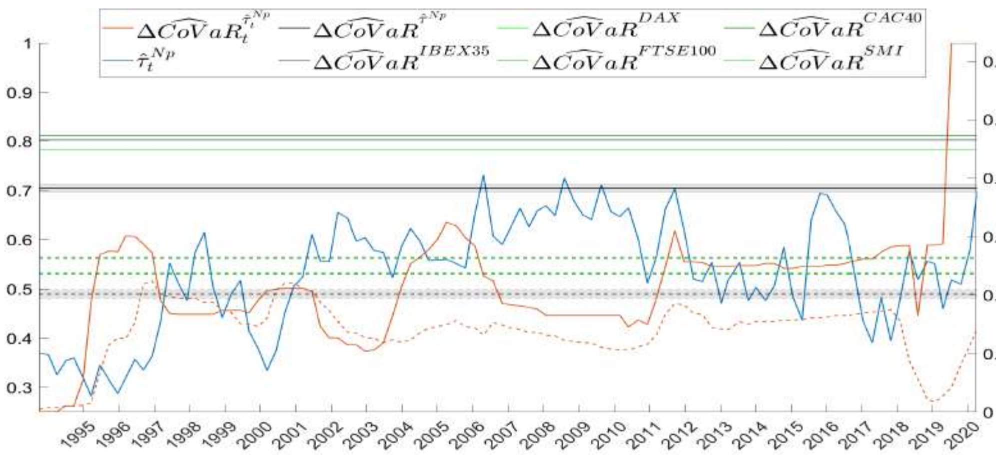


Figure 3.6: Euro Stoxx $\Delta CoVaR$ estimates. The solid blue line is the multivariate $\hat{\tau}_t^{Np}$ with the scale on the left Y-axis. The constant dashed and solid lines represent the static $\Delta CoVaR$ estimates for $q = 0.05$ and $q = 0.01$ summarized in Table 3.3. Confidence intervals of static $\Delta CoVaR$ for high $\hat{\tau}_t^{Np}$ quantiles are also shown, included with shaded gray bands. The dashed and solid red lines represent the estimated $\Delta CoVaR_{t,q}^{VaR_{t,1-q}^{\hat{\tau}_t^{Np}}}$ for $q = 0.05$ and $q = 0.01$ respectively.

axis shows the scale for the nonparametrically estimated multivariate time varying Kendall's tau (solid blue line). The time varying behavior is clear and significant, and larger effects appear to be related to periods of increasing dependence (the effect of high market dependence on market return quantiles is also supported by Figure 3.5). This feature is more evident in the $\Delta CoVaR_{t,0.01}$ estimation. We highlight a spike in the time varying $\Delta CoVaR$ in late 2019 and early 2020 during a period of increasing dependence, which coincides with the spread of the global health alert due to COVID-19. For the sake of comparison, we also estimate $\Delta CoVaR$ when other dependence measures are considered as risk factors. The results show similar patterns, with the proposed multivariate rank correlation having a greater impact on Euro Stoxx quantiles than the dependence measures based on linear correlation (see Appendix B.2.4).

The above results lead us to question causality, which we analyze by adapting the Granger causality test with the statistic provided by Jeong et al. (2012) to this nonparametric context. To that end, we denote the observations for Euro Stoxx returns as X_t. The statistic is $J_T = (T(T-1)h_1h_2)^{-1}\sum_{t=1}^{T}\sum_{s\neq t} K^*_{ts}\epsilon_t\epsilon_s$, where $\epsilon_t = I\{X_t \leq \hat{Q}_q(X_t|X_{t-1})\} - q$, $K^*_{ts} = k\Big((X_{t-1} - X_{s-1})/h_1\Big) \times k\left((\hat{\tau}^{Np}_{t-1} - \hat{\tau}^{Np}_{s-1})/h_2\right)$, and $\hat{Q}_q(X_t|X_{t-1})$ is the estimated nonparametric q-th conditional quantile of X_t given X_{t-1}, computed through $\hat{F}^{X_t|X_{t-1}}(x){=}\sum_{r=1}^{T} K_{h,rX_{t-1}} I\{X_r{\leq}x\}$ with $K_{h,rX_{t-1}} = k\Big((X_r - X_{t-1})/h\Big)$. Following Jeong et al. (2012) we find that $T(h_1h_2)^{1/2}J_T \longrightarrow N(0,\sigma_0^2)$, where a consistent estimator of the variance is $\hat{\sigma_0}^2{=}[2q^2(1{-}q)^2/(T(T{-}1)h_1h_2)]\sum_{s\neq t}K^{*2}_{ts}$. In practice, causality must be tested considering only the information available at each time, so each $\hat{\tau}^{Np}_t$ must be estimated using one-sided kernels. However, since the values for tau are estimated rather than observed, this estimation error affects the distribution of the statistic and the size of the test might become distorted. To avoid such distortion we use the whole sample T for estimation and a uniform subsample $T^* = T/10$ for the test, following Ferreira and Orbe (2018) and references therein.

Figure 3.7 presents the standardized J_T statistic results for different quantiles with the solid black line. We also plot the 1% quantile of the standard normal distribution to test for the significance of J_T. The results show that there is a significant causality effect of the past dependence for low and high Euro Stoxx quantiles. In fact, the evidence suggests that the causality effect is even greater for very extreme quantiles. These results reinforce the interpretation of dependence as a risk factor in financial systems.

Beyond comovements, some macroeconomic variables have also been studied related to bear market scenarios. Being aware of the policy uncertainty effect in financial markets, Baker et al. (2016) proposed the economic policy uncertainty (EPU) index, a measure based on newspaper coverage of policy-related issues. Specifically, EPU has been largely taken into consideration in the analysis of influential factors of market risk and more recently of systemic risk (Pástor and

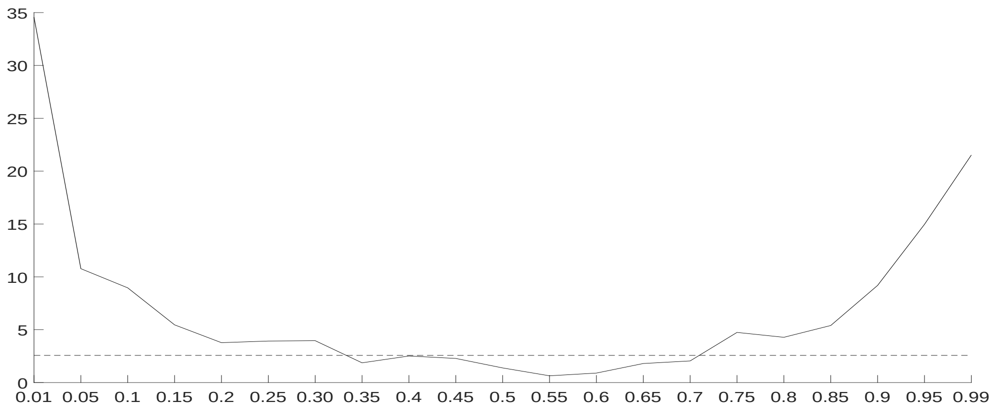


Figure 3.7: J-test statistic. The solid line in the figure shows the standardized J-test statistic for the quantiles in [0.01, 0.99]. The dashed line is the 1% quantile of the standard normal distribution.

Veronesi, 2013; Bernal et al., 2016; Stolbov et al., 2018; Su et al., 2019).

Therefore, we consider the logarithm of daily US EPU index data for the period of study fixed before and test for causality on the Euro Stoxx quantile. As for the comovements, we find causality evidence of the past policy uncertainty for low and high Euro Stoxx quantiles.

6 Conclusions

The dependence structure of financial assets is an area that is being analyzed increasingly in economics and finance, especially since the last global financial crisis. Furthermore, evidence of time variation in asset comovements makes the dependence structure even more important. We analyze the dependence underlying the five principal European indexes beyond Pearson's linear correlation. The relevant literature contains a great deal of research into possible improvements in capturing dependence in variables based on linear correlation, though linear correlation is intended to capture dependence in elliptically distributed variables. Unfortunately, financial variables have asymmetries and heavy-tailed distributions. One way to tackle this issue is to use copulas to describe these dependence structures. We are interested in non-Gaussian measures beyond linear correlation, so we consider Kendall's tau, a coefficient that detects monotonicity rather than linearity and which is not restricted to symmetric distributions.

To study time varying joint dependence between variables, we propose nonparametric estimators for a multivariate time varying copula and the corresponding Kendall's tau. The consistency and asymptotic normality of the two estimators are obtained under local stationary variables and α-mixing coefficients.

We conduct a Monte Carlo simulation study to analyze the performance of the estimator proposed, and compare the results with the estimated Kendall's tau computed from the DCC in the bivariate context. Different scenarios are simulated, allowing for complex joint distributions and considering sinusoidal, trimodal, and autoregressive behaviors. The proposed estimator performs better in all scenarios. Moreover, the improvement obtained in the case of a mixture of Archimedean copulas is even more substantial.

In an empirical application we use the nonparametric multivariate time varying measure to analyze the dependence of the returns of the five European indexes. The proposed Kendall's tau estimator provides a smoother estimation than the rank correlation computed from the pairwise average DCC, making it easier to interpret as an indicator of the dependence trend. An analysis of multivariate dependence shows that features observed in the pairwise Kendall's tau trend still remain. As expected, the multivariate measure proposed in this paper is more sensitive to new variables. Moreover, the trend estimated appears to be linked to major historical events such as the dot.com bubble, the September 11 attack, the global financial crisis, and Brexit. Very recently, the effect of COVID-19 has led to a very high increase in the level of dependence between European indexes.

We use the nonparametric multivariate time varying Kendall's tau as a comovements indicator to analyze the effect of high dependence in bear market trends through the $\Delta CoVaR$ risk measure, which is defined as the contribution to the market returns quantile when dependence changes from the median to high. The results show significant time varying behavior on the part of the $\Delta CoVaR$ risk measure. Finally, we study whether there is causality from the comovements to the risky tails of market returns. A Granger causality type test provides evidence that there is.

In practice, the nonparametric time varying Kendall's tau estimations adequately reflect the trend in multivariate dependence between European indexes. Moreover, this tau provides a source of information on the behavior of the economic system over time for the European Central Bank, which is also helpful for banking regulation. An interesting topic for future research would be to relate this proposed measure of dependence to macroeconomic variables.

Although it goes beyond the scope of this chapter, we introduce the EPU index as a macroeconomic variable that can negatively affect the Euro Stoxx quantiles. The results point towards a significant causal effect of the economic policy uncertainty on the Euro Stoxx risky quantiles. This result leaves for further research to analyze how policy uncertainty determines the comovements of financial markets.

References

Adrian, T. and Brunnermeier, M. K. (2016),"CoVaR," *American Economic Review*, 106(7), 1705–1741.

Baker, S. R., Bloom, N., and Davis, S. J. (2016),"Measuring Economic Policy Uncertainty," *The Quarterly Journal of Economics*, 131(4), 1593–1636.

Bernal, O., Gnabo, J.-Y., and Guilmin, G. (2016),"Economic Policy Uncertainty and Risk Spillovers in the Eurozone," *Journal of International Money and Finance*, 65, 24–45.

Bouezmarni, T., Rombouts, J. V., and Taamouti, A. (2012),"Nonparametric Copula-Based Test for Conditional Independence with Applications to Granger Causality," *Journal of Business & Economic Statistics*, 30(2), 275–287.

Capiello, L., Engle, R., and Sheppard, K. (2006),"Asymmetric Dynamics in the Correlations of Global Equity and Bond Returns," *Journal of Financial Econometrics*, 4(4), 537–572.

Casas, I., Ferreira, E., and Orbe, S. (2020),"Time-Varying Coefficient Estimation in SURE Models. Aplication to Portfolio Management," *Journal of Financial Econometrics*, Doi: 10.1093/jjfinec/nbz010.

Chu, Y., Deng, S., and Xia, C. (2019),"Bank Geographic Diversification and Systemic Risk," *The Review of Financial Studies*, 33(10), 4811–4838.

Dahlhaus, R. (1997),"Fitting Time Series Models to Nonstationary Processes," *The Annals of Statistics*, 25, 1–37.

Deheuvels, P. (1979),"La Fonction de Dépendance Empirique et ses Propriétés: Un Test Non Paramétrique d'Independence," *Bulletin de l'Académie Royal de Belgique, Classe des Sciences*, 65, 274–292.

Dias, A. and Embrechts, P. (2010),"Modeling Exchange Rate Dependence Dynamics at Different Time Horizons," *Journal of International Money and Finance*, 29(8), 1687–1705.

Doukhan, P. (1994),"Mixing," in *Mixing. Lecture Notes in Statistics*, New York: Springer, pp. 15–23.

Embrechts, P., Lindskog, F., and McNeil, A. (2003a),"Modelling Dependence with Copulas and Applications to Risk Management," in *Handbook of Heavy Tailed Distributions in Finance: Book 1*, ed. S.T. Rachev, Amsterdam, NL: Elsevier Science B.V., pp. 329–384.

Embrechts, P., Höing, A., and Juri, A. (2003b), "Using Copulae to Bound the Value-at-Risk for Functions of Dependent Risks," *Finance and Stochastics*, 7(2), 145–167.

Engle, R. (2002), "Dynamic Conditional Correlation: A Simple Class of Multivariate Generalized Autoregressive Conditional Heteroskedasticity Models," *Journal of Business & Economic Statistics*, 20(3), 339–350.

Fan, J. (1992), "Design-Adaptive Nonparametric Regression," *Journal of the American Statistical Association*, 87(420), 998–1004.

Fermanian, J. and Scaillet, O. (2003), "Nonparametric Estimation of Copulas for Time Series," *Journal of Risk*, 5, 25–54.

Fermanian, J., Radulović, D., and Wegkamp, M. (2004), "Weak Converge of Empirical Copula Processes," *Bernoulli*, 10(5), 847–860.

Ferreira, E., Gil-Bazo, J., and Orbe, S. (2011), "Conditional Beta Pricing Models: A Nonparametric Approach," *Journal of Banking & Finance*, 35(12), 3362–3382.

Ferreira, E. and Orbe, S. (2018), "Why are There Time-Varying Comovements in the European Stock Market?," *The European Journal of Finance*, 24(10), 828–848.

Fortin, I. and Kuzmics, C. (2002), "Tail-Dependence in Stock-Return Pairs," *Intelligent Systems in Accounting, Finance & Management*, 11, 89–107.

Gijbels, I., Veraverbeke, N., and Omelka, M. (2011), "Conditional Copulas, Association Measures and their Applications," *Computational Statistics & Data Analysis*, 55(5), 1919–1932.

Gijbels, I., Omelka, M., and Veraverbeke, N. (2012), "Multivariate and Functional Covariates and Conditional Copulas," *Electronic Journal of Statistics*, 6, 1273–1306.

Hafner, C. M. and Reznikova, O. (2010), "Efficient Estimation of a Semiparametric Dynamic Copula Model," *Computational Statistics & Data Analysis*, 54(11), 2609–2627.

Hafner, C. M. and Manner, H. (2012), "Dynamic Stochastic Copula Models: Estimation, Inference and Applications," *Journal of Applied Econometrics*, 27(2), 269–295.

Hart, J. D. and Vieu, P. (1990), "Data-Driven Bandwidth Choice for Density Estimation Based on Dependent Data," *The Annals of Statistics*, 18, 873–890.

Hull, J. C. (2018), *Options, Futures, and Other Derivatives*, (10th ed.), New York: Pearson Education.

Jeong, K., Härdle, W., and Song, S. (2012),"A Consistent Nonparametric Test for Causality in Quantile," *Econometric Theory*, 28, 861–887.

Joe, H. (1990),"Multivariate Concordance," *Journal of Multivariate Analysis*, 35, 12–30.

Jondeau, E. and Rockinger, M. (2006),"The Copula-GARCH Model of Conditional Dependencies: An International Stock Market Application," *Journal of International Money and Finance*, 25(5), 827–853.

Jorion, P. (2007), *Financial Risk Manager Handbook* (Vol. 406, revised ed.), Hoboken, NJ: John Wiley & Sons.

Kendall, M. G. and Smith, B. B. (1940),"On the Method of Paired Comparisons," *Biometrika*, 31(3/4), 324–345.

Krishnan, C., Petkova, R., and Ritchken, P. (2009),"Correlation Risk," *Journal of Empirical Finance*, 16(3), 353–367.

Kristensen, D. (2012),"Non-Parametric Detection and Estimation of Structural Change," *The Econometrics Journal*, 15(3), 420–461.

Loaiza-Maya, R., Smith, M. S., and Maneesoonthorn, W. (2018),"Time Series Copulas for Heteroskedastic Data," *Journal of Applied Econometrics*, 33(3), 332–354.

Longin, F. and Solnik, B. (1995),"Is the Correlation in International Equity Returns Constant: 1960-1990?," *Journal of International Money and Finance*, 14, 3–26.

McNeil, A., Frey, R., and Embrechts, P. (2015), *Quantitative Risk Management*, (revised ed.), Princeton, NJ: Princeton University Press.

Moskowitz, T. J. (2003),"An Analysis of Covariance Risk and Pricing Anomalies," *The Review of Financial Studies*, 16(2), 417–457.

Nelsen, R. B. (1992),"On Measures of Association as Measures of Positive Dependence," *Statistics & Probability Letters*, 14(4), 269–274.

——— (1996),"Nonparametric Measures of Multivariate Association," in *Distributions with Fixed Marginals and Related Topics, IMS Lecture Notes-Monograph Series* (Vol. 28), eds. L. Rüschendorf, B. Schweizer and M. D. Taylor, Hayward, CA: Institute of Mathematical Statistics, pp. 223–232.

Oh, D. H. and Patton, A. J. (2013), "Simulated Method of Moments Estimation for Copula-Based Multivariate Models," *Journal of the American Statistical Association*, 108(502), 689–700.

——— (2018), "Time-Varying Systemic Risk: Evidence From a Dynamic Copula Model of CDS Spreads," *Journal of Business & Economic Statistics*, 36(2), 181–195.

Palaro, H. P. and Hotta, L. K. (2006), "Using Conditional Copula to Estimate Value at Risk," *Journal of Data Science*, 4, 93–115.

Paparoditis, E. and Politis, D. N. (2002), "Local Block Bootstrap," *Comptes Rendus Mathematique*, 335(11), 959–962.

Pástor, L. and Veronesi, P. (2013), "Political Uncertainty and Risk Premia," *Journal of Financial Economics*, 110(3), 520–545.

Patton, A. J. (2006), "Modelling Asymmetric Exchange Rate Dependence," *International Economic Review*, 47(2), 527–556.

——— (2012), "A Review of Copula Models for Economic Time Series," *Journal of Multivariate Analysis*, 110, 4–18.

Pollet, J. M. and Wilson, M. (2010), "Average Correlation and Stock Market Returns," *Journal of Financial Economics*, 96(3), 364–380.

Rabhi, Y. and Bouezmarni, T. (2019), "Nonparametric Inference for Copulas and Measures of Dependence Under Length-Biased Sampling and Informative Censoring," *Journal of the American Statistical Association*, 00(531), 1–11.

Samitas, A. and Tsakalos, I. (2013), "How can a Small Country Affect the European Economy? The Greek Contagion Phenomenon," *Journal of International Financial Markets, Institutions and Money*, 25, 18–32.

Scarsini, M. (1984), "On Measures of Concordance," *Stochastica*, 8(3), 201–218.

Schmid, F. and Schmidt, R. (2007), "Multivariate Extensions of Spearman's Rho and Related Statistics," *Statistics & Probability Letters*, 77(4), 407–416.

Sklar, M. (1959), "Fonctions de Répartition à N Dimensions et Leurs Marges," *Publications de l'Institut de Statistique de L'Université de Paris*, 8, 229–231.

Stolbov, M., Karminsky, A., and Shchepeleva, M. (2018), "Does Economic Policy Uncertainty Lead Systemic Risk? A Comparative Analysis of Selected European Countries," *Comparative Economic Studies*, 60(3), 332–360.

Su, Z., Fang, T., and Yin, L. (2019), "Understanding Stock Market Volatility: What is the Role of US Uncertainty?," *The North American Journal of Economics and Finance*, 48, 582–590.

Taylor, M. (2007), "Multivariate Measures of Concordance," *Annals of the Institute of Statistical Mathematics*, 59(4), 789–806.

Veraverbeke, N., Omelka, M., and Gijbels, I. (2011), "Estimation of a Conditional Copula and Association Measures," *Scandinavian Journal of Statistics*, 38(4), 766–780.

Vieu, P. (1991), "Quadratic Errors for Nonparametric Estimates Under Dependence," *Journal of Multivariate Analysis*, 39(2), 324–347.

Wang, P. and Moore, T. (2012), "The Integration of the Credit Default Swap Markets During the US Subprime Crisis: Dynamic Correlation Analysis," *Journal of International Financial Markets, Institutions and Money*, 22, 1–15.

White, H. and Domowitz, I. (1984), "Nonlinear Regression with Dependent Observations," *Econometrica*, 52, 143–161.

Appendix B

B.1 Proofs. This section contains the proofs of theorems stated in Section 2.

B.1.1 Proof of Theorem 1

The multivariate version of Sklar's theorem can be expressed in a time varying context as $F_t(y_1,...,y_p) = C_t(F_{1t}(y_1),...,F_{pt}(y_p))$. This expression can in turn be inverted to obtain the time varying copula structure: $C_t(u_1,....,u_p) = F_t(F_{1t}^{-1}(u_1),...,F_{pt}^{-1}(u_p))$. First, we derive the asymptotic properties of $\hat{F}_t(\boldsymbol{y})$ at a general point $\boldsymbol{y} = (y_1,...,y_p)$,

$$\hat{F}_t(\boldsymbol{y}) = \frac{1}{Th}\sum_{r=1}^{T} K_{h,tr} I\{Y_{1r} \leq y_1,...,Y_{pr} \leq y_p\} = \frac{1}{Th}\sum_{r=1}^{T} K_{h,tr} I\{\boldsymbol{Y}_r \leq \boldsymbol{y}\},$$

where $\boldsymbol{Y}_r = (Y_{1r},...,Y_{pr})$ and $\{\boldsymbol{Y}_r \leq \boldsymbol{y}\} = \{Y_{1r} \leq y_1,...,Y_{pr} \leq y_p\}$. For the bias, standard tools for nonparametric estimation give

$$\begin{aligned} E\left(\hat{F}_t(\boldsymbol{y})\right) &= E\left[\frac{1}{Th}\sum_{r=1}^{T} K_{h,tr} I\{\boldsymbol{Y}_r \leq \boldsymbol{y}\}\right] = \frac{1}{Th}\sum_{r=1}^{T} K_{h,tr}\int_{-\infty}^{y_p}\cdots\int_{-\infty}^{y_1} f_r(\boldsymbol{w})d\boldsymbol{w} \\ &= \frac{1}{Th}\sum_{r=1}^{T} K_{h,tr} F_r(\boldsymbol{y}) = F_t(\boldsymbol{y}) + \frac{h^2 c_k}{2} F_t''(\boldsymbol{y}) + o(h^2). \end{aligned}$$

The variance $V(\hat{F}_t(\boldsymbol{y}))$ can be written as:

$$\begin{aligned} V(\hat{F}_t(\boldsymbol{y})) &= E\Big[\hat{F}_t(\boldsymbol{y}) - E(\hat{F}_t(\boldsymbol{y}))\Big]^2 = E\left[\frac{1}{Th}\sum_{r=1}^{T} K_{h,tr} I\{\boldsymbol{Y}_r \leq \boldsymbol{y}\} - F_t(\boldsymbol{y}) - \frac{h^2 c_k}{2} F_t''(\boldsymbol{y}) + o(h^2)\right]^2 \\ &= E\left[\frac{1}{Th}\sum_{r=1}^{T} K_{h,tr}\Big(I\{\boldsymbol{Y}_r \leq \boldsymbol{y}\} - F_t(\boldsymbol{y}) - \frac{h^2 c_k}{2} F_t''(\boldsymbol{y}) + o(h^2)\Big)\right]^2 \\ &= \sum_{r,s=1}^{T} \frac{K_{h,tr} K_{h,ts}}{T^2 h^2}\Big[E\Big(I\{\boldsymbol{Y}_r \leq \boldsymbol{y}\} I\{\boldsymbol{Y}_s \leq \boldsymbol{y}\}\Big) - \Big(F_t(\boldsymbol{y}) + \frac{h^2 c_k}{2} F_t''(\boldsymbol{y}) + o(h^2)\Big) \\ &\quad \times\Big(E(I\{\boldsymbol{Y}_r \leq \boldsymbol{y}\}) + E(I\{\boldsymbol{Y}_s \leq \boldsymbol{y}\})\Big) + \Big(F_t(\boldsymbol{y}) + \frac{h^2 c_k}{2} F_t''(\boldsymbol{y}) + o(h^2)\Big)^2\Big]. \end{aligned}$$

Use $V(\hat{F}_t(\boldsymbol{y})) = V_1^C + V_2^C$, where V_1^C accounts for the sum of equal terms $(r = s)$ and V_2^C for $(r \neq s)$. Then,

$$
\begin{aligned}
V_1^C &= \sum_{r=1}^{T} \frac{K_{h,tr}^2}{T^2h^2}\Big[\Big(1-2F_t(\boldsymbol{y})-h^2c_kF_t''(\boldsymbol{y})+o(h^2)\Big)E(I\{\boldsymbol{Y}_r \le \boldsymbol{y}\})+\Big(F_t(\boldsymbol{y})+\frac{h^2c_k}{2}F_t''(\boldsymbol{y})+o(h^2)\Big)^2\Big] \\
&= \sum_{r=1}^{T} \frac{K_{h,tr}^2}{T^2h^2}\Big[\Big(1-2F_t(\boldsymbol{y})-h^2c_kF_t''(\boldsymbol{y})\Big)F_r(\boldsymbol{y})+F_t(\boldsymbol{y})^2+h^2c_kF_t(\boldsymbol{y})F_t''(\boldsymbol{y})+o(h^2)\Big] \\
&= \Big(1-2F_t(\boldsymbol{y})-h^2c_kF_t''(\boldsymbol{y})\Big)V_{11}^C + \Big(F_t(\boldsymbol{y})^2+h^2c_kF_t(\boldsymbol{y})F_t''(\boldsymbol{y})+o(h^2)\Big)V_{12}^C, \qquad (4)
\end{aligned}
$$

where $V_{11}^C := (Th)^{-2}\sum_{r=1}^T K_{h,tr}^2 F_r(\boldsymbol{y})$ and $V_{12}^C := (Th)^{-2}\sum_{r=1}^T K_{h,tr}^2$. Now, kernel properties and Taylor expansion lead to

$$
V_{11}^C = \frac{1}{Th}\int_{-1}^{1} k(\eta)^2 F_{t-h\eta T}(\boldsymbol{y})d\eta + o\left(\frac{1}{T}\right) = \frac{d_k}{Th}F_t(\boldsymbol{y}) + o\left(\frac{1}{Th}\right).
$$

Similarly, $V_{12}^C = d_k(Th)^{-1} + o\left((Th)^{-1}\right)$. Summing up the above terms in (4), the term of the variance is

$$
V_1^C = \frac{d_k}{Th}F_t(\boldsymbol{y})(1-F_t(\boldsymbol{y})) + o\left(\frac{1}{Th}\right).
$$

For the cross terms, $r \neq s$, write

$$
\begin{aligned}
V_2^C = \sum_{\substack{r=1\\ s\neq r}}^{T} \frac{K_{h,tr}K_{h,ts}}{T^2h^2} \Big[& E\Big(I\{max(\boldsymbol{Y_r},\boldsymbol{Y_s}) \le \boldsymbol{y}\}\Big) - \Big(F_t(\boldsymbol{y}) + \frac{h^2c_k}{2}F_t''(\boldsymbol{y}) + o\left(h^2\right)\Big) \\
& \times\Big(E(I\{\boldsymbol{Y}_r \le \boldsymbol{y}\})+E(I\{\boldsymbol{Y}_s \le \boldsymbol{y}\})\Big) + \Big(F_t(\boldsymbol{y}) + \frac{h^2c_k}{2}F_t''(\boldsymbol{y}) + o\left(h^2\right)\Big)^2\Big],
\end{aligned}
$$

where $\max(\boldsymbol{Y_r},\boldsymbol{Y_s}) = (\max(Y_{1r},Y_{1s}),...,\max(Y_{pr},Y_{ps}))$ is the p-tuple of maximums. Note that

$$
E\Big(I\{\max(\boldsymbol{Y_r},\boldsymbol{Y_s}) \le \boldsymbol{y}\}\Big) = \int_{-\infty}^{y_p}\cdots\int_{-\infty}^{y_1} f_{r,s}^M(\boldsymbol{w})d\boldsymbol{w} = F_{r,s}^M(\boldsymbol{y}) = F_r(\boldsymbol{y})F_s(\boldsymbol{y}),
$$

where f^M and F^M denote the density and the distribution functions of the p-tuple of maximums respectively. The cross term can then be rewritten as

$$
\begin{aligned}
V_2^C &= \sum_{\substack{r=1\\ s\neq r}}^{T} \frac{K_{h,tr}K_{h,ts}}{T^2h^2}\Big[F_r(\boldsymbol{y})F_s(\boldsymbol{y}) - \Big(F_r(\boldsymbol{y})+F_s(\boldsymbol{y})\Big)\Big(F_t(\boldsymbol{y})+\frac{h^2}{2}c_kF_t''(\boldsymbol{y})+o\left(h^2\right)\Big) + \\
&\quad + \Big(F_t(\boldsymbol{y}) + \frac{h^2c_k}{2}F_t''(\boldsymbol{y}) + o(h^2)\Big)^2\Big] \\
&= V_{21}^C - \Big(F_t(\boldsymbol{y})+\frac{h^2}{2}c_kF_t''(\boldsymbol{y})+o\left(h^2\right)\Big)V_{22}^C + \Big(F_t(\boldsymbol{y})+\frac{h^2c_k}{2}F_t''(\boldsymbol{y})+o(h^2)\Big)^2 V_{23}^C, \qquad (5)
\end{aligned}
$$

where $V_{21}^C := (Th)^{-2} \sum_{\substack{r=1 \\ s\neq r}}^{T} K_{h,tr} K_{h,ts} F_r(\boldsymbol{y}) F_s(\boldsymbol{y})$. Changing variables gives

$$V_{21}^C = \int_{-1}^{1} k(\eta) F_{t-h\eta T}(\boldsymbol{y}) d\eta \int_{-1}^{1} k(v) F_{t-hvT}(\boldsymbol{y}) dv + o\left(\frac{1}{T}\right) = F_t(\boldsymbol{y})^2 + h^2 F_t(\boldsymbol{y}) F_t''(\boldsymbol{y}) c_k + o\left(h^2\right).$$

A second term can be developed in a similar way:

$$V_{22}^C := \frac{1}{T^2 h^2} \sum_{\substack{r=1 \\ s\neq r}}^{T} K_{h,tr} K_{h,ts} (F_r(\boldsymbol{y}) + F_s(\boldsymbol{y})) = 2F_t(\boldsymbol{y}) + h^2 c_k F_t''(\boldsymbol{y}) + o\left(h^2\right).$$

Defining a last term as $V_{23}^C := (Th)^{-2} \sum_{\substack{r=1 \\ s\neq r}}^{T} K_{h,tr} K_{h,ts}$, the cross term of the variance (5) is

$$V_2^C = -\frac{h^4 c_k^2}{2} F_t''(\boldsymbol{y}) + o\left(h^4\right).$$

Thus, the variance is

$$V\left(\hat{F}_t(\boldsymbol{y})\right) = V_1^C + V_2^C = \frac{d_k}{Th} F_t(\boldsymbol{y}) \Big(1 - F_t(\boldsymbol{y})\Big) + o\left(h^4\right) + o\left(\frac{1}{Th}\right).$$

Given the $Bias(\hat{F}_t(\boldsymbol{y}))$ and variance $V(\hat{F}_t(\boldsymbol{y}))$, the mean squared error of the estimator $\hat{F}_t(\boldsymbol{y})$ is

$$MSE(\hat{F}_t(\boldsymbol{y})) = \frac{h^4 c_k^2}{4} F_t''(\boldsymbol{y})^2 + \frac{d_k}{Th} F_t(\boldsymbol{y}) \left(1 - F_t(\boldsymbol{y})\right) + o\left(\frac{1}{Th}\right) + o\left(h^4\right).$$

Therefore, defining $\tilde{C}_t(\boldsymbol{u}) = \hat{F}_t\left(F_{1t}^{-1}(u_1), ..., F_{pt}^{-1}(u_p)\right)$, the consistency of $\tilde{C}_t$ follows straightforwardly. $\hat{C}_t(\boldsymbol{u}) = \hat{F}_t\left(\hat{F}_{1t}^{-1}(u_1), ..., \hat{F}_{pt}^{-1}(u_p)\right)$, so different smoothing parameters can be used for marginals than for the copula, with the marginal parameter being optimal and the copula parameter suboptimal.

The consistency results can now be expanded straightforwardly to estimated marginals. To that end, rewrite $\hat{C}_t - C_t = \tilde{C}_t - C_t + \hat{C}_t - \tilde{C}_t$. If the smoothing parameter for the marginal estimators $\hat{F}_{jt}$, $j = 1, ..., p$, is optimal and a suboptimal parameter is considered for the multivariate distribution, the leading terms in the MSE are the same for $\hat{C}_t$ and $\tilde{C}_t$. Using the same smoothing parameters will lead to a more cumbersome leading term but will still ultimately result in the same final order using Cauchy-Schwartz. For the first case, the leading terms have a clearer interpretation.

To prove consistency under dependency the idea is to quantify the difference between the MSE obtained for independent variables and that obtained for α-mixing variables. If the difference tends to zero at a rate $O((Th)^{-1})$ it can be considered negligible (see Hart and Vieu, 1990; Vieu, 1991) and the MSE for both scenarios would be asymptotically equal. $MSE(\cdot)$ is obtained under dependent variables, while $\overline{MSE}(\cdot)$ assumes independence. Thus, the difference analyzed is

$$|MSE(\hat{C}_t(\boldsymbol{u})) - \overline{MSE}(\hat{C}_t(\boldsymbol{u}))| = |E(\hat{C}_t(\boldsymbol{u}) - C_t(\boldsymbol{u}))^2 - \overline{E}(\hat{C}_t(\boldsymbol{u}) - C_t(\boldsymbol{u}))^2|.$$

Let $\hat{\boldsymbol{F}}_t^{-1}(\boldsymbol{u}) = (\hat{F}_{1t}^{-1}(u_1), ..., \hat{F}_{pt}^{-1}(u_p))$ be the p-tuple of the inverse marginal distribution functions. This difference can be rewritten as

$$\Big|E\Big[\sum_{r=1}^{T}\frac{K_{h,tr}}{Th}I\{\mathbf{Y}_r \leq \hat{\boldsymbol{F}}_t^{-1}(\boldsymbol{u})\} - C_t(\boldsymbol{u})\Big]^2 - \overline{E}\Big[\sum_{r=1}^{T}\frac{K_{h,tr}}{Th}I\{\mathbf{Y}_r \leq \hat{\boldsymbol{F}}_t^{-1}(\boldsymbol{u})\} - C_t(\boldsymbol{u})\Big]^2\Big|$$

$$= \Big|E\Big[\sum_{r,s=1}^{T}\frac{K_{h,tr}K_{h,ts}}{T^2h^2}\Big(I\{\mathbf{Y}_r \leq \hat{\boldsymbol{F}}_t^{-1}(\boldsymbol{u})\} - C_t(\boldsymbol{u})\Big)\Big(I\{\mathbf{Y}_s \leq \hat{\boldsymbol{F}}_t^{-1}(\boldsymbol{u})\} - C_t(\boldsymbol{u})\Big)\Big]$$

$$-\overline{E}\Big[\sum_{r,s=1}^{T}\frac{K_{h,tr}K_{h,ts}}{T^2h^2}\Big(I\{\mathbf{Y}_r \leq \hat{\boldsymbol{F}}_t^{-1}(\boldsymbol{u})\} - C_t(\boldsymbol{u})\Big)\Big(I\{\mathbf{Y}_s \leq \hat{\boldsymbol{F}}_t^{-1}(\boldsymbol{u})\} - C_t(\boldsymbol{u})\Big)\Big]\Big|.$$

For $r = s$, the difference is null. For the cross terms the difference is

$$\Big|\sum_{\substack{r=1\\ s\neq r}}^{T}\frac{K_{h,tr}K_{h,ts}}{T^2h^2}\Big[E\Big(I\{\mathbf{Y}_r \leq \hat{\boldsymbol{F}}_t^{-1}(\boldsymbol{u})\}I\{\mathbf{Y}_s \leq \hat{\boldsymbol{F}}_t^{-1}(\boldsymbol{u})\}\Big) - \overline{E}\Big(I\{\mathbf{Y}_r \leq \hat{\boldsymbol{F}}_t^{-1}(\boldsymbol{u})\}I\{\mathbf{Y}_s \leq \hat{\boldsymbol{F}}_t^{-1}(\boldsymbol{u})\}\Big)\Big]\Big|$$

$$\leq \sum_{\substack{r=1\\ s\neq r}}^{T}\Big|\frac{K_{h,tr}K_{h,ts}}{T^2h^2}\Big[E\Big(I\{\mathbf{Y}_r \leq \hat{\boldsymbol{F}}_t^{-1}(\boldsymbol{u})\}I\{\mathbf{Y}_s \leq \hat{\boldsymbol{F}}_t^{-1}(\boldsymbol{u})\}\Big) - \overline{E}\Big(I\{\mathbf{Y}_r \leq \hat{\boldsymbol{F}}_t^{-1}(\boldsymbol{u})\}\Big)\overline{E}\Big(I\{\mathbf{Y}_s \leq \hat{\boldsymbol{F}}_t^{-1}(\boldsymbol{u})\}\Big)\Big]\Big|$$

$$= \sum_{\substack{r=1\\ s\neq r}}^{T}\Big|\frac{K_{h,tr}K_{h,ts}}{T^2h^2}cov\Big(I\{\mathbf{Y}_r \leq \hat{\boldsymbol{F}}_t^{-1}(\boldsymbol{u})\}, I\{\mathbf{Y}_s \leq \hat{\boldsymbol{F}}_t^{-1}(\boldsymbol{u})\}\Big)\Big|$$

$$\leq \sum_{\substack{r=1\\ s\neq r}}^{T}\Big|\frac{\overline{K}^2}{T^2h^2}cov\Big(I\{\mathbf{Y}_r \leq \hat{\boldsymbol{F}}_t^{-1}(\boldsymbol{u})\}, I\{\mathbf{Y}_s \leq \hat{\boldsymbol{F}}_t^{-1}(\boldsymbol{u})\}\Big)\Big|,$$

where $\overline{K}$ is a the kernel bound following Assumption *iii)* in Theorem 1. Lemma 3 in Doukhan (1994, p. 10) states that for two random variables X, $\mathscr{F}_{1,k}$-measurable, and Y, $\mathscr{F}_{k+n,\infty}$-measurable, the following covariance inequality holds:

$$|Cov(X, Y)| \leq 4\alpha(n)(ess{-}sup|X|)(ess{-}sup|Y|), \tag{6}$$

where function $ess{-}sup|U|$ is defined as $ess{-}sup|U| = inf\{a \in \mathbb{R} | P(U > a) = 0\}$. If $\{\mathbf{Y}_t\}$ is an α-mixing process, then the random variable $I\{\mathbf{Y}_t \leq \boldsymbol{y}\}$ is also α-mixing for any $\boldsymbol{y} \in \mathbb{R}^p$. Indeed, for random variables $I\{\mathbf{Y}_t \leq \hat{\boldsymbol{F}}_t^{-1}(\boldsymbol{u})\}$, it holds that $ess{-}sup|I\{\mathbf{Y}_t \geq \hat{\boldsymbol{F}}_t^{-1}(\boldsymbol{u})\}| = 1$. Therefore, the covariance expression can be bounded as

$$\sum_{\substack{r=1\\s\neq r}}^{T}\left|\frac{\overline{K}^2}{T^2h^2}cov\Big(I\{\mathbf{Y}_r\leq \hat{\boldsymbol{F}}_t^{-1}(\boldsymbol{u})\},I\{\mathbf{Y}_s\leq \hat{\boldsymbol{F}}_t^{-1}(\boldsymbol{u})\}\Big)\right|\leq \frac{\overline{K}^2}{T^2h^2}\sum_{\substack{r=1\\s\neq r}}^{T}4\alpha(|r-s|)$$

$$\times(ess-sup|I\{\mathbf{Y}_r\leq \hat{\boldsymbol{F}}_t^{-1}(\boldsymbol{u})\}|)(ess-sup|I\{\mathbf{Y}_s\leq \hat{\boldsymbol{F}}_t^{-1}(\boldsymbol{u})\}|)\leq \frac{\overline{K}^2}{T^2h^2}\sum_{\substack{r=1\\s\neq r}}^{T}4\alpha(|r-s|).$$

Additionally, using Assumption $i)$ in Theorem 1,

$$\frac{\overline{K}^2}{T^2h^2}\sum_{\substack{r=1\\s\neq r}}^{T}4\alpha(|r-s|)=\frac{8\overline{K}^2}{T^2h^2}\sum_{s<r}\alpha(r-s)\leq \frac{8\overline{K}^2}{Th^2}\sum_{l=1}^{T-1}\alpha(l)=\frac{8\overline{K}^2}{h^2}o\left(T^{-6/5}\right),$$

which indicates that the difference in MSE from independent to dependent variables tends to zero at a faster rate than $o((Th)^{-1})$. It can thus be concluded that the mean squared errors of $\hat{C}_t$ for the two scenarios are asymptotically equivalent.

B.1.2 Proof of Theorem 2

Write $(Th)^{1/2}\left(\hat{C}_t(\boldsymbol{u})-C_t(\boldsymbol{u})\right)$ as

$$\begin{aligned}(Th)^{1/2}\Big(\hat{C}_t(\boldsymbol{u})-C_t(\boldsymbol{u})\Big) &= (Th)^{-1/2}\sum_{r=1}^{T}\Big[K_{h,tr}\Big(I\{\mathbf{Y}_r\leq \hat{\boldsymbol{F}}_t^{-1}(\boldsymbol{u})\}-E(I\{\mathbf{Y}_r\leq \hat{\boldsymbol{F}}_t^{-1}(\boldsymbol{u})\})\Big)\Big]\\ &\quad+(Th)^{1/2}\Big[E\Big(\sum_{r=1}^{T}\frac{K_{h,tr}}{Th}I\{\mathbf{Y}_r\leq \hat{\boldsymbol{F}}_t^{-1}(\boldsymbol{u})\}\Big)-C_t(\boldsymbol{u})\Big]\\ &= (Th)^{-1/2}\sum_{r=1}^{T}\Big[K_{h,tr}\Big(I\{\mathbf{Y}_r\leq \hat{\boldsymbol{F}}_t^{-1}(\boldsymbol{u})\}-E(I\{\mathbf{Y}_r\leq \hat{\boldsymbol{F}}_t^{-1}(\boldsymbol{u})\})\Big)\Big]\\ &\quad+(Th)^{1/2}\Big[E(\hat{C}_t(\boldsymbol{u}))-C_t(\boldsymbol{u})\Big]\\ &= (Th)^{-1/2}\sum_{r=1}^{T}\Big[K_{h,tr}\Big(I\{\mathbf{Y}_r\leq \hat{\boldsymbol{F}}_t^{-1}(\boldsymbol{u})\}-E(I\{\mathbf{Y}_r\leq \hat{\boldsymbol{F}}_t^{-1}(\boldsymbol{u})\})\Big)\Big]+O(Th^5)^{1/2}.\end{aligned}$$

Theorem 2.4 in White and Domowitz (1984) is used for the asymptotic normality of the term $(Th)^{-1/2}\sum_{r=1}^{T}\Big[K_{h,tr}I\{\mathbf{Y}_r\leq \hat{\boldsymbol{F}}_t^{-1}(\boldsymbol{u})\}-E(K_{h,tr}I\{\mathbf{Y}_r\leq \hat{\boldsymbol{F}}_t^{-1}(\boldsymbol{u})\})\Big]$. Consider $\alpha(T^{-\delta})$ random variables sequence Z_r, where $\delta=r/(r-1)$ for any $r>1$. Then define $S_T=\sum_{r=1}^{T}Z_r$ and $D_a(Th)=(Th)^{-1/2}\sum_{r=a+1}^{a+T}Z_r$. If $i)E(Z_r)=0$, $ii)E|Z_r|^{2r}\leq\delta<\infty$, and $iii)$ there exists D

finite and nonzero such that $E(D_a^2(Th)) - D \to 0$ as $Th \to \infty$, uniformly in a, then

$$(Th)^{-1/2} D^{-1/2} S_T \xrightarrow{d} N(0,1).$$

Take $Z_r = K_{h,tr} I\{\boldsymbol{Y}_r \le \hat{\boldsymbol{F}}_t^{-1}(\boldsymbol{u})\} - E(K_{h,tr} I\{\boldsymbol{Y}_r \le \hat{\boldsymbol{F}}_t^{-1}(\boldsymbol{u})\})$. The first condition $i)$ is satisfied by definition. The second condition $ii)$ is also satisfied by the assumptions in Theorem 2. For the third condition $iii)$ the following holds:

$$\underset{a}{sup}\Big|(Th)^{-1} E\Big[\sum_{r=a+1}^{a+T}\big(K_{h,tr} I\{\boldsymbol{Y}_r \le \hat{\boldsymbol{F}}_t^{-1}(\boldsymbol{u})\} - E\big(K_{h,tr} I\{\boldsymbol{Y}_r \le \hat{\boldsymbol{F}}_t^{-1}(\boldsymbol{u})\}\big)\big)\Big]^2 - d_k C_t(\boldsymbol{u})(1 - C_t(\boldsymbol{u}))\Big|$$

$$= (Th)^{-1} \underset{a}{sup}\Big|\sum_{\substack{r=a+1\\ r\neq s}}^{a+T} K_{h,tr} K_{h,ts} E\Big[\big(I\{\boldsymbol{Y}_r \le \hat{\boldsymbol{F}}_t^{-1}(\boldsymbol{u})\} - E(I\{\boldsymbol{Y}_r \le \hat{\boldsymbol{F}}_t^{-1}(\boldsymbol{u})\})\big)\big(I\{\boldsymbol{Y}_s \le \hat{\boldsymbol{F}}_t^{-1}(\boldsymbol{u})\}$$

$$-E(I\{\boldsymbol{Y}_s \le \hat{\boldsymbol{F}}_t^{-1}(\boldsymbol{u})\})\big)\Big| = (Th)^{-1} \underset{a}{sup}\Big|\sum_{\substack{r=a+1\\ r\neq s}}^{a+T} K_{h,tr} K_{h,ts}\Big[E\big(I\{\boldsymbol{Y}_r \le \hat{\boldsymbol{F}}_t^{-1}(\boldsymbol{u})\} I\{\boldsymbol{Y}_s \le \hat{\boldsymbol{F}}_t^{-1}(\boldsymbol{u})\}\big)$$

$$-E\big(I\{\boldsymbol{Y}_r \le \hat{\boldsymbol{F}}_t^{-1}(\boldsymbol{u})\}\big) E\big(I\{\boldsymbol{Y}_s \le \hat{\boldsymbol{F}}_t^{-1}(\boldsymbol{u})\}\big)\Big]\Big| \le \frac{\overline{K}^2}{Th} \underset{a}{sup} \sum_{\substack{r=a+1\\ r\neq s}}^{a+T} \Big|cov\big(I\{\boldsymbol{Y}_r \le \hat{\boldsymbol{F}}_t^{-1}(\boldsymbol{u})\}, I\{\boldsymbol{Y}_s \le \hat{\boldsymbol{F}}_t^{-1}(\boldsymbol{u})\}\big)\Big|,$$

where $\overline{K}$ is a the kernel bound following Assumption $iii)$ in Theorem 1. On the other hand, the covariance inequality (6) is used as above, taking into account that $ess\text{–}sup|I\{\boldsymbol{Y}_t \ge \hat{\boldsymbol{F}}_t^{-1}(\boldsymbol{u})\}| = 1$ still holds. Therefore, the following bound is established for condition $iii)$:

$$\frac{\overline{K}^2}{Th} \underset{a}{sup} \sum_{\substack{r=a+1\\ r\neq s}}^{a+T} \Big|cov\big(I\{\boldsymbol{Y}_r \le \hat{\boldsymbol{F}}_t^{-1}(\boldsymbol{u})\}, I\{\boldsymbol{Y}_s \le \hat{\boldsymbol{F}}_t^{-1}(\boldsymbol{u})\}\big)\Big| \le \frac{\overline{K}^2}{Th} \underset{a}{sup} \sum_{\substack{r=a+1\\ r< s}}^{a+T} 4\alpha(r-s) \le \frac{8\overline{K}^2}{h} \sum_{l=1}^{T-1} \alpha(l) = o(1),$$

when $h = o(T^{-1/5})$ as $T \to \infty$. Thus,

$$(Th)^{-1/2} \sum_{r=1}^{T}\Big[K_{h,tr} I\{\boldsymbol{Y}_r \le \hat{\boldsymbol{F}}_t^{-1}(\boldsymbol{u})\} - E(K_{h,tr} I\{\boldsymbol{Y}_r \le \hat{\boldsymbol{F}}_t^{-1}(\boldsymbol{u})\})\Big] \xrightarrow{d} N\Big(0, d_k C_t(\boldsymbol{u})(1 - C_t(\boldsymbol{u}))\Big).$$

B.1.3 Proof of Theorem 3

To analyze the consistency of the proposed Kendall's tau estimator, the first step is to derive the asymptotic bias of the estimator. Since $\sum_{r=1}^{T} (Th)^{-2} K_{h,tr}^2 = O((Th)^{-1})$, as $T \to \infty, h \to 0$, and $Th \to \infty$, the asymptotic property of the following modified estimator is derived:

$$\hat{\tau}_t^* = \frac{1}{2^{p-1} - 1}\Big(2^p \sum_{r=1}^{T} \sum_{s=1}^{T} \frac{1}{T^2 h^2} K_{h,tr} K_{h,ts} I\{\boldsymbol{Y}_r < \boldsymbol{Y}_s\} - 1\Big).$$

For the expected value of the estimator $\hat{\tau}_t^*$ it holds that

$$E\big(\hat{\tau}_t^*\big) \;=\; \frac{1}{2^{p-1}-1}\Big(\frac{2^p}{T^2h^2}\sum_{r=1}^{T}\sum_{s=1}^{T}K_{h,tr}K_{h,ts}E\left[I\{\boldsymbol{Y}_r<\boldsymbol{Y}_s\}\right]-1\Big).$$

Applying the law of iterated expectations, the expected value can be rewritten as follows:

$$\begin{aligned}
E\big(\hat{\tau}_t^*\big) &= \frac{1}{2^{p-1}-1}\left[2^p\sum_{r,s=1}^{T}\frac{K_{h,tr}K_{h,ts}}{T^2h^2}E_s\Big[E_r\Big[I\{\boldsymbol{Y}_r<\boldsymbol{Y}_s\}|\boldsymbol{Y}_s\Big]\Big]-1\right]\\
&= \frac{1}{2^{p-1}-1}\left[2^p\sum_{r=1}^{T}\sum_{s=1}^{T}\frac{K_{h,tr}K_{h,ts}}{T^2h^2}E_s\left[F_r(\boldsymbol{Y}_s)\right]-1\right]\\
&= \frac{1}{2^{p-1}-1}\left[2^p\int_1^T\int_1^T\frac{K_{h,tr}K_{h,ts}}{T^2h^2}\int_{\mathbb{R}^p}F_r(\boldsymbol{y})f_s(\boldsymbol{y})d\boldsymbol{y}drds-1+o\left(\frac{1}{T}\right)\right],
\end{aligned}$$

and Taylor expansion leads to

$$\begin{aligned}
E\big(\hat{\tau}_t^*\big) &= \frac{1}{2^{p-1}-1}\left[2^p\int_{-1}^{1}\int_{-1}^{1}k(\eta)k(v)\int_{\mathbb{R}^p}F_{t-h\eta T}(\boldsymbol{y})f_{t-hvT}(\boldsymbol{y})d\boldsymbol{y}dvd\eta-1+o\left(\frac{1}{T}\right)\right]\\
&= \frac{1}{2^{p-1}-1}\left(2^p\int_{\mathbb{R}^p}F_t(\boldsymbol{y})f_t(\boldsymbol{y})d\boldsymbol{y}-1\right)+\frac{2^{p-1}h^2c_k}{2^{p-1}-1}\int_{\mathbb{R}^p}\Big(F_t(\boldsymbol{y})f_t''(\boldsymbol{y})+F_t''(\boldsymbol{y})f_t(\boldsymbol{y})\Big)d\boldsymbol{y}+o\big(h^2\big)\\
&= \tau_t+\frac{2^{p-1}h^2c_k}{2^{p-1}-1}\int_{\mathbb{R}^p}\Big(F_t(\boldsymbol{y})f_t''(\boldsymbol{y})+F_t''(\boldsymbol{y})f_t(\boldsymbol{y})\Big)d\boldsymbol{y}+o\left(h^2\right).
\end{aligned}$$

Define $g_t(\boldsymbol{y})=\Big(F_t(\boldsymbol{y})f_t''(\boldsymbol{y})+F_t''(\boldsymbol{y})f_t(\boldsymbol{y})\Big)$. Thus, for the expected value it holds that

$$E\big(\hat{\tau}_t^*\big) \;=\; \tau_t+\frac{2^{p-1}h^2c_k}{2^{p-1}-1}\int_{\mathbb{R}^p}g_t(\boldsymbol{y})d\boldsymbol{y}+o\left(h^2\right).$$

We now analyze the asymptotic behavior of the nonparametric multivariate time varying Kendall's tau through differences in mean squared error as in the proof of Theorem 1. Thus, the difference between the mean squared errors in the dependent and independent cases is

$$|MSE(\hat{\tau}_t^*)-\overline{MSE}(\hat{\tau}_t^*)|=|E(\hat{\tau}_t^*-\tau_t)^2-\overline{E}(\hat{\tau}_t^*-\tau_t)^2|,$$

which can be developed as

$$\begin{aligned}
&\Big|E\Big[\frac{2^p}{2^{p-1}-1}\sum_{r,s=1}^{T}\frac{K_{h,tr}K_{h,ts}}{T^2h^2}I\{\boldsymbol{Y}_r<\boldsymbol{Y}_s\}-1-\tau_t\Big]^2-\overline{E}\Big[\frac{2^p}{2^{p-1}-1}\sum_{r,s=1}^{T}\frac{K_{h,tr}K_{h,ts}}{T^2h^2}I\{\boldsymbol{Y}_r<\boldsymbol{Y}_s\}-1-\tau_t\Big]^2\Big|\\
&=\Big|\sum_{r,s=1}^{T}\sum_{r',s'=1}^{T}2^{2p}\frac{K_{h,tr}K_{h,tr'}K_{h,ts}K_{h,ts'}}{(2^{p-1}-1)^2T^4h^4}\Big[E\Big(I\{\boldsymbol{Y}_r<\boldsymbol{Y}_s,\boldsymbol{Y}_{r'}<\boldsymbol{Y}_{s'}\}\Big)-\overline{E}\Big(I\{\boldsymbol{Y}_r<\boldsymbol{Y}_s,\boldsymbol{Y}_{r'}<\boldsymbol{Y}_{s'}\}\Big)\Big]\Big|.
\end{aligned}$$

The leading term in this difference corresponds to the cases in which only two of the subindexes coincide. The rest of the terms are of smaller order or null. Here we consider one of the four terms that make up the leading term, since the remaining three terms can be analyzed in a similar manner.

When $r = r' \neq s \neq s'$, fix $\boldsymbol{Y}_r$, which gives

$$\Big|\frac{2^{2p}}{(2^{p-1}-1)^2}\sum_{\substack{r=1\\ s\neq s'\neq r}}^{T}\frac{K_{h,tr}^2 K_{h,ts}K_{h,ts'}}{T^4h^4}\Big[E\big(I\{\boldsymbol{Y}_r<\boldsymbol{Y}_s\}I\{\boldsymbol{Y}_r<\boldsymbol{Y}_{s'}\}\big)-\overline{E}\big(I\{\boldsymbol{Y}_r<\boldsymbol{Y}_s\}\big)\overline{E}\big(I\{\boldsymbol{Y}_r<\boldsymbol{Y}_{s'}\}\big)\Big]\Big|$$

$$=\Big|\frac{2^{2p}}{(2^{p-1}-1)^2}\sum_{\substack{r=1\\ s\neq s'\neq r}}^{T}\frac{K_{h,tr}^2 K_{h,ts}K_{h,ts'}}{T^4h^4}cov\big(I\{\boldsymbol{Y}_r<\boldsymbol{Y}_s\},I\{\boldsymbol{Y}_r<\boldsymbol{Y}_{s'}\}\big)\Big]\Big|$$

$$\leq\Big|\frac{2^{2p}\overline{K}^2}{(2^{p-1}-1)^2}\sum_{\substack{r=1\\ s\neq s'\neq r}}^{T}\frac{K_{h,tr}^2}{T^4h^4}cov\big(I\{\boldsymbol{Y}_r<\boldsymbol{Y}_s\},I\{\boldsymbol{Y}_r<\boldsymbol{Y}_{s'}\}\big)\Big| \tag{7}$$

with a kernel bound $\overline{K}$ following Assumption iii) in Theorem 1. Thus, it suffices to consider the covariance inequality (6) stated in Doukhan (1994, p. 10) and that $ess{-}sup|I\{\boldsymbol{Y}_t\geq\boldsymbol{y}\}|=1$ holds to bound (7),

$$\Big|\frac{2^{2p}\overline{K}^2}{(2^{p-1}-1)^2}\sum_{\substack{r=1\\ s\neq s'\neq r}}^{T}\frac{K_{h,tr}^2}{T^4h^4}cov\big(I\{\boldsymbol{Y}_r<\boldsymbol{Y}_s\},I\{\boldsymbol{Y}_r<\boldsymbol{Y}_{s'}\}\big)\Big|\leq\frac{2^{2p}\overline{K}^2}{(2^{p-1}-1)^2}\sum_{\substack{r=1\\ s\neq s'\neq r}}^{T}\frac{K_{h,tr}^2}{T^4h^4}(ess{-}sup|I\{\boldsymbol{Y}_r<\boldsymbol{Y}_s\}|)$$

$$\times(ess{-}sup|I\{\boldsymbol{Y}_r<\boldsymbol{Y}_{s'}\}|)\leq\frac{2^{2p}\overline{K}^2}{(2^{p-1}-1)^2T^4h^4}\sum_{\substack{r=1\\ s\neq s'\neq r}}^{T}K_{h,tr}^2 4\alpha(|s-s'|)$$

for all $\boldsymbol{Y}_r$, since the bound does not depend on the value $\boldsymbol{Y}_r$. Considering Assumption i) in Theorem 1, the following is obtained:

$$\frac{2^{2p}\overline{K}^2}{(2^{p-1}-1)^2T^4h^4}\sum_{\substack{r=1\\ s\neq s'}}^{T}K_{h,tr}^2 4\alpha(|s-s'|)=\frac{2^{2p}\overline{K}^2}{(2^{p-1}-1)^2T^4h^4}\sum_{\substack{r=1\\ s< s'}}^{T}K_{h,tr}^2 8\alpha(s-s')$$

$$=\frac{2^{2p+3}\overline{K}^2d_k}{(2^{p-1}-1)^2T^3h^3}\sum_{s<s'}\alpha(s-s')\leq\frac{2^{2p+3}\overline{K}^2d_k}{(2^{p-1}-1)^2T^2h^3}\sum_{l=1}^{T-1}\alpha(l)=\frac{2^{2p+3}\overline{K}^2d_k}{(2^{p-1}-1)^2h^3}o\Big(T^{-11/5}\Big).$$

The combination of all the terms shows that the $MSE(\hat{\tau}_t^*)$ is asymtotically equivalent to that with independent variables, $\overline{MSE}(\hat{\tau}_t^*)$. Since $(Th)^{-2}\sum_{s=1}^{T}K_{h,ts}^2=O((Th)^{-1})$, the above results also hold for $\hat{\tau}_t$ and consistency is guaranteed.

Finally, we derive the asymptotic distribution of the $\hat{\tau}_t$ estimator through the convergence of the multivariate copula in a way similar to that already used for conditional copulas by Veraverbeke et al. (2011) and for right-censored length-biased data by Rabhi and Bouezmarni (2019), both in the bivariate case. Actually, it holds that $\int_{I^p} \hat{C}_t(\boldsymbol{u}) d\hat{C}_t(\boldsymbol{u}) = (Th)^{-1} \sum_{s=1}^{T} K_{h,ts} \hat{C}_t(\hat{\boldsymbol{F}}_t(\boldsymbol{Y}_s))$ since the left-hand side is the estimated mean of $\hat{C}_t(\boldsymbol{u})$. Thus, substituting the copula estimator $\hat{C}_t(\boldsymbol{u})$ in the latter expression leads to the following expression: $\sum_{r,s=1}^{T} w_{h,tr} w_{h,ts} I\{\boldsymbol{Y}_r < \boldsymbol{Y}_s\} + \sum_{s=1}^{T} w_{h,ts}^2$, where $w_{h,ts} = (Th)^{-1} K_{h,ts}$. Hence, assuming that $\sum_{s=1}^{T} w_{h,ts}^2 = O\left((Th)^{-1}\right)$ holds, $\hat{\tau}_t$ can be written as

$$\begin{aligned} \hat{\tau}_t &= \frac{1}{2^{p-1}-1}\Big(\frac{2^p}{1-\sum_{s=1}^{T} w_{h,ts}^2}\int_{I^p} \hat{C}_t d\hat{C}_t - \frac{1+(2^p-1)\sum_{s=1}^{T} w_{h,ts}^2}{1-\sum_{s=1}^{T} w_{h,ts}^2}\Big) \\ &= \frac{1}{2^{p-1}-1}\Big(2^p \int_{I^p} \hat{C}_t d\hat{C}_t - 1\Big) + O\left(\frac{1}{Th}\right). \end{aligned}$$

Moreover, provided that Kendall's tau can be written as a functional of the copula, the following map is defined $\phi : C_t \to 1/(2^{p-1}-1)(2^p \int_{I^p} C_t dC_t - 1)$. Similarly to Lemma 1 in Veraverbeke et al. (2011), ϕ is Hadamard differentiable at C_t tangentially to the set of functions on $[0,1]^p$ and its derivative is given by

$$\phi'(\xi) = \frac{2^p}{2^{p-1}-1}\left(\int_{I^p} C_t d\xi + \int_{I^p} \xi dC_t\right).$$

Hence, the delta method establishes that the asymptotic distribution of $\hat{\tau}_t$ is

$$(Th)^{1/2}(\hat{\tau}_t - \tau_t) = (Th)^{1/2}(\phi(\hat{C}_t) - \phi(C_t)) \to \phi'(C_t^L),$$

where C_t^L is the limiting distribution of $(Th)^{-1/2}(\hat{C}_t - C_t)$ so that the variance for $\hat{\tau}_t$ is determined by $\sigma^2(\phi'(C_t^L))$.

B.2 Additional results

B.2.1 Computational cost

Table B.1: Computational cost for time varying Kendall's tau with $S = 100$ replicates in minutes or hours.

Copula	$\hat{\tau}_t^{Np}$	$\hat{\tau}_t(\hat{\rho}_t^{DCC})$	$\hat{\tau}_t(\hat{\theta}_t^{SDC})$
$Clayton(\theta_{1t})$	$23'$	$10'$	$7.5h$
$Clayton(\theta_{2t})$	$24'$	$8'$	$4.5h$
$Clayton(\theta_{3t})$	$24'$	$12'$	$7.5h$

The results in Table B.1 show the computational cost of time varying Kendall's tau for data generated from a Clayton copula with S=100 replicates, expressed in minutes and hours. The results for Kendall's tau computed from the SDC model are those obtained implementing non-consistent estimates, i.e. the first step of the method in question. To obtain consistent estimates more implementation is needed, which would result in an even higher computational cost.

B.2.2 Multivariate versus average pairwise Kendall's tau

It is proved that the multivariate and average pairwise Kendall's tau coincide for $p = 3$ variables (Nelsen, 1996). Nevertheless, we want to compare the performance of the two estimators in higher dimensions. To that end, we conduct a simulation study in a multivariate context of $p = 4$. Four Gaussian variables with zero mean and unit variance correlated through a correlation matrix are generated. We consider several cases for the correlations matrix: $C_1 = I_4$,

$$C_2 = \begin{pmatrix} 1 & 0.85 & 0.5 & 0.7 \\ & 1 & 0.5 & 0.75 \\ & & 1 & 0.55 \\ & & & 1 \end{pmatrix}, \ C_3 = \begin{pmatrix} 1 & 0.85 & -0.5 & -0.7 \\ & 1 & -0.5 & -0.75 \\ & & 1 & 0.55 \\ & & & 1 \end{pmatrix}, \ C_4 = \begin{pmatrix} 1 & 0.85 & -0.5 & -0.7 \\ & 1 & -0.5 & -0.75 \\ & & 1 & 0 \\ & & & 1 \end{pmatrix},$$

$$C_5 = \begin{pmatrix} 1 & 0.85 & 0.5 & 0.7 \\ & 1 & 0.5 & 0.75 \\ & & 1 & 0 \\ & & & 1 \end{pmatrix}, \quad C_6 = \begin{pmatrix} 1 & 0.83 & 0.88 & 0.7 \\ & 1 & 0.75 & 0.93 \\ & & 1 & 0.78 \\ & & & 1 \end{pmatrix}, \quad C_7 = \begin{pmatrix} 1 & 0.85 & 0 & 0.7 \\ & 1 & 0 & 0.75 \\ & & 1 & 0 \\ & & & 1 \end{pmatrix}.$$

Table B.2: Nonparametric multivariate and pairwise average unconditional Kendall's tau for a set of four variables.

τ/*Correlation matrix*	C_1	C_2	C_3	C_4	C_5	C_6	C_7
$\hat{\tau}^{Np}$	-0.012	0.426	-0.059	-0.129	0.376	0.606	0.240
$\bar{\hat{\tau}}^{Np}$	-0.013	0.446	-0.119	-0.177	0.407	0.606	0.277

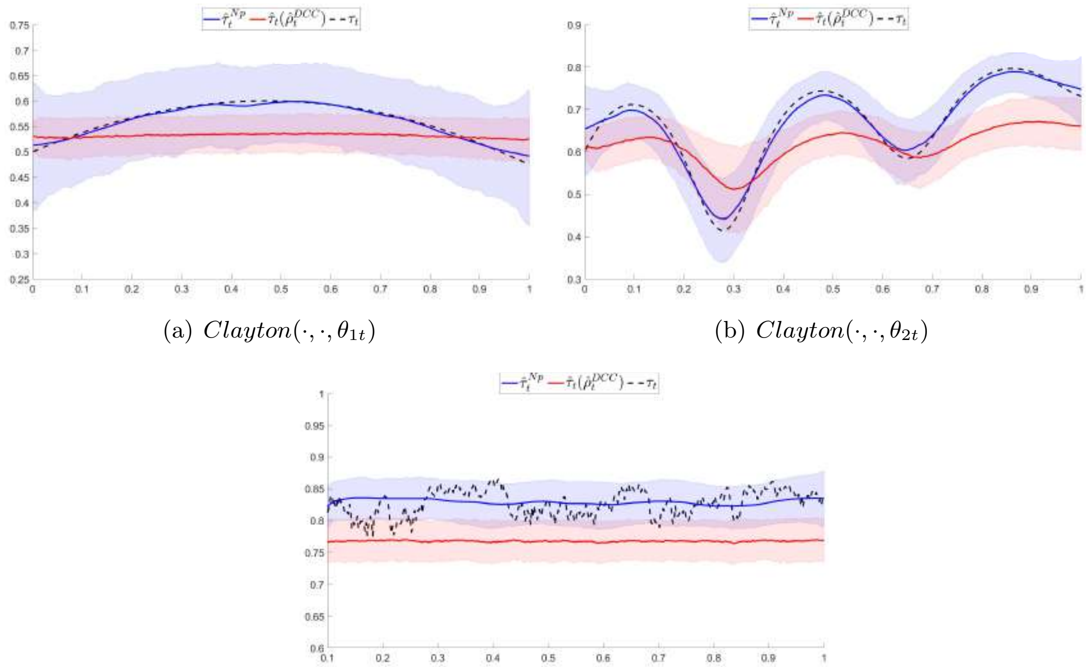
Table B.2 summarizes the results for dependence measured by the nonparametric multivariate Kendall's tau $\hat{\tau}^{Np}$ and the nonparametric pairwise average Kendall's tau $\bar{\hat{\tau}}^{Np}$ for a set of four variables. The first column indicates the measure of dependence. The rest of the columns show the dependence results obtained for variables generated with each correlations matrix. The results show that in a multivariate setting with more than three variables the multivariate Kendall's tau $\left(\hat{\tau}^{Np}\right)$ and the average pairwise Kendall's tau $(\bar{\hat{\tau}}^{Np})$ values are different. Moreover, it is noticeable that the multivariate coefficient is always lower than the average pairwise in absolute value.

B.2.3 Complementary simulation study

Consider a sample of T=500 observations randomly drawn from a Clayton copula $(z_{1,t}, z_{2,t}) \sim C(F(z_{1,t}), F(z_{2,t}), \theta_t)$ with a time varying dependence parameter $\theta_t = \theta(t/T)$ based on two identical univariate GARCH models, where $\sigma^2_{i,t} = 5 \cdot 10^{-6} + 0.05 z^2_{i,t-1} + 0.9\sigma^2_{i,t-1}$, $i = 1,2$ with $z_{i,t} = \sigma_{i,t}\varepsilon_{i,t}$ and $\varepsilon_{i,t} \sim N(0,1)$. Three functional forms for the dependence parameter are simulated: A sinusoidal $\theta_{1t} = 2 + sin(10t/(3T))$, a trimodal $\theta_{2t} = 3 + 1.9 sin(50t/(3T)) + 4(t/T)^2$, and an autoregressive one $\theta_{3t} = 0.9\theta_{3,t-1} + \epsilon_t$, where $\epsilon_t \sim N(1, 0.5)$ and the initial value is $\theta_{3,0} = 2$.

Figure B.1 presents the true values of Kendall's tau (dashed black lines), and the estimated time varying Kendall's tau and the pointwise confidence intervals obtained via both the nonparametric multivariate estimator and the DCC procedure (blue and red lines respectively). The results show that the DCC based Kendall's tau is able to capture the general trend, but shows a slight oversmoothing and a degree of lag in the estimates. Nevertheless, the nonparametric multivariate time varying estimator performs better when the dependence structure presents a quadratic (Figure B.1(a)) or a more sinusoidal form (Figure B.1(b)), where it seeks to mimic the trend of the real value of Kendall's tau. Finally, for autoregressive dependence (Figure B.1(c)) it can be seen that the proposed estimator clearly outperforms the Kendall's tau relying on the DCC since it is not centered on the real values.

Table B.3 shows the squared bias, variance, and MSE of the above estimated Kendall's taus corresponding to the data generated by the copula and dependence structure indicated in the first column. The results in the three-column second block are obtained with the nonparametric



(a) $Clayton(\cdot, \cdot, \theta_{1t})$ (b) $Clayton(\cdot, \cdot, \theta_{2t})$

(c) $Clayton(\cdot, \cdot, \theta_{3t})$

Figure B.1: Real and estimated time varying Kendall's tau with 95% confidence intervals for data generated from a Clayton copula.

estimator and those in the three-column third block are calculated from the DCC. In this single copula setting it can also be observed that all variances are of the same order but that the nonparametric estimator shows less bias. Frank and Gumbel single parametric copulas were also considered and similar results were obtained. Thus, the nonparametric multivariate time varying Kendall's tau estimator is an interesting alternative estimator of dependence not only for mixtures of copulas but even for simple dependence structures.

Table B.3: Kendall's tau squared bias, variance, and MSE (1000 replications).

Copulas	$\hat{\tau}_t^{Np}$			$\hat{\tau}_t(\hat{\rho}_t^{DCC})$		
	$Bias^2$	Var.	MSE	$Bias^2$	Var.	MSE
$Clayton(\cdot, \cdot, \theta_{1t})$	$0.2 \cdot 10^{-4}$	$4.5 \cdot 10^{-3}$	$4.5 \cdot 10^{-3}$	$1.8 \cdot 10^{-3}$	$1.1 \cdot 10^{-3}$	$3.0 \cdot 10^{-3}$
$Clayton(\cdot, \cdot, \theta_{2t})$	$1.8 \cdot 10^{-4}$	$3.0 \cdot 10^{-3}$	$3.2 \cdot 10^{-3}$	$6.8 \cdot 10^{-3}$	$2.5 \cdot 10^{-3}$	$9.3 \cdot 10^{-3}$
$Clayton(\cdot, \cdot, \theta_{3t})$	$7.4 \cdot 10^{-4}$	$7.9 \cdot 10^{-4}$	$1.5 \cdot 10^{-3}$	$4.6 \cdot 10^{-3}$	$9.8 \cdot 10^{-4}$	$5.6 \cdot 10^{-3}$

NOTE: The table shows squared bias, variance, and MSE values for the dependence structures considered. Results for the nonparametric time varying Kendall's tau estimator are given in columns two to four, while the last three columns contain the results obtained from the DCC correlation transformed into Kendall's tau.

B.2.4 Additional results

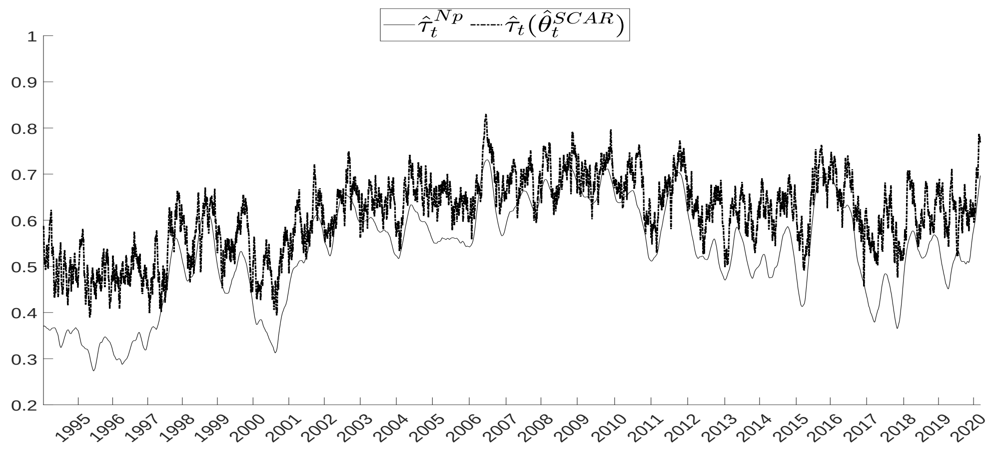


Figure B.2: Multivariate nonparametric time varying Kendall's tau and pairwise average Kendall's tau computed from the SCAR model with a Gumbel copula. The solid line is the multivariate nonparametric time varying Kendall's tau. The dashed line is the pairwise average Kendall's tau computed from the SCAR model.

Figure B.2 shows the multivariate Kendall's tau estimated through the proposed estimator (solid line) and the pairwise average of the SCAR model rank correlations with a Gumbel copula (dash-dotted line).

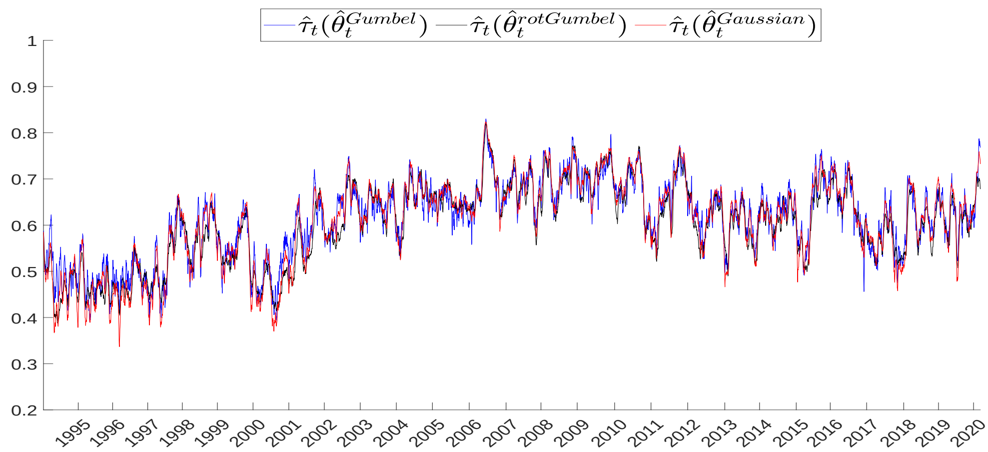


Figure B.3: Parametric multivariate dependence estimates obtained from the SCAR model. The blue line is the pairwise average Kendall's tau when a Gumbel copula is considered, while the black and red lines are for rotated-Gumbel and Gaussian copulas respectively.

Figure B.3 contains the pairwise average Kendall's tau obtained from the SCAR model rank correlation by considering Gumbel (black line), rotated-Gumbel (blue line) and Gaussian (red line) copulas respectively. The results show the same trend for the three selected copulas and similar rank correlation estimates from one copula to another.

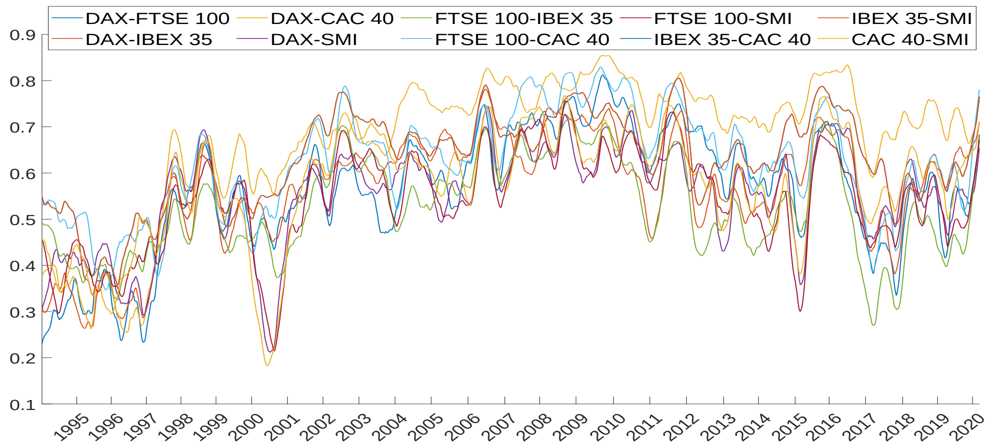


Figure B.4: Nonparametric pairwise time varying Kendall's tau.

Figure B.4 shows the nonparametric pairwise time varying Kendall's tau estimates between indexes over the period of study.

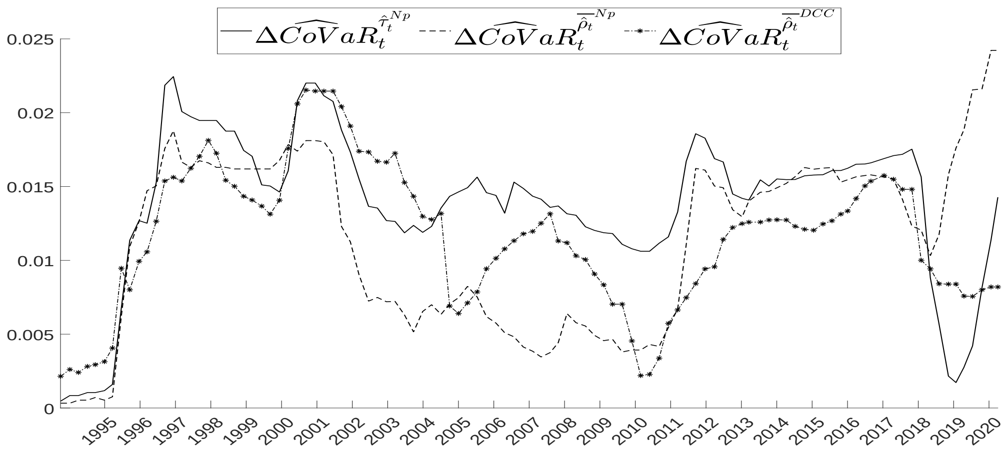


Figure B.5: Euro Stoxx $\Delta CoVaR$ estimates for several risk factors.

In Figure B.5 we show the time varying Euro Stoxx $\Delta CoVaR$ estimates for different risk factors. The results for the nonparametric multivariate time varying Kendall's tau are plotted with the solid line. The dashed and dash-dotted lines show the $\Delta CoVaR$ estimates when the pairwise average nonparametric correlation and the pairwise average DCC linear correlation, respectively, are considered as risk factors.

The patterns are consistent and show similar periods. Moreover, in light of the observed results, the multivariate rank correlation is more influential in the Euro Stoxx quantile than the measures based on linear correlations. In fact, the average $\Delta CoVaR$ for the multivariate nonparametric time varying Kendall's tau is 0.0136, while the pairwise average DCC and the pairwise average nonparametric correlation generate average figures for $\Delta CoVaR$ of 0.0116 and 0.0113 respectively. However, this prompts further research to test whether the differences are significant and whether these measures can predict future market downslides.

4

Sare teoria

1 Sarrera

Aurreko kapituluan indize nazionalen artean dauden antzekotasunak neurtu ditugu arrisku faktore bezala ulertu daitekeen aldagai anitzeko menpekotasun faktore baten bidez. Hala ere, indize nazionalak dagokien herrialdean garrantzitsuenak diren aktiboekin osatuta daude eta hori dela eta, indize nazional horien konposizioa aldatzen joaten da denboran zehar. Beste alde batetik, azken finantza-krisi globalak merkatuen aparteko ezaugarri bat utzi zuen agerian, arrisku sistemikoa izendatutakoarekin erlazionatuta dagoena. Izan ere, arrisku sistemikoaz hitz egiterakoan hiru kontzeptu gailentzen dira (Scott, 2012): konektibitatea, kutsadura eta menpekotasuna. Bereziki, aldagaien arteko elkar-konektibitateak pisu handia hartu du finantza-merkatuak aztertzerako orduan.

Azterketa ugari aurki daitezke finantza-indizeak eta haien arteko erlazioak aztertzen dituztenak. Hala ere, herrialde bakoitzeko aktiboen arteko gora-beheren inguruan azterketa gutxi daude. Ondorioz, finantza-merkatu ezberdinetako aktiboek sortzen dituzten egituren azterketa sakonago batean jarri nahi dugu arreta. Egitura horiek grafo edo sareen bidez adierazi daitezke matematikoki, non egiturak osatzen dituzten aktiboak grafoaren erpinak diren eta ertzek hoien artean ezartzen diren erlazioak adierazten dituzten. Ikusiko dugun bezala, adierazteko modu horrek egiturak aztertzeko aljebra matrizialak ematen dituen tresnez baliatzea baimentzen du.

Sareen artean sare konplexuak deitzen direnak daude eta hauek sistema konplexuen egitura adierazten dute. Sistema konplexuak elkar konektatuta dauden atalez osatzen dira. Hala ere, haien propietateak ez dira berehalakoak, hau da, sistemak bere osotasunean duen portaera atal bakoitzaren portaeraren menpekoa da baina baita atalen erlazioek eragiten dutenaren menpekoa ere. Hasieran esan bezala, sare konplexuek garrantzia handia hartu dute azken urteetan genetika, biokimika, kimika, ekologia eta soziologiaren hainbat esparrutan besteak beste (ikusi adibidez Ratkaj et al., 2010; Futschik et al., 2007; Grzybowski et al., 2009; Croes et al., 2006; Espinoza-Valdez et al., 2010; Montoya et al., 2006; Borgatti et al., 2009; Estrada, 2011).

Finantza-merkatuak sare konplexuen adibide oso argia diren arren, ekonomia eta finantzen munduan hauen erabilera askoz ere berriagoa da. Izan ere, gaur egun finantza-merkatuek izaera oso konplexua dute merkatuetako akzioen arteko loturak ere gero eta konplexuagoak direlako (ikusi Caldarelli et al., 2013; da Cruz eta Lind, 2012; Galbiati et al., 2013; Varela et al., 2015; Bougheas eta Kirman, 2015; Tang et al., 2018; Lai eta Hu, 2021, beste batzuen artean). Hurrengo kapituluan finantza-merkatuak sare konplexuen bitartez aztertuko ditugu, baina bete-betean finantza-sareak eratu baino lehen sare-teoriaren oinarriak aurkeztea beharrezkoa da.

2 Sareak eta sare motak

Lehenik eta behin sareen oinarrizko definizio eta ezagutza batzuk aurkeztuko ditugu.

Definizioa 21. Izan bitez $V = \{v_1, \cdots, v_p\}$ zehaztatu gabeko elementuen multzoa eta $V \otimes V$, V multzoko elementuen $[v_i, v_j]$ motako bikote ordenatuen multzoa, non $\otimes$ eragiketa bilineal bati lotutako biderketa tentsoriala den. Orduan, V multzoko erlazio bat $E \subseteq V \otimes V$ formako edozein azpimultzo da. Gainera, esango dugu

- E erlazioa simetrikoa dela baldin eta $[v_i, v_j] \in E$ eta $[v_i, v_j] \in E$ betetzen badira.
- E erlazioa erreflexiboa dela baldin eta $v \in V$ guztietarako $[v, v] \in E$ betetzen bada.
- E erlazioa anti-erreflexiboa dela baldin eta edozein $[v_i, v_j] \in E$ bikoterako $v_i \neq v_j$ betetzen bada.

V elementuen multzoa eta hauek lotzen dituen E erlazioen multzoarekin grafo bat sortu dezakegu. Grafo horretan V multzoko osagaiei *erpinak* esaten zaie eta E multzoko osagaiei *ertzak*. Hurrengo definizioak grafoen definizio formal bat ematen du.

Definizioa 22. Sare bat $G = (V, E)$ bikote bat bezala definitzen da, non V eta E erpinez eta ertzez osatutako bi multzo finitu diren. Sareak *grafoak* bezala ere ezagutzen dira.

Sare mota desberdinak aurki ditzakegu ertzek betetzen dituzten propietateen arabera. E erlazioa simetrikoa baldin bada, G-ri *zuzendugabeko sarea* esaten zaio. Aldiz, E erlazioa simetrikoa eta anti-erreflexiboa baldin bada eta errepikaturiko ertzik ez badu, orduan G *sare sinplea* dela esango dugu. Beraz, sare sinple batean erpin bikote bakoitza segmentu lerro baten bitartez konektatuta dago. Gainera, sareetako loturek norabidea ere izan dezakete, hain zuzen ere, erlazio ez simetrikoek adierazten dutena. Kasu horretan G *sare zuzendua* da, non zuzendutako lotura ertz batean hasi eta beste ertz batean bukatzen den.

Sare edo grafoek badituzte baita zenbait ezaugarri berezi. Sare bateko *begizta* bat (v, v) motako ertz bat bezala definitzen da. Beraz, sare sinple batek ez du inolako begiztarik. $V_z \subseteq V$ eta $E_z \subseteq V_z \otimes V_z \cap E$ azpimultzoez osatutako $G_z = (V_z, E_z)$ grafoa, G-ren *azpi-grafo* bat da. Bestalde, $\overline{G} = (V, \overline{E})$ sare sinplea G-ren *osagarria* da, G-ren erpin multzo bera duena eta $(u, v) \in E \iff (u, v) \notin \overline{E}$. Gainera, sareko erpin guztiak ertzen bidez lotuta badaude grafoa *osoa* dela esaten da, K_p adierazten dena. Horrela, p erpineko sare batean $p(p-1)/2$ lotura daude. Aldiz, loturarik gabeko sareari *sare hutsa* esaten zaio eta $\overline{K}_p$ bezala adierazten da K_p sare osoaren osagarria delako. Beste propietate garrantzitsu bat grafoen arteko isomorfismoari dagokiona da. G_1 eta G_2 grafoak isomorfoak direla esaten da haien ertzen artean banan-banako korrespondentzia $f : V_1 \to V_2$ baldin badago, non $u, v \in V_1$ guztietarako u eta v-ren arteko ertz kopurua $f(u)$ eta $f(v)$-ren arteko ertz kopuruarekin bat datorren.

Sare edo grafo teorian badaude garrantzitsuak diren beste hainbat kontzeptu ere, grafoaren elementuen auzokidetasunarekin lotura dutenak. u eta v bi erpin *auzokideak* direla esaten da baldin eta $e = \{u, v\}$ lotura baten bitartez konektatuta badaude. Hori horrela izanik, u eta v erpinak e loturarekin inzidenteak dira eta hortaz, e lotura u eta v erpinekin ere inzidentea da. Testuinguru honetan loturen auzokidetasuna ere definitu dezakegu: $e_1 = \{u, v\}$ eta $e_2 = \{v, w\}$ loturak *auzokideak* direla esango dugu bi loturak gutxienez ertz batekin inzidenteak baldin badira. Erpinen inzidentziari lotuta, erpin jakin batek sarean duen garrantzia erpin horrekin inzidenteak diren lotura kopuruaren bitartez aztertu daiteke, *erpinaren gradua* bezala ezagutzen dena. Beraz, erpin baten graduak erpin horri lotuta dauden ertz kopurua neurtuko du.

Ertzen arteko loturek pisuak ere izan ditzakete, zenbaki errealez adierazita daudenak (Barrat et al., 2004; Newman, 2004). Pisu horiek gauza ezberdinak adierazi ditzakete, hala nola, interakzioen indarra, informazio fluxua, une jakin bateko kontaktu kopurua, etab. Hala ere, aurretik emandako sareen definizioak ez ditu lotura anizkoitzak ezta autobegiztak izaterik kontuan hartzen. Hori dela eta, sareen definizio orokorrago bat beharrezkoa da:

Definizioa 23. Izan bedi V erpinen multzo finitu bat eta $E \subseteq V \otimes V = \{e_1, \ldots, e_m\}$ loturen multzo bat. Orduan, sare bat $G = (V, E, f)$ hirukote bat bezala definitzen da, f funtzioa E multzoko elementuak V multzoko elementu bikoteekin erlazionatzen dituen funtzioa izanik, non $v_i, v_j \in V$ bada, orduan $f : e_r \to [v_i, v_j]$ eta $f : e_s \to [v_j, v_i]$. Ertzen arteko loturek pisuak baldin badituzte, E loturen multzoaren ordez loturen pisuez sortutako $W = \{w_1, \ldots, w_m\}$ multzoa kontsideratzen da, non $w_\ell \in \mathbb{R}$, $\ell = 1, \ldots, m$. Kasu honetan, sarea $G = (V, W, f)$ hirukotearen bidez definitzen da eta *sare haztatua* esaten zaio.

Batzuetan sareek *klike* izeneko azpi-grafo edo azpi-sare berezi batzuk dituzte. Azpi-grafo hauek grafo osoak dira, hau da, erpin bakoitza beste erpin guztiekin konektatuta dago. Klike baten tamaina azpi-sare horretan dagoen erpin kopuruak ematen du eta klike-zenbakia sareko klike handienaren tamaina da, *klike maximoa* bezala ezagutzen dena. Ertz kopuru nahikoa duten sareetarako tamaina jakin bateko klikeak existitzen direla ziurtatu daiteke. Hain zuzen ere, Turán (1941) matematikariaren teoremaren arabera p erpineko sare sinple batek m ertz baldin baditu, orduan k gutxieneko tamaina duen klike bat existitzen da baldin eta $m > \frac{p^2(k-2)}{2(k-1)}$ bada.

Sare-egituren propietate batzuk erpinen arteko auzokidetasun erlazioekin lotuta daude eta matrizeen bidez adierazi daitezke. Zentzu honetan, *auzokidetasun matrizea* bezala ezagutzen den matrizea definitzen da.

Definizioa 24. Demagun $G = (V, E)$ sare sinple bat dela, non $V = \{1, \ldots, p\}$ eta izan bedi $\boldsymbol{A}$, p tamainako matrize karratua. G grafoaren *auzokidetasun matrizea* $\boldsymbol{A} = (a_{ij})$ bezala definitzen da, non $1 \leq i, j \leq p$ guztientzat

$$a_{ij} = \begin{cases} 1, & (i,j) \in E \text{ bada}, \\ 0, & (i,j) \notin E \text{ bada}. \end{cases}.$$

Beraz, i eta j edozein bi aldagai adierazten dituzten erpinak nolabait erlazionatuta baldin badaude auzokidetasun matrizean dagokion a_{ij} elementuak 1 balioa hartuko du eta, aldagai horiek inolako erlaziorik ez baldin badute aldiz, 0 balioa. Gainera, sareak auto-begiztarik ez baldin badu $\boldsymbol{A}$ auzokidetasun matrizeko diagonaleko elementuak zero izango dira. Beraz, sare batean $\boldsymbol{A}$ matrizeko m-garren errenkada edo zutabeak k_m osagai ez nulu ditu. k_m zenbaki hori aurretik definitutako erpinaren gradua da eta esan bezala, m-garren erpinak dituen auzokide kopurua erakusten du.

Edozein matrize karratu auzokidetasun matrize bezala interpretatu daiteke. $\boldsymbol{A}$ matrizeko edozein a_{ij} elementu binario bat ezik ertz baten pisu bezala ulertzen baldin badugu, $\boldsymbol{A}$ matrizeak eragindako sarea jadanik aurretik definitutako sare haztatua izango da. Auzokidetasun matrize honek beraz, mota desberdinetako erlazioak adieraz ditzake. Sareek erpinen arteko konekzioak erakusten dituzte baina inplizituki sareetan dauden konekzio ez zuzenak ere hain garrantzitsuak edo are garrantzitsuagoak izan daitezke. Nahiz eta erpinak zuzenki ertz baten bitartez lotuta egon ez, ibilbidean zehar erpinen bat edo gehiago partekatzen badute, pentsa genezake informazioa erpin batetik bestera pasatzen dela. Hori dela eta, ibilbide mota ezberdinak aurki ditzakegu sare batean.

Definizioa 25. Sare bateko *ibilaldi* edo *paseo* bat $(u_1, v_1), \ldots, (u_r, v_r)$ erpin eta ertzen segida bat da, non $v_i = u_{i+1}$, $i = 1, \ldots r-1$. Ibilaldia *itxia* izango da $v_r = u_1$ bada eta hauek *zirkuitu* edo *ziklo* bezala ezagutzen dira. Kasu berezi bat 3 luzerako zirkuitua da, *triangelua* esaten zaiona. Ibilaldiaren luzera bide horretako ertz kopuruak zehazten du. Gainera, ibilaldi batean ertz guztiak ez dute ezberdinak izan behar, izan ere erpin eta ertz guztiak ezberdinak dituen ibilaldiari *ibilbidea* esaten zaio.

Bi erpinen arteko ibilaldirik laburrena $\boldsymbol{A}$ auzokidetasun matrizearen ondoz ondoko berreketen bitartez aurki daiteke. Zehazki informazio hau q balioak ematen du, zeinarentzat $\boldsymbol{A}^q$ matrizeko (i, j) elementua lehen aldiz elementu ez nulua den. Bestalde, grafo zuzenduek *iturri-sareak* izaten dituzte, zuzendutako sare batean ertz guztiei norabidea kentzean lortzen den sarea da.

Konektibitatea

Orain arte konekzioez aritu garenean edozein erpin bikote lotzen dituen ertzez aritu izan gara. Hala ere, konektibitate kontzeptua sareetara orokortu daiteke. Sare bat *konektatuta* dagoela esaten da edozein bi erpin lotzen dituen ibilbide bat baldin badago. Bestela, sarea *deskonektatuta* dagoela esaten da. Jatorrizko sareari konektatuta dauden azpigrafo guztiak sarearen osagai konektatuak dira. Konektibitatearen barruan erpin eta ertz konektibitateak bereizi daitezke. Konektatutako sare baten *erpin konektibitatea* (*ertz konektibitatea*) sarea deskonektatua izatea eragiten duen erpin (ertz) kopuru txikiena da. Zuzendugabeko sare batean konektibitatea identifikatzea erraza da. Izan ere, sarea konektatuta ez baldin badago haren zati bat sareko gainerako osagaietatik banatuta egongo da. Bestalde, sare zuzenduetan bereizketa edo zatiketa hau konplexuagoa izan ohi da eta horregatik, azpi-sareak aztertzen dira konektibitatea errazago identifikatzeko.

Sare zuzenduetarako ere erpinen auzokidetasuna definitu daiteke. Kasu honetan, u erpina v erpinaren *auzokidea* dela esaten da u-tik v-ra zuzendutako lotura bat baldin badago, $e = \{u, v\}$. Kasu honetan, erpinen gradua bi osagaitan banatzen da: erpin baten *barne-graduak* erpin horretara inzidenteak diren lotura kopurua neurtzen du eta *kanpo-graduak* erpinetik irteten diren lotura kopurua.

Auzokidetasun matrizeak konektibitatea aztertzeko balio dute. Konkretuki, auzokidetasun matrizearen permutazio matrizeak erabiltzen dira. $\boldsymbol{P} \in \mathbb{R}^{p \times p}$ matrize bat *permutazio matrizea* dela esaten da baldin eta $\boldsymbol{P}$ matrizea $\boldsymbol{A}$ matrize batekin biderkatzean eragina $\boldsymbol{A}$ matrizearen errenkada edota zutabeen permutazioa soilik bada. Permutazio matrize bat lortzeko nahikoa da identitate matrizearen errenkadak permutatzearekin.

Sare bat aztertzerakoan erpin nagusiak zeintzuk diren jakiteak egitura topologikoari buruzko informazio oso baliagarria ematen du eta zentralitate neurriek zehaztapen hori egiten laguntzen duten tresnak dira. Erpin baten garrantzia neurtzeko ezaugarri ezberdinak erabili daitezke, hain zuzen ere, zuzenean beste erpin batekin komunikatzeko ahalmena, beste erpin batzuekiko hurbiltasuna edo sareko hainbat atal ezberdin lotzen dituen ezinbesteko bitartekaria izatea. Hauetako ezaugarri bakoitzak zentralizazio neurri ezberdin bat sortzen du.

Erpinen komunikazio ahalmena *zentralitate graduaren* bitartez neurtzen da eta aurretik definitutako erpinen graduari egiten dio erreferentzia. Izan ere, erpin bat beste bat baino zentralagoa izango da baldin eta lehenengo erpinaren gradua bigarrenarena baino handiagoa bada eta ondorioz, zentralitate gradu altueneko erpinak beste erpinak baino eragin handiagoa izango du sarean. Horregatik, erpinen gradua zentralizazio neurri bezala erabili daiteke.

Aipatu dugun bezala, p tamainako sare sinple bateko edozein i erpinaren gradua auzokidetasun matrizea erabiliz kalkulatu daiteke dagokion errenkada edo zutabeari erreparatuz. Beraz, i erpinaren gradua $k_i = \sum_{j=1}^{p} a_{ij} = (\mathbf{1}'\boldsymbol{A})_i = (\boldsymbol{A}\mathbf{1})_i$ da, non $\mathbf{1}$ p tamaina duen 1 balioez osatutako bektorea den. Beraz, zentralizazio graduaren arabera i erpina j erpina baino zentralagoa dela esango dugu $k_i > k_j$ bada. Zuzendugabeko sare batean parte hartzen duten erpin guztien gradua sareko gradu-bektorea bezala ezagutzen den $\mathbf{k} = (k_1, \ldots, k_p) = (\mathbf{1}'\boldsymbol{A})' = \boldsymbol{A}\mathbf{1}$ bektorearen bidez adierazten da. Bereahala ondorioztatzen da $\overline{k} = p^{-1} \sum_{i=1}^{p} k_i$ kopuruak sarearen batazbesteko erpin-gradua ematen duela. Honekin erlazionatuta Eulerren *bostekoaren* lema aurkitu dezakegu. Lema honek erpin-gradu guztien batura lotura kopuruaren bikoitza dela frogatzen du, $\sum_{i=1}^{p} k_i = 2m$, non m sareko ertz kopurua den. Ondorioz, batazbesteko erpin-gradua $\overline{k} = 2p^{-1}m$ izango da. Sare haztatuen kasuan erpinen gradua antzerako era batean kalkulatzen da auzokidetasun matrize haztatuaren bitartez.

Erpin baten graduak 1 luzeerako ibilaldi kopurua zenbatzen duenez, zentralitate graduak sarean gertatzen diren berehalako eraginak neurtzen ditu. Adibidez, demagun sarean erpin batetik bestera pasatzen den informazioa adierazten duen prozesu bat kontsideratzen dugula. Orduan, sarean zehar informazio hori jasotzeko aukera dagokion erpinaren graduaren menpekoa da. Hau da esaterako, sare bateko infekzio arriskuaren kasua, non erpin batzuk kutsatuta dauden eta erpin horiekin lotuta egoteak kutsadura hedatzea eragiten duen. Kasu honetan, gradu altua duten erpinek gradu baxua dutenek baino arrisku handiagoa izango dute modu esanguratsuan beste erpin batzuk kutsatu eta baita kutsatuak izateko ere.

Zentralitate graduak duen bereizgarri handienetako bat erpin baten hurbileko auzokideen eragina soilik neurtzea da. Askotan beharrezkoa da hurbileko auzokideen eraginaz gain urrunago daudenen efektua ere neurtzea. Tesi honetan mota honetako neurriak landuko ez ditugun arren, aipatu beharra dago literaturan badaudela zenbait proposamen, hala nola, Katz-en zentralitatea, bektore propioen zentralitatea, azpi-grafoen zentralitatea edo bibrazio zentralitatea bezalako neurriak (Estrada, 2012).

Zuzendutako sareetan auzokidetasun matrizeak ez dira simetrikoak izaten eta kasu horietan, $\boldsymbol{A1}$ eta $\boldsymbol{1A}$ ez dira beti berdinak izango. Hortaz, barne-gradua ($k_i^{barne} = \sum_{j=1}^{p} a_{ji} = (\boldsymbol{1'A})_i$) eta kanpo-gradua ($k_i^{kanpo} = \sum_{j=1}^{p} a_{ij} = (\boldsymbol{A1})_i$) zentralitate graduak bereizten dira, erpin konkretu batera doazen eta erpinetik irteten diren lotura kopurua neurtzen dutenak, hurrenez hurren. Kasu honetan, $\sum_{i=1}^{p} k_i^{barne} = \sum_{i=1}^{p} k_i^{kanpo} = m$.

Bigarren zentralitate neurria *hurbiltasun zentralitatea* da eta erpin batek gainontzeko erpinenganako duen hurbiltasuna zehazten du. Hurbiltasuna ibilbide laburreneko distantzian oinarrituta neurtzen da. Zuzendugabeko sare batean i erpinaren hurbiltasuna $HZ(i) = (p-1)/s(i)$ bezala definitzen da, non $s(i) = \sum_{j \in V(G)} d(i,j)$ distantzia-batura ibilbide laburreneko $d(i,j)$ distantziekin kalkulatzen den. Zentralitate graduarekin gertatzen den bezala, zuzendutako sareetan erpinek barne eta kanpo hurbiltasun zentralitateak dituzte. Barne hurbiltasunak erpin batek informazioa jasotzen duen gainontzeko erpinetatik duen hurbiltasuna neurtzen du. Kanpo hurbiltasunak aldiz, erpin jakin batek informazioa bidaltzen dien beste erpinekiko duen hurbiltasuna.

Hirugarren neurria *bitartekaritza zentralitatea* da eta erpin batek beste edozein erpin bikoteren arteko komunikazio edo loturan duen garrantzia zehazten du. Hurbiltasun zentralitateak bezala, bitartekaritzak informazioa erpinak lotzen dituen ibilbide laburrenaren bitartez bidaltzen dela hartzen du bere gain. Zuzendugabeko sare batean i erpinaren bitartekaritza $BZ(i) = \sum_{i \neq k \neq j} \sum_{k \neq j} \rho(j,i,k)/\rho(j,k)$ bezala definitzen da, non $\rho(j,k)$, j eta k erpinak lotzen dituen ibilbide laburrenen kopurua eta $\rho(j,i,k)$ laburrenak diren ibilbide hoietatik i erpinetik igarotzen diren ibilbide kopurua diren. Zuzendutako sareetan $\rho(j,k)$ eta $\rho(j,i,k)$ terminoak j erpinetik k erpinera doazen ibilbide zuzenduetan oinarritzen dira.

Hasieran esan bezala, hurrengo kapituluan finantza-sareak sortuko ditugu finantza-merkatu ezberdinen egitura aztertzeko. Finantza-merkatu bateko akzio bakoitza merkatu horri dagokion sarearen erpin bat izango da eta akzio ezberdinen arteko erlazioak ertzen bitartez adieraziko ditugu. Beraz, akzioek eta haien arteko erlazioek merkatu ezberdinetako finantza-sareak sortuko dituzte. Gainera, tesi honetan, zuzendutako sareak alde batera utziko ditugu eta zuzendugabeko sareak erabiliko ditugu burtsa merkatuetako egitura topologikoak aztertzeko.

3 Aljebra sare teorian

Sareak matrizeen bitartez adierazi daitezkeenez, aljebra matriziala erabili dezakegu beraien propietateak aztertzeko. Testuinguru honetan, erabilgarriak izango diren auzokidetasun matrizeen balio eta bektore propioei buruzko emaitza batzuk aurkeztuko ditugu.

Definizioa 26. Izan bitez p dimentsioko $\boldsymbol{A}$ matrize bat, λ balio eskalarra eta $\mathbf{x}$ bektorea non $\boldsymbol{A}\mathbf{x} = \lambda\mathbf{x}$ betetzen den. Aurreko ekuazioa betetzen duen edozein λ baliori *balio propioa* esaten zaio eta baldintza hori betetzen duen edozein $\mathbf{x}$ bektore ez nulu *bektore propio* bezala ezagutzen da.

Kontuan izan λ-ren balio guztietarako $\boldsymbol{A}\mathbf{0} = \lambda\mathbf{0}$ propietatea betetzen dela, nahiz eta zero bektorea bektore propio bat ez izan. $\boldsymbol{A}\mathbf{x} = \lambda\mathbf{x}$ erlazioa beste modu batean idatziz, honako hau lortu dezakegu: $\boldsymbol{A}\mathbf{x} = \lambda\mathbf{x} \Leftrightarrow (\boldsymbol{A} - \lambda\boldsymbol{I})\mathbf{x} = 0$, non $\boldsymbol{I}$ identitate matrizea den. Adierazpen honek $\boldsymbol{A} - \lambda\boldsymbol{I}$ matrizea singularra dela adierazten du. Hori dela eta, λ balio propioak $det(\boldsymbol{A} - \lambda\boldsymbol{I}) = 0$ ekuazioaren erroak direla ondorioztatzen da. $det(\boldsymbol{A} - \lambda\boldsymbol{I})$ determinantea p graduko polinomio bat da, $det(\boldsymbol{A} - \lambda\boldsymbol{I}) = \alpha_0 + \alpha_1\lambda + \ldots + \alpha_p\lambda^p$, eta $\chi_A(\lambda) =$ adierazten den *polinomio karakteristiko* izenaz ezagutzen da. Bestalde, $\chi_A(\lambda) = 0$ ekuazioari *ekuazio karakteristikoa* esaten zaio. Beraz, λ, $\boldsymbol{A}$ matrizearen balio propioa da baldin eta soilik baldin $\chi_A(\lambda)$ polinomioaren erroa bada. $\chi_A(\lambda)$ polinomioaren $\chi_A(0)$ koefizientea $\boldsymbol{A}$-ren determinantea da eta λ^{p-1}-ren koefizienteak $\alpha_{p-1} = (-1)^{p-1}tr(\boldsymbol{A})$ betetzen du. Kasu honetan, zuzendugabeko sareak kontsideratzen ditugunez, auzokidetasun matrizeak simetrikoak dira eta ondorioz bere balio propio guztiak errealak izango dira.

Behin $\boldsymbol{A}$ matrizearen balio propioak kalkulatu ditugula, polinomio karakteristikoa $\chi_A(\lambda) = (\lambda_1 - \lambda)^{q_1}(\lambda_2 - \lambda)^{q_2} \cdots (\lambda_k - \lambda)^{q_k}$ eran faktorizatu daiteke, non λ_i balioak bakarrak diren. Polinomio karakteristikoaren q_i berretzaileak λ_i balio propioaren errepikakortasuna adierazten du eta balio propioaren *anizkoiztasun aljebraikoa* bezala definitzen da. Ohartu anizkoiztasun aljebraikoen baturak auzokidetasun matrizearen dimentsioa ematen duela, $\sum_i q_i = p$. Bestalde, balio propio bati lotutako bektore propioen kopuruari balio propioaren *anizkoiztasun geometrikoa* esaten zaio. Gainera, polinomio karakteristikoaren koefizienteak $\boldsymbol{A}$ matrizearen minore printzipalekin lotuta daude: $(-1)^k\alpha_{p-k}$, $\boldsymbol{A}$-ren $k \times k$ minore printzipal guztien batura da. Kasu berezia da matrizearen traza, $tr(\boldsymbol{A})$, 1×1 minore printzipalen batura dena. Kontuan izan, traza matrizearen balio propioen bitartez ere kalkulatu dezakegula, $tr(\boldsymbol{A}) = \sum_{i=1}^{k} \lambda_i$. Azken hau kontuan hartuta, $-tr(\boldsymbol{A})$ polinomio karakteristikoaren λ^{p-1} koefizientea izango da. Beraz, polinomio karakteristikoa $\chi_A(\lambda) = \lambda^p - tr(\boldsymbol{A})\lambda^{p-1} + \ldots + (-1)^p det(\boldsymbol{A})$ ere idatzi daiteke.

Matrize bati lotutako polinomio karakteristikoaren balio propioek sortzen duten multzoari izen berezi bat ematen zaio.

Definizioa 27. A matrize baten *espektrua* bere balio propioen multzoa da, hots $\varphi(\boldsymbol{A}) = \{\lambda : det(\boldsymbol{A} - \lambda \boldsymbol{I}) = 0\}$. $\boldsymbol{A}$-ren balio propio handienaren moduluari *erradio espektrala* esaten zaio, hau da, $\rho(\boldsymbol{A}) = max_{\lambda \in \varphi(\boldsymbol{A})} |\lambda|$.

Sareak aztertzeko auzokidetasun matrizea zuzenean erabili baharrean matrizearen informazio espektrala erabiltzen dugu. Funtzio matrizialen portaera gainera, balore eta bektore propio hauen interakzioekin zuzenki lotuta dago eta antzekotasun transformazioak interakzio hauek ulertzeko garrantzitsuak dira.

Definizioa 28. $\boldsymbol{X}$ matrize ez-singularra baldin bada, $S_X : \boldsymbol{A} \rightarrow \boldsymbol{X}^{-1}\boldsymbol{A}\boldsymbol{X}$ aplikazioa *antzekotasun eraldaketa* bat da. Ondorioz, $\boldsymbol{A}$ eta $\boldsymbol{X}^{-1}\boldsymbol{A}\boldsymbol{X}$ *antzekoak* direla esaten da $\boldsymbol{X}$ matrize alderanzgarri bat existitzen bada.

Antzekotasun eraldaketek betetzen duten propietate garrantzitsu bat hurrengo hau da: matrize baten antzekotasun eraldaketa baten eta jatorrizko matrizearen determinanteak berdinak dira, hots, $det(\boldsymbol{X}^{-1}\boldsymbol{A}\boldsymbol{X}) = det(\boldsymbol{A})$. Era berean, edozein $\lambda \in \mathbb{R}$ baliorako, $det(\boldsymbol{X}^{-1}\boldsymbol{A}\boldsymbol{X} - \lambda \boldsymbol{I}) = det(\boldsymbol{A} - \lambda \boldsymbol{I})$. Ondorioz, antzekoak diren bi matrizek polinomio karakteristiko bera dute eta beraz, balio eta bektore propio berdinak izango dituzte. Hala ere, ohartu kontrakoak ez duela zertan egia izan behar, polinomio karakteristiko bera duten bi matrizek ez dute zertan antzekoak izan. Testuinguru honetan, $\boldsymbol{A}\mathbf{x} = \lambda \mathbf{x}$ eta $\boldsymbol{B} = \boldsymbol{C}^{-1}\boldsymbol{A}\boldsymbol{C}$ badira, orduan $\boldsymbol{B}\boldsymbol{C}^{-1}\mathbf{x}$. Beraz, $\boldsymbol{A}$ matrizearen bektore eta balio propioa $\mathbf{x}$ eta λ dira baldin eta $\boldsymbol{C}^{-1}\mathbf{x}$ λ balio propio bezala duen $\boldsymbol{B}$ matrizearen bektore propioa bada. Ondorioz, banan banako korrespondentzia dago $\boldsymbol{A}$ eta $\boldsymbol{B}$ matrizeen balio eta bektore propioen artean. Emaitza honek erakusten du matrize baten egitura propioa berehalakoa izan daitekela antzekotasun eraldaketa aproposa egiten bada eta orduan hurrengoa izango litzateke matrize bat antzekotasun eraldaketen bidez zenbateraino erreduzitu daitekeen jakitea. Erantzuna teorema espektralak ematen du. Horretarako matrizeen antzekotasun eraldaketak ematen dituen tresna definitzen dugu lehenik eta behin.

Orokorrean, p-dimentsioko edozein $\boldsymbol{A}$ matrizek antzekotasun eraldaketa bat du baldin eta $\boldsymbol{X}^{-1}\boldsymbol{A}\boldsymbol{X} = \mathcal{J}$ betetzen bada, non $\mathcal{J} = diag(\mathcal{J}_1, ..., \mathcal{J}_k)$ blokeka diagonala den matrize bat den eta bloke bakoitza $\boldsymbol{A}$-ren λ_i balio propioa duen

$$\mathcal{J}_i = \begin{pmatrix} \lambda_i & 1 & & \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{pmatrix}, \quad i = 1, ..., k,$$

matrizea den. Aurreko $\mathcal{J}$ matrizeari $\boldsymbol{A}$-ren *Jordanen forma kanonikoa* esaten zaio eta $\boldsymbol{A} = \boldsymbol{X}\mathcal{J}\boldsymbol{X}^{-1}$ adierazpenari Jordanen deskonposizioa. Jordanen forma kanonikoak matrizeen antzekotasuna ere zehazten du. Izan ere, edozein matrizeren deskonposizioan $\mathcal{J}$ matrizearen blokeak bakarrak direnez, bi matrize antzekoak direla esaten da baldin eta soilik baldin dagozkien $\mathcal{J}$ matrizeek bloke diagonal berdinak badituzte. $\mathcal{J}$ matrizearen bloke guztiak 1-dimentsioko matrizeak baldin badira matrizeari matrize sinplea esaten zaio. Guk erabiliko ditugunak matrize sinpleak izango dira. Matrize ez sinpleak zuzendutako sareetan agertzen dira. Ondorioz, Jordan-en deskonposaketak teorema espektrala zehazten du.

Teorema 16. ***(Teorema espektrala)*** *$\boldsymbol{A} \in \mathbb{R}^{p \times p}$ matrize simetrikoa baldin bada, orduan $\boldsymbol{Q}$ matrize ortogonal bat eta $\boldsymbol{D}$ matrize diagonal errela existitzen dira $\boldsymbol{A} = \boldsymbol{Q}\boldsymbol{D}\boldsymbol{Q}'$ betetzen delarik, non $\boldsymbol{D} = diag(\lambda_1, ..., \lambda_p)$, $\boldsymbol{A}$-ren balio propioez osatutako matrize diagonala den eta $\boldsymbol{Q}$-ren zutabeak haren bektore propioak diren.*

Auzokidetasun matrizeen espektrua

Fenomeno fisiko asko haien propietate espektralak aztertuz uler daitezke. Zuzendugabeko sareekin soilik lan egingo dugunez, kontsidera ditzagun sare sinpleak eta erabili dezagun auzokidetasun matrizea sarearen espektrua aztertzeko.

Grafo baten inbarianteak grafoaren isomorfismoekiko inbarianteak diren propietateak dira. Izan ere, auzokidetasun matrize baten espektrua grafoaren inbariantearen adibide bat da. Propietate hauek grafoaren egituraren menpekoak dira soilik eta ez grafoaren irudikapen konkretu baten menpekoak, hala nola, erpin eta ertz kopurua, klikearen zenbakia edota ertz zein erpin konektibitateak. Beraz, sarea izendatzeko modua edozein dela ere, lortuko ditugun balio propioak berdinak izango dira. Kontuan izan berrizendapena egitea $\boldsymbol{P}$ permutazio matrize baten bitartez antzekotasun eraldaketa bat egitea dela, $det(\boldsymbol{PAP'} - \lambda\boldsymbol{I}) = det(\boldsymbol{P}(\boldsymbol{A} - \lambda\boldsymbol{I})\boldsymbol{P}') = det(\boldsymbol{A} - \lambda\boldsymbol{I})det(\boldsymbol{P})det(\boldsymbol{P}') = det(\boldsymbol{A} - \lambda\boldsymbol{I})$, eta ondorioz permutazioak ez dituela polinomio karakteristikoaren erroak aldatzen.

Sare baten espektrua ibilaldi kopurua zenbatzeko erabili dezakegu. Bereziki, $\boldsymbol{A}^q$ potentziaren elementuek erpinen artean dauden q luzerako ibilaldiak adierazten dituztenez, ibilaldi itxien kopurua matrize horren diagonalean aurki daiteke. Hortaz, q luzerako ibilaldi itxien batura

$tr(\boldsymbol{A}^q) = \sum_{i=1}^{p} \lambda_i^q$ da, non $\boldsymbol{A}$ matrizearen balio propioak λ_i diren.

4 Sare egituren azterketa

Sare konplexuen funtzionamenduak erlazio estua du fenomeno fisikoekin. Fisikaren adarretako bat termodinamika da eta, estatistikarekin batera, sareen portaera ikertzeko tresna oso erabilgarria da. Termodinamikak sistema bat beti unibertso fisiko baten parte dela suposatzen du. Ohartu kasu honetan sareak sistemaren papera jokatzen duela. Beste alde batetik, eta fisika klasikoaren arabera, sistemaren egoera une jakin batean sistema hori osatzen duten partikulen posizio eta momentuetan oinarritzen den deskribapenaren araberakoa da. Egoera honetan bi osagai bereiztu daitezke; mikroegoera, sistemako partikulen parametroak zehazten dituen egoera, eta makroegoera, sistemaren egoera bat zeinetan partikulen banaketa energia mailetan zehar zehaztuta dagoen.

Demagun erpinak bi multzotan zoriz banatuta dituen sare bat dugula. Suposa dezagun baita sarea isolatuta dagoela, ez duela loturarik beste sistema batekin eta sarea existitzen den egoera konstantea dela denboran zehar. Sistema bat mikroegoera batean aurkitzeko probabilitatea denboran zehar konstantea denean sistema orekan dagoela esaten dugu. Beraz, sareen egitura topologikoaren azterketa prozesuan beraien oreka maila zein den ikertu nahi dugu. Hori dela eta, lehenik eta behin oreka kontzeptua matematikoki nola adierazi daitekeen ulertu behar da. Testuinguru honetan oreka erpinen antolaketarekin egongo da erlazionatuta.

Izan ere, sareetan erpinak nola antolatuta dauden ulertzeak asko lagundu dezake haien egitura topologikoa ulertzeko. Errealitatean askotan sareetako erpinak multzokatuta egoten dira beste erpin batzuekin dituzten loturen arabera. Sare topologiaren azterketan lotura hoiek ematen dituzten partiketak bilatzen ditugu. Formalki, helburua sare bat erpinez osatutako q azpimultzo disjuntutan zatitzea da, non V_i azpimultzo horiek hurrengo propietateak betetzen dituzten:

i) $\bigcup_{i=1}^{q} V_i = V$ eta $V_i \bigcap V_j = \emptyset$, $i \neq j$.

ii) Azpimultzoen arteko lotura kopurua (*ebakidura tamaina*) minimoa da.

iii) $|V_i| \sim p/q$, non $1 \leq i \leq q$ eta $|\cdot|$ multzoaren kardinalitatea den.

Sare haztatuen kasuan *ebakidura tamaina* amankomunean dauden ertzen pisuen batura bezala definitzen da. Beraz, *iii*) baldintza $W_i \sim W/q$ bezala zehazten da, non W_i, V_i-garren azpimultzoko erpinen pisuen batura eta W sare osoko erpinen pisuen batura diren, hurrenez hurren. Aurrekoetako *iii*) baldintza edo sare haztatuetarako baliokidea dena betetzen denean, sarearen partiketa *orekatua* dela esango dugu. Bestela, sarea *ez-orekatua* da. Sare konplexuen kasuan, orekaren kontzeptua edozein ziklotara hedatu daiteke. Izan ere, sare bat orekatua dela esaten da baldin eta bere ziklo guztiak orekatuak badira. Zikloen oreka maila neurtzeko luzera guztietako ibilaldiak zenbatzen dira, auzokidetasun matrizearen balio propioen bidez egin daitekena, aurreko atal batean azaldu den bezala.

Praktikan sareen oreka maila haien egonkortasunarekin erlazionatzen da. Horrela, orekatua den sare batek sistema egonkorra dela adierazten du. Zentzu honetan, egonkortasuna sistemek izan dezaketen propietate onuragarri bat da. Ohartu sistema egonkorra izateak ez duela sareko ertz guztiak positiboak izatea inplikatzen. Ez-egonkortasunak aldiz, sistemaren hauskortasuna erakusten du. Finantza-merkatuak aztertzerakoan merkatuen egonkortasuna ezaugarri bereziki garrantzitsua da bai inbertsio zein erregulatzaileen ikuspuntutik. Hori dela eta, finantza-merkatuek orekan dauden edo ez ikustea baino, orekatik zenbatean aldentzen diren neurtzea interesatzen zaigu.

Sistemaren orekarekin lotuta, termodinamika eta estatistika mekanikoaren oinarrizko kontzeptu bat entropiaren kontzeptua da, sistemaren tenperatura unitate bat gehitzean aprobetxagarria ez den sarearen barne-energiaren gehikuntza neurtzen duen magnitude fisiko bat. Beraz, entropia sistemaren zorizko antolaketaren neurri bat bezala interpretatu daitekenez, sistemaren anabasaren parametro bat da. Zentzu honetan, sistema batek guztiz ausazkoa den antolaketa bat baldin badu entropia altua duela esaten da eta bere balio maximoa sistema orekara hurbiltzen denean hartuko du. Entropia eta tenperaturarekin erlazionatutako propietateak izango dira hain zuzen ere sareen topologiari buruzko informazioa emango digutenak.

Sare konplexuak etengabe daude kanpo tentsioen eraginpean, non kanpo tentsio hauek sarearen topologiaren independenteak diren. Horrela, nahiz eta faktore horiek sare egituraren independenteak izan, eginkizun erabakigarria izan dezakete sare egituren bilakaeran. Adibidez, erlazio sozialen sareak gizartearen ingurugiro sozial, ekonomiko eta politikoan daude murgilduta eta giro hauetako aldaketek zuzenki alda ditzakete erlazio sozialen loturak. Antzerako egoeran daude finantza-sareak, ekonomikoki edo politikoki ematen diren aldaketek sarearen egituran eragina izan dezaketelarik. Termodinamikaren arloan kanpo eragile horiek sistema tenperatura jakin bat duen bainuontzi batean murgilduta egotearekin parekatzen da. Izan ere, sistemen anabasak erlazio handia du tenperatura honekin. Beraz, demagun orain sareak bainuontzi termal batean sartuta daudela. Bainuontzi edo tanga termal bat sareak baino dimentsio askoz handiagoak dituen sistema bat da eta sarea eta tangaren arteko interakzio energia mezprezagarria dela suposatzen da. Hau da, E_{tanga} tangaren energia eta E_{sarea} sarearen energia baldin badira, hurrenez hurren, orduan E_{tanga} eta energia osoa ($E_{osoa} = E_{tanga} + E_{sarea}$) E_{sarea} baino askoz handiagoak izango dira.

Aipatutako kanpo estresore horiek handiak direnean, krisi ekonomiko batean esaterako, akzioen arteko erlazio batzuk indartu eta beste batzuk ahuldu egin daitezke fusio edo aliantza estrategikoak direla eta. Kanpo faktore horiek β parametro baten bidez adierazten dira eta sistemaren alderantzizko tenperaturarekin du erlazioa. Hori dela eta, sarea murgilduta dagoen ingurunean edozein aldaketa gertatzen bada, β parametroak edozein sarek jasaten duen kanpoko tentsioaren eginkizuna betetzen du eta sareko ertz guztiak β-gatik haztatuta daude aldi berean. Sarea haztatua bada ere, ertz bakoitza β parametroarekin biderkatuko da. Izan ere, kanpo eraginak kontuan hartzen dituen edozein sare jadanik sare haztatu bat da, non ertz guztiek pisu berdina duten.

Ondorioz, tenperaturak infiniturantz jotzen duenean, β-k zerorantz jotzen du eta ertz guztiek zero pisua izatera joko dute, sare deskonektatua bilakatuz (sare hutsa). Ohartu egoera honetan sareak tentsio egoera oso larria jasaten duela eta ondorioz, sarea disfuntzionala bihurtzen dela. $\beta = 1$ denean aldiz, sare sinple bat dugu, non konektatutako erpin bikote bakoitzak lotura bakarra duen. Tenperaturak zerora jotzen duenean, $\beta \to \infty$, ertz guztiak infinitu dira eta ertzek balio oso altuak hartzen dituzte erpinen arteko konekzioen sendotzea adierazten dutenak.

Hurrengo kapituluan sare egiturak haien oreka mailaren bitartez aztertuko ditugu. Horretarako, Estrada eta Benzi (2014)-k proposatutako *sarearen oreka gradua* edo *sarearen balantzea* bezala definitzen den koefizientea erabiliko dugu, benetako sarea guztiz orekatua izatetik zenbatean urruntzen den neurtzen duena. Esan bezala, sareak kanpo faktoreen eragina ere jasaten dute eta neurri honek faktore exogenoen eragina ere kontutan hartzen du jadanik definitutako β parametroaren bitartez.

Sareetan sakonduz

Errealitatean sare bat dugunean, garrantzitsua da sare hori nola sortu den eta egitura topologikoa sortzen duten erpin eta ertz taldearen eboluzioaren atzean zer dagoen ikustea. Horretarako, sareak eraikitzeko eredu ezberdinak kontsideratu eta konparatu ditzakegu. Hasiera batean pentsatu genezake erpinak probabilitate jakin batekin lotzen dituen eredu batean. Hau da, demagun p erpin ditugula eta dauden $p(p-1)/2$ lotura posible bakoitzerako i eta j edozein bi erpin probabilitate batekin lotzen ditugula. Beraz, sarearen zenbait propietate finkatuta daudela suposatzen baldin badugu, parametro horiek sarearen egituran duten eragina aztertzeko aukera ematen duten eredu ezberdinak sortu ditzakegu.

Horietako bat, eta praktikan gehien erabilienetarikoa dena, zorizko sareen *Erdös-Rényi* eredua da (Erdős eta Rényi, 1960). Eredu honetan erpin multzo bat hautatzen da eta multzo horretako erpin bikote bakoitzaren arteko lotura bikote horrentzat ezarritako probabilitate baten araberakoa izango da. Hau da, izan bitez p erpin eta finkatu dezagun aldez aurretik $prob_{ER} \geq 0$ parametro bat edozein (i, j) erpin lotzeko probabilitatearen muga adierazten duena. Gainera, erpin bikote bakoitzerako r^{ij} zorizko zenbaki bat kontsideratzen dugu $[0, 1]$ tartean uniformeki hautatua. Orduan, finkatutako $prob_{ER}$ mugarako $prob_{ER} > r^{ij}$ betetzen baldin bada, ertz bat ezarriko dugu i eta j erpinak lotzen dituena. $prob_{ER} = 0$ hautatu badugu, orduan sarea guztiz deskonektatuta egongo da, hau da, ez dago sareko erpin bikoteren bat lotzen duen ertzik. Aldiz, $prob_{ER} = 1$ bada, erpin guztiak haien artean konektatuta egotea onartzen dugu, grafo oso bat sortzen dutelarik. Erdös-Rényi ereduarekin lortutako p erpin eta m ertzeko sarea $G_{ER}(p, prob_{ER})$ edo $G_{ER}(p, m)$ adierazten da. Eredu honetan erpin bakoitzerako batazbesteko ertz kopurua $\overline{m} = p(p-1)prob_{ER}/2$ da eta hori dela eta, batazbesteko erpin gradua $\overline{k} = (p-1)prob_{ER}$ da.

Erdös-Rényi ereduaz gain, literaturan badaude beste eredu batzuk ere, hots, Barabási-Albert (Barabási eta Albert, 1999) edo Watts-Strogatz (Watts eta Strogatz, 1998) bezalako ereduak (eredu hauei buruzko xehetasun gehiagorako ikusi Estrada eta Knight (2015)).

5 Ondorioak

Finantza aldagaiek sortzen dituzten egiturak aztertzeko sare konplexuen teoria erabili daiteke. Kapitulu honetan finantza sareak eraiki ahal izateko sare eta sare konplexuei buruzko oinarrizko tresnak azaldu ditugu. Horien artean, auzokidetasuna eta auzokidetasun matrizea oso garrantzitsuak dira. Izan ere, auzokidetasun matrizeek sareen egitura topologikoak matematikoki adierazteko balio dute, non matrizeko elementu bakoitzaren posizioak sareko bi ertz adierazten dituen eta elementuaren balioak bi ertzen erlazioa. Hurrengo kapituluko azterketa zuzendugabeko sareetara mugatuko da eta ondorioz, erabiliko ditugun auzokidetasun matrizeak simetrikoak izango dira.

Aljebra matrizialak auzokidetasun matrizeen bidez adierazitako sareen propietateak aztertzen laguntzen du. Esaterako, propietateetako bat sarearen oreka da, aztertzen ari garen sistemaren egonkortasunari buruzko informazioa ematen duena. Propietate hau bereziki interesgarria izan daiteke, batez ere finantzen arloan. Gainera, kontuan izan behar dugu sareak kanpo eragileen menpekoak direla. Hain zuzen ere, edozein sistema testuinguru biologiko, ekonomiko edo politiko batean mugitzen da. Horregatik, garrantzitsua da sareak aztertzerako orduan hauei exogenoak diren faktoreen eragina aintzat hartzea.

Kapitulu honetan aurkezten den sare konplexuei buruzko teoria literaturan aurkitu daitekeenaren zati txiki bat besterik ez da. Izan ere, azken urteetan sare konplexuen teoriak izugarrizko garapena jasan du, praktikan izan dezakeen erabilgarritasuna dela eta. Zentzu honetan, teoria horren azterketa sakonago batek finantza sistemak aztertzeko tresna gehiago emango lituzke eta beraz, etorkizunerako helburu interesgarri bat izango litzateke sare konplexuen emaitza teoriko berritsuenak praktikara eramatea.

Bibliografia

Barabási, A.-L. eta Albert, R. (1999),"Emergence of Scaling in Random Networks," *Science*, 286(5439), 509–512.

Barrat, A., Barthelemy, M., Pastor-Satorras, R. eta Vespignani, A. (2004),"The Architecture of Complex Weighted Networks," *Proceedings of the National Academy of Sciences*, 101(11), 3747–3752.

Borgatti, S. P., Mehra, A., Brass, D. J. eta Labianca, G. (2009),"Network Analysis in the Social Sciences," *Science*, 323(5916), 892–895.

Bougheas, S. eta Kirman, A. (2015),"Complex Financial Networks and Systemic Risk: A Review," *Complexity and Geographical Economics*, 19, 115–139.

Caldarelli, G., Chessa, A., Pammolli, F., Gabrielli, A. eta Puliga, M. (2013),"Reconstructing a Credit Network," *Nature Physics*, 9(3), 125–126.

Croes, D., Couche, F., Wodak, S. J. eta van Helden, J. (2006),"Inferring Meaningful Pathways in Weighted Metabolic Networks," *Journal of Molecular Biology*, 356, 222–236.

da Cruz, J. P. eta Lind, P. G. (2012),"The Dynamics of Financial Stability in Complex Networks," *The European Physical Journal B*, 85(8), 1–9.

Erdős, P. eta Rényi, A. (1960),"On the Evolution of Random Graphs," *Publications of the Mathematical Institute of the Hungarian Academy of Sciences*, 5, 17–60.

Espinoza-Valdez, A., Femat, R. eta Ordaz-Salazar, F. C. (2010),"A Model for Renal Arterial Branching Based on Graph Theory," *Mathematical Biosciences*, 225, 36–43.

Estrada, E. (2011),"Community Detection Based on Network Communicability," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 21, 016103.

——— (2012), *The Structure of Complex Networks: Theory and Applications* , New York: Oxford University Press.

Estrada, E. eta Benzi, M. (2014),"Walk-Based Measure of Balance in Signed Networks: Detecting Lack of Balance in Social Networks," *Physical Review E*, 90(4), 042802.

Estrada, E. eta Knight, P. A. (2015), *A First Course in Network Theory* , New York: Oxford University Press.

Futschik, M. E., Chaurasia, G., Tschaut, A., Russ, J., Babu, M. M. eta Herzel, H. (2007),"Functional and Transcriptional Coherency of Modules in the Human Protein Interaction Network," *Journal of Integrative Bioinformatics*, 4(3), 198–207.

Galbiati, M., Delpini, D. eta Battiston, S. (2013),"The Power to Control," *Nature Physics*, 9(3), 126–128.

Grzybowski, B. A., Bishop, K. J., Kowalczyk, B. eta Wilmer, C. E. (2009),"The "Wired" Universe of Organic Chemistry," *Nature Chemistry*, 1, 31–36.

Lai, Y. eta Hu, Y. (2021),"A Study of Systemic Risk of Global Stock Markets Under COVID-19 Based on Complex Financial Networks," *Physica A: Statistical Mechanics and its Applications*, 566, 125613.

Montoya, J. M., Pimm, S. L. eta Solé, R. V. (2006),"Ecological Networks and Their Fragility," *Nature*, 442(7100), 259–264.

Newman, M. E. (2004),"Analysis of Weighted Networks," *Physical Review E*, 70(5), 056131.

Ratkaj, I., Stajduhar, E., Vucinic, S., Spaventi, S., Bosnjak, H., Pavelic, K. eta Pavelic, S. K. (2010),"Integrated Gene Networks in Breast Cancer Development," *Functional & Integrative Genomics*, 10, 11–19.

Scott, H. (2012),"Interconnectedness and Contagion," *Research Report by Committee on Capital Market Regulation.*

Tang, Y., Xiong, J. J., Jia, Z.-Y. eta Zhang, Y.-C. (2018),"Complexities in Financial Network Topological Dynamics: Modeling of Emerging and Developed Stock Markets," *Complexity*, 2018(468014).

Turán, P. (1941),"On an Extremal Problem in Graph Theory," *Matematikai és Fizikai Lapok*, 48, 436–452.

Varela, L. M., Rotundo, G., Ausloos, M. eta Carrete, J. (2015),"Complex Network Analysis in Socioeconomic Models," in *Complexity and Geographical Economics*, Springer, pp. 209–245.

Watts, D. J. eta Strogatz, S. H. (1998),"Collective Dynamics of 'Small-World' Networks," *Nature*, 393(6684), 440–442.

5

The Role of Connectivity in Stock Market Networks

Published in Scientific Reports under the title "Loss of Balance in Stock Markets", DOI: 10.1038/s41598-021-91266-4

1 Introduction

The 2007-08 financial crisis triggered unprecedented distress of the financial markets worldwide and arose a new type of risk, the systemic risk. Consequently, the European sovereign debt crisis that followed the global financial crisis has been largely studied in the financial literature.

Since the credit default swaps (CDS) are financial derivatives commonly used to swap credit risk, sovereign and corporate CDS have been commonly studied to understand the mechanism of contagion between bank and sovereign default risk and therefore, to be able to quantify the financial systemic risk. Acharya et al. (2014) document that post-bailout changes in sovereign CDS explain changes in bank CDS even after controlling for aggregate and bank-level determinants of credit spreads, confirming the sovereign-bank loop. Alter and Beyer (2014) use CDS to

Article published in collaboration with Ernesto Estrada and Brais Álvarez Pereira.

quantify spillovers between sovereign credit markets and banks in the euro area. De Bruyckere et al. (2013) investigate contagion between bank and sovereign default risk in Europe over the period 2007–2012.

Using other instruments as returns of hedge funds, banks, broker/dealers, or insurance companies, Billio et al. (2012) conclude that banks play a much more important role in transmitting shocks than other financial institutions. To do that, they use networks with measures of connectedness based on principal-components analysis and Granger-causality. The use of networks in finance was previously highlighted by Allen et al. (2010), related to propagation and generation of systemic risk. Battiston et al. (2012b) construct networks for banks based on borrowing/lending relationships and analyze the effect of the diversification of credit risk on systemic risk and financial market stability. In a related work, Corsi et al. (2018) define and construct Granger-causality tail risk networks between 33 systemically important banks (G-SIBs) and 36 sovereign bonds worldwide. They explode the network structure to identify periods of distress in financial markets and possible channels of systemic risk propagation. Acemoglu et al. (2015) also remark the importance of identifying the SIBs in financial networks and find that at a certain point a densely interconnected financial system becomes more fragile to propagations of shocks. Betz et al. (2016) study the evolution of tail risk dependence and systemic risk contributions as well as spillover effects between European banks and sovereigns networks, and showed that financial market fragmentation during the sovereign crisis coincides with an increasing interconnectedness between banks and sovereigns.

However, as Gross and Siklos (2020) point out, beyond the studies of Augustin et al. (2018) and Bedendo and Colla (2015) that find significant risk spillovers from sovereigns to corporations in Europe; there is little empirical evidence of how the nonfinancial corporate sector has been affected by the rise in the sovereign and bank credit risk. Gross and Siklos (2020) use a CDS network and show that bank and sovereign risk are important drivers of corporate credit risk. On the one hand, the Lehman bankruptcy and the European debt crisis fundamentally transformed the network structure and, on the other hand, the transmission of risk within the nonfinancial sector remained largely unchanged during the crisis events.

Liu et al. (2014) study the downside risk and the portfolio diversification with special consideration of the crisis period. They divided the euro-zone equity markets into the sadly denominated GIIPS (Greece, Italy, Ireland, Portugal, and Spain) and the Core (Austria, Finland, France, Germany, and the Netherlands) and analyzed the Value at Risk for the domestic indexes and for the two groups, with mixed results.

Based on all the above, we are interested in a better understanding of the changes that occurred in the national stock network structure for the European countries. Particularly, we want to study whether these changes have been different between countries and how the financial sector is related to the nonfinancial sector. Moreover, we want to evaluate the differences, if this the case, between the European countries that belong to the GIIPS and other European and non-European core countries. Specifically, we consider Germany and France as two of the core European countries and the US and Japan as non-European world leaders.

In complex systems such as financial networks there are different ways to define the connectivity between network constituents. In a previously mentioned study, Corsi et al. (2018) consider Granger-causality tail risk as the criteria to determine the interconnectivity between banks and sovereign bonds. Nevertheless, other ways of connecting financial institutions such as the borrowing/lending relations have also been considered (Battiston et al., 2012a,b, 2016; Bardoscia et al., 2021). In the case of stock markets, where stocks are represented by the nodes of the network, we establish the edges such that they capture the dependence between stocks. In this sense, Mantegna (1999) proposes to quantify the similarity between two stocks i and j through a distance function $d_{ij} = (2(1 - \rho_{ij}))^{1/2}$ based on their pairwise linear correlation ρ_{ij}. Many related studies use distances to construct national and international stock market networks (Birch et al., 2016; Brida et al., 2016; Kauê Dal'Maso Peron et al., 2012; Wang et al., 2018; Zhao et al., 2018; Guo et al., 2018). An alternative approach proposed by Chi et al. (2010) includes another linear correlation-based criterion that establishes the existence of the link only if $\rho_{ij} > |z|$, for a given threshold z. Using any of the previous approaches implies that the weight of the links or edges between stocks is nonnegative, and this is at the same time its main disadvantage since there is an inherent loss of information (Heiberger, 2014).

Beyond previous approaches, Stavroglou et al. (2019) remark the importance of considering the sign of dependence and they construct separated stock networks for positive and negative relationships. In this sense, the theory of signed graphs (Harary et al., 1953) has recently become very popular since it allows to analyze networks with the simultaneous presence of positive and negative links (Leskovec et al., 2010; Facchetti et al., 2011; Estrada and Benzi, 2014; Kirkley et al., 2019; Shi et al., 2019).

Therefore, we use signed networks to analyze the stock markets' topology for the previously mentioned nine countries (Ferreira et al., 2021). The links between stocks are established in terms of rank correlations, which provide a measure of similarity better suited for financial distributions than those based on linear correlation. Moreover, note that in finance the time varying behavior of the networks is an important characteristic since, as Ascorbebeitia Bilbatua et al. (2021) emphasize, the time varying comovements become a risk factor. Therefore, we fill a gap pointed out in a recent study (Bardoscia et al., 2021) by analyzing the evolution of each national equity network for the period between January 2005 and September 2020, which takes into account several convulse periods such as the global financial crisis of 2007, the European debt crisis, Brexit, the US elections, and the first months of the COVID-19 crisis. To account for the time variation, we use time varying Kendall's tau as proposed by Ascorbebeitia Bilbatua et al. (2021).

In this context, we introduce the concept of balance in financial networks as an indicator of the stability of stock markets. The results show a strong decay in the balance degree from high to poorly balanced markets in six of the countries, which points out a clear loss of market stability. The balance-unbalance transition (BUT) occurs around September 2011 in the US, Portugal, Ireland, Greece, and Spain, just after the Black Monday in August 2011. The transition in France is delayed until September 2012. Nevertheless, the rest of the stock markets do not show such a transition pattern. In a more detailed study, we find that the loss of stability in stock markets is a consequence of a topological reorganization of some small-cap stocks that belong to the nonfinancial sector. Therefore, the results support the importance of the effect of too interconnected to fail companies on financial stability and the evaluation of systemic risk. This is a novel and overwhelming result that has a clear impact on asset allocation and hedging strategies, which opens promising avenues for future research.

2 Methodology

To study the similarity between stock price changes we consider the time series for stock returns. As mentioned in the Introduction, there are different ways to define the connectivity between stocks. Beyond other common approaches that use distance or linear correlation matrices, we use rank correlations. To establish how the similarity between stocks is studied in terms of dependence, let us consider the returns of three stocks $\{Y_{jt}\}_{t=1}^{T}$, $j = 1, 2, 3$. Then, four different cases can emerge from the triad that these three stocks and their mutual connectivities form: (i) the three pairs of stocks are correlated, (ii) one pair of stocks is correlated and two pairs are anticorrelated, (iii) two pairs of stocks are correlated and one pair is anticorrelated, and (iv) the three pairs of stocks are anticorrelated. If we represent the three stocks at the vertices of a triangle and the rank correlations as the corresponding edges, we have that the four cases analyzed before can be represented as in Figure 5.1.

A triangle like those illustrated in Figure 5.1 is known as a signed triangle. Beyond triangles, we have a set of $p > 3$ nodes representing the stocks in a market and a set of edges connecting such nodes, so that we can construct networks. Since each edge in the network accounts for the dependence between the nodes to which it is connected, the constructed network is a weighted

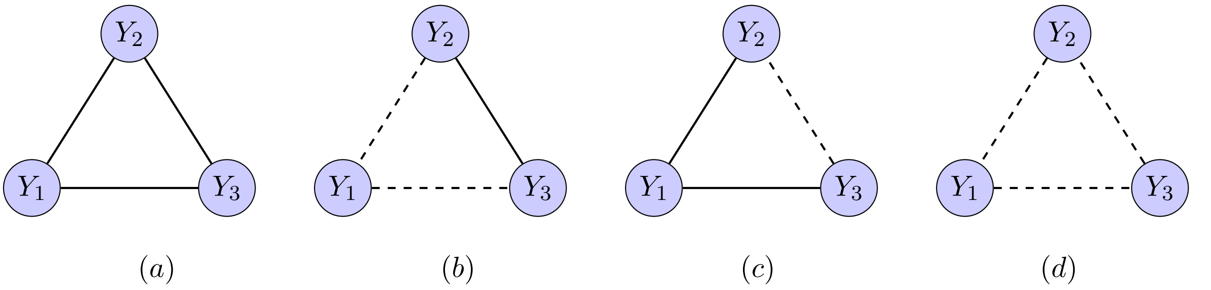


Figure 5.1: Representation of interdependencies between stocks Y_1, Y_2 and Y_3 in a triad. Stocks are represented as the vertices (or nodes) of a triangle and edges represent the rank correlation between them as accounted by Kendall's tau. Positive Kendall's taus between stocks are represented as solid lines and the negative ones as dashed lines.

signed network.

Let V be a set of nodes that represent the stocks of a given market in a graph $G = (V, E, W)$, where $E \subseteq V \times V$ is a set of edges and $W : E \rightarrow [-1, +1]$ is a mapping that assigns a value between -1 and $+1$ to each edge. Note that in this case, the edges represent the rank correlation between each pair of nodes. Thus, these graphs are known as Weighted Signed Stock Networks (WSSNs).

Networks can be represented mathematically using matrix algebra. Let $\mathbf{A}$ be a p-size square matrix. Then, it is possible to transfer the information in the network to matrix $\mathbf{A}$. The link between any pair of nodes (i, j) is identified with the corresponding a_{ij} element in $\mathbf{A}$, such that a_{ij} is the weight of the edge connecting such nodes. The square matrix $\mathbf{A}$ whose entries are a_{ij} for $1 \leq i, j \leq p$ is called the adjacency matrix of G. In this case the adjacency matrix contains the rank correlation coefficients for each pair of stocks (i, j).

As pointed out in several studies and just remarked previously, the time varying behavior of distributions is an important characteristic in finance. Then, our interest is to investigate if the relationship between stocks changes with time and its effect on the network structure. To analyze the evolution of the WSSNs over time, the edges of the networks are defined on the basis of the time varying pairwise Kendall's taus between stock returns. Such coefficients are estimated with the nonparametric estimator proposed in Chapter 3 for time varying Kendall's tau under α-mixing local stationary variables,

$$\hat{\tau}_t^{ij} = \frac{4}{(Th)^2 - \sum_{r=1}^{T} K_{h,tr}^2} \sum_{r,s=1}^{T} K_{h,tr} K_{h,ts} \mathbf{I} \left\{ Y_{ir} < Y_{is}, Y_{jr} < Y_{js} \right\} - 1, \quad (1)$$

where $K_{h,tr} = k\left((t - r)/(Th)\right)$, $k(\cdot)$ denotes the kernel weight used to introduce smoothness into the distribution of the variables, $h > 0$ is the bandwidth that regulates the degree of smoothness, and $I\{\cdot\}$ is the indicator function. Note that expression (1) is the nonparametric estimator for the bivariate time varying Kendall's tau, the particular case of the estimator (3) proposed in Chapter 3 when $p = 2$.

For each pair of stocks $\{(i,j)\}_{\substack{i=1 \\ j\neq i}}^{p}$ in a given market, we calculate $\{\hat{\tau}_t^{ij}\}_{t=1}^{T}$ and construct the rank correlations matrix $\mathbf{M}_t$ whose entries are the values $\hat{\tau}_t^{ij}$. Moreover, let $\varepsilon \in \mathbb{R}^+$ be a given threshold. Based on each matrix $\mathbf{M}_t$, we create a matrix $\mathbf{A}_t$ whose entries are $a_t^{ij} = \hat{\tau}_t^{ij}\mathbf{I}\{\left|\hat{\tau}_t^{ij}\right| \geq \varepsilon\}$. The threshold determines the amount below which the correlation is considered to be negligible. We consider $\varepsilon = 0.3$ a reasonable value. A higher threshold decreases the number of connections that remain alive in the network. Moreover, the main findings are robust to alternative specifications such as $\varepsilon = 0.2$ and $\varepsilon = 0.4$. Then, the matrix $\mathbf{A}_t$ is defined to be the adjacency matrix of a WSSN at time t. Here we construct weighted adjacency matrices from January 2005 until September 2020. Therefore, we have a set of WSSNs, $\{G_t\}_{t=1}^{T}$, obtained with matrices $\mathcal{A} = \{\mathbf{A}_t\}_{t=1}^{T}$.

Balance in Weighted Signed Stock Networks

Signed triangles such as the ones represented in Figure 5.1 account for situations in which positive comovements are represented by positive edges and inverse comovements are represented with negative edges. According to Heider's hypothesis of balance (Heider, 1946), some combinations of positive and negative edges give stability to the networks while other combinations provide instability. Since Heider identifies the concept of stability with balance, the degree of balance would show how stable is the system we are analyzing. The most stable state is known as a balanced state and those states far from equilibrium are known as unbalanced. According to the interpretation in Heider (1946) about the balance concept, a signed triangle for which the product of its edge signs is positive is known as a balanced triangle (cases (a) and (b) in Figure 5.1). Otherwise, it is said to be unbalanced (cases (c) and (d) in Figure 5.1). Note that when considering rank correlations, two aspects are considered: the strength of the dependence and its sign.

The previous concepts related to the triangle balance can be naturally extended to networks of any size using signed graphs. Actually, Cartwright and Harary (1956) propose to use signed networks to model the balance/unbalance in social systems and prove the *structure theorem*. This theoretical result states that a balanced network can be split up into two disjoint sets of nodes such that all positive connections are between nodes in the same group and any negative connection implies nodes from different groups. In this sense, a signed graph where all the triangles are balanced is said to be balanced. Beyond triangles, this result can be generalized to cycles, closed sequences of nodes and edges where such nodes and edges are all different. Hence, a signed graph whose all cycles are balanced is also balanced. For any cycle in a WSSN, we will say that it is positive if the product of all Kendall's taus representing the edges of the cycle is positive. Therefore, we can determine that a WSSN is balanced if all its cycles are positive. Note that a completely correlated network is always a balanced network. On this basis, theoretical arguments can lead one to think that a completely correlated stock network is an ideal scenario to aim for, but from a financial point of view, it is not the most promising scenario. In fact, negative correlations between stocks are essential for hedging strategies. We remark that networks with negative edges can also be balanced, but it depends on the quantity of such negative edges. A related theoretical result states that networks whose all cycles have an even number of negative edges are balanced (Estrada, 2019; Kirkley et al., 2019).

However, the key point here is not to reduce the problem of whether the network is balanced or not, but to quantify how close or far a WSSN is from balance. To do so, we consider a hypothetical scenario where a given WSSN $G_t = (V, E, W_t)$ has evolved from a balanced network G'_t and we quantify the departure of the WSSN from its balanced version. For simplicity, we consider $G'_t = (V, E, |W_t|)$, a completely correlated version of G_t where all the edges are positive, i.e. a network with an adjacency matrix where $a_{ij} > 0$ for any pair of stocks (i, j).

As explained in Chapter 4, complex networks are always exposed to a certain economic and political environment or to external stress factors that are independent of the topology of the network. Specifically, in finance, many external factors can influence the stock market structure. Such type of changes or shocks surrounding the network affect all the connections in the network at the same time and consequently, they can change the structure of the network. Therefore, it is necessary to assess their effect on the topology of the WSSNs. In physics, and more concretely in thermodynamics, the environment that surrounds the networks is associated with a thermal bath with a temperature $\mathcal{T}_t$ where any complex network is submerged at each moment t. To account for such external factors' affections in the network a β_t parameter is considered. β_t is identified as the inverse temperature of the networks that represents a level of "agitation" of the market at a given moment t.

In this context, a measure of structural balance for a WSSN that characterizes the equilibrium $G'_t \rightleftarrows G_t$ is defined by Estrada and Benzi (2014) as

$$K_t = \frac{tr\left(\exp\left(\beta_t A_t\left(G_t\right)\right)\right)}{tr\left(\exp\left(\beta_t A_t\left(G'_t\right)\right)\right)}, \tag{2}$$

where $\mathbf{A}_t(G_t)$ is the adjacency matrix of G_t and $tr(\mathbf{X})$ is the trace of a matrix $\mathbf{X}$. According to expression (2), when the level of global risk is very low, $\beta_t \to 0$, we have that $K_t \to 1$ and the corresponding WSSN tends to be balanced independently of its topology and edge weights.

Moreover, $K_t = 1$ if and only if the WSSN is balanced. Thus, this equilibrium constant is bounded as $0 < K_t \leq 1$, with the lower bound indicating a large departure of the WSSN from its balanced analogous. Therefore, the departure of K_t from unity characterizes the degree of unbalance of the WSSN.

3 Data

We study the stock markets for some European countries, the US and Japan. The European countries include Greece, Italy, Ireland, Portugal, and Spain which are known as GIIPS, plus Germany and France as core countries. For comparison purposes, the US and Japan stock markets are considered as they belong to important drivers of the world economy. We work with equities' daily closing price and volume data from 01-03-2005 to 09-15-2020, obtained from the Morningstar database.

The database of each stock market is created according to the next procedure. The criteria used to consider a company as a candidate to form the database for each country is the trading volume at the end of April 2020. Initially, a set of around 250 companies with the highest trading volume on the 28th of April 2020 is selected. Exceptions include the US and Japan, for which the initial sets are extended to 794 and 427 companies respectively because their stock markets are very diversified in terms of volume. Then, companies with a big amount of missing values according to the daily closing prices series (more than three market years of consecutive missing values or more than 30 % of non-consecutive missing values) are dropped out. There are many possible reasons for those missing values, such as trading cancellation for a while, stock market exit, late entry into the stock market, merger of companies, bankruptcy, etc. We find that allowing for 30% of missing values in a stock price series is a reasonable threshold. Nevertheless, there are some exceptions, companies with more than 30% of missings but with a leading trading volume such as Abengoa S.A Class B (ABG.P) from the Spanish stock market, which constitutes 39% of the total trading volume on the reference date, are excluded from this screening. After the cleaning process, the database is composed of a set of representative companies according to the sector to which they belong (under the Russell Global Sector classification) and their trading volume.

Table 5.1 shows the number of companies considered and the percentage of the total trading volume that constitute such assets for each country. Figure C.1 in Appendix C shows the unconditional Kendall's tau between companies in each sectors combination. The colormaps show that the US, Portugal, Ireland, and Greece are countries with many negative intersectoral

Table 5.1: Number of assets and the percentage of the total trading volume.

	Germany	France	Greece	Italy	Ireland	Portugal	Spain	US	Japan
no. of assets	81	78	73	83	32	36	78	119	120
volume	70.86%	76.9%	97.41%	90.56%	85.79%	98.75%	86.62%	45.66%	54.58%

relations, while Japan stands out for having a positive and strong dependence between sectors

Throughout this chapter, and to construct the adjacency matrices described in Section 2, daily returns of the equities' prices are considered.

Since the stock market structures are submitted to external factors and motivated by results in Chapter 3, we consider the policy uncertainty surrounding each country as such an external factor. Hence, we use the Economic Policy Uncertainty (EPU) index introduced in Chapter 3 as a proxy of the inverse temperature β_t of the networks from January 2005 to September 2020. Portugal is an exception since there is no specific EPU index available for it, so the European EPU index is considered instead. EPU index data is publicly available on the Economic Policy Uncertainty website[1]. To unify the scales of the national EPU indexes, the parameter β_t is defined as the relative EPU, $EPU_t/\max(EPU)$, such that $\beta_t \in (0, 1]$ for all the countries. Hence, we say that the risk is high when the temperature of the system at a given time is approaching the maximum temperature of the country in the period of study, $\beta_t \to 1$. On the contrary, the risk is minimum when the temperature at a given time is much smaller than the maximum temperature of the period, $\beta_t \to 0$.

4 Results

We construct the stock networks of the nine countries for the period between January 2005 and September 2020 based on the adjacency matrices described in Section 2 and we analyze the evolution of the balance of such networks. Top panels in Figures 5.2 and 5.3 present the evolution of the balance degree K_t (blue lines) together with its detrended cumulative sum (DCS), represented with red lines. The DCSs allow for quantitative analysis of these balance-unbalance transitions (BUT) as well as for the identification of the dates at which they occur. For the periods in which the balance does not decay the DCS is increasing, and it has an inflection point when the balance degree drops abruptly. Hereafter a decreasing balance causes a negative slope of the DCS.

[1] `https://www.policyuncertainty.com/index.html`

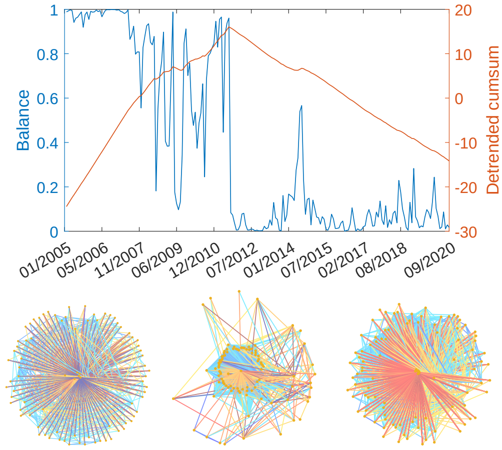
Bottom panels in Figures 5.2 and 5.3 illustrate the snapshots of three representative networks over the period of study using the degree of the nodes as a proxy for their location. Nodes with the highest degrees are located at the center and those with a low degree are located at the periphery of the graph. The colors of the edges indicate their strength based on Kendall's tau estimates, going from dark red for the most negative links to blue for the most positive ones. Appendix C.2 provides the snapshots of all the networks where node labels are incorporated.

The first interesting observation is the large balance observed for all markets between January 2005 and August 2007 with $K_t \geq 0.98$, indicating that in such period all the stock markets are stable. However, from August 2007 the temporal evolution shows fluctuations in the balance degree in some countries (see top panels of Figure 5.2(a-e)). In five of those countries, there is a strong drop in the balance around September/October 2011. Specifically, there is a transition from high to poorly balanced markets in Ireland in September 2011 and the US, Portugal, Greece, and Spain in October 2011. It is striking that these countries are those who present negative rank correlations among sectors (see Figure C.1 in Appendix C.1). Attending to the DCSs, a positive slope can be appreciated in all five markets with BUT between January 2005 and September/October 2011, the moment at which the slope changes indicating an abrupt decay of the balance. Therefore, these change points are identified as the date of the transitions. After the BUT, the US and Portugal display quite similar behavior. Slightly different behavior is appreciated for Spain, whose balance decay was interrupted in April 2017. At this point, the balance started to recover at a rate similar to that in the period between January 2005 and October 2011. Note also that in the latter case, as well as for Ireland, the balance degree drops dramatically again after the crisis produced by COVID-19. The results for the French stock market are also included (Figure 5.2(f)) because, although the BUT occurs significantly later, i.e. in September 2012, it displays a behavior similar to that of Spain.

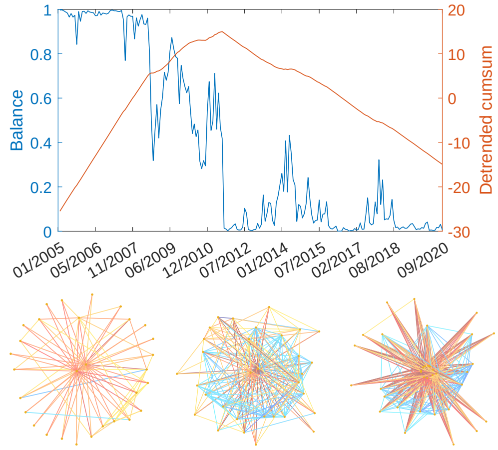
In the opposite situation we find Germany, Italy, and Japan, whose stock markets always display a very constant balance. Even if some fluctuations can be appreciated, their DCSs have an almost zero slope supporting that these markets are highly balanced between 2005 and 2020 (see Figure 5.3). The colormaps in Figure C.1 show that stocks in those three countries have always a positive dependence among sectors, especially in Japan.

Different factors may influence the BUTs in the six countries. These factors can be external to the WSSN structures such as the level of risk to which a given market is exposed. Other exogenous factors include financial and economic shocks that occurred in the period of study, not necessarily in the country whose stock market is analyzed. This factor is accounted for through the parameter β_t introduced before.

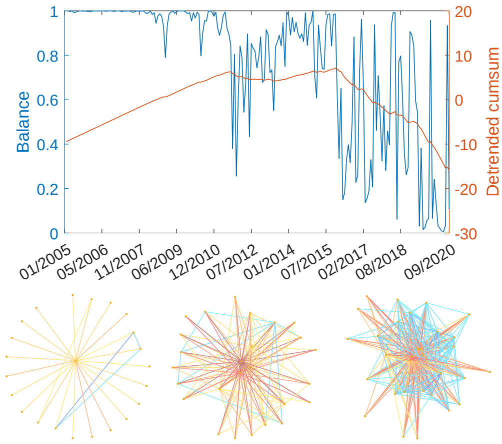
Among many possible exogenous risk factors, the EPU is one candidate to be a proxy of



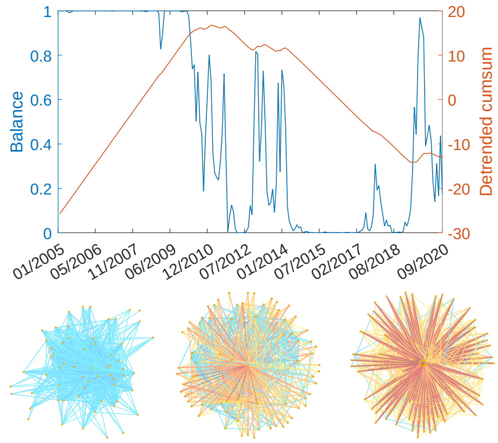
(a) The US



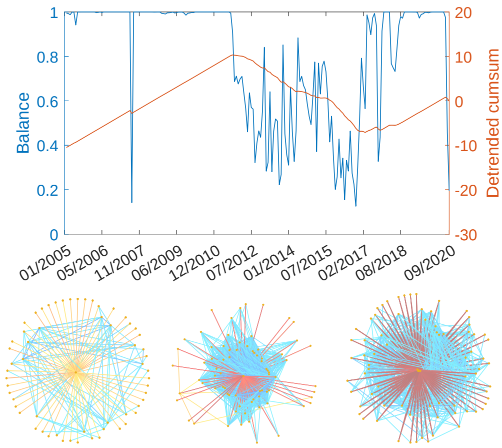
(b) Portugal



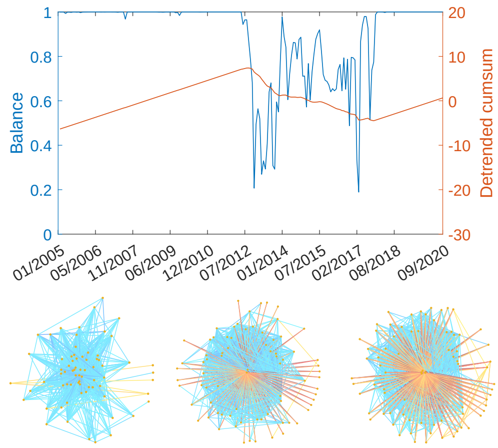
(c) Ireland



(d) Greece

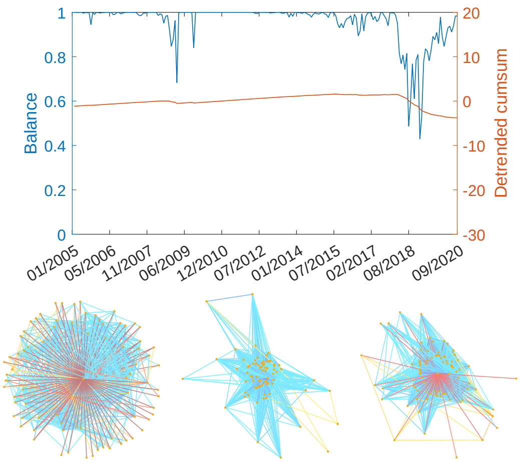


(e) Spain

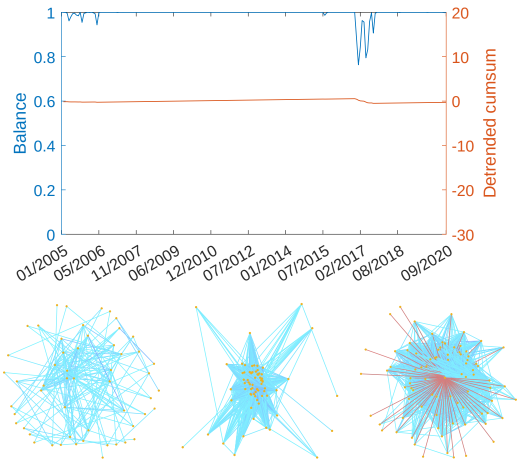


(f) France

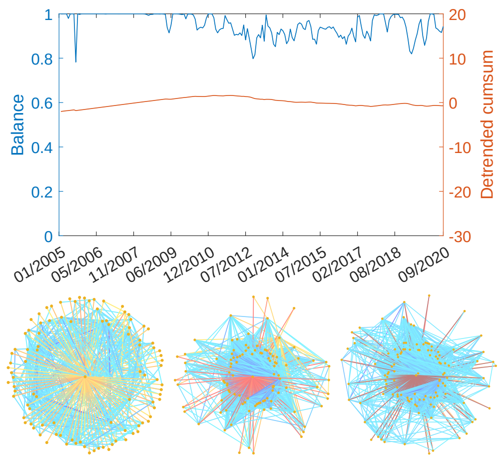
Figure 5.2: (Top panels) Evolution of the balance (blue line) and of its DCS (red line) between January 2005 and September 2020. (Bottom panels) Snapshots for three representative networks before, during, and after the BUT. The colors of the edges go from dark red for the most negative Kendall's tau estimates to blue for the most positive ones.



(a) Germany



(b) Italy



(c) Japan

Figure 5.3: (Top panels) Evolution of the balance (blue line) and of its DCS (red line) between January 2005 and September 2020. (Bottom panels) Snapshots for three representative networks before, during, and after the BUT. The colors of the edges go from dark red for the most negative Kendall's tau estimates to blue for the most positive ones.

β_t. The analysis of the EPU index for all countries indicates that the level of uncertainty was relatively low from January 2005 to August 2007, a period that displayed a high balance in all markets. However, we cannot determine that the national EPU indexes cause the BUTs observed around September/October 2011 by themselves. In fact, the national EPU indexes peaked at a different moment for each country: Portugal in June 2016, the US in May 2020, Spain in October 2017, France in April 2017, Greece in November 2011, and Ireland in May 2020. In this sense, Mei et al. (2018) find evidence that in European stock markets the national EPU indexes do not have a significant impact in volatility forecasting and determine that the

US EPU index does contain useful predictive information of the European stock markets, mostly during recessions.

Moreover, the association between an increasing EPU and a loss of balance conflicts also with the lack of strong drop of balance during the financial crisis of 2008. The US stock market is an exception in this sense since it suffered from significant balance fluctuations in this period. A possible explanation might be an interruption of the negative time varying correlations between policy uncertainty and stock market returns during the financial crisis, probably as a consequence of the bailout package for the US banking sector of 2008 and the stimulus package of 2009, which pushed the market into positive returns even if the policy uncertainty remained high (Antonakakis et al., 2013).

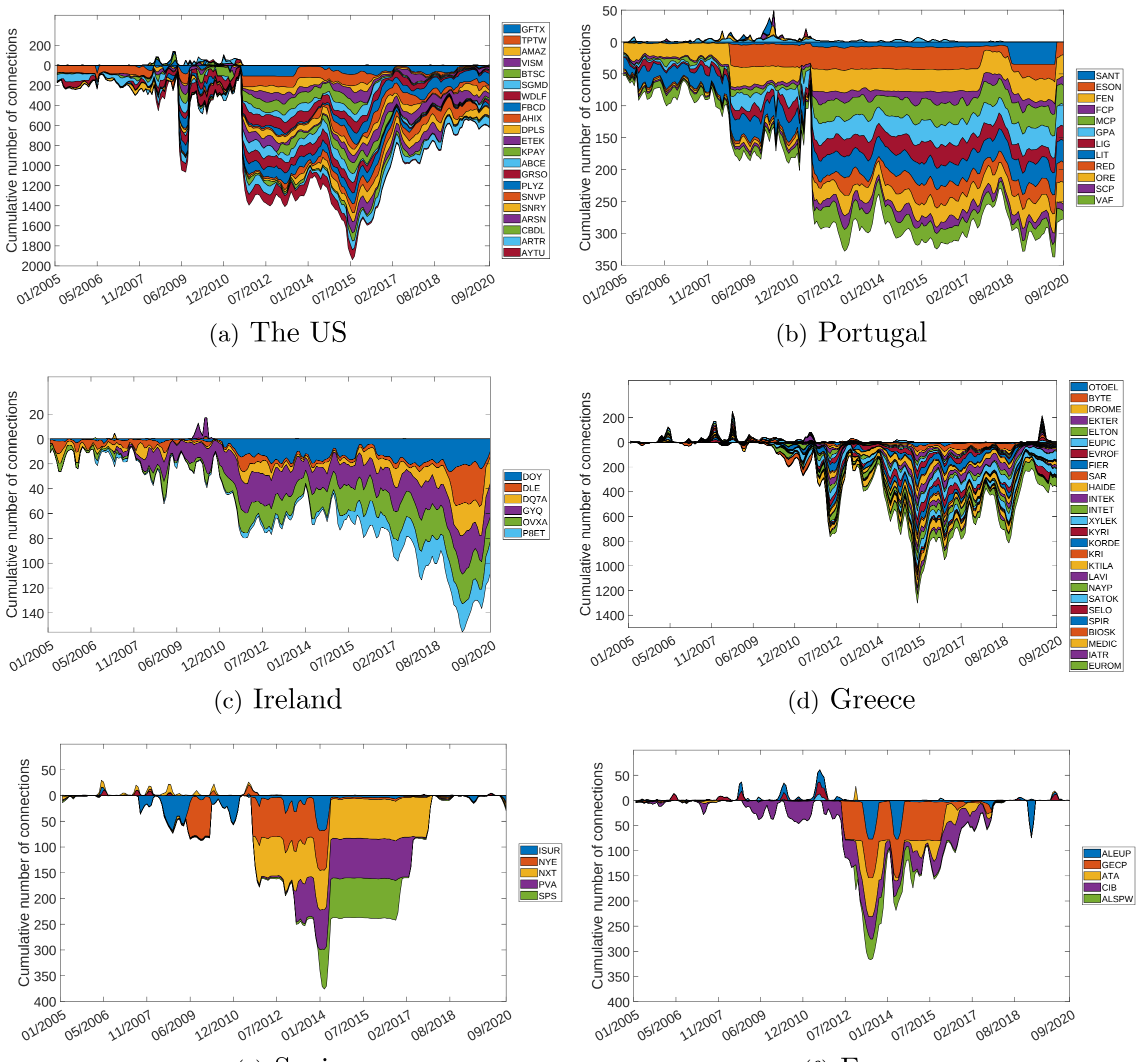
It is interesting that between 2005 and the COVID-19 crisis, the US EPU index peaked in August 2011, coinciding with the Black Monday[2]. Moreover, the US (VIX), European (VSTOXX), and Asian (VKOSPI) stock market volatility indexes reached extremely high values in September 2011, which in turn coincides with the balance transition found in the five stock markets. These findings point towards the events associated with Black Monday 2011 as the most likely triggers for those BUTs.

Apart from external factors, we search for any possible triggering of the BUTs inherent to the WSSNs topology. Endogenous factors involve the sign and magnitude of the interactions between the stocks if any. This information is accounted for in the adjacency matrices through Kendall's tau rank correlation estimates, $\hat{\tau}_t^{ij}$. The evidence shows that, although negative dependencies are necessary for unbalance, neither the number of negative edges nor the ratio of negative to total edges explains it. Consider for instance of the US WSSN in September 2010, which has 950 negative edges and $K_t = 0.245$, indicating its lack of balance. The US WSSN in February 2011 has also 950 negative edges, but it is balanced ($K_t = 0.946$). Something similar happens with Portugal, whose WSSNs in July 2012 and in May 2019 have 720 negative edges but the balance degrees are $K_t = 0.700$ and $K_t = 0.407$ respectively. Moreover, the WSSNs in the Spanish stock market in August 2005 and September 2007 have the same proportion of negative to total edges (1.4%), but even if the network was perfectly balanced in August 2005 ($K_t = 1$), it was unbalanced ($K_t = 0.142$) in September 2007.

In a more detailed exploration of the WSSNs after BUT (see bottom panels in Figure 5.2) we find central small groups of stocks negatively connected with the rest of the stocks in the network. These subgraphs, defined as small fully-negative cliques (FNC), are highly unbalanced and can be identified in Figure 5.2 by the naked eye because of their star appearance. The relations of

[2]The Black Monday 2011 refers to August 8, when the US and the global stock market crashed, following the credit rating downgrade of the US sovereign debt from AAA to AA+ for the first time.

the stocks in these cliques are investigated and the temporal evolution results for the cumulative number of positive and negative connections are presented over and below the horizontal line in the panels in Figure 5.4, respectively. The panels show that each of the stocks in the FNC increases significantly its number of negative connections with the rest of the stocks after the BUT and is negatively connected to almost every other stock in the WSSN. These FNCs look



(a) The US
(b) Portugal
(c) Ireland
(d) Greece
(e) Spain
(f) France

Figure 5.4: Evolution of the cumulative number of connections of the stocks in the FNC for the six countries where BUT occurred.

similar to complete split graphs (CSG) (Estrada and Benzi, 2017), graphs that can be partitioned into a clique and an independent set (a set of nodes not connected among them) such that every node in the independent set is linked to every node in the clique. Note that this structure is found in all markets where BUT occurs (see the snapshots of Figure 5.2). Therefore, we conclude that the unbalance is mainly due to the presence of these cliques of a small group of mutually anticorrelated stocks. This finding supports that the stock markets' stability is related to the interdependences in the network. Actually, the results point towards the dependence structure underlying the cliques as the triggerer of market instability. In a related work, Markose et al. (2012) remark the consequences of too interconnected to fail financial firms in systemic risk and its effect on the instability of CDS networks. The results are also in line with the conclusions in Acemoglu et al. (2015), who highlight that some features such as the interdependencies in financial networks work as a source of systemic risk and cause instability.

Table 5.2: Analysis of the WSSNs with BUT.

Country	s	ΔK	Companies in clique
Greece	26	0.728	BIOSK, BYTE, DROME, EKTER, ELTON, **EUPIC**, EUROM, EVROF, FIER, HAIDE, IATR, INTEK, INTET, KORDE, KRI, KTILA, KYRI, LAVI, MEDIC, NAYP, SAR, OTOEL, SATOK, SELO, SPIR, XYLEK
The US	21	0.724	ABCE, **AHIX**, AMAZ, ARSN, ARTR, AYTU, **BTSC**, CBDL, DPLS, ETEK, FBCD, GFTX, GRSO, **KPAY**, PLYZ, SGMD, SNRY, SNVP, TPTW, VISM, WDLF
Portugal	12	0.721	ESON, FCP, **FEN**, GPA, LIG, LIT, MCP, ORE, RED, **SANT**, SCP, VAF
Ireland	6	0.368	DEL, DOY, DQ7A, GYQ, OVXA, P8ET
Spain	5	0.324	**ISUR**, NXT, **NYE**, SPS, PVA
France	5	0.200	ALEUP, ALSPW, ATA, CIB, GECP

NOTE: Columns two and four are the number of companies s forming a central negative clique for each country and the acronym for the companies in such clique. Column three indicates the difference in the average balance. Stock tickers in bold belong to the financial sector.

Columns two and four in Table 5.2 contain the number of companies and the tickers of such companies in the negative cliques for each country. The tickers in bold represent the companies that belong to the financial sector. Column three shows the difference between the average balance before and after the BUT, $\Delta K = \overline{K}_{Before} - \overline{K}_{After}$.

The study of these FNCs reveals some other characteristics about the constituent companies of the cliques. First, note that the size of the FNC is highly correlated with the intensity of the drop in the balance. The FNCs are formed by small and micro-caps (the mid-cap SAR in Greece is an exception) that mainly belong to the nonfinancial sector as it is shown in Table 5.2. Based on these results, we split up the WSSNs into financial and nonfinancial sectors and analyze them separately. The results show that the BUT is observed mainly for the networks containing nonfinancial-nonfinancial interactions (see Figures C.3 to C.5 in Appendix C). Moreover, note that the colormaps of many countries with BUT in Figure C.1 (Appendix C) also show that the most negative rank correlations are found between stocks in nonfinancial sectors. Concretely, negative relationships prevail between stocks related to technology, health, industry, and consumer goods.

5 Conclusions

In previous chapters, we have highlighted the importance of the dependence structure between financial assets and its evolution over time. Therefore, the stocks listed in the markets and their mutual interdependences establish the topological structure of such markets. In the literature, network analyses based on transforming linear correlations into distances are very common, mainly because they allow to use classical network techniques. Nevertheless, we understand that not considering either the positive or the negative dependence is a mistake since an important part of the information is ignored. In this study, we construct weighted signed graphs for stock networks and the interdependencies are determined in terms of Kendall's tau rank correlations, which have clear benefits over the linear correlation coefficients.

Attending to the evidence of time variation in comovements, we are interested in the evolution of stock networks where the stocks are represented by nodes and the links between nodes are determined using the nonparametric estimator proposed in Chapter 3 for time varying Kendall's tau. Thus, the rank correlation matrix at time t allows creating the corresponding adjacency matrix which leads in turn to construct a weighted stock signed network for that time t.

We study signed stock networks of nine European and non-European countries between January 2005 and September 2020. The use of signed graphs allows introducing the concept of balance in financial networks for the first time. The balance is an equilibrium constant and it can be interpreted as a coefficient that quantifies the stability of the networks over time. The results show a strong drop in the balance degree in six of the nine countries considered. As far as we know, this pattern was not identified in stock markets before. In five of those countries the transition from high to poorly balanced networks occurs around September 2011, following the Black Monday in August 2011. The only country belonging to the GIIPS that does not show such transition is Italy, and together with the German and Japanese stock markets, exhibits a high balance over the entire sample period. Moreover, we find that the Spanish and Irish stock markets were hit hard by the COVID-19 crisis, after which they suffer a further increase of instability.

To shed light on the determinants of the BUTs, different factors have been considered. The results show that the balance transition is triggered by a topological reorganization of a small group of stocks forming a FNC. Specifically, there is a movement of the stocks in the cliques from the periphery to the center of the network. These cliques are highly unbalanced and their stocks are also negatively connected with the rest of the stocks in the network. A more detailed study reveals that the FNCs are constituted by nonfinancial companies with low capitalization. These findings go in line with the conclusions obtained in Chapter 3 that connectedness is a risk source associated with systemic risk and also with the markets' instability. Such a way to construct networks shows the presence of unstable stock markets in the last decade and reveals the vulnerability of the markets to suffer from contagion effects of too interconnected to fail companies' distress.

We think that this study opens a new area of research about market stability due to the strong impacts it has on asset allocation and hedging strategies. In this sense, further studies are necessary to dig into the causes and consequences of the topological reorganization led by those small-cap stocks.

References

Acemoglu, D., Ozdaglar, A., and Tahbaz-Salehi, A. (2015),"Systemic Risk and Stability in Financial Networks," *American Economic Review*, 105(2), 564–608.

Acharya, V., Drechsler, I., and Schnabl, P. (2014),"A Pyrrhic Victory? Bank Bailouts and Sovereign Credit Risk" *The Journal of Finance*, 69, 2689–2739.

Allen, F., Babus, A., and Carletti, E. (2010), "Financial Connections and Systemic Risk,". Technical report, National Bureau of Economic Research.

Alter, A. and Beyer, A. (2014),"The Dynamics of Spillover Effects During the European Sovereign Debt Turmoil," *Journal of Banking & Finance*, 42, 134–153.

Antonakakis, N., Chatziantoniou, I., and Filis, G. (2013),"Dynamic Co-Movements of Stock Market Returns, Implied Volatility and Policy Uncertainty," *Economics Letters*, 120(1), 87–92.

Ascorbebeitia Bilbatua, J., Ferreira García, E., and Orbe Mandaluniz, S. (2021),"The Effect of Dependence on European Market Risk. A Nonparametric Time Varying Approach," *Journal of Business & Economic Statistics*, Doi: 10.1080/07350015.2021.1883439.

Augustin, P., Boustanifar, H., Breckenfelder, J., and Schnitzler, J. (2018),"Sovereign to Corporate Risk Spillovers," *Journal of Money, Credit, and Banking*, 50(5), 857–891.

Bardoscia, M., Barucca, P., Battiston, S., Caccioli, F., Cimini, G., Garlaschelli, D., Saracco, F., Squartini, T., and Caldarelli, G. (2021),"The Physics of Financial Networks," ArXiv: 2103.05623.

Battiston, S., Puliga, M., Kaushik, R., Tasca, P., and Caldarelli, G. (2012a),"Debtrank: Too Central to Fail? Financial Networks, the Fed and Systemic Risk," *Scientific Reports*, 2(1), 1–6.

Battiston, S., Gatti, D. D., Gallegati, M., Greenwald, B., and Stiglitz, J. E. (2012b),"Default Cascades: When Does Risk Diversification Increase Stability?," *Journal of Financial Stability*, 8(3), 138–149.

Battiston, S., Farmer, J. D., Flache, A., Garlaschelli, D., Haldane, A. G., Heesterbeek, H., Hommes, C., Jaeger, C., May, R., and Scheffer, M. (2016),"Complexity Theory and Financial Regulation," *Science*, 351(6275), 818–819.

Bedendo, M. and Colla, P. (2015),"Sovereign and Corporate Credit Risk: Evidence from the Eurozone," *Journal of Corporate Finance*, 33, 34–52.

Betz, F., Hautsch, N., Peltonen, T. A., and Schienle, M. (2016),"Systemic Risk Spillovers in the European Banking and Sovereign Network," *Journal of Financial Stability*, 25, 206–224.

Billio, M., Getmansky, M., Lo, A. W., and Pelizzon, L. (2012),"Econometric Measures of Connectedness and Systemic Risk in the Finance and Insurance Sectors," *Journal of Financial Economics*, 104(3), 535–559.

Birch, J., Pantelous, A. A., and Soramäki, K. (2016),"Analysis of Correlation Based Networks Representing DAX 30 Stock Price Returns," *Computational Economics*, 47(4), 501–525.

Brida, J. G., Matesanz, D., and Seijas, M. N. (2016),"Network Analysis of Returns and Volume Trading in Stock Markets: The Euro Stoxx Case," *Physica A: Statistical Mechanics and its Applications*, 444, 751–764.

Cartwright, D. and Harary, F. (1956),"Structural Balance: A Generalization of Heider's Theory," *Psychological Review*, 63(5), 277.

Chi, K. T., Liu, J., and Lau, F. C. (2010),"A Network Perspective of the Stock Market," *Journal of Empirical Finance*, 17(4), 659–667.

Corsi, F., Lillo, F., Pirino, D., and Trapin, L. (2018),"Measuring the Propagation of Financial Distress with Granger-Causality Tail Risk Networks," *Journal of Financial Stability*, 38, 18–36.

De Bruyckere, V., Gerhardt, M., Schepens, G., and Vander Vennet, R. (2013), "Bank/Sovereign Risk Spillovers in the European Debt Crisis," *Journal of Banking & Finance*, 37(12), 4793–4809.

Estrada, E. and Benzi, M. (2014), "Walk-Based Measure of Balance in Signed Networks: Detecting Lack of Balance in Social Networks," *Physical Review E*, 90(4), 042802.

——— (2017), "Core-Satellite Graphs: Clustering, Assortativity and Spectral Properties," *Linear Algebra and its Applications*, 517, 30–52.

Estrada, E. (2019), "Rethinking Structural Balance in Signed Social Networks," *Discrete Applied Mathematics*, 268, 70–90.

Facchetti, G., Iacono, G., and Altafini, C. (2011), "Computing Global Structural Balance in Large-Scale Signed Social Networks," *Proceedings of the National Academy of Sciences*, 108(52), 20953–20958.

Ferreira, E., Orbe, S., Ascorbebeitia, J., Pereira, B. Á., and Estrada, E. (2021), "Loss of Structural Balance in Stock Markets," *Scientific Reports*, 11(1), 1–10.

Gross, C. and Siklos, P. L. (2020), "Analyzing Credit Risk Transmission to the Nonfinancial Sector in Europe: A Network Approach," *Journal of Applied Econometrics*, 35(1), 61–81.

Guo, X., Zhang, H., and Tian, T. (2018), "Development of Stock Correlation Networks Using Mutual Information and Financial Big Data," *PloS one*, 13(4), e0195941.

Harary, F. et al. (1953), "On the Notion of Balance of a Signed Graph," *Michigan Mathematical Journal*, 2(2), 143–146.

Heiberger, R. H. (2014), "Stock Network Stability in Times of Crisis," *Physica A: Statistical Mechanics and its Applications*, 393, 376–381.

Heider, F. (1946), "Attitudes and Cognitive Organization," *The Journal of Psychology*, 21(1), 107–112.

Kauê Dal'Maso Peron, T., da Fontoura Costa, L., and Rodrigues, F. A. (2012), "The Structure and Resilience of Financial Market Networks," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 22(1), 013117.

Kirkley, A., Cantwell, G. T., and Newman, M. (2019), "Balance in Signed Networks," *Physical Review E*, 99(1), 012320.

Leskovec, J., Huttenlocher, D., and Kleinberg, J. (2010), "Signed Networks in Social Media," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 1361–1370.

Liu, T., Hammoudeh, S., and Santos, P. A. (2014), "Downside Risk and Portfolio Diversification in the Euro-Zone Equity Markets with Special Consideration of the Crisis Period," *Journal of International Money and Finance*, 44, 47–68.

Mantegna, R. N. (1999), "Hierarchical Structure in Financial Markets," *The European Physical Journal B-Condensed Matter and Complex Systems*, 11(1), 193–197.

Markose, S., Giansante, S., and Shaghaghi, A. R. (2012), ""Too Interconnected to Fail" Financial Network of US CDS Market: Topological Fragility and Systemic Risk," *Journal of Economic Behavior & Organization*, 83(3), 627–646.

Mei, D., Zeng, Q., Zhang, Y., and Hou, W. (2018), "Does US Economic Policy Uncertainty Matter for European Stock Markets Volatility?," *Physica A: Statistical Mechanics and its Applications*, 512, 215–221.

Shi, G., Altafini, C., and Baras, J. S. (2019), "Dynamics Over Signed Networks," *SIAM Review*, 61(2), 229–257.

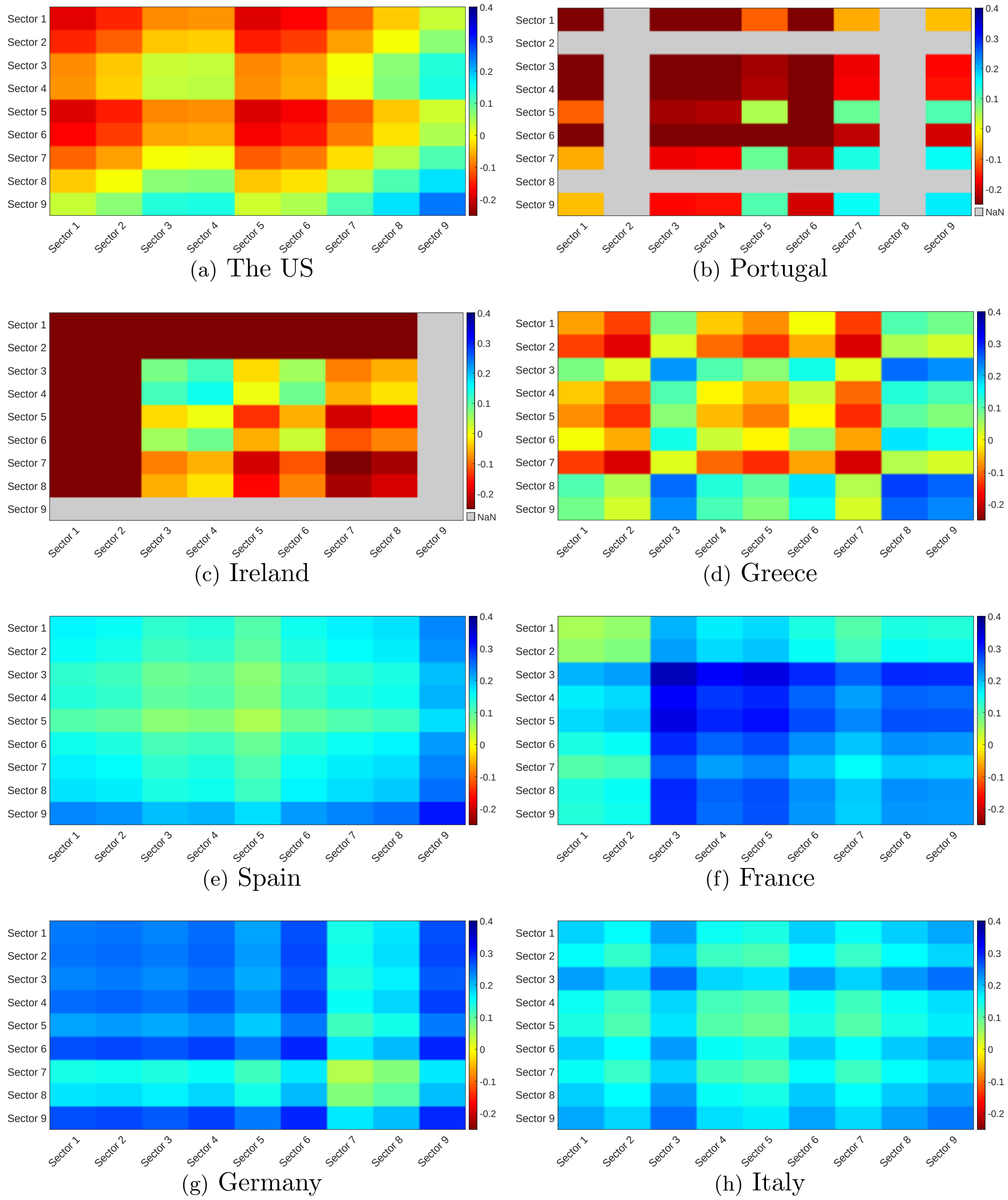
Stavroglou, S. K., Pantelous, A. A., Stanley, H. E., and Zuev, K. M. (2019), "Hidden Interactions in Financial Markets," *Proceedings of the National Academy of Sciences*, 116(22), 10646–10651.

Wang, G.-J., Xie, C., and Stanley, H. E. (2018), "Correlation Structure and Evolution of World Stock Markets: Evidence from Pearson and Partial Correlation-Based Networks," *Computational Economics*, 51(3), 607–635.

Zhao, L., Wang, G.-J., Wang, M., Bao, W., Li, W., and Stanley, H. E. (2018), "Stock Market as Temporal Network," *Physica A: Statistical Mechanics and its Applications*, 506, 1104–1112.

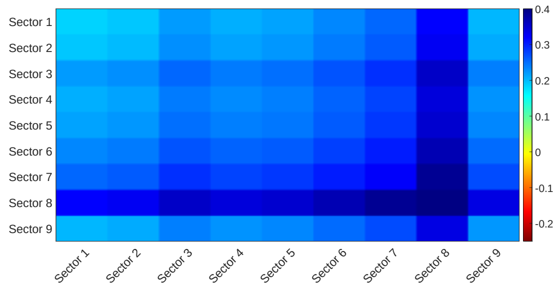
Appendix C

C.1 Kendall's tau colormaps



(a) The US (b) Portugal (c) Ireland (d) Greece (e) Spain (f) France (g) Germany (h) Italy

Figure C.1: Dependence colormap between companies classified by sector for each country.



(i) Japan

Figure C.1: Dependence colormap between companies classified by sector for each country.

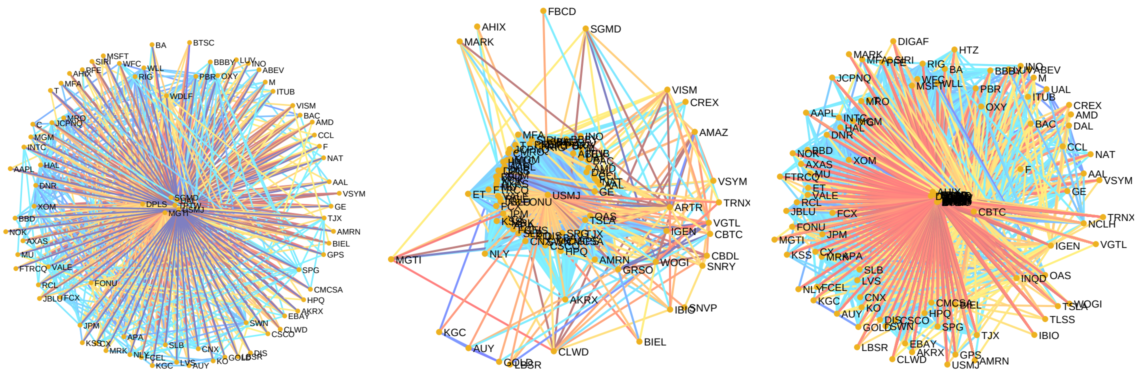
Figure C.1 contains the colormaps for the average unconditional Kendall's tau between the companies that belong to each pairwise sector combination. Kendall's tau estimates are represented by colors that go from dark red for the most negative values to blue for the most positive ones as indicated in the color range on the right of each panel. Panel columns/rows in grey indicate that the data set does not contain companies from at least one of the sectors involved. Sectors 1 to 9 are Technology, Health Care, Financial Services, Consumer Discretionary, Consumer Staples Producer Durables, Materials and Processing, Energy, and Utilities.

C.2 WSSNs by country

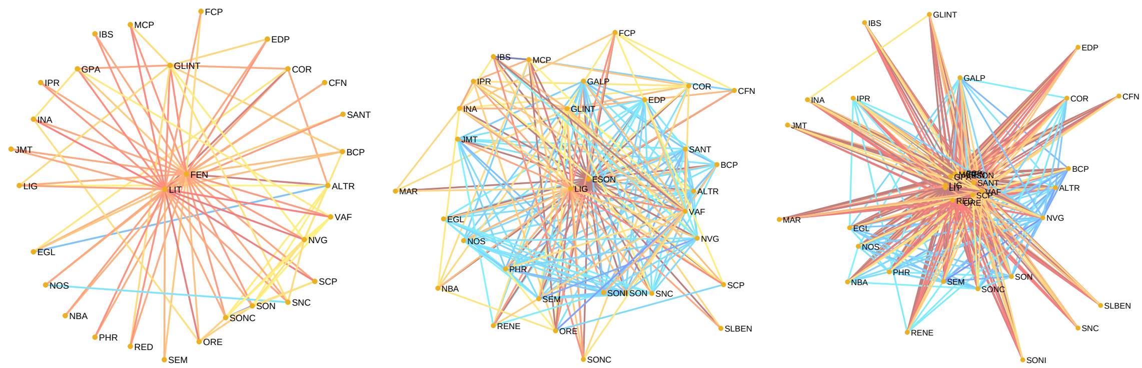
Figure C.2 illustrates the snapshots of three representative networks before, during, and after September 2011.

C.3 WSSNs by financial and non-financial sectors

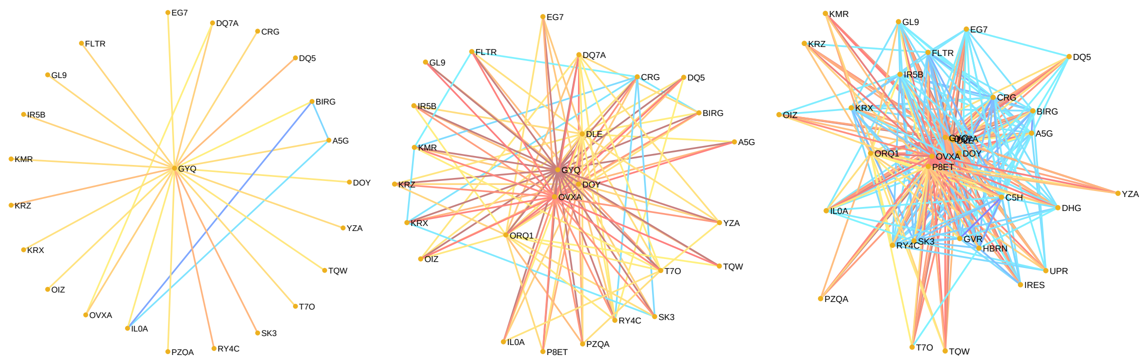
Figure C.3 and Figure C.4 show the evolution of the balance when we split the WSSNs by financial and nonfinancial sectors. Figure C.5 presents the balance evolution for the WSSNs with cross interactions between financial and nonfinancial sectors' stocks.



(a) The US

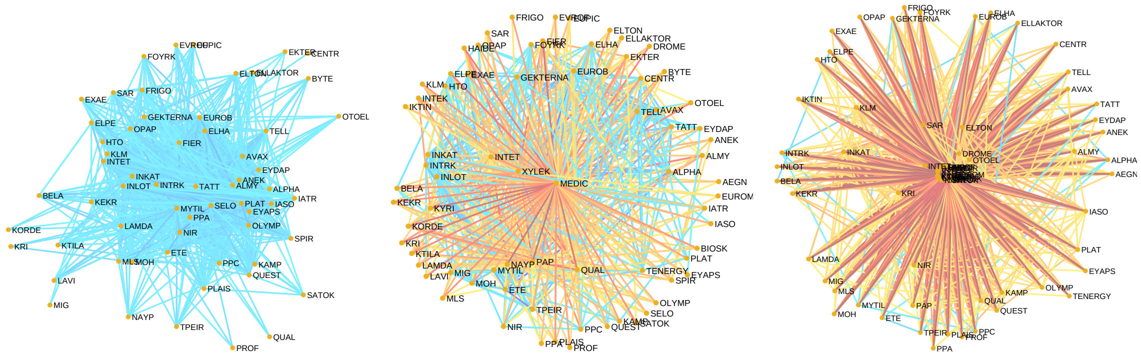


(b) Portugal

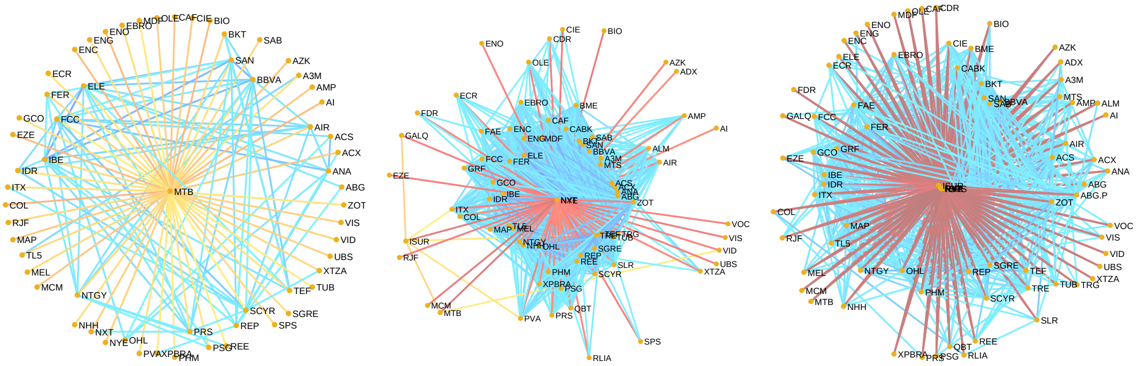


(c) Ireland

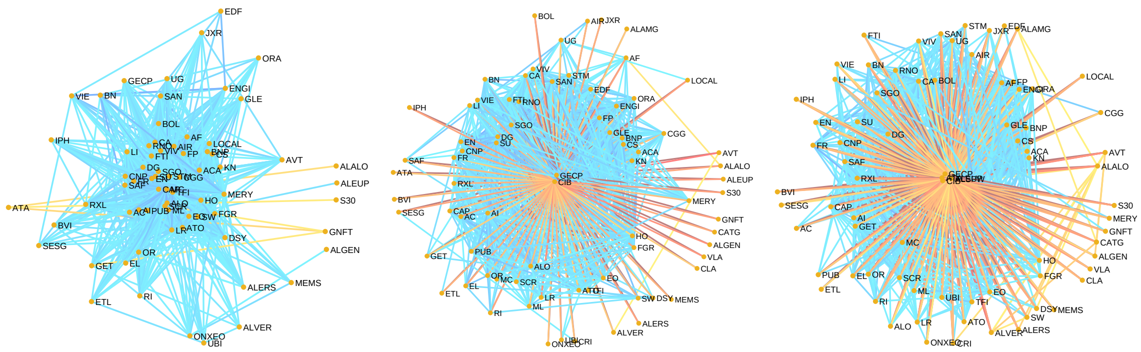
Figure C.2: Snapshots for three representative networks before, during, and after the BUT. The colors of the edges go from dark red for the most negative Kendall's tau estimates to blue for the most positive ones.



(d) Greece

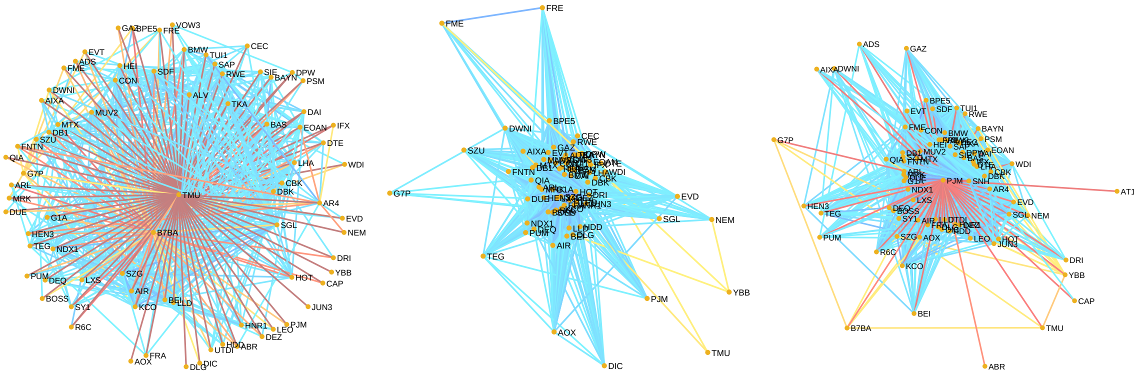


(e) Spain

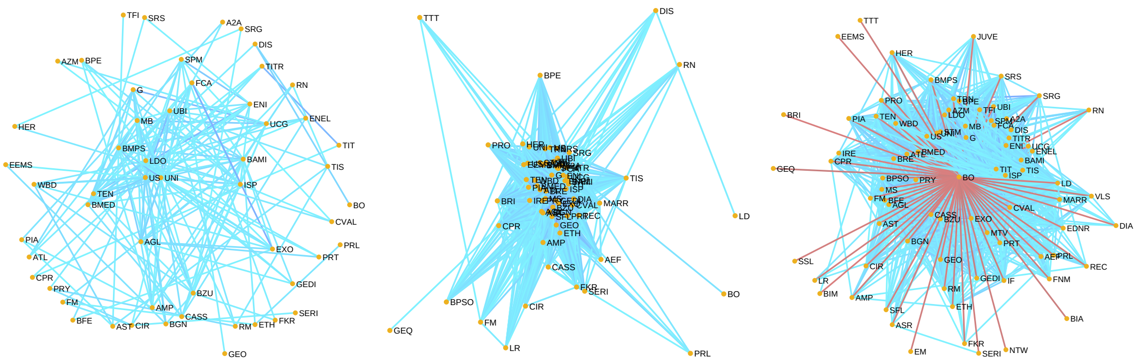


(f) France

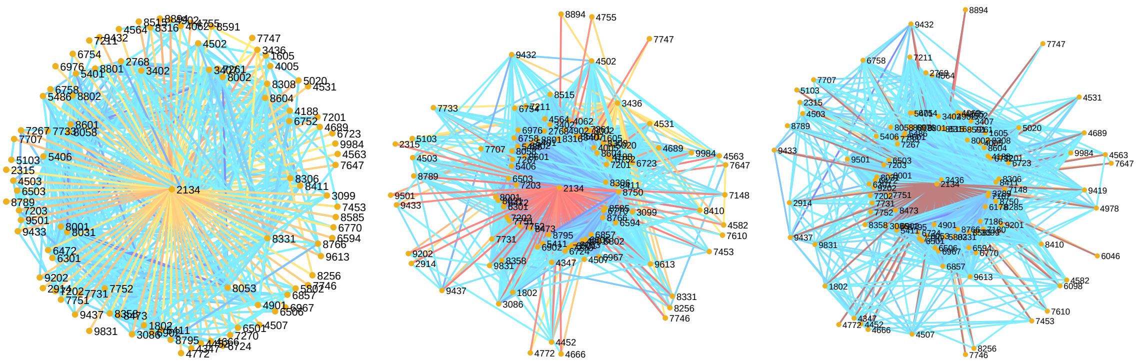
Figure C.2: Snapshots for three representative networks before, during, and after the BUT. The colors of the edges go from dark red for the most negative Kendall's tau estimates to blue for the most positive ones.



(g) Germany

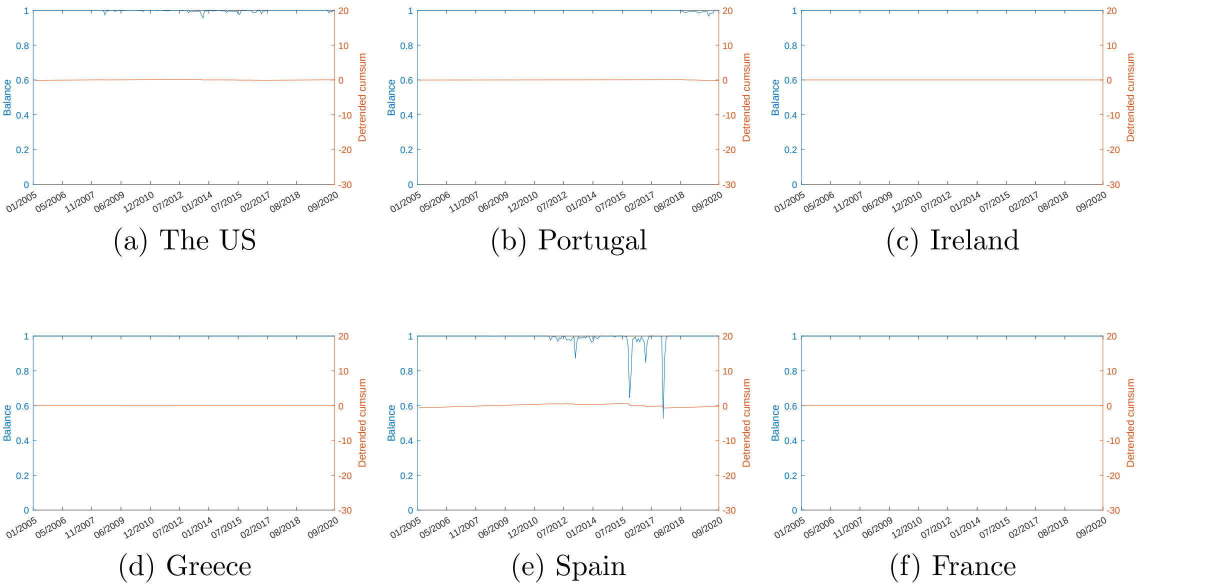


(h) Italy



(i) Japan

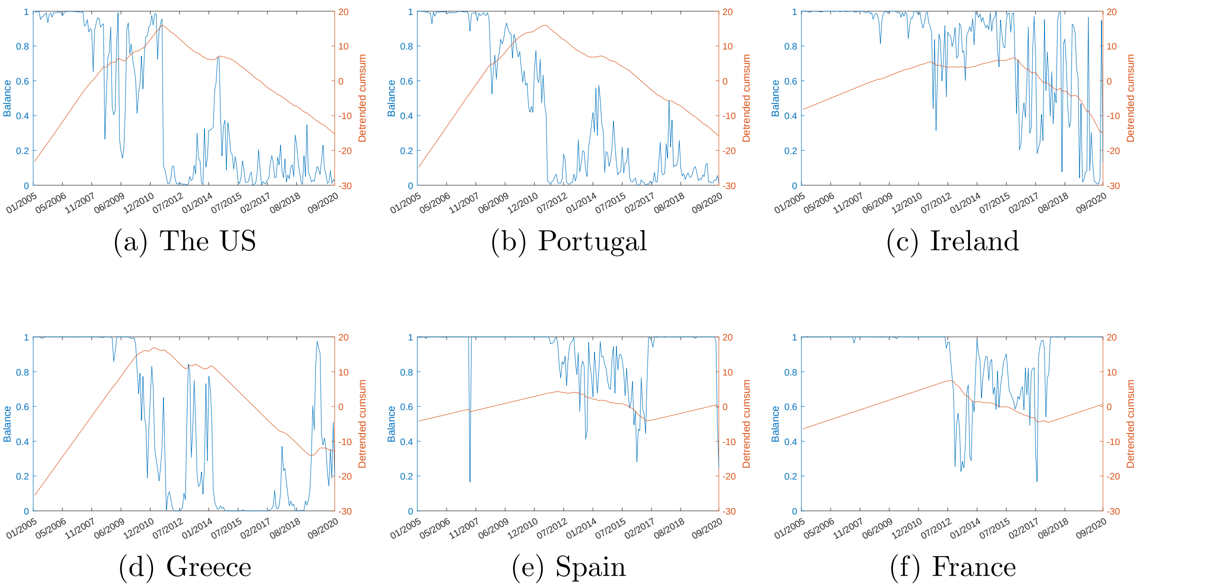
Figure C.2: Snapshots for three representative networks before, during, and after the BUT. The colors of the edges go from dark red for the most negative Kendall's tau estimates to blue for the most positive ones.



(a) The US (b) Portugal (c) Ireland

(d) Greece (e) Spain (f) France

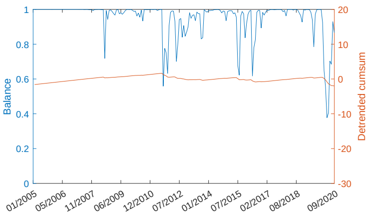
Figure C.3: Evolution of the balance for stocks in the financial sector between January 2005 and September 2020 in the WSSN (blue line) and of its DCS (red line).



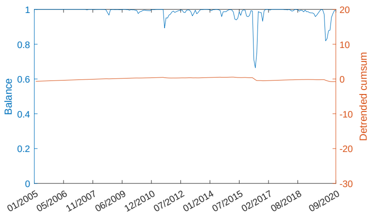
(a) The US (b) Portugal (c) Ireland

(d) Greece (e) Spain (f) France

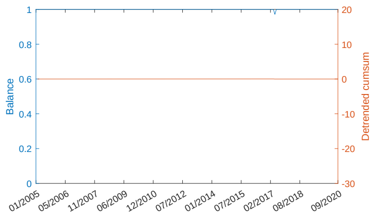
Figure C.4: Evolution of the balance for stocks in the non-financial sector between January 2005 and September 2020 in the WSSN (blue line) and of its DCS (red line).



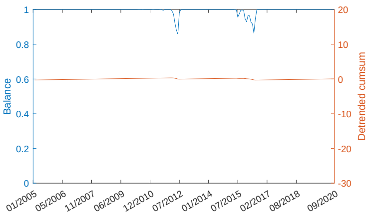
(a) The US



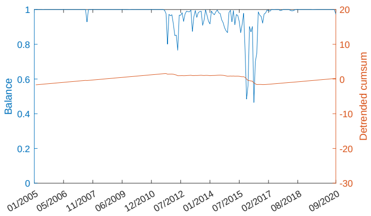
(b) Portugal



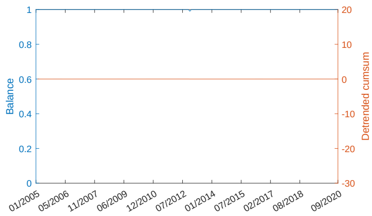
(c) Ireland



(d) Greece



(e) Spain



(f) France

Figure C.5: Evolution of the balance in the WSSN for interactions between financial and non-financial sectors' stocks from January 2005 until September 2020 (blue line), and of its DCS (red line).

Ondorioak

Tesi honetan aldagai anitzeko menpekotasun egiturak aztertzen dira. Aldagaien baterako egiturek informazio asko ematen dute; hala nola, menpekotasun gradua edo aldagai baldintzatzaile batek baterako egituretan duen eragina, beste batzuen artean. Zentzu honetan, hainbat aldagairen baterako banaketa modelizatzeko kopulak erabiliko ditugu.

Lehenengo kapituluan kopulen oinarrizko teoria azaltzen da. Lehenik eta behin, bi aldagaien kasurako emaitzak aurkezten dira eta bigarren zatian emaitza horien aldagai anitzetarako orokorpenak ematen dira. Aurkezten diren emaitza teorikoen artean, kopulen inguruko definizioak eta aldagai osagarrien eragina neurtzeko modua daude. Kapitulu honetan literaturan aurkitu daitezkeen kopula parametriko ezberdinak ere deskribatzen dira. Hala ere, tesiko hurrengo kapituluetan estimazio metodo ez-parametrikoak erabiltzen dira menpekotasun egiturak aztertzeko, forma funtzional zehatzik ez izateak malgutasun handiagoa ematen duelako. Gainera, zenbat eta aldagai gehiago izan, haien arteko erlazioak konplexuagoak bihurtzen dira eta zehaztapen parametrikoek analisia asko mugatzen dute zentzu honetan.

Askotan, aldagai osagarriek aldagai multzoko egituretan nola eragiten duten aztertzea interesatzen zaigu, estatistikaren arloan batez ere. Hori dela eta, bigarren kapituluan aldagai anitzeko menpekotasun egitura baldintzatuak modelizatzen dira berdinki banatutako aldagai independenteen testuinguruan definitutako estimatzaile ez-parametrikoak erabiliz. Menpekotasun anizkoitza neurtzeko hein korrelazioa erabiltzen da eta beste aldagai baten eragina Kendall-en tauaren estimatzaile baldintzatu batekin estimatzen dugu. Metodo ez-parametrikoek abantaila asko dituzten arren, leuntasun parametroaren aukeraketak menpekotasunaren estimazioa asko alboratu dezake. Hortaz, parametroa estimatzeko prozedura bat proposatzen dugu. Simulazio ikerketek prozedura balioztatzen duten arren, parametroarekiko Kendall-en tauaren sentsibilitatea handia dela erakusten dute.

Menpekotasun egituren azterketaren barnean, aldagai baldintzatzaileek murrizketa lineal esanguratsurik eragiten duten interesatzen zaigu. Horretarako, Wald motako estatistiko bat proposatzen dugu, zeinak murrizketa anitzen kontrasteak onartzen dituen; hala nola, baldintzatutako menpekotasun konstantea, tau ezberdinen arteko murrizketak, nahiz populazio ezberdinetako hein korrelazio baldintzatuen konparaketak. Simulazio ikerketa batean, proposatutako estatistikoa jadanik literaturan aurkitu daitekeen beste estatistiko batekin jartzen dugu lehian. Wald motako estatistikoa, prozedura bien konparaketa posible den kasuetan gailendu egiten da, egitura konplexuak dituzten kasuetan barne. Gainera, kostu konputazionalari dagokionez, gure proposamenaren abantaila oso esanguratsua da. Kontsideratutako gainerako kasuetan ere proposamenak emaitza onak erakusten ditu.

Aplikazio praktikoan, 2019. urteko RCI indizea osatzen duten oinarrien menpekotasun anizkoitza aztertzen da, konkretuki, Europa mailan gizarteak erakundeen kalitatearen inguruan duen pertzepzioak efizientzia eta berrikuntza taldeetako oinarrietan duen eragina. Espero genuen bezala, oinarrien artean erlazio positiboa aurkitzen dugu. Gainera, berrikuntzaren eta derrigorrezko bigarren hezkuntza mailatik gorako hezkuntzaren arteko erlazioa erakundeen kalitatearen okerragotzearekin ahuldu egiten dela ikusten da. Etorkizunari begira, eskualdeetako erakundeen kalitateak beren giza kapitala berrikuntza emaitzetara bideratzeko duten gaitasunean eraginik duen ikertzea interesgarria izango litzatekeela uste dugu.

Ekonomia eta finantzetan aldagaiek denborarekiko menpekotasuna dute. Gainera, azken finantza-krisiaren ondorioz finantzetan ematen diren menpekotasun egituren ikerketak asko aregotu dira. Hirugarren kapituluan finantza aldagaien menpekotasun egituretan eta hauek denboran zehar duten bilakaeran zentratzen gara. Hori dela eta, denborarekiko aldakortasuna neurtzen duten kopularen eta Kendall-en tauaren estimatzaile ez-parametrikoak proposatu eta dagozkien propietate asintotikoak garatzen ditugu inguruneka egonkorrak diren aldagaietarako.

Simulazio ikerketen bitartez, Kendall-en taurako proposatu den estimatzaileak eredu ezberdinetan duen portaera aztertzen dugu eta ezaguna den DCC korrelaziotik eratorritako tauak eredu horietan duen portaerarekin konparatzen dugu. Lortzen diren emaitzen arabera, Kendall-en tauaren estimatzaileak portaera hobea du eta hobekuntza are nabariagoa da egitura konplexuagoak zehazten dituzten ereduetarako.

Ondorioz, denborarekin aldakorra den Kendall-en tauaren estimatzailea Europako bost indizeren errentagarritasunek duten menpekotasuna aztertzeko erabiltzen dugu. Menpekotasuna estimatzeko erabilitako prozedura guztiek antzerako ezaugarriak nabarmentzen dituzten arren, aldagai anitzeko Kendall-en tauak menpekotasunak duen tendentziaren seinale garbiago bat ematen du. Horien artean, irailak 11ko dorre bikien erasoa, mundu mailako finantza krisia edo oraintsu gertatutako COVID-19-aren krisia bezalako gertaerak nabarmentzen dira. Bereziki aipagarria da COVID-19-aren krisiak indizeen baterako menpekotasunean eragin duen gorakada.

Aurreko emaitzak kontuan izanik, menpekotasun maila altuek merkatuaren beheranzko joeretan nola eragiten duten aztertzen dugu $\Delta CoVaR$ arrisku neurriaren bitartez. Horretarako, Kendall-en taurako proposatutako estimatzailea indizeen baterako mugimenduen adierazle gisa erabiltzen dugu. Emaitzek arriskuak denboran zehar aldaketa nabarmena duela erakusten dute. Izan ere, iraganeko mepekotasunak merkatuko errentagarritasunaren mutur arriskutsuetan duen eraginaren ebidentzia aurkitzen dugu.

Proposatutako menpekotasun neurriak aldagai makroekonomikoekiko duen erlazioa ikustea etorkizunerako aztergai interesgarri bat litzateke. Testuinguru honetan, arrisku sistemikoarekin lotura estua duela ikusi den politika ekonomikoaren ziurgabetasun (EPU) indizean pentsatu dugu aldagai makroekonomiko posible bezala. Menpekotasun adierazlearekin bezala, politika ziurgabetasunak Euro Stoxx-aren kuantil arriskutsuetan izan dezakeen kausalitate efektua aztertu dugu eta azken honek eragin esanguratsua duela aurkitu dugu. Ildo honetatik, uste dugu garrantzitsua dela, beste adierazle batzuen artean, politika ziurgabetasunak finantza merkatuak nola zehazten dituen ikertzea eta erregulazio politikari begira, arrisku iturri bezala duen eragina kuantifikatzea.

Hirugarren kapituluan ikusi dugu Europako indize nazional garrantzitsuenek antzerako bilakaera dutela denboran zehar. Herrialde bakoitzeko merkatuek aldiz, portaera oso desberdinak izaten dituzte. Bestalde, azken urteetan indarra hartu duen arrisku sistemikoak finantza merkatuetako portaeraren ardatzetako bat menpekotasuna dela dio. Hori dela eta, akzioek sortzen dituzten egitura topologikoetan jarri nahi dugu fokua.

Finantza merkatuak aztertzeko modu bat egitura topologikoak sare edo grafoen bitartez adieraztea da. Beraz, laugarren kapituluan sareak eraiki ahal izateko grafo teoriari buruzko oinarrizko kontzeptu eta emaitzak azaltzen dira. Garrantzitsuak dira bereziki auzokidetasuna eta auzokidetasun matrizea, hauek akzioen arteko erlazioak matematikoki adierazteko aukera ematen dutelako. Sareei lotuta, nahiko berritsua den eta sistemaren egonkortasunarekin lotuta dagoen orekaren ideia aurkezten da. Gainera, sareen izaera nolakoa den ikusteko erabiltzen diren konparazio eredu batzuk ere aipatzen dira.

Bosgarren kapituluan munduko bederatzi finantza merkatutako akzio sareak eraikitzen ditugu aurreko kapituluan aurkeztutako tresnez baliatuz. Hautatutako herrialdeen artean GIIPS bezala ezagutzen direnez gain, Alemania, Frantzia, EEBB eta Japonia hartzen dira Europako eta mundu mailako ekonomia indartsuen ordezkari gisa. Akzioen arteko loturak zehaztatzeko hein korrelazioak erabiltzen ditugu. Bestalde, akzio sareek denboran zehar jasotzen dituzten aldaketak aztertzea garrantzitsua dela iruditzen zaigu, batez ere finantza, zor publikoa eta COVID-19-aren krisi garaietan. Horregatik, sareek 2005etik 2020ko irailera arte duten bilakaera aztertzen dugu hirugarren kapituluan proposatutako Kendall-en taua erabiliz. Gainera, aurreko emaitzak ikusita, akzioen merkatuek kanpo faktoreen eragina izan dezaketela pentsa dezakegu. Hortaz, akzio sareak eraikitzerakoan politika ziurgabetasunaren efektua kontuan hartu dugu aldagai makroekonomiko bezala.

Zeinuetako sareak eraikitzeak orekaren kontzeptua finantza sareetan lehen aldiz erabiltzeko aukera ematen du. Zentzu honetan, oreka maila neurtzen duen koefiziente bat erabiltzen da, merkatuen egongortasunaren adierazle bezala interpretatu daitekena. Koefiziente honetarako lortzen diran emaitzek orain arte aurkitu gabeko orekaren galera nabarmen bat erakusten dute aztertutako bost herrialdetan 2011. urteko iraila inguruan. Gertaera hau urte horretako abuztuko *astelehen beltza* eta gero ematen da. Alemania, Italia eta Japoniako merkatuek berriz oreka nahiko konstantea dute denboran zehar eta ez da halako trantsiziorik hautematen.

Oreka galeraren eragileak aztertu nahian, egiturei eragin diezaketen faktore ezberdinak kontsideratu dira eta eragile nagusia akzio multzo txiki bat dela ondorioztatu da. Izan ere, multzo horretako akzioek azpi-sare oso desorekatua sortzen dute, sare osoaren desoreka ekartzen duena. Azpi-sare horien ikerketa sakonago batek desoreka sortzen duten akzioak enpresa txikiak direla erakusten du, gehien bat finantzazkoak ez diren sektoreetatik datozenak.

Emaitza hauek hirugarren kapituluko emaitzekin bat egiten dute menpekotasunaren garrantziari dagokionez, akzio gutxi batzuek dituzten erlazioak baitira akzio saree osoaren egonkortasuna apurtzen dutenak. Honela, finantza merkatuen hauskortasuna agerian geratzen da, batez ere modu estuan konektatuta dauden akzioek izan ditzaketen estualdien kutsadurarekin zerikusia duena. Egindako ikerketa hasiera bat besterik ez da, oraindik azterketa gehiago behar dira azpi-sare horiek nola sortzen diren ikusteko eta, arrisku sistemikoari begira, egitura topologikoaren berantolaketa honek izan ditzakeen ondorioak aztertzeko. Uste dugu emaitza hauek estaldura eta akzio sorroen aukeraketa estrategiak hobetzeko bidean aurrera pauso txiki bat ematen dutela.

Laburbilduz, tesi honek aldagai anitzeko menpekotasun egiturak hobeto ulertzeko tresnak lantzen ditu. Lehenengo kapituluetan aldagai osagarri bati baldintzatutako menpekotasun egiturak aztertu eta aldagai osagarriaren balio ezberdinek duten eragina ikusteko estatistiko bat proposatzen dugu. Hortik aurrera, menpekotasun anizkoitzaren denborarekiko bilakaera ikertzeko tresnak proposatzen ditugu eta praktikan, bereziki finantzen arloan, duten erabilgarritasuna aurkezten dugu. Aplikazio praktikoak dira erregulatzaile, inbertitzaile, nahiz finantza erakundeetako profesionalentzat emaitza teoriko nahiz praktikoen erakargarritasuna erakusten dutenak. Interesgarria izango litzateke bigarren kapituluan proposatutako estatistiko motak askeak ez diren aldagaien testuinguruan lantzea finantza merkatuen azterketetan erabili ahal izateko. Bestalde, lehenengo eta laugarren kapituluei dagokienez, aipatu beharra dago literaturan ez dugula kopula ezta sare teoriari buruzko halako euskarazko materialik aurkitu. Zentzu honetan beraz, bi kapitulu hauek goi-mailako ikasketetarako euskarazko eskuliburu bezala erabili ahalko litzatekez.

eman ta zabal zazu

Universidad del País Vasco

Euskal Herriko Unibertsitatea