

Prognoziranje realizirane volatilnosti: empirijski nalazi referentnog europskog dioničkog indeksa

JOSIP ARNERIĆ

Ekonomski fakultet Sveučilišta u Zagrebu

Trg J. F. Kennedyja 6, 10000 Zagreb

Hrvatska

jarneric@net.efzg.hr

ORCID iD: <https://orcid.org/0000-0002-2901-2609>

Izvorni znanstveni rad / *Original scientific paper*

UDK / UDC: 336.76:519.2=163.42

Primljeno / Received: 23. siječnja 2026. / January 23rd, 2026.

Prihvaćeno za objavu / Accepted for publishing: 23. ožujka 2026. / March 23rd, 2026.

DOI: 10.15291/oc.5015

Sažetak: Realizirana volatilnost temelji se na intradnevnom visokofrekventnim prinosima i standardna je mjera stvarne, premda nepoznate, integrirane volatilnosti financijske imovine. Iako je realizirana volatilnost determinirana ex-post, u posljednjem desetljeću razvijeno je nekoliko modela za njezino predviđanje, koji se razlikuju prema svojstvima koja nastoje obuhvatiti, kao što su duga memorija, heteroskedastičnost, cjenovni skokovi, asimetrična reakcija na pozitivne i negativne šokove te mikro strukturni šum, kao i prema tome modeliraju li realiziranu volatilnost izravno ili posredno. Uspješnost takvih modela još je uvijek nedovoljno istražena, osobito na europskim tržištima kapitala, za razliku od američkih tržišta koja su zastupljenija u dosadašnjim studijama. Stoga je cilj ovoga rada usporediti uspješnost odabranih modela u prognoziranju realizirane volatilnosti indeksa DAX, koji se smatra referentnim europskim dioničkim indeksom. U tu svrhu uspoređuju se HAR, MEM, HEAVY i realizirani GARCH modeli ne samo radi utvrđivanja modela s najvećom prognostičkom točnošću, već i radi ispitivanja ovisi li njihova učinkovitost o frekvenciji uzorkovanja i izboru realizirane mjere koja se prognozira. Time rad doprinosi literaturi u kojoj ne postoji konsenzus o najprikladnijim modelima za prognoziranje pojedinih realiziranih mjera volatilnosti. Empirijski nalazi pružaju implikacije za upravljanje rizicima i vrednovanje financijskih instrumenata na europskim tržištima kapitala, osobito u razdoblju neizvjesnosti obuhvaćenom analizom. Istraživanje se temelji na jednominutnim zaključnim cijenama indeksa DAX, dok se prognostička točnost vrednuje primjenom više kriterija.

Ključne riječi: realizirana volatilnost, prognostički modeli, referentni dionički indeks

JEL klasifikacija: C53, C58, D53, G17

1. Uvod

Implementacijom visokofrekventnih podataka, opaženih u vrlo kratkim vremenskim intervalima, u postojeće modele znatno je poboljšana preciznost procjene i prognoze volatilnosti financijske imovine. U empirijskim istraživanjima najzastupljeniji je GARCH model (engl. *Generalized Autoregressive Conditional Heteroskedasticity*) za prognozu dnevne volatilnosti na temelju niskofrekventnih zaključnih cijena te se, zbog svoje superiornosti, često koristi kao referentni model u usporednim analizama

(Hansen i Lunde, 2005). Riječ je o parametarskom modelu uvjetne volatilnosti (engl. *conditional volatility*), koja je vremenski promjenjiva i uvjetovana informacijama iz prošlosti (Bollerslev, 1986). Osnovna ideja GARCH modela je da trenutna varijanca prinosa ovisi o varijanci prinosa iz prethodnog razdoblja, čime se opisuje autoregresijsko ponašanje volatilnosti. Zbog takve dinamike ona se može jednostavno prognozirati. Razvoj GARCH metodologije izravno je povezan s Robertom Engleom, koji je 1982. predstavio ARCH model (Engle, 1982), preteču današnjih GARCH modela te je za svoj pionirski doprinos 2003. godine nagrađen Nobelovom nagradom za ekonomiju. GARCH modeli od tada su postali standardni alat, ne samo u prognoziranju i upravljanju tržišnim rizikom, već i u optimizaciji portfelja, vrednovanju financijskih izvedenica, modeliranju učinaka zaraze i prelijevanja volatilnosti te u utvrđivanju svojstava različitih vrsta imovine, poput sigurnog utočišta ili diverzifikatora. S vremenom su formulirane brojne specifikacije GARCH modela, od asimetričnih (EGARCH, APARCH, TGARCH, GJR-GARCH i dr.) do multivarijatnih (CCC, DCC, BEKK, DVEC i dr.), s ciljem opisivanja dinamike volatilnosti u različitim tržišnim uvjetima, ne ograničavajući se samo na jednu imovinu.

Uz parametarske modele, koji djeluju *ex-ante*, razvijeni su i neparametarski modeli volatilnosti, no njihova je prognostička moć ograničena jer zbog svoje konstrukcije djeluju *ex-post*. Realizirana volatilnost, odnosno realizirana varijanca RV (engl. *realized variance*), neparametarska je mjera dnevne integrirane volatilnosti utemeljena na intradnevnom opažanjima (Andersen et al., 2001) i kao takva ne modelira izravno svoju buduću dinamiku. Stoga se nameće istraživačko pitanje: Može li se realizirana volatilnost, iako je *ex-post* determinirana, uspješno prognozirati? Odgovor na to pitanje nije neposredan jer u radu s visokofrekventnim podacima potrebnima za procjenu RV postoji niz izazova koje treba svladati, kao što su mikrostrukturni šum (engl. *microstructure noise*), cjenovni skokovi i nesinkroniziranost opažanja. Prema Arnerić i Matković (2019), visokofrekventni podaci često su kontaminirani mikrostrukturnim šumom, koji stvara prividnu volatilnost zbog oscilacija između kupovne (engl. *bid*) i prodajne (engl. *ask*) cijene, nesinkronog trgovanja, cjenovnih skokova i niske likvidnosti, osobito na izranjajućim tržištima (engl. *emerging markets*). Te anomalije precjenjuju RV i narušavaju učinkovitost prognostičkih modela. Autori nadalje ističu važnost filtriranja visokofrekventnih podataka, optimalnog odabira frekvencije uzorkovanja i uporabe realiziranih mjera robusnih na mikrostrukturni šum i cjenovne skokove. Pronalazak „najbolje“ realizirane mjere, ne nužno realizirane varijance RV, preduvjet je za njezino modeliranje i prognoziranje. Prvi iskorak u tom smjeru dao je Corsi (2009), koji je uveo heterogeni autoregresijski model realizirane volatilnosti HAR-RV (engl. *Heterogeneous Autoregressive Realized Volatility*) i pokazao da se dinamika volatilnosti može opisati linearnom kombinacijom prošlih vrijednosti na dnevnom, tjednom i mjesečnom horizontu, prema pretpostavci heterogene strukture tržišta (različite skupine tržišnih sudionika reagiraju na informacije u različitim vremenskim horizontima). Nadalje, razvijen je model multiplikativne pogreške MEM (engl. *Multiplicative Error Model*), koji omogućuje fleksibilnije modeliranje volatilnosti polazeći od umnoška uvjetne sredine realizirane mjere i slučajne komponente koja poprima isključivo nenegativne vrijednosti s očekivanjem jednakim jedan (Engle, 2002; Engle i Gallo, 2006). Na njega se nadovezuje HEAVY model (engl. *High-frequency-based Volatility*) Shepharda i Sheparda (2010) u kojem se uvjetna varijanca i uvjetno očekivanje realizirane mjere zasebno modeliraju kao latentni, ali povezani dinamički procesi. Budući da HEAVY model smanjuje perzistentnost volatilnosti, implicira da je duga memorija GARCH modela posljedica zanemarivanja visokofrekventnih informacija (Nguyen et al., 2024). Primjene HEAVY modela doživjele su i multivarijatna proširenja (Noureldin et al., 2012; Shepard i Xu, 2019; Bauwens i Xu, 2023; Girardi i Caporin, 2025), analizirajući kovarijance prinosa više imovina istodobno.

Sljedeći ključni iskorak predstavili su Hansen i suradnici (2011) realiziranim GARCH modelom (engl. *Realized GARCH*), koji povezuje dnevne prinose, uvjetnu varijancu i realiziranu mjeru volatilnosti. Za razliku od standardnog GARCH modela, realizirani GARCH model uključuje i visokofrekventne podatke kroz realiziranu mjeru volatilnosti u dodatnoj mjernoj jednadžbi (engl. *measurement equation*). Empirijske studije pokazuju da realizirani GARCH sustavno nadmašuje standardne GARCH modele u predviđanju volatilnosti, osobito na kratkim horizontima gdje je realizirana volatilnost najinformativnija (Sharma i Vipul, 2016). Model su dodatno proširili Hansen i Huang (2016) realiziranim EGARCH modelom, što omogućuje uključivanje asimetričnih volatilnih reakcija i više realiziranih mjera (Naimoli et al., 2022). Paralelna istraživačka struja nastoji postići slične ciljeve kombiniranjem postojećih modela, pri čemu su Huang i suradnici (2016) predložili realizirani HAR-GARCH model, koji spaja HAR komponente duge memorije s GARCH modelom.

Dosadašnja istraživanja ponajprije su usmjerena na razvijena američka tržišta kapitala, dok su europska slabije zastupljena i nisu sustavno analizirana. Nedovoljno je istraženo koliko su modeli koji integriraju visokofrekventne podatke doista uspješni u prognoziranju realizirane volatilnosti na europskim tržištima unatoč njihovu progresivnom razvoju. Stoga se nameće potreba za popunjavanjem tog istraživačkog jaza evaluacijom i usporedbom različitih prognostičkih modela realizirane volatilnosti, što je u radu i učinjeno na primjeru njemačkog burzovnog indeksa DAX, jednog od najlikvidnijih predstavnika europskih dioničkih tržišta zbog visoke tržišne kapitalizacije i intenzivne trgovinske aktivnosti. U tu se svrhu uspoređuju HAR, MEM, HEAVY i realizirani GARCH modeli primjenom RMSE, MAPE i QLIKE kriterija dok se statistička značajnost razlika u prognostičkim sposobnostima formalno testira Diebold-Mariano testom. Nadalje, ostaje otvoreno pitanje ovisi li i u kojoj mjeri uspješnost odabranih modela o izboru realizirane mjere volatilnosti i frekvenciji uzorkovanja. Rasvjetljavanje tog pitanja dodatni je znanstveni doprinos, s važnim implikacijama za institucionalne investitore i praktičare u financijskom sektoru, jer pruža smjernice za učinkovitije modeliranje tržišnog rizika. Rezultati istraživanja korisni su i nositeljima ekonomskih politika u oblikovanju makroprudencijalnih mjera usmjerenih na očuvanje financijske stabilnosti, jer omogućuju rano prepoznavanje početka razdoblja povećane volatilnosti. Tržišni rizik, utemeljen na budućoj realiziranoj volatilnosti, izravno utječe i na minimalne kapitalne zahtjeve koje regulatorna tijela propisuju financijskim institucijama.

Ovaj članak, uz uvodni dio, obuhvaća četiri dodatna poglavlja. U drugom poglavlju analizira se relevantna literatura i ističu otvorena istraživačka pitanja. U trećem poglavlju opisuju se korišteni podaci i metode, nakon čega slijedi empirijska analiza i rasprava o rezultatima u četvrtom poglavlju. Rad završava petim poglavljem, koje donosi glavne zaključke, ograničenja i smjernice za buduća istraživanja.

2. Pregled literature i istraživački jaz

U suvremenoj literaturi razvijeno je mnogo pristupa koji nastoje unaprijediti modeliranje i prognoziranje volatilnosti uključivanjem visokofrekventnih podataka. Najjednostavniji među njima je GARCH-X, koji proširuje standardni GARCH model dodavanjem realizirane mjere volatilnosti kao egzogene varijable u jednadžbu uvjetne varijance. Međutim, GARCH-X model ne prognozira realiziranu volatilnost, već uvjetnu volatilnost, iako se koristi informacijama iz realizirane mjere (Engle, 2002). Mnogi radovi uspoređuju učinkovitost realiziranih GARCH modela i njihovih varijanti sa standardnim GARCH modelima, dok drugi uspoređuju alternativne modele, poput HEAVY i MEM modela. Sharma i Vipul (2016) analiziraju više razvijenih tržišta i tržišta u razvoju (16 međunarodnih dioničkih indeksa) te pronalaze da realizirani GARCH nadmašuje standardne GARCH modele u kratkoročnim

prognozama i procjeni rizične vrijednosti VaR (engl. *Value-at-Risk*), osobito u razdobljima povišene volatilnosti. Do sličnog zaključka dolaze Frömmel i suradnici (2014) pri modeliranju dnevne volatilnosti spot cijena električne energije od veljače 2005. do travnja 2013. te Xiao i suradnici (2023) primjenom frakcijski integriranog realiziranog GARCH modela, tzv. FIGARCH-RV. Watanabe (2012) dodatno pokazuje da uključivanje realiziranih mjera poboljšava kvantilne prognoze prinosa i procjenu repnog rizika (engl. *tail risk*) S&P500 indeksa, pri čemu realizirani GARCH ostvaruje stabilnije rezultate u kombinaciji s pristupima temeljenima na teoriji ekstremnih vrijednosti EVT (engl. *Extreme Value Theory*).

Važan dio literature bavi se pitanjem složenosti modela i robusnosti prognoza. Xie i Yu (2020) empirijski pokazuju da su mnoge proširene specifikacije realiziranog GARCH modela prekomjerno parametrizirane te da jednostavnije varijante pružaju jednake ili čak bolje rezultate izvan uzorka (engl. *out-of-sample*). Autori uspoređuju tri varijante realiziranog GARCH modela na podacima 28 pojedinačnih dionica i zaključuju da parsimonijski modeli imaju stabilnije procjene i manje su osjetljivi na mikrostrukturni šum. Novija literatura produbljuje ovu spoznaju analizom informacijskog sadržaja različitih realiziranih mjera. Xu i suradnici (2024) pokazuju da realizirane mjere hvataju različite komponente volatilnosti, uključujući kontinuiranu komponentu i prekonocne skokove (engl. *overnight jumps*) te upućuju na zaključak da je selektivno uključivanje informativnijih mjera korisnije od neselektivnog agregiranja više njih. Time se dodatno potvrđuje važnost parsimonije i promišljenog odabira realizirane mjere. Wu i suradnici (2024) naglašavaju da realizirani GARCH može korigirati pristranost realiziranih mjera, što su dokumentirali na primjeru dva glavna indeksa kineskog tržišta dionica (kompozitni indeks šangajske i šenženske burze). Fang i Han (2025) primjenjuju realizirani GARCH model u kontekstu vrednovanja opcija, koristeći podatke o američkim tržišnim indeksima S&P500, NASDAQ100 i Dow Jones Industrial Average (DJIA) te nalaze da uključivanje realizirane mjere smanjuje pogreške u određivanju cijena opcija u odnosu na standardne GARCH i BS modele (engl. *Black and Scholes*). Paralelno se primjenjuju i podatkovno utemeljeni pristupi za prognoziranje realizirane volatilnosti, poput neuronskih mreža, pri čemu Arnerić i suradnici (2018) pokazuju da jednosmjerne neuronske mreže FNN (engl. *Feedforward Neural Networks*) konstruirane kao nelinearna proširenja standardnih HAR-RV modela, uključujući komponente skokova i asimetrije, postižu bolju prilagodbu podacima i veću prognostičku točnost na temelju MSE i MAE kriterija, analizirajući njemački DAX.

Za razliku od realiziranih GARCH modela, HEAVY modeli polaze od pretpostavke da realizirane mjere izravno upravljaju dinamikom uvjetne volatilnosti i kvadriranih prinosa, čime se izbjegava potreba za zasebnom mjernom jednadžbom. Temeljni okvir HEAVY modela razvili su Shephard i Sheppard (2010), koji pokazuju da korištenje realizirane varijance dobivene iz intradnevni podataka značajno poboljšava prognoze volatilnosti u odnosu na klasične GARCH modele, što je potvrđeno analizom 34 financijska instrumenta u razdoblju od 1996. do 2009. godine. Noureldin i suradnici (2012) proširuju taj pristup uvođenjem multivarijantnih HEAVY modela koji omogućuju modeliranje dinamike realizirane kovarijance i realizirane korelacije, uz znatno smanjene dimenzionalne zahtjeve naspram standardnih multivarijantnih GARCH specifikacija (MGARCH). Nadalje, Sheppard i Xu (2019) uvode faktorske HEAVY modele, čime omogućuju parsimonijsko modeliranje volatilnosti velikog broja financijskih instrumenata putem zajedničkih latentnih faktora. Njihov uzorak obuhvaća 93 likvidne dionice američkog S&P100 indeksa u razdoblju od siječnja 2000. do prosinca 2014. godine, a realizirane mjere volatilnosti izračunavaju se iz 5-minutnih intradnevni prinosa. Konačno, Bauwens i Xu (2023) integriraju HEAVY pristup sa strukturama dinamičke korelacije, razvijajući DCC-HEAVY i DECO-

HEAVY modele te empirijski pokazuju njihovu superiornu prognostičku sposobnost na primjeru 29 dionica, sastavnica DJIA indeksa, u razdoblju koje uključuje dot-com krizu i globalnu financijsku krizu. HEAVY modeli često se povezuju s MEM modelima s obzirom na to da oba pristupa polaze od multiplikativne strukture volatilitnosti, iako se u potonjima realizirane mjere tretiraju latentnim varijablama. S jedne strane MEM modeli primjenjuju se za proširenje GARCH-X specifikacija, a s druge strane za pojednostavljenje realiziranog GARCH modela, čime se reducira parametarska složenost i poboljšava prognostička učinkovitost (Nguyen et al., 2024). MEM modeli posebno su prikladni za kratkoročno i srednjoročno prognoziranje volatilitnosti te se mogu lako integrirati s drugim modelima, poput MIDAS (engl. *Mixed Data Sampling*) modela (Amendola et al., 2024). Engle i Gallo (2006) provode empirijsku analizu koristeći intradnevne podatke visoke frekvencije za S&P500 indeks i devizni tečaj USD/DEM, s 5-minutnim prinosima u razdoblju od 2. siječnja 1996. do 31. prosinca 2002. Autori konstruiraju više realiziranih mjera volatilitnosti i ugrađuju ih u MEM model, pokazujući da se njihovom kombinacijom značajno poboljšava procjena latentne dnevne volatilitnosti. Nguyen i suradnici (2024) koriste američke dionice iz indeksa S&P500 u razdoblju od 1. siječnja 2000. do 31. prosinca 2020., uspoređuju klasične GARCH-X i MEM modele s njihovim proširenjima temeljenima na dubokom učenju (engl. *deep learning*) te zaključuju da takva proširenja postižu točniju kratkoročnu i srednjoročnu prognozu. Amendola i suradnici (2024) empirijski primjenjuju dvostruke MEM modele, analizirajući realiziranu volatilitnost S&P500 indeksa, a rezultati impliciraju da razdvajanje volatilitnosti na dugoročnu i kratkoročnu komponentu dovodi do boljih prognoza naspram standardnih MEM i GARCH specifikacija.

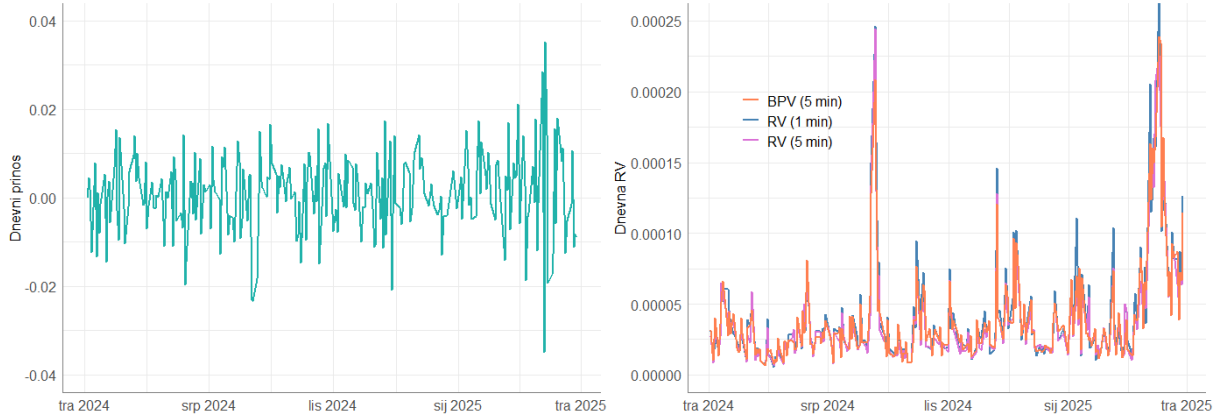
Među najjednostavnijima je HAR model, široko prepoznat po svojoj interpretabilnosti i računskoj učinkovitosti (Corsi, 2009), koji često postiže jednaku prognostičku točnost ili čak nadmašuje složenije specifikacije, poput tehnika strojnog učenja (Chassot i Audrino, 2025), osobito kada se u obzir uzmu vremenski promjenjivi parametri (Wang et al., 2016). Postoje brojne varijante i proširenja HAR modela, poput HARJ, HARCJ, HARQ, HARQJ, CHAR i CHARQ, koje se razlikuju prema načinu uključivanja skokova i kontinuirane komponente volatilitnosti, upotrebi kvantilnih mjera umjesto realiziranih mjera, te proširenju osnovne HAR strukture na zajedničko modeliranje prinosa i volatilitnosti. Corsi (2009), koji uvodi HAR-RV model, primjenjuje tri agregirane komponente realizirane volatilitnosti (dnevnu, tjednu i mjesečnu) te pokazuje da ova jednostavna struktura uspješno replicira empirijsko svojstvo duge memorije i postiže zadovoljavajuću prognozu dnevne realizirane volatilitnosti S&P500 indeksa. Na to se nadovezuju Corsi i Reno (2012), koji proširuju HAR model realizirane volatilitnosti S&P500 indeksa iz 5-minutnih cijena, opaženih od 1996. do 2009. godine, pokazujući da eksplicitno modeliranje asimetrije poboljšava kratkoročne prognoze. Daljnje unaprjeđenje HAR modela donose Bollerslev i suradnici (2016), koji predlažu da se prognoze volatilitnosti korigiraju korištenjem pogrešaka prethodnih prognoza i pokazuju da takvi modeli imaju veću prediktivnu sposobnost u odnosu na standardne HAR i GARCH modele.

Zaključno, prethodne studije koje integriraju visokofrekventne podatke u prognoziranje volatilitnosti konvergiraju prema nekoliko zajedničkih spoznaja. Prvo, gotovo svi radovi potvrđuju da uključivanje realiziranih mjera volatilitnosti konstruiranih iz intradневnih cijena poboljšava prognozu u odnosu na standardne GARCH modele, neovisno o tržišnom režimu i vrsti imovine, to jest realizirani GARCH modeli dosljedno nadmašuju standardne GARCH modele u kratkoročnim prognozama, osobito tijekom nestabilnosti tržišta. Drugo, sve više autora slaže se da složenost modela ne jamči bolje prognoze, dok parsimonijski modeli pružaju stabilnije i robusnije *out-of-sample* rezultate. Alternativni pristupi poput HEAVY i MEM modela pokazuju da izravno inkorporiranje realizirane mjere u dinamiku uvjetne (latentne) volatilitnosti osigurava značajna prognostička poboljšanja, naročito u multivarijatnom

kontekstu. Treće, unatoč primjeni metoda strojnog učenja, najjednostavniji HAR ostaje iznimno konkurentan. Time se u recentnoj literaturi oblikuje konsenzus da je ključ uspješnog prognožiranja realizirane volatilnosti u selekciji informativnih i robusnih realiziranih mjera i jednostavnosti modela (načelo parsimonije). Pored fragmentiranih nalaza, ostaje izražen istraživački jaz. Većina empirijskih istraživanja provedena je na američkim i kineskim tržištima, dok su europska tržišta slabije zastupljena i nisu sustavno analizirana. Relativna uspješnost pojedinih modela rijetko se ispituje s obzirom na izbor realizirane mjere dnevne volatilnosti i frekvencije uzorkovanja, iako HAR, MEM, HEAVY i realizirani GARCH modeli pretežno koriste realiziranu varijancu RV izračunatu na 5-minutnoj frekvenciji, uz eventualna proširenja na realizirane mjere otporne na cjenovne skokove, poput *bipower* realizirane varijance BPRV i realizirane mjere otporne na mikrostrukturni šum, poput realiziranog kernela RK. Također, 5-minutna frekvencija uzorkovanja u referiranim radovima smatra se razumnom kao balans između pristranosti i efikasnosti realiziranih mjera, ali osim tog obrazloženja autori ne nude empirijske dokaze da je ista optimalna, i ne analiziraju osjetljivost spomenutih modela pri drugim frekvencijama uzorkovanja. Jedan od rijetkih radova koji preispituje je li zaista 5-minutna frekvencija optimalna, uzimajući u obzir 104 realizirane mjere ne analizira uspješnost HAR, MEM, HEAVY i realiziranog GARCH modela koji bi te mjere inkorporirali, a empirijski rezultati ograničeni su na energetske tržište odnosno na volatilnost cijene sirove nafte (Lyu et al., 2025). Ova ograničenja upućuju na potrebu za empirijskom usporedbom različitih modela i realiziranih mjera na europskim dioničkim tržištima, uz eksplicitnu analizu robusnosti prognoza na izbor frekvencije uzorkovanja.

3. Podaci i metode

Od 2. travnja 2024. do 28. ožujka 2025. opaženo je 120 231 zadnja cijena u jednako udaljenim, nepraznim i nepreklapajućim intervalima od jedne minute, što u prosjeku odgovara približno 8,5 sati dnevnom trgovanju na frankfurtskoj burzi s pomoću elektroničkog sustava Xetra tijekom 253 radna dana, uzimajući u obzir skraćeno radno vrijeme u pojedinim danima. Dnevni prinosi i odabrane realizirane mjere volatilnosti DAX indeksa prikazani su na Slici 1. Dnevni prinosi osciliraju oko nule, u rasponu od -3,48 % do 3,53 % (Tablica 1.), s povremenim periodima pojačane varijabilnosti (fenomen grupiranja volatilnosti). Realizirane mjere volatilnosti jasno odražavaju vremensku promjenjivost, pri čemu je dnevna realizirana varijanca temeljena na jednominutnim prinosima osjetljivija na mikrostrukturni šum u odnosu na petominutnu RV mjeru (Slika 1.). Istodobno, *bipower* realizirana varijanca (BPRV) izračunata na petominutnoj frekvenciji ima blaže izražene ekstreme, što je čini otpornijom na cjenovne skokove, ali ne nužno i na mikrostrukturni šum. Dnevna RV (1 min) kreće se u rasponu od gotovo 0 % do 2,59 %. Uz uobičajene pokazatelje deskriptivne statistike, u fazi prije procjene odabranih modela, koji su detaljno opisani u nastavku, provedena je dijagnostička provjera kako bi se ispitali preduvjeti za njihovu primjenu (Tablica 1.).



Slika 1. Dnevni prinosi i dnevne realizirane mjere volatilnosti DAX indeksa

Izvor: izrada autora u programu RStudio na temelju podataka iz Refinitiv Eikon baze (2025)

Za dnevnu RV pri najvišoj frekvenciji uzorkovanja od jedne minute potvrđena je prisutnost mikrostrukturnog šuma (Box-Ljung test autokorelacije i udio šumovite varijance), cjenovnih skokova (AJ test prema Ait-Sahalia i Jacod (2009)) i duge memorije (GPH test prema Geweke i Porter-Hudak (1983)), dok su za dnevne prinose dokazani heteroskedastičnost (ARCH test), teški repovi distribucije (mjera zaobljenosti i Jarque-Bera test) te asimetrija informacija (CCF test).

Tablica 1. Deskriptivna statistika i dijagnostička provjera svojstava dnevnih prinosa i dnevne realizirane varijance DAX indeksa pri najvišoj frekvenciji uzorkovanja

Dnevni prinosi		Dnevna realizirana varijanca (1 min)	
Najmanja vrijednost	-0,034835	Najmanja vrijednost	0,000006
Najveća vrijednost	0,035284	Najveća vrijednost	0,000259
Srednja vrijednost	0,000856	Srednja vrijednost	0,000037
Standardna devijacija	0,009237	Standardna devijacija	0,000035
ARCH(5) test	62,41***	ACF(1) test	-0,0443**
Zaobljenost	7,082***	AJ test	8,69%
JB test	12,992**	NV omjer	23,54%
CCF(1) test	-0,0624**	GPH test	0,592***

Napomena: *, ** i *** označavaju statističku značajnost na razini signifikantnosti od 10 %, 5 % i 1 %

Izvor: izračun autora u programu RStudio na temelju podataka iz Refinitiv Eikon baze (2025)

Značajnost ARCH testa pri 5 vremenskih pomaka potvrđuje uvjetnu heteroskedastičnost, to jest vremenski promjenjivu varijancu dnevnih prinosa, dok visoka mjera zaobljenosti (znatno viša od 3) i značajan JB test dodatno upućuju na fenomen teških repova i odstupanje od normalnosti. Negativan i statistički značajan kros-korelacijski koeficijent prvog reda CCF(1), kojim se mjeri korelacija između tekućih kvadriranih dnevnih prinosa i prinosa prethodnog dana, ukazuje na prisutnost asimetrije informacija, tzv. efekt poluge, jer su negativni prinosi iz prethodnog dana povezani s povećanom volatilnošću tekućeg dana. Nadalje, negativna i statistički značajna autokorelacija prvog reda jednominutnih prinosa ACF(1) upućuje na prisutnost mikrostrukturnog šuma. AJ test proveden je za svaki trgovinski dan i pokazuje da su statistički značajni cjenovni skokovi prisutni u približno 8,7 % trgovinskih dana, što sugerira da su oni rijetki, ali snažni (slučaj konačne aktivnosti skokova – engl. *finite activity jumps*). Omjer NV (engl. *noise variance ratio*) oko 23,5 % je udio realizirane varijance

induciran šumom pri jednodnevnoj frekvenciji uzorkovanja i dodatno potvrđuje utjecaj mikrostrukturnog šuma, zbog čega je dnevna RV pristrana. Konačno, GPH procjena frakcijskog integracijskog parametra $d \approx 0,59$ statistički je značajna i ukazuje na izraženu dugu memoriju realizirane varijance (volatilitnost je visoko perzistentna jer se sporo vraća na svoju prosječnu razinu nakon šoka). Shodno navedenom, empirijski je opravdana primjena odabranih modela koji uvažavaju navedena svojstva, imajući na umu da realizirane mjere volatilitnosti tek treba pažljivo odabrati iako će se isti modeli koristiti za prognoziranje različitih realiziranih mjera i pri različitim frekvencijama uzorkovanja s ciljem ispitivanja njihove robusnosti i prediktivne sposobnosti.

Odabrane su četiri realizirane mjere volatilitnosti: RV, BPRV, TTSRV i RTTSRV (Tablica 2.) i četiri prognostička modela: HAR, MEM, HEAVY i realizirani GARCH (Tablica 3.), koji će se, uz standardne apsolutne i relativne kriterije točnosti prognoze (RMSE i MAPE), vrednovati i s pomoću QLIKE (engl. *Quasi-Likelihood*) kao dodatnog kriterija koji je robusan na pogreške u realiziranim mjerama, čime omogućuje pouzdaniju usporedbu modela (Patton, 2011). Naposljetku, primjenjuje se Diebold-Mariano (DM) test za formalnu usporedbu prognostičke točnosti u testnom razdoblju, kojim se ispituje jesu li razlike u prognostičkim pogreškama između dvaju konkurentskih modela statistički značajne (Diebold i Mariano, 1995). Za svaku realiziranu mjeru provodi se šest DM testova s obzirom na to da se četiri modela uspoređuju u parovima.

Tablica 2. Realizirane mjere dnevne volatilitnosti

Formulacija realizirane mjere	Otpornost na šumove	Otpornost na skokove
$RV_t^d = \sum_{i=1}^{M-1} r_{ti}^2$		
$BPRV_t^d = \frac{\pi}{2} \sum_{i=2}^{M-1} r_{t_{i-1}} r_{ti} $		X
$TTSRV_t^d = \frac{1}{1 - \frac{\bar{n}}{M-1}} \left(\frac{1}{S} \sum_{s=1}^S \sum_{i=1}^{n_s} r_{ti,s}^2 - \frac{\bar{n}}{M-1} \sum_{i=1}^{M-1} r_{ti}^2 \right)$	X	
$RTTSRV_t^d = \frac{1}{1 - \frac{\bar{n}}{M-1}} c \left(\frac{1}{S} \sum_{s=1}^S \sum_{i=1}^{n_s} r_{ti,s}^2 I_i - \frac{\bar{n}}{M-1} \sum_{i=1}^{M-1} r_{ti}^2 I_i \right)$	X	X

Napomena: intradnevni prinos r_{ti} je prinos i -te frekvencije u danu t , M je broj intradnevnih intervala unutar 8,5 satnog radnog dana, $r_{ti,s}$ je prinos i -te frekvencije u danu t spore vremenske skale s , dok je I_i indikator funkcija.

Izvor: formulacija autora prema navedenoj literaturi

Najjednostavnija mjera je realizirana varijanca RV, kao zbroj kvadrata intradnevnihih prinosa, koja konvergira prema integriranoj varijanci kako se frekvencija uzorkovanja povećava (Andersen i et al., 2001). Međutim, u praksi je RV kontaminirana mikrostrukturnim šumom, zbog čega je pristrana, a pristranost raste s porastom frekvencije uzorkovanja (Hansen i Lunde, 2005). Izvori šuma uključuju nesinkrono trgovanje, diskretna opažanja cijena i *bid-ask bounce* efekt, što dovodi do negativne autokorelacije u visokofrekventnim prinosima (Aït-Sahalia et al., 2011). Dodatno, RV nije otporna na cjenovne skokove, zbog čega može značajno precijeniti volatilitnost u razdobljima naglih tržišnih promjena. Kako bi se izolirao utjecaj cjenovnih skokova, uvedena je *bipower* realizirana varijanca BPRV (Barndorff-Nielsen i Shephard, 2006), iako ostaje osjetljiva na mikrostrukturni šum. BPRV se

računa kao zbroj umnožaka apsolutnih vrijednosti susjednih intradnevni prinosa, koji se normalizira za konstantu $\pi/2$. Problem mikrostrukturnog šuma adresira dvostruko skalirana realizirana varijanca TTTSRV (engl. *Two-Times Scaled Realized Variance*), koju su razvili Zhang i suradnici (2005). Ključna ideja TTTSRV je razdvajanje vremenskih skala na brzu i sporu. Brza vremenska skala odgovara najvišoj dostupnoj frekvenciji uzorkovanja, dok spora vremenska skala podrazumijeva nižu frekvenciju uzorkovanja. TTTSRV se dobiva tako da se najprije izračuna prosjek realiziranih volatilnosti na sporijoj vremenskoj skali, a zatim se od tog prosjeka oduzme korekcija koja se temelji na realiziranoj volatilnosti izračunanoj iz svih visokofrekventnih prinosa na brzom vremenskoj skali. Izbor spore vremenske skale određen je kompromisom između pristranosti i varijance procjenitelja. Arnerić i Matković (2019) predlažu praktičan pristup u kojem se brza vremenska skala fiksira na najvišu dostupnu frekvenciju, a optimizira se samo spora skala, što omogućuje zadržavanje svih intradnevni podataka uz istodobno uklanjanje pristranosti. Iako je TTTSRV otporna na šum, nije otporna na cjenovne skokove (Tablica 2.), odnosno TTTSRV je konzistentan i asimptotski nepristran procjenitelj integrirane varijance samo u uvjetima odsutnosti cjenovnih skokova. Stoga Boudt i Zhang (2013) uvode robusnu varijantu dvostruko skalirane realizirane varijance RTTSRV (engl. *Robust Two-Times Scaled Realized Variance*), koja dodatno koristi indikatorsku binarnu funkciju $I_i \in \{0, 1\}$ za identifikaciju ekstremnih intradnevni prinosa koji se množe s nulom i na taj se način uklanjaju iz izračuna realizirane mjere, pri čemu uobičajena razina praga odgovara trima standardnim devijacijama od prosjeka.

Tablica 3. Prognostički modeli realizirane volatilnosti

Model	Specifikacija
HAR	$RV_t^d = \omega + \alpha RV_{t-1}^d + \beta \left(\frac{1}{5} \sum_{j=1}^5 RV_{t-j}^d \right) + \gamma \left(\frac{1}{22} \sum_{j=1}^{22} RV_{t-j}^d \right) + \varepsilon_t$
MEM	$RV_t^d = \mu_t \varepsilon_t$ $\mu_t = \omega + \alpha RV_{t-1}^d + \beta \mu_{t-1}$
HEAVY	$r_t = \mu + z_t \sqrt{\sigma_t^2}$ $\sigma_t^2 = \omega + \alpha RV_{t-1}^d + \beta \sigma_{t-1}^2$ $RV_t^d = \mu_t \varepsilon_t$ $\mu_t = \gamma + \lambda RV_{t-1}^d + \theta \mu_{t-1}$
RGARCH	$r_t - \mu = z_t \sqrt{\sigma_t^2}$ $\sigma_t^2 = \omega + \alpha (r_{t-1} - \mu)^2 + \beta \sigma_{t-1}^2$ $\ln(RV_t^d) = \lambda + \theta \ln(\sigma_t^2) + \tau_1 z_t + \tau_2 (z_t^2 - 1) + u_t$

Izvor: formulacija autora prema navedenoj literaturi

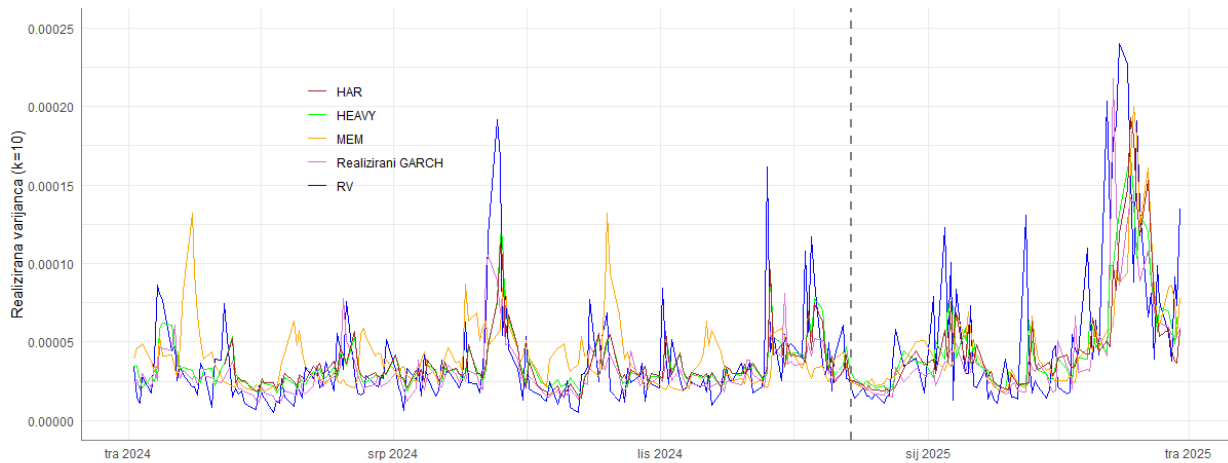
HAR model izravno modelira dnevnu realiziranu varijancu RV_t^d regresirajući je na dnevnu, tjednu i mjesečnu realiziranu varijancu iz prethodnog razdoblja, pri čemu je tjedna realizirana varijanca peteročlani pomični prosjek, a mjesečna dvadesetdvočlani pomični prosjek dnevnih realiziranih mjera (Tablica 3). Osnovni HAR model u Tablici 3 zahtjeva procjenu 4 parametra (ω , α , β i γ), pri čemu su ε_t nezavisno i jednako distribuirani inovacijski članovi. HAR model nema latentnu varijancu i ne modelira dnevne prinose. MEM model uvodi multiplikativnu strukturu realizirane varijance množeći

uvjetne sredine μ_t s inovacijskim članovima, te osigurava pozitivnost realizirane varijance, premda i dalje ne modelira dnevne prinose. HEAVY model čini iskorak jer istodobno modelira prinose te koristi RV za modeliranje latentne uvjetne varijance prinosa σ_t^2 . Realizirani GARCH zauzima srednju poziciju: zadržava klasičnu GARCH dinamiku uvjetne varijance na temelju centriranih dnevnih prinosa ($r_t - \mu$), dok RV ulazi kroz mjernu jednadžbu u logaritamskoj formi (koja osigurava nenegativnost realizirane mjere) s uključenom asimetrijom kroz standardizirane inovacijske članove z_t s obzirom na to da je $z_t = (r_t - \mu)/\sigma_t$. Broj parametara u navedenim modelima varira od 3 (MEM model zahtijeva najmanje parametara) do 9 (realizirani GARCH zahtijeva najviše parametara, to jest 7 strukturnih parametara ω , α , β , λ , θ , τ_1 i τ_2 2 parametra asimetrične Studentove t-distribucije). Kako bi se obuhvatile asimetrija i teški repovi distribucije prinosa umjesto normalne distribucije inovacijskih članova često se pretpostavlja asimetrična Studentova t-distribucija čiji oblik ovisi ne samo o stupnjevima slobode df (engl. *degrees of freedom*), već i o parametru ξ . U svim će se modelima pretpostaviti navedena distribucija, pri čemu će vrijednost $\xi < 1$ upućivati na ljevostranu asimetriju, a mali broj stupnjeva slobode, uz uvjet da je veći od 4, na dobru prilagodbu teškim repovima. Pretpostavka o distribuciji inovacijskih članova nužna je za procjenu parametara metodom najveće vjerodostojnosti MLE (engl. *Maximum Likelihood Estimation*). Također, HAR, MEM i HEAVY modeli omogućuju izravno prognoziranje realizirane mjere volatilnosti, dok realizirani GARCH to omogućuje neizravno s pomoću prognoze uvjetne (latentne) varijance prinosa σ_t^2 .

4. Empirijski nalazi i rasprava

Uobičajena je sekvencijalna podjela opažanja na podskupove za treniranje i testiranje u omjeru 70:30 kako bi se očuvala vremenska struktura podataka i osigurala stabilna procjena modela s dovoljno velikim brojem opažanja unutar uzorka (engl. *in-the-sample*). Podskup opažanja za testiranje, to jest izvan uzorka (engl. *out-of-sample*), poslužit će za usporedbu ex-ante procijenjene realizirane varijance i njene ex-post prognostičke vrijednosti, dobivenih na temelju HAR, MEM, HEAVY i realiziranog GARCH modela. Ovaj se omjer fiksira radi konzistentnosti rezultata, dok će se osjetljivost točnosti prognoza iste realizirane mjere analizirati u odnosu na frekvenciju uzorkovanja kroz sve uspoređivane modele, a naknadno i u odnosu na različite realizirane mjere, od kojih svaka ima vlastitu optimalnu frekvenciju uzorkovanja. Nadalje, RMSE (engl. *Root Mean Square Error*) je odabran kao standardna mjera prognostičke pogreške izražena u jedinicama mjere varijable koja se predviđa, dok MAPE (engl. *Mean Absolute Percentage Error*) omogućuje interpretaciju prognostičke pogreške u relativnom (postotnom) iznosu i olakšava usporedbu učinkovitosti među modelima, a QLIKE se koristi kao dodatni kriterij jer je robustan na mjernu pogrešku realizirane volatilnosti. Stoga ovaj rad doprinosi novim empirijskim spoznajama o modelu koji konzistentno pokazuje prognostičku učinkovitost, odnosno najmanju osjetljivost kako na frekvenciju uzorkovanja, tako i na izbor realizirane mjere volatilnosti, čime se popunjava istraživački jaz sustavnom i komparativnom analizom navedenih modela.

Premda prognoze HAR, HEAVY, MEM i realiziranog GARCH modela u periodu od 5. prosinca 2024. do 28. ožujka 2025. (podskup za testiranje) uglavnom prate dinamiku realizirane varijance, one se razlikuju od modela do modela, osobito u testnom periodu u kojem su prisutni nagli skokovi i povećana realizirana volatilnost, pri čemu je frekvencija uzorkovanja fiksirana na 10 minuta samo radi ilustracije tih razlika (Slika 2), dok se u Tablici 4 uspoređuju procijenjeni modeli u prognoziranju iste realizirane mjere volatilnosti pri različitim frekvencijama uzorkovanja.



Slika 2. Procijenjene i prognostičke vrijednosti realizirane varijance dobivene različitim modelima za fiksnu frekvenciju uzorkovanja ($k=10$)

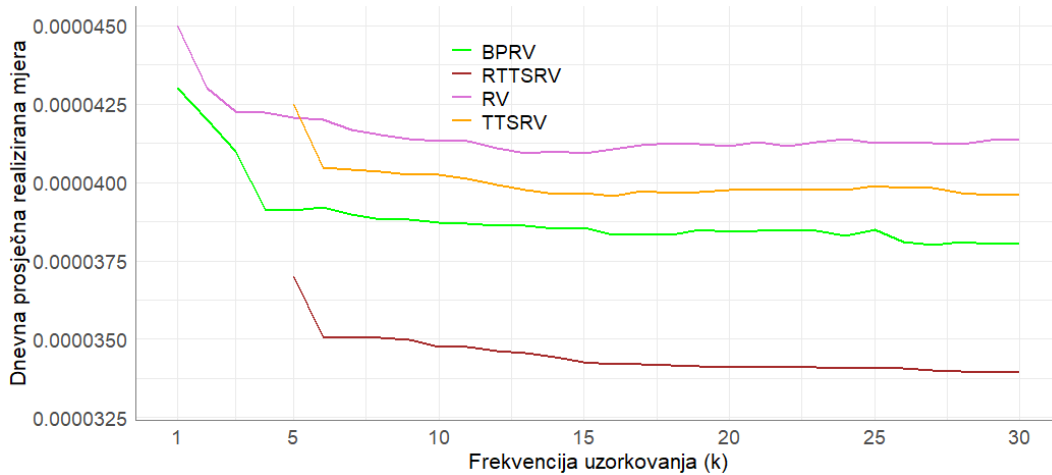
Izvor: izrada autora u programu RStudio na temelju podataka iz Refinitiv Eikon baze (2025)

Tablica 4. Prognostička sposobnost različitih modela realizirane mjere volatilnosti u ovisnosti o frekvenciji uzorkovanja

Frekvencija uzorkovanja	Prognostička sposobnost izvan uzorka	Prognostički model realizirane varijance RV			
		HAR	MEM	HEAVY	Realizirani GARCH
$k = 1$	RMSE	0,0000048	0,0000044	0,0000046	0,0000037
	MAPE	20,74%	18,13%	22,60%	15,31%
	QLIKE	-9,8169	-9,8177	-9,8166	-9,8212
$k = 5$	RMSE	0,0000375	0,0000407	0,0000351	0,0000364
	MAPE	5,58%	5,48%	5,32%	4,33
	QLIKE	-8,9163	-8,9082	-8,9198	-8,9034
$k = 10$	RMSE	0,0000433	0,0000473	0,0000406	0,0000421
	MAPE	5,84%	6,17%	5,91%	4,77%
	QLIKE	-8,8732	-8,8437	-8,8709	-8,8346
$k = 15$	RMSE	0,0000442	0,0000500	0,0000417	0,0000419
	MAPE	6,38%	6,55%	6,24%	5,01%
	QLIKE	-8,9309	-8,8961	-8,9336	-8,8973

Izvor: izračun autora u programu RStudio na temelju podataka iz Refinitiv Eikon baze (2025)

Tablica 4. pokazuje da realizirani GARCH konzistentno ima najbolju prognostičku sposobnost prema MAPE kriteriju za sve frekvencije uzorkovanja, iako su svi modeli manje učinkoviti u prognoziranju realizirane volatilnosti pri najvišoj frekvenciji od jedne minute. S nižom frekvencijom uzorkovanja prognostičke pogreške se smanjuju, a smanjuju se i razlike među modelima, pri čemu QLIKE pokazuje vrlo slične vrijednosti i potvrđuje stabilnost njihovog relativnog poretka.



Slika 3. Realizirane mjere RV, BPRV, TTSRV i RTTSRV pri različitim frekvencijama
Izvor: izrada autora u programu RStudio na temelju podataka iz Refinitiv Eikon baze (2025)

Dnevne prosječne realizirane mjere volatiliteta (RV, BPRV, TTSRV i RTTSRV) s porastom k (smanjenjem frekvencije uzorkovanja) postupno se smanjuju i stabiliziraju, što upućuje na reduciranje mikrostrukturnog šuma, pri čemu se optimalna zajednička frekvencija može približno odrediti u rasponu od 10 do 15 minuta (Slika 3.). Dvostruko skalirane mjere TTSRV i RTTSRV nisu računane za više frekvencije (niže od $k = 5$) jer zahtijevaju barem $10 \times k$ opažanja u jednom danu. Bolje je odabrati frekvenciju od 15 minuta ako se favorizira stabilnost uz zanemariv gubitak informacija (prelaskom s 10-minutne na 15-minutnu frekvenciju dnevno se gubi 17 intradnevni prihoda pri izračunu mjera RV i BPRV) ili bez gubitka informacija (mjere TTSRV i RTTSRV zadržavaju sve podatke). Ovakav pristup primijenjen je u nastavku analize. Tablice 5 i 6 prezentiraju rezultate procijenjenih parametara 4 modela za prognozirane 4 realizirane mjere volatiliteta pri istoj frekvenciji uzorkovanja od 15 minuta, koja je utvrđena kao zajednička optimalna frekvencija. Za testno razdoblje predočene su i vrijednosti kriterija RMSE, MAPE i QLIKE te rezultati Diebold-Maraino testa.

Na temelju Tablica 5. i 6. može se donijeti nekoliko zaključaka o relativnoj uspješnosti pojedinih modela ovisno o vrsti realizirane mjere volatiliteta koja se prognozira, pri čemu se osjetljivost na skokove i mikrostrukturni šum izravno reflektira u tome koji model pruža bolje prognoze u testnom razdoblju. Za standardnu realiziranu volatilitet RV, koja je istovremeno osjetljiva i na skokove i na mikrostrukturni šum, realizirani GARCH se pokazuje kao najprikladniji. Iako HEAVY model ima nešto niži RMSE, realizirani GARCH ima znatno niži MAPE, što je važno jer relativna pogreška bolje odražava kvalitetu prognoze u razdoblju povišene volatiliteta. Njegova prednost proizlazi iz mjerne jednadžbe koja eksplicitno povezuje latentnu uvjetnu varijancu prihoda s realiziranom mjerom (parametar θ je pozitivan i statistički značajan), čime se učinkovito filtrira šum iz RV-a. Značajni parametri τ_1 i τ_2 ukazuju na prisutnost asimetrije i nelinearnog učinka inovacija, odnosno negativan τ_1 implicira jači utjecaj negativnih šokova, dok pozitivan τ_2 potvrđuje da veliki šokovi, bez obzira na predznak, dodatno povećavaju realiziranu volatilitet. Stupnjevi slobode t-distribucije upućuju na teške repove (niska vrijednost stupnjeva slobode), dok relativno velika varijanca pogreške u mjernoj jednadžbi σ_u^2 , premda nesigifikantna, potvrđuje da je RV mjera opterećena šumom, ali ga model uspješno filtrira. Kod BPRV mjere, koja je otporna na skokove najbolju prognostičku uspješnost pokazuje HEAVY model prema svim kriterijima, dok realizirani GARCH ostaje konkurentan.

Tablica 5. Rezultati procjena modela realizirane volatilnosti za mjere RV i BPRV

	Realizirana mjera RV ($k = 15$)				Realizirana mjera BPRV ($k = 15$)			
	HAR	MEM	HEAVY	RGARCH	HAR	MEM	HEAVY	RGARCH
ω	0,0000*** (0,0000)	0,0412 (0,0957)	0,4339*** (0,1233)	-1,7960 (1,2769)	0,0000*** (0,0000)	0,0001 (0,0787)	0,4339*** (0,0910)	-1,5922 (1,2415)
α	0,3422** (0,1591)	0,2847*** (0,0630)	0,2166*** (0,0697)	0,2758** (0,1111)	0,4060*** (0,1241)	0,3205*** (0,0519)	0,2166*** (0,0664)	0,2549** (0,1152)
β	0,2827* (0,1573)	0,0412 (0,0957)	0,1491 (0,1449)	0,5101*** (0,1124)	0,2130* (0,1255)	0,7397*** (0,0830)	0,1490 (0,1385)	0,5519*** (0,1202)
γ	-0,2161 (0,1330)		0,1032** (0,0490)		-0,1734 (0,1177)		0,0791* (0,0468)	
λ			0,4201*** (0,1156)	1,2102*** (0,4381)			0,3782*** (0,1174)	1,2426** (0,4835)
θ			0,2786 (0,2049)	0,5306*** (0,0283)			0,3671 (0,2247)	0,5339*** (0,0284)
τ_1				-0,1386*** (0,0478)				-0,1574*** (0,0478)
τ_2				0,1303*** (0,0355)				0,1213*** (0,0351)
ξ				0,7637*** (0,0912)				0,7599*** (0,0933)
df				6,0000*** (2,0010)				6,0000*** (1,6335)
σ_u^2				1,0872 (4,2032)				1,3094 (4,6421)
RMSE	0,0000442	0,0000500	0,0000417	0,0000419	0,0000422	0,0000457	0,0000381	0,0000406
MAPE	6,38%	6,55%	6,24%	5,01%	6,18%	6,27%	4,68%	5,65%
QLIKE	-8,9309	-8,8961	-8,9336	-8,8973	-8,9647	-8,9382	-8,9344	-8,9762
DM test	HAR	-0,1293	0,2759	0,2477	HAR	-0,1036	0,2078	0,1722
	MEM		1,4587*	1,5963*	MEM		1,6943**	1,4771*
	HEAVY			-0,2355	HEAVY			-0,2901

Napomena: *, ** i *** označavaju statističku značajnost na razini signifikantnosti od 10%, 5% i 1%, dok su u zagradama standardne pogreške procijenjenih parametara

Izvor: izračun autora u programu RStudio na temelju podataka iz Refinitiv Eikon baze (2025)

U realiziranom GARCH modelu i dalje su značajni parametri asimetrije, što sugerira da čak i kod mjera robusnih na cjenovne skokove informacija o predznaku i veličini inovacija ostaje relevantna. Nešto veća σ_u^2 u odnosu na RV upućuje na to da BPRV, iako robusna na skokove, sadrži i veći šum. Za TTSRV, mjeru prvenstveno dizajniranu za uklanjanje mikrostrukturnog šuma, realizirani GARCH ponovno daje najniži MAPE, dok se HEAVY i HAR vrlo blisko prate prema RMSE kriterij. Ovdje je posebno važno uočiti da se varijanca pogreške u mjernoj jednadžbi σ_u^2 realiziranog GARCH modela značajno smanjuje u odnosu na RV, što potvrđuje da je TTSRV manje pristrana realizirana mjera volatilnosti. Unatoč tome, parametri τ_1 i τ_2 ostaju statistički značajni i očekivanog predznaka. Naime, negativan τ_1 sugerira da negativni šokovi povećavaju realiziranu volatilnost više nego pozitivni šokovi iste magnitude, čime je zahvaćena asimetrija informacija, dok pozitivan τ_2 sugerira da veliki šokovi, neovisno o predznaku, dodatno povećavaju realiziranu volatilnost. HAR model u svim primjenama nije reprezentativan zbog

neznačajne mjesečne komponentom, dok su kratkoročna i srednjoročna komponente statistički signifikantne na razinama značajnosti od 5% odnosno 10%.

Tablica 6. Rezultati procjena modela realizirane volatilnosti za mjere TTSRV i RTTSRV

	Realizirana mjera TTSRV ($k = 15$)				Realizirana mjera RTTSRV ($k = 15$)			
	HAR	MEM	HEAVY	RGARCH	HAR	MEM	HEAVY	RGARCH
ω	0,0000*** (0,0000)	0,0001 (0,0940)	0,4339*** (0,1189)	-1,8563 (1,3060)	0,0000*** (0,0000)	0,0002 (0,0686)	0,4339*** (0,1100)	-1,9578 (1,7016)
α	0,3887** (0,1754)	0,3232*** (0,0572)	0,2166*** (0,0821)	0,3050** (0,1430)	0,4068*** (0,1558)	0,3440*** (0,0413)	0,2166*** (0,0664)	0,3222*** (0,1249)
β	0,2303 (0,1469)	0,7062*** (0,1128)	0,1491 (0,1558)	0,4714*** (0,1494)	0,2111 (0,1311)	0,7231*** (0,0747)	0,1491 (0,1469)	0,4379*** (0,1113)
γ	-0,1295 (0,1215)		0,4339*** (0,1189)		-0,1926 (0,1370)		0,0673** (0,0309)	
λ			0,1051*** (0,0389)	1,1697*** (0,4230)			0,4933*** (0,1468)	1,2699** (0,5085)
θ			0,4076*** (0,1009)	0,5062*** (0,0270)			0,2794 (0,1912)	0,4637*** (0,0247)
τ_1				-0,1343*** (0,0466)				-0,1616*** (0,0413)
τ_2				0,1399*** (0,0342)				0,1009*** (0,0306)
ξ				0,7591*** (0,0926)				0,7477*** (0,0879)
df				6,0000** (3,8898)				6,0000** (3,6736)
σ_u^2				0,6842 (4,0652)				0,5223 (4,8941)
RMSE	0,0000371	0,0000418	0,0000363	0,0000376	0,0000381	0,0000427	0,0000377	0,0000403
MAPE	5,93%	5,83%	5,86%	4,73%	6,17%	6,03%	6,23%	5,07%
QLIKE	-8,9845	-8,9461	-8,9681	-8,9393	-9,0083	-8,9647	-8,9896	8,9436
DM test	HAR	-0,5228	0,1924	0,0847	HAR	-0,6117	0,1742	-0,3674
	MEM		0,7687	0,7355	MEM		0,6633	0,2965
	HEAVY			-0,1321	HEAVY			-0,2022

Napomena: *, ** i *** označavaju statističku značajnost na razini signifikantnosti od 10%, 5% i 1%, dok su u zagradama standardne pogreške procijenjenih parametara)

Izvor: izračun autora u programu RStudio na temelju podataka iz Refinitiv Eikon baze (2025)

Kod RTTSRV, mjere koja je robusna i na skokove i na šum, razlike među modelima dodatno se smanjuju, ali realizirani GARCH ponovno pokazuje prednost u odnosu na ostale modele. HEAVY model u ovom slučaju gubi prednost jer dodatna robusnost mjere smanjuje potrebu za simultanim modeliranjem prinosa i volatilnosti s pomoću multiplikativne strukture. MEM model u svim slučajevima pokazuje visoku perzistentnost kroz β parametar, osobito kod robusnijih mjera, što je tipično za multiplikativne modele i ukazuje na to da je dugoročna razina volatilnosti gotovo u potpunosti određena prošlim vrijednostima. Iako je stabilan i MEM u pravilu nije prognostički učinkovitost. Realizirani GARCH se pokazuje kao najfleksibilniji i najrobusniji model, osobito za mjere osjetljive na šum te na

šum i skokove, dok je HEAVY model posebno učinkovit za mjere koje su robusne isključivo na cjenovne skokove. HAR i MEM modeli zaostaju i slabije su alternative, premda su jednostavniji i zahtijevaju manje parametara. Nultom hipotezom Diebold-Mariano testa pretpostavlja se da su prosječne kvadratne pogreške između dvaju modela jednake, što implicira jednaku prognostičku sposobnost. Predznak testove veličine sugerira koji je model bolji, to jest ako DM test ima negativnu vrijednost i statistički je značajan tada model u retku Tablica 5 i 6 ima bolju prognostičku sposobnost od modela u stupcu (test na donju granicu), a obratno ako DM test ima pozitivnu vrijednost i statistički je značajan tada model u stupcu ima bolju prognostičku sposobnost od modela u retku (test na gornju granicu). Shodno navedenom, HEAVY i realizirani GARCH modeli imaju bolju prognostičku sposobnost u odnosu na MEM model na razini signifikantnosti od 10% kada se prognozira RV, dok među ostalim modelima nema značajne razlike. Do istog se zaključka dolazi kada se prognozira i BPRV, ali je tada razlika između HEAVY i MEM modela značajna na razini signifikantnosti od 5 %. Kod dvostruko skaliranih realiziranih mjera volatilnosti razlike nisu statistički značajne.

5. Zaključak

Ovaj rad pruža sustavnu i empirijski utemeljenu usporedbu prognostičke učinkovitosti HAR, MEM, HEAVY i realiziranog GARCH modela na primjeru referentnog europskog dioničkog indeksa DAX, koristeći visokofrekventne intradnevne podatke i više realiziranih mjera volatilnosti. Glavni doprinos rada proizlazi iz istodobne evaluacije utjecaja frekvencije uzorkovanja i izbora realizirane mjere na relativnu uspješnost istoimenih modela, čime se popunjava istraživački jaz u literaturi koja je do sada bila dominantno usmjerena na američka tržišta i standardnu realiziranu varijancu RV.

Empirijski nalazi jasno pokazuju da ne postoji univerzalno najbolji model za prognoziranje svih realiziranih mjera volatilnosti, već da je njihova relativna učinkovitost uvjetovana statističkim svojstvima mjere koja se prognozira. Realizirani GARCH model konzistentno pokazuje najveću robusnost i stabilnost prognostičke sposobnosti, osobito kod mjera koje su osjetljive na mikrostrukturni šum ili kombinaciju šuma i cjenovnih skokova, poput standardne realizirane varijance RV. Njegova prednost proizlazi iz mjerne jednadžbe koja omogućuje eksplicitno razdvajanje signalne komponente volatilnosti od mjerne pogreške, što je empirijski potvrđeno relativno velikom varijancom pogreške u mjernoj jednadžbi kod šumovitih mjera. Značajni parametri asimetrije (τ_1 i τ_2) dodatno potvrđuju da realizirani GARCH uspješno hvata nelinearne i asimetrične reakcije volatilnosti, što je osobito važno u razdobljima tržišnog stresa, dok su relativno mali procijenjeni stupnjevi slobode indikator dobre prilagodbe Studentove t-distribucije teškim repovima. Nadalje, prognoze volatilnosti pri najvišoj frekvenciji (jedna minuta) znatno su lošije za sve modele, zbog dominantnog utjecaja mikrostrukturnog šuma. Optimalna frekvencija uzorkovanja nalazi se u rasponu od 10 do 15 minuta, pri čemu se postiže kompromis između gubitka informacija i redukcije pristranosti realiziranih mjera.

HEAVY model pokazuje najbolju prognostičku učinkovitost kod realiziranih mjera robusnih isključivo na skokove, poput BPRV, gdje njegova multiplikativna struktura i simultano modeliranje prinosa i volatilnosti dolaze do izražaja. Međutim, prelaskom na mjere kod kojih je mikrostrukturni šum reduciran (TTSRV i RTTSRV), njegova relativna prednost slabi, što upućuje na to da HEAVY model najveću korist ostvaruje upravo kada realizirana mjera još uvijek sadrži informaciju o ekstremnim kretanjima, ali ne i dominantan šum. HAR model nije konkurentan, posebice zbog neznačajnosti dugoročne realizirane mjere volatilnosti. To potvrđuje da jednostavna heterogena struktura ne može potpuno replicirati dugoročnu memoriju volatilnosti. MEM model, iako stabilan i parsimonijski, u pravilu zaostaje u prognostičkoj učinkovitosti, što se može pripisati njegovoj ograničenoj sposobnosti

razdvajanja šuma i signala te izostanku mehanizama za hvatanje asimetrije. Diebold-Mariano test pokazuje da su modeli HEAVY i realizirani GARCH statistički značajno superiorniji u odnosu na MEM model pri prognoziranju realiziranih mjera volatilnosti koje nisu robusne na cjenovne skokove i/ili mikrostrukturni šum, iako između HEAVY i realiziranog GARCH modela nije utvrđena statistički značajna razlika. S druge strane, kod mjera koje su robusne na šumove i/ili skokove nisu utvrđene statistički značajne razlike ni između jednog para modela, što implicira da su u tom slučaju svi odabrani modeli podjednako uspješni, uz uvjet pravilnog odabira optimalne frekvencije uzorkovanja, iako prema kriterijima MAPE i QLIKE blagu prednost ima realizirani GARCH model unatoč velikom broju parametara.

Preciznije prognoze dnevne realizirane volatilnosti, osobito na kratkom horizontu, važne su za svakodnevno upravljanje portfeljima i alokaciju kapitala. Za praktičare u financijskom sektoru ovi rezultati pružaju jasne smjernice o izboru realizirane mjere, frekvencije uzorkovanja i prognostičkog modela u internim sustavima upravljanja rizicima. Na makroprudencijalnoj razini, mogućnost ranog prepoznavanja nadolazećih razdoblja povišene volatilnosti omogućuje regulatornim tijelima pravovremeno prilagođavanje instrumenata financijske stabilnosti. S obzirom na to da se procjene buduće realizirane volatilnosti izravno ugrađuju u izračun regulatornih mjera rizika, uključujući minimalne kapitalne zahtjeve, unapređenje njihove prognostičke točnosti smanjuje vjerojatnost sustavnih pogrešaka u procjeni rizika te doprinosi stabilnijem i otpornijem financijskom sustavu.

Rad ima nekoliko ograničenja. Analiza je provedena samo na jednom dioničkom indeksu, stoga dobiveni rezultati imaju implikacije isključivo za razvijena europska tržišta kapitala, poput njemačkog. Izbor referentnog europskog dioničkog indeksa DAX uvjetovan je činjenicom da kvalitetni visokofrekventni podaci, potrebni za izračun realiziranih mjera volatilnosti, postoje samo na tržištima s učestalim i intenzivnim trgovanjem. Na manje likvidnim tržištima, zbog izostanka trgovanja u svakoj minuti, potrebno je primijeniti nižu početnu frekvenciju uzorkovanja, dok optimalnu tek treba odrediti, čime se gubi na količini informacija i preciznosti procjene realiziranih mjera. Zato se takve analize najčešće provode na najlikvidnijim tržištima. Buduća istraživanja mogla bi proširiti analizu na multivarijatan kontekst, ispitati vremenski promjenjive parametre, kombinirati realizirane mjere s pristupima strojnog učenja ili uvesti dodatne egzogene varijable te obuhvatiti dulje razdoblje promatranja. Zaključno, rad pruža nove empirijske spoznaje da uspješno prognoziranje realizirane volatilnosti ne ovisi o složenosti modela, već o usklađenosti modela sa statističkim svojstvima realizirane mjere, koja izravno ovise o optimalnoj frekvenciji uzorkovanja. Time se daje jasan znanstveni doprinos i nude konkretne smjernice za primjenu prognostičkih modela realizirane volatilnosti na europskim tržištima kapitala.

Literatura:

Aït-Sahalia, Y., Jacod, J. (2009). Testing for jumps in a discretely observed process. *The Annals of Statistics*, 37(1), 182-222. <https://doi.org/10.1214/07-AOS568>

Aït-Sahalia, Y., Mykland, P. A., Zhang, L. (2011). Ultra high frequency volatility estimators with dependent microstructure noise. *Journal of Econometrics*, 160(1), 160-175. <https://doi.org/10.1016/j.jeconom.2010.03.003>

- Amendola, A., Candila, V., Cipollini, F., Gallo, G. M. (2024). Doubly multiplicative error models with long- and short-run components. *Socio-Economic Planning Sciences*, 91, 101764. <https://doi.org/10.1016/j.seps.2023.101764>
- Andersen, T. G., Bollerslev, T., Diebold, F. X., Labys, P. (2001). The distribution of realized exchange rate volatility. *Journal of the American Statistical Association*, 96(453), 42-55. <https://doi.org/10.1198/016214501750332965>
- Arnerić, J., Matković, M. (2019). Challenges of integrated variance estimation in emerging stock markets. *Proceedings of Rijeka Faculty of Economics: Journal of Economics and Business*, 37(2), 713-739. <https://doi.org/10.18045/zbefri.2019.2.713>
- Arnerić, J., Poklepović, T., Teai, J. W. (2018). Neural network approach in forecasting realized variance using high-frequency data. *Business Systems Research Journal*, 9(2), 18-34. <https://doi.org/10.2478/bsrj-2018-0016>
- Barndorff-Nielsen, O. E., Shephard, N. (2006). Econometrics of testing for jumps in financial economics using bipower variation. *Journal of Financial Econometrics*, 4(1), 1-30. <https://doi.org/10.1093/jjfinec/nbi022>
- Bauwens, L., Xu, Y. (2023). DCC- and DECO-HEAVY: Multivariate GARCH models based on realized variances and correlations. *International Journal of Forecasting*, 39(2), 938-955. <https://doi.org/10.1016/j.ijforecast.2022.03.005>
- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3), 307-327. [https://doi.org/10.1016/0304-4076\(86\)90063-1](https://doi.org/10.1016/0304-4076(86)90063-1)
- Bollerslev, T., Patton, A., Quaedvlieg, R. (2016). Exploiting the errors: A simple approach for improved volatility forecasting. *Journal of Econometrics*, 192, 1-18. <https://doi.org/10.1016/j.jeconom.2015.10.007>
- Boudt, K., Zhang, J. (2013). Jump robust two time scale covariance estimation and realized volatility budgets. *Quantitative Finance*, 15(6), 1041-1054. <https://doi.org/10.1080/14697688.2012.741692>
- Chassot, J., Audrino, F. (2025). HARd to Beat: The Overlooked Impact of Rolling Windows in the Era of Machine Learning. *International Journal of Forecasting*. In Press. <https://doi.org/10.1016/j.ijforecast.2025.06.003>
- Cipollini, F., Gallo, G. M. (2025). Multiplicative Error Models: 20 years on. *Econometrics and Statistics*, 10, 100123. <https://doi.org/10.1016/j.ecosta.2025.100123>
- Corsi, F. (2009). A simple approximate long-memory model of realized volatility. *Journal of Financial Econometrics*, 7(2), 174-196. <https://doi.org/10.1093/jjfinec/nbp001>
- Corsi, F., Reno R. (2012). Discrete-time volatility forecasting with persistent leverage effect and the link with continuous-time volatility modeling. *Journal of Business and Economic Statistics*, 30, 368-380. <https://doi.org/10.1080/07350015.2012.663261>

- Diebold, F. X., Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business and Economic Statistics*, 13(3), 253-263. <https://doi.org/10.1080/07350015.1995.10524599>
- Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, 50(4), 987-1007. <https://doi.org/10.2307/1912773>
- Engle, R. F. (2002). New Frontiers for ARCH Models. *Journal of Applied Econometrics*, 17(5), 425-446. <https://doi.org/10.1002/jae.683>
- Engle, R. F., Gallo, G. M. (2006). A multiple indicators model for volatility using intra-daily data. *Journal of Econometrics*, 131(1-2), 3-27. <https://doi.org/10.1016/j.jeconom.2005.01.018>
- Fang, Z., Han, J.-Y. (2025). Realized GARCH model in volatility forecasting and option pricing. *Computational Economics*, 66(5), 3637-3657. <https://doi.org/10.1007/s10614-024-10826-8>
- Frömmel, M., Han, X., Kratochvil, S. (2014). Modeling the daily electricity price volatility with realized measures. *Energy Economics*, 44, 492-502. <https://doi.org/10.1016/j.eneco.2014.03.001>
- Geweke, J., Porter-Hudak, S. (1983). The estimation and application of long memory time series models. *Journal of Time Series Analysis*, 4(4), 221-238. <https://doi.org/10.1111/j.1467-9892.1983.tb00371.x>
- Girardi, M., Caporin, M. (2025). Conditional autoregressive G model for common factor detection in the stock market. *Econometrics and Statistics*. In Press. <https://doi.org/10.1016/j.ecosta.2025.05.002>
- Hansen, P. R., Huang, Z., Shek, H. H. (2011). Realized GARCH: A joint model for returns and realized measures of volatility. *Journal of Applied Econometrics*, 27(6), 877-906. <https://doi.org/10.1002/jae.1234>
- Hansen, P. R., Lunde, A. (2005). A Forecast Comparison of Volatility Models: Does Anything Beat a GARCH(1,1)? *Journal of Applied Econometrics*, 20(7), 873-889. <https://doi.org/10.1002/jae.800>
- Hansen, P. R., Huang, Z. (2016). Exponential GARCH modeling with realized measures of volatility. *Journal of Business and Economic Statistics*, 34(2), 269-287. <https://doi.org/10.1080/07350015.2015.1038543>
- Huang, Z., Liu, H., Wang, T. (2016). Modeling long memory volatility using realized measures of volatility: A realized HAR GARCH model. *Economic Modelling*, 52, 812-821. <https://doi.org/10.1016/j.econmod.2015.10.018>
- Lyu, Y., Yang, Z., Luo, Y., Qin, Z., Yi, H., Ke, R. (2025). Forecasting the volatility of crude oil futures market: Does the simple 5-minute RV hold up? *Energy Economics*, 146, 108509. <https://doi.org/10.1016/j.eneco.2025.108509>
- Nguyen, H. T., Nguyen, H., Tran, M.-N. (2024). Deep learning enhanced volatility modeling with covariates. *Finance Research Letters*, 69, 106145. <https://doi.org/10.1016/j.frl.2024.106145>

- Naimoli, A., Gerlach, R. H., Storti, G. (2022). Improving the accuracy of tail risk forecasting models by combining several realized volatility estimators. *Economic Modelling*, 107, 105701. <https://doi.org/10.1016/j.econmod.2021.105701>
- Noureldin, D., Shephard, N., Sheppard, K. (2012). Multivariate high-frequency-based volatility (HEAVY) models. *Journal of Applied Econometrics*, 27(6), 907-933. <https://doi.org/10.1002/jae.1260>
- Patton, A. J. (2011). Volatility forecast comparison using imperfect volatility proxies. *Journal of Econometrics*, 160(1), 246-256. <https://doi.org/10.1016/j.jeconom.2010.03.034>
- Sharma, P., Vipul. (2016). Forecasting stock market volatility using realized GARCH model: International evidence. *The Quarterly Review of Economics and Finance*, 59, 222-230. <https://doi.org/10.1016/j.qref.2015.07.005>
- Shephard, N., Sheppard, K. (2010). Realizing the future: Forecasting with high-frequency-based volatility (HEAVY) models. *Journal of Applied Econometrics*, 25(2), 197-231. <https://doi.org/10.1002/jae.1158>
- Sheppard, K., Xu, W. (2019). Factor high-frequency-based volatility (HEAVY) models. *Journal of Financial Econometrics*, 17(1), 33-65. <https://doi.org/10.1093/jjfinec/nby028>
- Wang, Y., Ma, F., Wei, Y., Wu, C. (2016). Forecasting realized volatility in a changing world: A dynamic model averaging approach. *Journal of Banking and Finance*, 64, 136-149. <https://doi.org/10.1016/j.jbankfin.2015.12.010>
- Watanabe, T. (2012). Quantile forecasts of financial returns using realized GARCH models. *Japanese Economic Review*, 63(1), 68-80. <https://doi.org/10.1111/j.1468-5876.2011.00548.x>
- Wu, X., Zhao, A., Wang, Y., Han, Y. (2024). Forecasting Chinese stock market volatility with high-frequency intraday and current return information. *Pacific-Basin Finance Journal*, 86, 102458. <https://doi.org/10.1016/j.pacfin.2024.102458>
- Xiao, M., Tao, Z., Gu, Z., Li, Z., Chen, X. (2023). New practice for investors in Chinese stock market: From perspective of fractionally integrated realized GARCH model. *Heliyon*, 9(7), e17845. <https://doi.org/10.1016/j.heliyon.2023.e17845>
- Xie, H., Yu, C. (2020). Realized GARCH models: Simpler is better. *Finance Research Letters*, 33, 101221. <https://doi.org/10.1016/j.frl.2019.06.019>
- Xu, W., Aschakulporn, P., Zhang, J. E. (2024). Heterogeneous volatility information content for the realized GARCH modeling and forecasting volatility. *Studies in Nonlinear Dynamics and Econometrics*. <https://doi.org/10.1515/snde-2024-0013>
- Zhang, L., Mykland, P. A., Aït-Sahalia, Y. (2005). A tale of two time scales. *Journal of the American Statistical Association*, 100(472), 1394-1411. <https://doi.org/10.1198/016214505000000169>

Forecasting Realized Volatility: Empirical Evidence from a Benchmark European Equity Index

JOSIP ARNERIĆ

University of Zagreb, Faculty of Economics and Business
Square J. F. Kennedy 6, 10000 Zagreb
Croatia

jarneric@net.efzg.hr

ORCID iD: <https://orcid.org/0000-0002-2901-2609>

Summary: Realized volatility is based on intraday high-frequency returns and serves as a standard measure of the true, but unobserved, integrated volatility of a financial asset. Although realized volatility is determined ex-post, several models have been developed over the past decade to forecast it. These models differ in the features they capture, such as long memory, heteroskedasticity, price jumps, asymmetric responses to positive and negative shocks, and microstructure noise, as well as in whether they model realized volatility directly or indirectly. The performance of these models remains insufficiently explored, particularly in European capital markets, in contrast to U.S. markets, which dominate existing studies. Therefore, the objective of this paper is to compare the forecasting performance of selected models for the realized volatility of the DAX, which is considered a benchmark European equity index. For this purpose, HAR, MEM, HEAVY, and realized GARCH models are compared not only to identify the model with the highest predictive accuracy, but also to examine whether their performance depends on the sampling frequency and the choice of the realized volatility measure being forecast. In doing so, the paper contributes to the literature, in which no consensus has yet been established regarding the most appropriate models for forecasting specific realized volatility measures. The empirical findings provide relevant implications for risk management and asset pricing in European capital markets, especially during the period of uncertainty covered by the analysis. The study uses one-minute closing prices of the DAX index, while forecasting accuracy is evaluated using multiple criteria.

Keywords: realized volatility, forecasting models, benchmark equity index

JEL codes: C53, C58, D53, G17