

VILNIAUS UNIVERSITETAS  
MATEMATIKOS IR INFORMATIKOS FAKULTETAS

Magistro baigiamasis darbas

**MIDAS IR LASSO REGRESIJŲ  
PALYGINIMAS**

**Comparison of MIDAS and LASSO Estimators**

Agnė Pelanytė

VILNIUS 2015

MATEMATIKOS IR INFORMATIKOS FAKULTETAS  
EKONOMETRINĖS ANALIZĖS KATEDRA

Darbo vadovas Lekt. dr. Vaidotas Zemlys

Darbo recenzentas Doc. Remigijus Lapinskas

Darbas apgintas 2016 m. sausio 14 d.

Darbas įvertintas \_\_\_\_\_

Registravimo NR. \_\_\_\_\_

Atidavimo į katedrą data 2016-01-04

# MIDAS ir LASSO regresijų palyginimas

## Anotacija

Skirtingo dažnio duomenų (MIDAS) regresija paremta apribota ankstinių pasiskirstymo struktūra, kuriai dažniausiai reikalingi tam tikri apribojimai, tokie kaip glodumas ar teigiamumas, kurie padeda gauti tikslesnę prognozę. Lasso metodu paremtos neapribotos MIDAS regresijos struktūra neapribojama, tačiau prognozės tikslumo siekiama mažinant modelio dimensiją. Todėl šio darbo tikslas yra palyginti, kaip veikia abu modeliai prognozės atžvilgiu kai struktūros apribojimai neatitinka realių sąryšių, tačiau modelių specifikacijos adekvatumas nėra atmetamas. Naudojant imitacinę analizę parodoma, kad apriboto MIDAS prognozės paklaida įprastai yra mažesnė nei neapriboto modelio su Lasso. Tačiau nežinant tikrosios ankstinių pasiskirstymo struktūros Lasso prognozė tikslumu aplenkia MIDAS augant imties dydžiui. Šis darbas tiria pasiskirsčiusių ankstinių modelio apribojimų sukeltas problemas naujame MIDAS kontekste. Tikimasi, kad tyrimo rezultatai leis geriau suprasti griežtų apribojimų poveikį modelio prognozės tikslumui ir suteiks pagalbos sprendžiant modelio parinkimo problemą taikomajame darbe.

**Raktiniai žodžiai:** MIDAS, LASSO, ankstinių pasiskirstymas, prognozė, skirtingi dažniai, U-MIDAS, glodumas, apribojimas, svoriai.

# Comparison of MIDAS and LASSO Estimators

## Abstract

Mixed Frequency Sampling (MIDAS) regression is based on a restricted lag distribution structure which requires some restrictions, such as smoothness or positivity, to achieve more accurate forecasts. An unrestricted MIDAS regression with Lasso (Least Absolute Shrinkage and Selection Operator), however, aims for better forecasts by simply reducing the dimension of a model. That is why the aim of this paper is to compare forecasting errors of these two regressions, when the restrictions do not represent reality correctly, but the adequacy of model specifications are not rejected. Using the imitation analysis it is shown that while the restricted MIDAS usually gives a better forecast than unrestricted with Lasso, when the real process is unknown and specified inaccurately (but similar enough), the Lasso forecast error becomes smaller than MIDAS while the sample size is increasing. This paper considers problems caused by the restrictions forced on the lag distribution in the new MIDAS context. The results of the research allow a better understanding of the effects of restrictions in forecasting accuracy and offer aid in choosing models for applied work.

**Key words:** MIDAS, LASSO, distributed lag, forecast, mixed frequency, U-MIDAS, smooth, restriction, weights.

# Turinys

Įvadas	4
<b>1 Literatūros apžvalga</b>	<b>5</b>
<b>2 MIDAS struktūros apribojimai bei alternatyvūs modeliai</b>	<b>6</b>
2.1 MIDAS su pasiskirsčiusių ankstinių struktūra . . . . .	6
2.2 Neapribotas skirtingo dažnio duomenų modelis U-MIDAS . . . . .	8
2.3 U-MIDAS su Lasso kintamųjų atrinkimu . . . . .	8
2.4 Tvirtinimas apie apribojimų poveikį prognozės tikslumui . . . . .	9
<b>3 Prognozės tikslumo palyginimas: imitacinė analizė</b>	<b>11</b>
3.1 Tyrimo metodologija . . . . .	11
3.2 Modelių palyginimas, kai svoriai aprašomi eksponentiniu Almon ankstinių polinomu . . . . .	12
3.3 Modelių palyginimas, kai svoriai aprašomi Gompertz polinomu . . . . .	17
3.4 Modelių palyginimas, kai svoriai aprašomi Beta polinomu . . . . .	18
3.5 Palyginimas kelių kintamųjų atveju . . . . .	19
3.6 Rezultatai . . . . .	20
<b>4 Išvados</b>	<b>21</b>
<b>5 Literatūra</b>	<b>23</b>
<b>A Priedas</b>	<b>24</b>
A.1 Almon polinomo atvejis. Įverčių atstumas nuo tikrųjų parametrų reikšmių . .	24
A.2 Almon polinomo atvejis. Konkretus įverčių palyginimas . . . . .	26
A.3 Gompertz polinomo atvejis. MSE rezultatai . . . . .	28
A.4 Beta polinomo atvejis. MSE rezultatai . . . . .	32

## Įvadas

Skirtingo dažnio duomenų modeliai (angl. Mixed Data Sampling, toliau - MIDAS) pastarąjį dešimtmetį tapo plačiai naudojamu įrankiu ekonominių rodiklių prognozavimui, ypač pasitvirtinus, kad jų prognozė tikslesnė nei agreguotų duomenų modelių. Tačiau pastaraisiais metais literatūroje susidomėta, kaip galima pagerinti arba tiesiog aplenkti MIDAS prognozę kitais metodais, kurie galbūt apimtų didesnę kintamųjų skaičių arba nereikalautų tokių griežtų apribojimų.

Atsižvelgiant į tai, kad šie MIDAS struktūros apribojimai praktikoje gali dažnai nepasitvirtinti, tačiau skirtingo dažnio duomenų nauda yra akivaizdi, norima atrasti būdą susitvarkyti su visomis išvardintomis problemomis. Taigi, kyla klausimas, ar nenaudingiau nagrinėti tiesinį MIDAS modelio analogą su silpnesniais apribojimais, tačiau, nenorint prarasti dimensijos mažinimo naudos, taikyti kintamųjų atrinkimo metodą. Vienas iš tokių galimų metodų, nagrinėjamų ir šiame darbe, yra Lasso (angl. Least Absolute Shrinkage and Selection Operator), leidžiantis tam tikrus įvertinius prilyginti nuliui.

Todėl darbo tikslas yra ištirti, kaip veikia korektiškas MIDAS modelis prognozės atžvilgiu lyginant su Lasso modeliu, nors tikri kintamųjų sąryšiai netenkina apribojimų. Tiksliau, nėra aprašomi glodžia apribota svorių funkcija su teigiamomis reikšmėmis.

Darbe siekiama išspręsti toliau išvardintus uždavinius. Pirma, siekiama teoriškai paaiškinti, kaip pastarieji apribojimai veikia visus pasirinktus modelių tipus. Antra, taikant Monte Karlo (angl. Monte Carlo, toliau - MC) metodą generuojamiems duomenims siekiama parodyti, kaip apribojimų naudojimas nežinant tikrojo procesų sąryšio veikia MIDAS, U-MIDAS bei U-MIDAS su Lasso prognozių tikslumą taikant pagrindines svorių funkcijas (Almon, Beta ir Gompertz polinomų šeimos). Tam lyginamos modelių vidutinės kvadratinės prognozės paklaidos (angl. Mean Squared Forecast Error, toliau - MSFE/MSE) augančiam imties dydžiui.

Apie pasiskirsčiusių ankstinių modelio apribojimų sukeltas problemas rašė Sims (1971, 1973) vienodo dažnio duomenų kontekste, tačiau daugiau dėmesio ši tema nesulaukė. Todėl šiame darbe grįžtama prie autoriaus minėtų problemų, tačiau naujame MIDAS kontekste. Šis tyrimas leidžia pažvelgti į griežtų apribojimų poveikį modelio prognozės tikslumui ir pritaikyti šią informaciją taikomajame darbe sprendžiant modelio parinkimo problemą.

# 1 Literatūros apžvalga

MIDAS modelis pasiūlytas Ghysels, Santa-Clara ir Valkanov (2002, 2004). Autoriai pristatė modelį kaip pasiskirsčiusių ankstinių modelio (angl. Distributed Lag Model, toliau - DLM) struktūrą turintį skirtingų dažnių modelį bei įrodė, kad jiems galioja panašios savybės. Viena iš jų - tai, kad ankstinių pasiskirstymas aprašomas funkcija, tenkinančia tam tikrus *a priori* apribojimus. Tariaama, kad tikroji funkcija yra absoliučiai integruojama<sup>1</sup>, nykstanti, dažnai ir glodi<sup>2</sup>, o jos reikšmės - teigiamos. Tokie apribojimai, nors suteikia naudos, yra ir kritikuojami.

Plačiau apie minėtus apribojimus rašyta dar Sims (1971, 1973) vienodo dažnio ir vieno kintamojo DLM kontekste. Pasak autoriaus, funkcijos glodumas ir reikšmių teigiamumas ankstinių pasiskirstymo struktūrai yra ne tik per griežta, bet praktikoje dažniausiai ir neteisinga prielaida, kuri nėra paremiama ekonominės teorijos išvadomis. Daugiau nei vieno kintamojo atveju apie tai užsiminta ir Geweke (1978) tyrime, kuriame pateikiamos panašios išvados. Tačiau plačiau šie apribojimai nebuvo tyrinėti nei teoriškai, nei empiriškai.

Tuo tarpu MIDAS modeliai per pastarąjį dešimtmetį tapo dažna ekonominių rodiklių prognozavimo priemone, kuri duoda tikslesnius rezultatus nei agreguotų duomenų modeliai (pvz., Clements, Galvão (2007) parodo, kad mėnesiniai duomenys naudingi ketvirtinio JAV BVP prognozei). Tačiau pastaraisiais metais susidomėta ir tuo, kaip didelis kintamųjų skaičius gali veikti paties MIDAS modelio prognozės tikslumą bei kaip jį pagerinti. Pavyzdžiui, Marsilli (2014) pasiūlyta sumažinti modelio dimensiją įprastam MIDAS modeliui pritaikius kintamųjų atrinkimo metodus, tokius kaip Lasso ar Bayeso. Tačiau čia nekreipiama dėmesio į pačią modelio struktūrą, o visas dėmesys skiriamas empiriniam prognozių tikslumo palyginimui prieš ir po kintamųjų atrinkimo.

Kitą vertus, Uematsu ir Tanaka (2015) pasiūlo alternatyvų būdą - kiekvieno kintamojo ankstinys laikomas atskiru kintamuoju ir tuomet vykdomas kintamųjų atrinkimas papildomai neapribojant modelio struktūros, o toks modelis pavadinamas P-MIDAS. Tačiau ir čia autorių tikslas yra rasti geresnį modelį esant itin dideliame kintamųjų skaičiui. Be to, tyrimo išvados, kaip ir Marsilli (2014), stipriai remiasi konkrečiau empirinio pavyzdžio rezultatais.

Todėl šiame darbe nagrinėjama dar Sims (1971) aprašyta *a priori* ankstinių pasiskirstymo funkcijos apribojimų įtaka MIDAS modelių prognozės tikslumui. Tačiau alternatyvus modelis, su kuriuo lyginama, remiantis Uematsu ir Tanaka (2015), pasirenkamas toks, kurio struktūra nėra apribota, tačiau neprarandama ir dimensijos sumažinimo nauda. Tai yra, tiesinis neapribotas MIDAS modelis, kuriam taikomas Lasso operatorius, tam tikrus parametrus prilyginantis nuliui. Lyginimas remiasi ir imitacine analize.

<sup>1</sup>Funkcija  $f : X \rightarrow \mathbb{R}$  absoliučiai integruojama, jei  $\int |f|d\mu < \infty$ .

<sup>2</sup>Funkcija  $f$  yra glodi, jei  $f \in C^\infty$ . T.y., ji be galo diferencijuojama ir kiekviena išvestinė  $f^{(1)}, f^{(2)}, \dots$  yra tolydi. Funkcija  $f : A \rightarrow \mathbb{R}$  yra tolydi, jei  $\lim_{x \rightarrow y} f(x) = f(y)$ , su  $x, y \in A \subset \mathbb{R}$ . T.y., su  $x$  norimai arti  $y$ ,  $f(x)$  ir  $f(y)$  taip pat arti.

## 2 MIDAS struktūros apribojimai bei alternatyvūs modeliai

Sims (1971, 1973) rašė apie galimas nekorektiškų arba per griežtų ankstinių pasiskirstymo funkcijų apribojimų pasekmes DLM kontekste. Todėl šiame skyriuje pristatoma, kaip ši problema atrodo MIDAS regresijos kontekste. Be to, iškeliamas pagrindinis tvirtinimas, kurio teisingumą siekiama parodyti šiame tyrime.

### 2.1 MIDAS su pasiskirsčiusių ankstinių struktūra

MIDAS modelio sąryšis su DLM buvo pristatytas Ghysels et al. (2004). Tarkime, turime du tolydaus laiko procesus  $y_t$  ir  $x_{tm_i}^{(i)}$ ,  $i = 1, \dots, k$ , kurie yra stacionarūs<sup>3</sup>. Tegul jų realizacijos diskrečiais laiko momentais yra žymimos  $Y_t$  ir  $X_{tm_i}^{(i)}$ . Čia  $t \in \mathbb{R}$  žymi laiką,  $i = 1, \dots, k$  - nepriklausomus kintamuosius, o  $m$  yra dažnių santykis (pvz.,  $m = 3$ , kai  $y$  yra ketvirtiniai, o  $x$  - mėnesiniai duomenys).

Pagrindiniai apribojimai, kurie sukelia MIDAS problemas, yra tie patys, kurie sukelia problemas DLM modelyje (kaip jau buvo minėta, Ghysels et al. (2004) straipsnyje įrodoma, kad abu modeliai turi tokias pačias savybes). Iš tikro, modelius galime užrašyti panašia forma. Tegul tolydaus laiko DL modelis vienodo dažnio duomenims užrašomas kaip

$$y_t = \sum_{i=1}^k b^{(i)} * x_t^{(i)} + u_t,$$

čia  $b(t)$  yra funkcija, aprašanti  $y$  ir  $x$  sąryšį ankstinių pasiskirstymo atžvilgiu (kitai sakant, svorių funkcija). Reikalinga prielaida, kad  $x$  yra egzogeniškas, t.y.,  $E[x_t u_s] = 0$ ,  $\forall t, s$ . Paklaidos neprivalo būti i.i.d. Ženklas  $*$  žymi funkcijų sąsūką (angl. convolution)<sup>4</sup>, kuri tolydaus laiko atveju užrašoma kaip

$$(f * g)(t) = \int_{-\infty}^{\infty} f(\tau)g(t - \tau)d\tau = \int_{-\infty}^{\infty} g(\tau)f(t - \tau)d\tau,$$

o diskretaus laiko atveju

$$(f * g)(t) = \sum_{\tau=-\infty}^{\infty} g(\tau)f(t - \tau).$$

Iš čia išplaukia ir atitinkamai užrašytas diskretaus laiko MIDAS modelis

$$y_t = \sum_{i=1}^k \beta_i^{(i)} \left( b^{(i)} * X_{tm_i}^{(i)} \right) + \varepsilon_t = \sum_{i=1}^k \beta_i^{(i)} \sum_{j=0}^{l_i} b_j^{(i)} X_{tm_i-j}^{(i)} + \varepsilon_t.$$

<sup>3</sup>Procesas  $X$  yra stacionarus, kai jo vidurkis  $\mu := EX_t = EX$  nepriklauso nuo laiko ir kovariacija  $\gamma(k) := E(X_t - \mu)(X_{t+k} - \mu)$  priklauso tik nuo  $k$ .

<sup>4</sup>T.y., svertinis  $f(\tau)$  vidurkis momentu  $t$ , kai svoriai duoti  $g(-\tau)$  pastūmus per  $t$ .



Abiems modeliams (tiek DLM, tiek MIDAS) dažniausiai taikomi panašūs apribojimai. Dažniausiai sutinkami iš jų yra apribojimai, teigiantys, kad svorių funkcija, aprašanti tikrąjį modelio ankstinių pasiskirstymą, yra glodi ir sumuojasi į 1-ą bei jos reikšmės yra teigiamos. Pavyzdžiui,

$$\omega_j(\delta; k) = \frac{\psi(\delta; k)}{\sum_{k=1}^K \psi(\delta; k)},$$

čia  $k$  rodo ankstinių skaičių, oiekiant sumažinti modelio dimensiją ir vertinti mažiau parametru, svorių funkcija normalizuojama, tai yra, jos reikšmės sumuojasi į 1. Dažniausiai literatūroje ir praktiniuose tyrimuose naudojami tokie svoriai:

1. Eksponentinis Almon ankstinių polinomas:  $\psi(\delta; k) = \exp\{\sum_{j=1}^p \delta_j k^j\}$ ,  $p \in \mathbb{N}$ ,  $\delta \in \mathbb{R}^p$ .
2. Beta polinomas:  $\psi(\delta_1, \delta_2; k) = x_k^{\delta_1-1} (1-x_k)^{\delta_2-1}$ ,  $x_k = \xi + (1-\xi)h(k)$ ,  $h(k) = \frac{k-1}{m-1}$ ,  $\xi > 0$ .
3. Gompertz:  $\psi(\delta_1, \delta_2; k) = e^{-\delta_1 e^{\delta_2 k}} e^{\delta_2 k} = \exp\{\delta_2 k - \delta_1 e^{\delta_2 k}\}$ .

Aptartos prielaidos naudingos tam tikrais atvejais. Kadangi modeliuojamas priklausomas kintamasis nuo kito kintamojo ankstinių, galima multikolinearumo problema, kai ankstiniai stipriai koreliuoja tarpusavyje. Tuomet įvertinai gali turėti dideles standartines paklaidas. Tai reiškia, kad įvertiniai gali būti ne tik netikslūs, bet ir įvertinami kaip nereikšmingi, turėti priešingą ženklą, nei tikrojo parametro reikšmės ženklas. Todėl  $b_j$  apribojant kaip glodžią funkciją su teigiamomis reikšmėmis išvengiama galimos problemos.

Be to, kaip jau minėta šiame skyriuje, toks apribojimas yra ir dimensijos mažinimo būdas. Aprašius  $b_j$  funkcija, kuri sumuojasi į 1-ą, pakanka vertinti tik atitinkamus hiperparametrus, kas pagreitina ir palengvina modelio vertinimą, ypač turint daug kintamųjų (eksponentinis Almon polinomas ir kt.).

Tačiau jei multikolinearumo nėra (arba jis silpnas) tuomet gauti įvairūs įvertinių ženklai bei nereikšmingumas gali būti tiesiog todėl, kad tokie ir yra tikrieji parametrai. Tokiu atveju glodumo apribojimas ne tik nereikalingas, bet gali ir pakenkti prognozės tikslumui. Pagal Sims (1973), standartinės įvertinių paklaidos, gautos vertinant regresiją paremtą tiksliais, bet neteisingais a priori apribojimais, nėra naudingos. Negana to, neteisinga yra ir tarti, kad ankstinių pasiskirstymo glodumas išplaukia iš ekonominės teorijos, nes toks teiginys taip pat yra labai silpnas.

Taigi, jei multikolinearumo problemos nėra, o ekonominė teorija nesiūlo stiprių argumentų, paremiančių griežtus apribojimus, tuomet jų taikymas turėtų būti nenaudingas prognozavimui. Tokiu atveju regresija, kuri gali vertinti skirtingų dažnių kintamuosius ir sugeba sumažinti modelio dimensiją, tačiau taip griežtai neapriboja ankstinių struktūros, turėtų prognozuoti tiksliau. Būtent tokia regresija siūloma naudoti neapribotą tiesinį U-MIDAS modelį su Lasso kintamųjų atrinkimu. Šios alternatyvos nauda pagrindžiama tolimesniuose skyreliuose.

## 2.2 Neapribotas skirtingo dažnio duomenų modelis U-MIDAS

Jei ankstiniai tarpusavyje stipriai nekoreliuoja bei ekonominė teorija neturi stiprių argumentų dėl ankstinių pasiskirstymo struktūros, tuomet glodumo bei kitos aptartos prielaidos nėra reikalingos. Tokiu atveju galima naudoti paprasčiausią neapribotą modelį U-MIDAS (angl. unrestricted). Kaip ir siūlo jo pavadinimas, šis modelis neapriboja ankstinių struktūros, parametrų ženklų ar reikšmingumo, ir vertinamas tiesiog kaip tiesinis modelis.

Pasiūlytas Foroni, Marcellino bei Schumacher (2011), šis modelis ypač naudingas kai dažnių skirtumai nėra dideli. Tačiau didžiausias pranašumas, lyginant su apribotu MIDAS, yra būtent tai, kad ankstinių pasiskirstymas neapriojamas:

$$Y_t = \sum_{i=1}^k \sum_{j=0}^{l_i} \beta_j^{(i)} X_{tm_i-j}^{(i)} + U_t,$$

čia  $\beta_j^{(i)} \in \mathbb{R}$ . Kiekvienas ankstinys vertinamas kaip atskiras kintamasis.

Tačiau lyginant su apribotu MIDAS modeliu, U-MIDAS turi trūkumą - nesumažinama dimensija, kuri verčia mus vertinti didesnę parametrų skaičių. Tai gali atsilipti modelio prognozės tikslumui, ypač turint didelį kintamųjų skaičių. Todėl verta taikyti papildomą dimensijos mažinimo metodą. Šiame tyrime pasirinktas kintamųjų atrinkimo metodas, kuris leidžia tam tikrus parametrus paprasčiausiai priskirti nuliui. Tai yra Lasso, apie kurio pranašumą rašoma 2.3. skyrelyje.

## 2.3 U-MIDAS su Lasso kintamųjų atrinkimu

Lasso (angl. Least Absolute Shrinkage and Selection Operator) pasiūlytas Tibshirani (1996). Šio operatoriaus tikslas yra modelio parametrų dimensijos mažinimas tikslesnei prognozei gauti. Įprasto tiesinio modelio įvertiniai yra nepaslinkti<sup>5</sup>, bet nebūtinai tikslūs. Naudojant Lasso kai kurie parametrai, kurie yra pakankamai arti 0, yra priskiriami 0. Tokiu būtu gaunamas paslinktumas, tačiau sumažinama įvertinių variacija, o kartu ir prognozės paklaida (bias - variance trade off)<sup>6</sup>.

Kadangi taikomas tiesiniam U-MIDAS modeliui, jis neapriboja parametrų ar svorių struktūros ar parametrų ženklų. Todėl gali būti taikomas nepaisant to, ar tikrasis sąryšis turėtų būti aprašomas glodžia funkcija, ar ne.

Galutinis modelis (toliau paprastumo dėlei vadinamas tiesiog Lasso)

$$Y_t = \sum_{i=1}^k \sum_{j=0}^{l_i} \beta_j^{(i)} X_{tm_i-j}^{(i)} + \varepsilon_t,$$

su įvertinio radimo problema

---

<sup>5</sup>Parametro  $\theta$  įvertinys  $\hat{\theta}$  yra nepaslinktas, jei  $E[\hat{\theta}] = \theta$ .

<sup>6</sup>Siekiant sumažinti prognozės paklaidą kartais pasirenkamas vertinimo metodas, kuriuo gaunama mažesnė įvertinių variacija, bet jie paslinkti. Jei neigiama mažesnės variacijos įtaka prognozės paklaidai viršija teigiamą poslinkio įtaką, tuomet pasirinkto metodo naudojimas pasiteisina.

$$\hat{\beta} = \operatorname{argmin}_{\beta} \sum_{t=1}^T \left( Y_t - \sum_{i=1}^k \sum_{j=0}^{l_i} \beta_j^{(i)} X_{tm_i-j}^{(i)} \right)^2$$

su sąlyga  $\sum_{i=0}^k \sum_{j=0}^{l_i} |\beta_j^{(i)}| \leq \lambda,$

su  $\lambda \geq 0$ . Čia  $\lambda$  yra parametras, kuris reguliuoja, kiek įvertinių bus prilyginama nuliui (kuo didesnis, tuo daugiau įvertinių priartinami bei prilyginami 0-iui). Vienas dažniausiai naudojamų būdų parinkti  $\lambda$  yra kryžminė patikra<sup>7</sup> (angl. cross-validation).

Kadangi toks modelis ne tik neapriboja ankstinių pasiskirstymo struktūros dėl to, kad taikomas pagal U-MIDAS, bet ir Lasso dėka sumažina dimensiją, tai tikimasi, kad jo prognozė bus tikslesnė tuomet, kai MIDAS apribojimai per griežti realiai situacijai.

## 2.4 Tvirtinimas apie apribojimų poveikį prognozės tikslumui

Kaip aprašyta šiame skyriuje, ankstinių multikolinearumas bei stiprios ekonominės teorijos implikacijos yra vieninteliai argumentai paremiantys ankstinių pasiskirstymo a priori apribojimus bei jų naudą prognozavimui. Remiantis Sims (1973), antrasis argumentas yra itin mažai tikėtinas. Taigi, jei ankstiniai nėra stipriai koreliuoti, tuomet apribota MIDAS regresija neturėtų būti geresnė prognozės tikslumo atžvilgiu už siūlomą alternatyva - neapribotą U-MIDAS su Lasso.

---

<sup>7</sup>Kryžminė patikra yra modelio vertinimas daliai imties. Tuomet įvertinus modelį imties poaibiui tikrinama, kaip jis veikia pridėjus naują (t.y., likusią) dalį duomenų. Toks metodas leidžia įvertinti modelio veikimą norint prognozuoti ateitį su naujais duomenimis.

Apibendrinus šias idėjas, galima suformuluoti pagrindinį šio darbo tvirtinimą:

*Neturint multikolinearumo problemos ir nežinant tikrosios ankstinių pasiskirstymo funkcijos formos, neapribotos U-MIDAS regresijos su Lasso prognozė tampa tikslesnė nei apribotos MIDAS regresijos.*

Būtent šio tvirtinimo tikrinimas ir yra pagrindinė tyrimo užduotis, kurios eiga ir rezultatai aprašomi tolimesniuose skyriuose.

### 3 Prognozės tikslumo palyginimas: imitacinė analizė

Šiame skyriuje tiriama, kokį poveikį ankstinių pasiskirstymo funkcijos apribojimai turi MIDAS modelio prognozės tikslumui lyginant su U-MIDAS bei Lasso modeliais, kurie buvo aprašyti 2 skyriuje. Todėl naudojama Monte Karlo (MC) metodu besiremianti simuliacinė analizė, kuri leidžia vertinti atitinkamus modelius didelį skaičių kartų bei gauti apibendrintus rezultatus augančiam imties dydžiui. Toliau pateikiama tikslesnė tyrimo metodologija bei gauti rezultatai.

#### 3.1 Tyrimo metodologija

Pirmiausia, kadangi norima tirti tokias situacijas, kai tikroji svorių funkcija (nesvarbu, žinoma, ar nežinoma) yra modeliuojama kaip įprastinė (glodi, teigiamų reikšmių) svorių funkcija MIDAS regresijos atveju, tai pirmas procedūros etapas yra tikrųjų duomenų ir ankstinių pasiskirstymo generavimas vieno kintamojo atveju. Šiuo atveju pasirinkti stacionarūs duomenys, pasiskirstę pagal normalųjį skirstinį. Tuomet kiekvienam atvejui taikoma atitinkama svorių funkcija, kurios forma aprašoma atskirai kiekvienam konkrečiam atvejui.

Vertinami trys aptarti modeliai: MIDAS, U-MIDAS ir U-MIDAS su Lasso, kuris toliau bus vadinamas tiesiog Lasso. Bendrai vertinamas modelis su  $l$  ankstinių atrodo taip

$$Y_t = \alpha + \varphi t + \beta \sum_{q=0}^l \omega_q X_{tm-q} + \varepsilon.$$

Svarbu pabrėžti, kad pradinis MIDAS modelis kiekvieną kartą aprašomas taip, kad jo funkcinės formos adekvatumas nebūtų atmetamas atitinkamų testų<sup>8</sup>. Tai yra, nors įverčiai nėra tikslūs, jie pakankamai arti tikrųjų reikšmių, kad modelio funkcinė forma būtų neatmesta kaip korektiška. Vertinti nekorektišką modelį būtų beprasmiška, nes jis visuomet prognozuos prasčiau nei korektiškas. Be to, tokia situacija atspindi realų tyrimą, kai tikslūs parametrai negali būti žinomi, bet juos norima įvertinti kuo tiksliau.

Modeliai vertinami augančiam tiek imties dydžiui ( $N = 100, 250, \dots, 1250, 1500$ ), tiek santykinai augančiam ankstinių skaičiui ( $l = \sqrt{N}$ ), kad būtų kuo tiksliau atspindimos asimptotinės modelių prognozių savybės. Tuomet vykdoma  $h$  žingsnių prognozė ( $h$  parenkamas kaip dalis nuo imties dydžio) ir skaičiuojamas kriterijus, leidžiantis palyginti kiekvieno modelio prognozės tikslumą - vidutinė kvadratinė paklaidos prognozė MSFE (toliau tiesiog MSE)<sup>9</sup>. Kiekvienam imties dydžiui procedūra kartojama 1000 kartų ir imami vidutiniai rezultatai. Tai leidžia susidaryti vieningą bendrą vaizdą apie modelių elgesį.

Be to, įsitikinama, kad kintamieji ir ankstiniai tarpusavyje stipriai nekoreliuotų. Tam

---

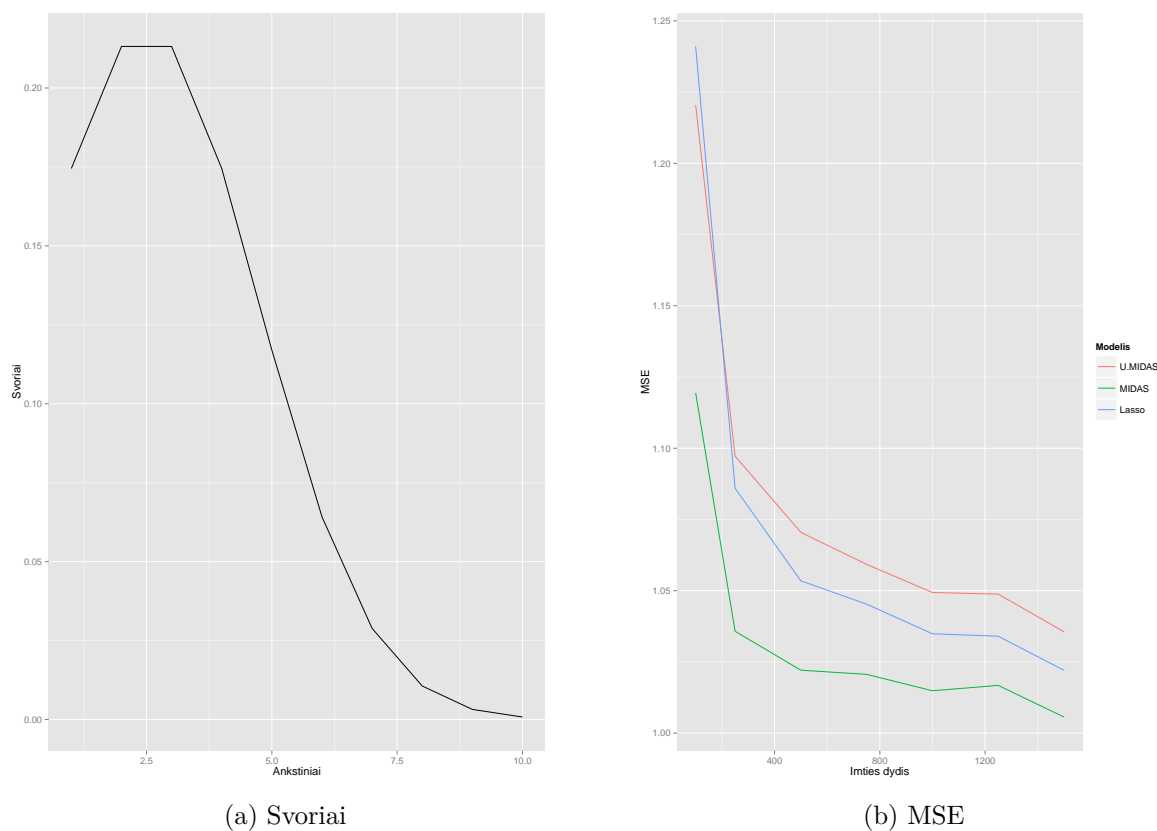
<sup>8</sup>Specifikacijos adekvatumo testas  $(\hat{\theta} - \tilde{\theta})A(\hat{\theta} - \tilde{\theta}) \sim \chi^2(d - q)$ , čia  $A$  yra normalizuojanti matrica,  $d = \dim(\theta)$  ir  $q = \dim(\gamma)$ , o  $\hat{\theta}$  yra neapriboto modelio įvertiniai, o  $\tilde{\theta} = \hat{f}_\gamma$  - apriboto. Plačiau skaityti [4].

<sup>9</sup> $MSFE = E[(\hat{Y}_i - Y_i)^2]$

naudojamas VIF kriterijus<sup>10</sup> (angl. Variation Inflation Factor), kuris leidžia įvertinti, kiek įvertinių variacijos yra per didelės ir ar yra multikolinearumas. Taip užtikrinamos teisingos sąlygos iškeltam tvirtinimui tirti.

### 3.2 Modelių palyginimas, kai svoriai aprašomi eksponentiniu Almon ankstinių polinomu

Tyrimas pradedamas nuo literatūroje dažniausiai sutinkamo Almon ankstinių polinomo šeimos. Čia vertinami 6 skirtingi procesai, atspindintys skirtingą su apribojimais susijusią situaciją.

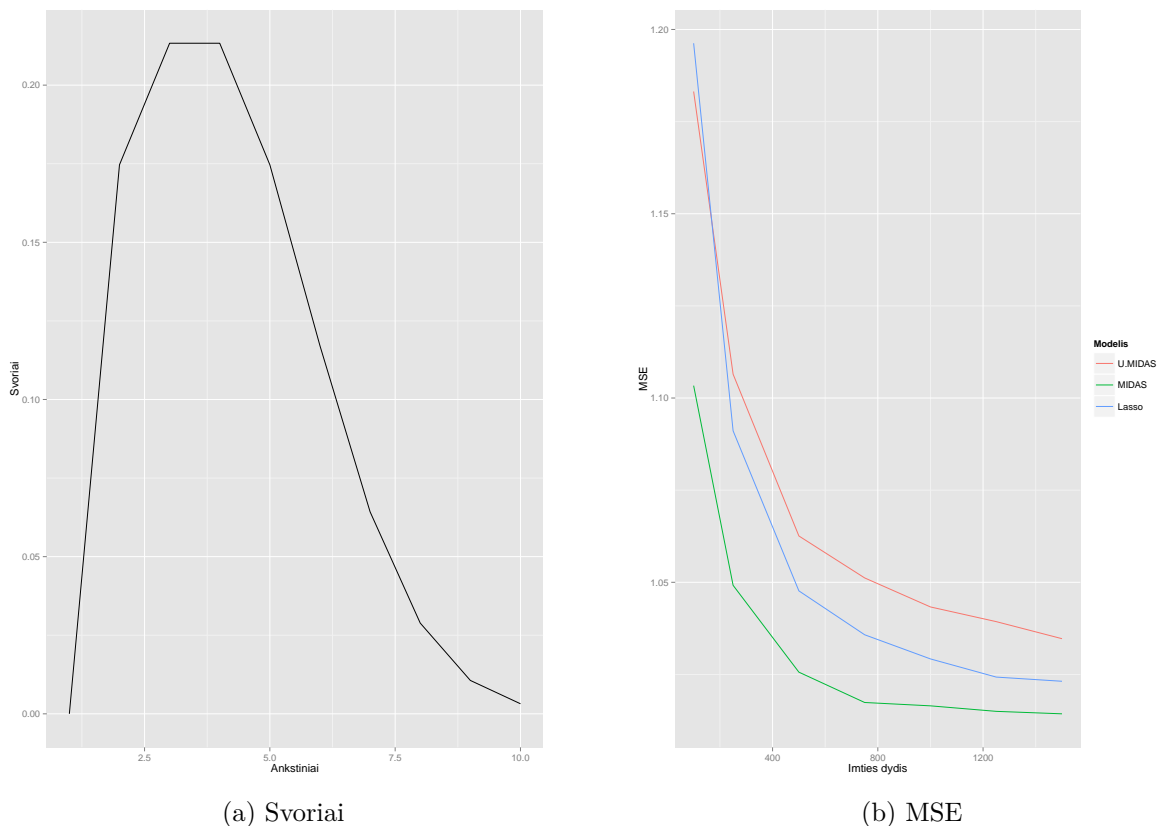


1 pav.: 1-ojo proceso svorių funkcija ir gauta prognozės paklaida MSE (raudona - U-MIDAS, žalia - MIDAS, mėlyna - Lasso)

<sup>10</sup> $VIF_k = \frac{1}{1-R_k^2}$ , čia  $R_k^2$  yra  $R^2$ , kai k-asis kintamasis regresuojamas nuo likusiųjų kintamųjų. Įprastai literatūroje, jei  $VIF < 4$ , tai multikolinearumo nėra ( $VIF = 1$  - kintamieji visiškai nekoreliuoja)

Pirmasis procesas aprašomas pagal svorių funkciją, kurią generuojanti funkcija yra glodi (žr. 1(a) pav.). Žinant, kad tikroji svorių funkcija iš tikro yra glodi, ją galima gana tiksliai aprašyti pagal tą patį eksponentinį Almon ankstinių polinomą, kuris ir buvo naudojamas tikrajam procesui gauti. Tačiau net jei tariama, kad tikroji funkcijos forma nežinoma, dėl jos savybių apribota svorių funkcija ją aprašys neblogai. Todėl nenuostabu, kad, nors ir visų vertintų modelių prognozė imčiai augant mažėja, MIDAS prognozuoja tiksliausiai (žr. 1(b) pav.).

Antrasis procesas panašus, tik pirmoji svorių funkcijos reikšmė - 0 (galimas atvejis, kai įtraukiamas nulinis ankstinys  $X_{tm}$ ). Tačiau toks staigus funkcijos pokytis taip pat nekeičia prognozių tikslumo rezultatų - MIDAS veikia geriausiai, Lasso kiek prasčiau, U-MIDAS prasčiausiai (žr. 2 pav.).



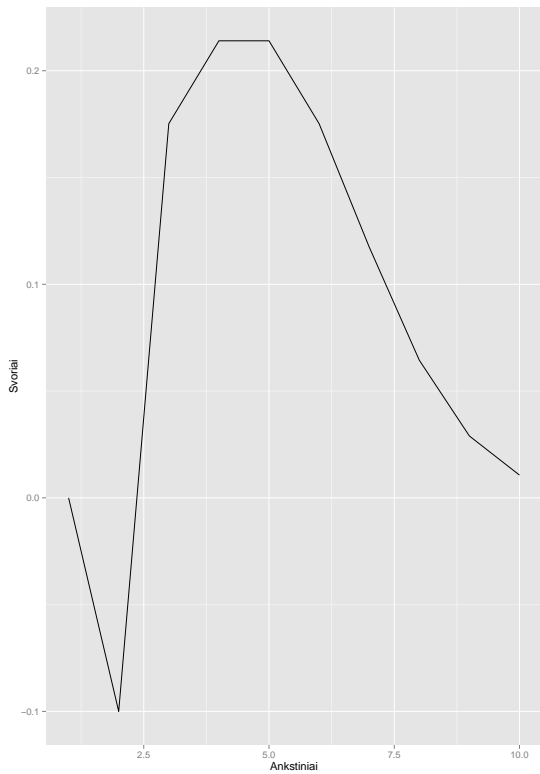
2 pav.: 2-ojo proceso svorių funkcija ir gauta prognozės paklaida MSE (raudona - U-MIDAS, žalia - MIDAS, mėlyna - Lasso)

Generuojant trečią procesą pridedamas "lūžis" bei neigiama svorių reikšmė (žr. 3(a) pav.). Taigi, tikrasis ankstinių pasiskirstymas pažeidžia abi prielaidas apie jo formą - ir glodumo, ir teigiamumo. Todėl ir rezultatai gaunami atitinkami: nors MIDAS lenkia kitus modelius pradžioje (iki  $N \approx 700$ ), toliau jo prognozės tikslumas tampa ne tik prastesnis už Lasso, bet ir pradeda nežymiai augti (žr. 3(b) pav.).

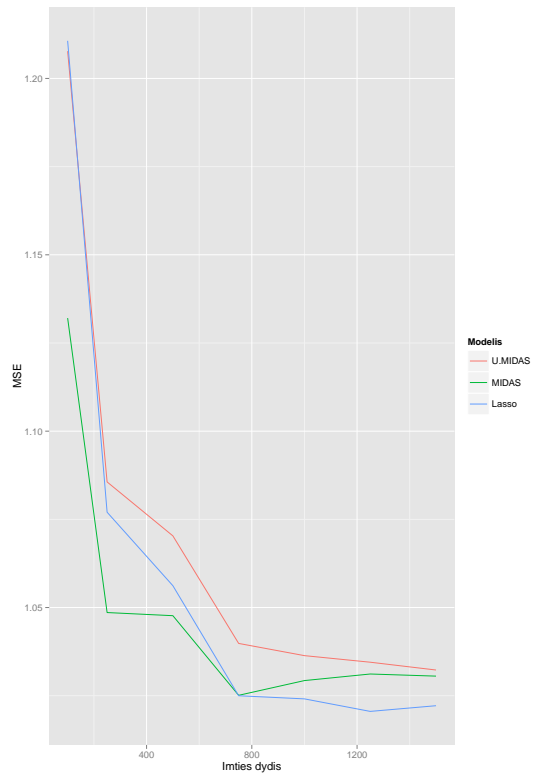
Kitiems 2-iems procesams pridedama dar daugiau "lūžių" ir neigiamų svorių, todėl gauti panašūs rezultatai (žr. 4-5 pav.). Augant imties dydžiui MIDAS prognozė ne tik tampa ne tokia tiksli kaip Lasso, tačiau tam tikrais atvejais paklaida ima augti.

Rezultatus patvirtina ir kitas modelio gerumo lyginimo kriterijus - vidutinis kvadratinis įverčių atstumas nuo tikrųjų parametrų reikšmių, kuris leidžia matyti, ar augant imties dydžiui įvertiniai artėja prie tikrųjų reikšmių. Šiuo atveju matomi rezultatai yra analogiški MSE rezultatams - apribojimų pažeidimai neigiamai veikia MIDAS įvertinių tikslumą augant imties dydžiui (žr. Priedas A.1. 8-10 pav.). Be to, Priede A.2. pateikiamas įverčių palyginimas konkrečioms atvejais, kaip pavyzdys, kaip MIDAS geriau vertina parametrus pirmiems procesams, tačiau Lasso sugeba įvertinti neigiamas reikšmes (žr. 11-13 pav.).

Visi pastarieji atvejai rėmėsi tuo, kad nepaisant tikrosios svorių funkcijos formos MIDAS svoriai buvo aprašomi apribota svorių funkcija. Tai yra, dažniausiai aprašomi nekorektiškai, nors ir pakankamai arti. T.y., naudotos svorių funkcijos, dažniausiai aptinkamos literatūroje.



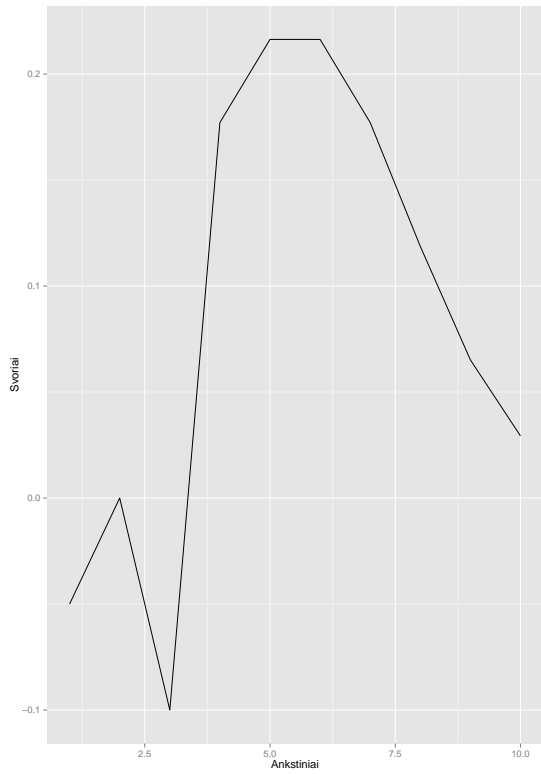
(a) Svoriai



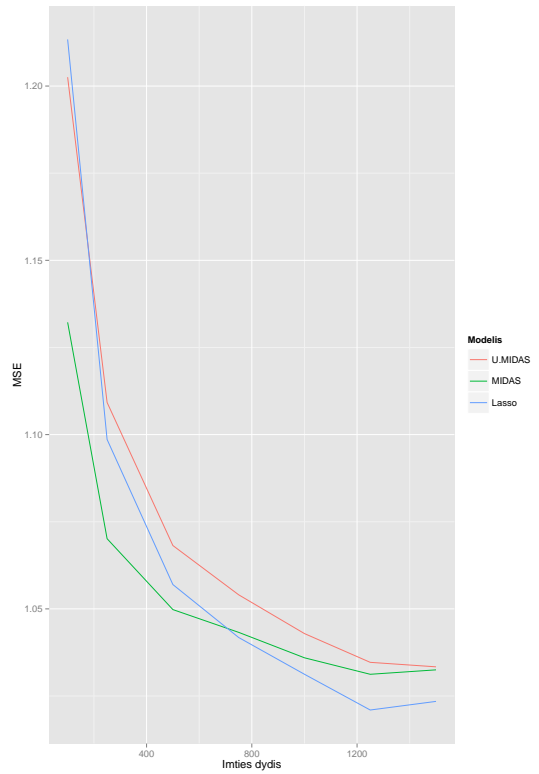
(b) MSE

3 pav.: 3-iojo proceso svorių funkcija ir gauta prognozės paklaida MSE (raudona - U-MIDAS, žalia - MIDAS, mėlyna - Lasso)



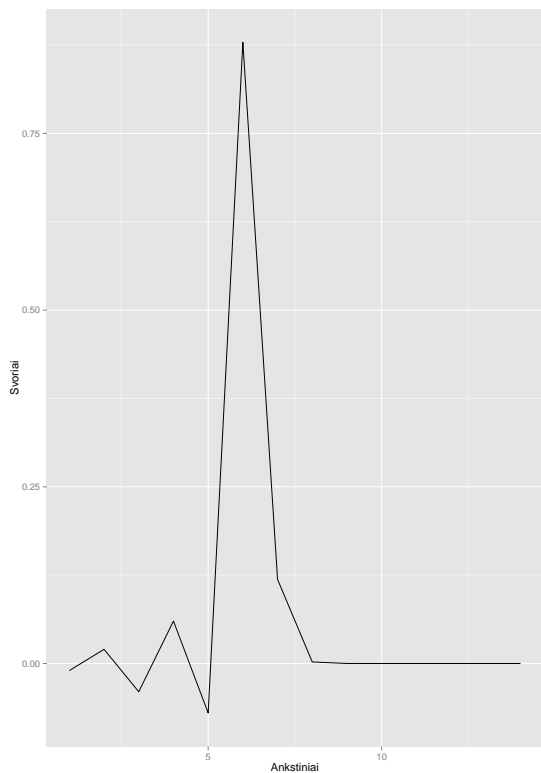


(a) Svoriai

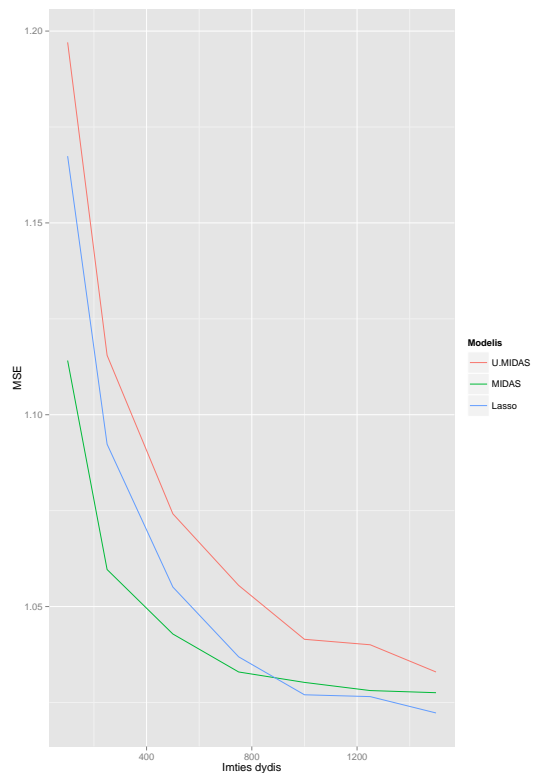


(b) MSE

4 pav.: 4-ojo proceso svorių funkcija ir gauta prognozės paklaida MSE (raudona - U-MIDAS, žalia - MIDAS, mėlyna - Lasso)

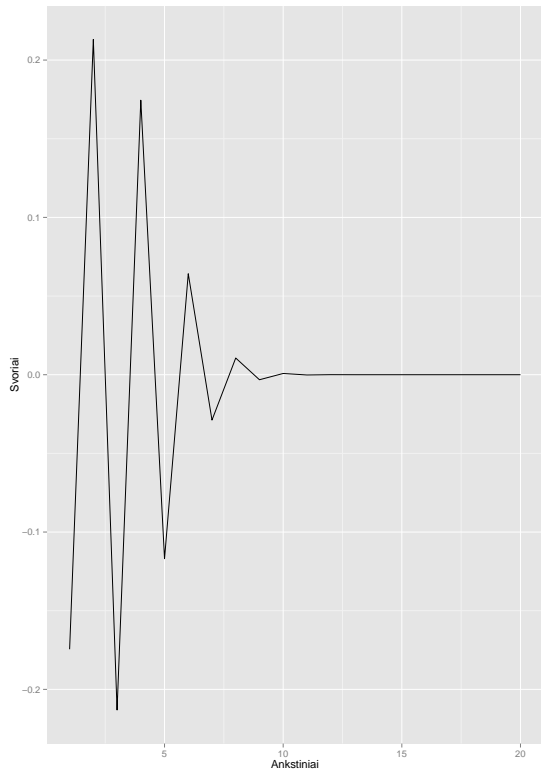


(a) Svoriai

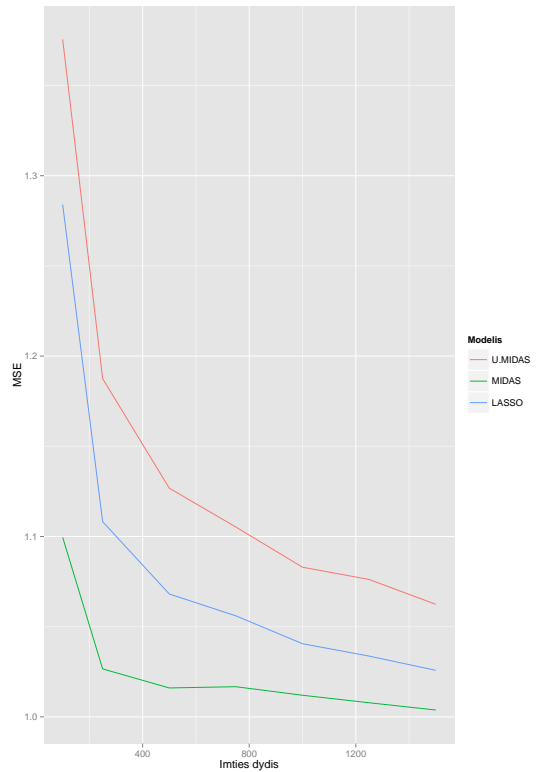


(b) MSE

5 pav.: 5-ojo proceso svorių funkcija ir gauta prognozės paklaida MSE (raudona - U-MIDAS, žalia - MIDAS, mėlyna - Lasso)



(a) Svoriai



(b) MSE

6 pav.: Alternuojanti svorių funkcija ir gauta prognozės paklaida MSE (raudona - U-MIDAS, žalia - MIDAS, mėlyna - Lasso)

Tačiau žinant tikrąjį ankstinių pasiskirstymą galima apsirašyti ir konkrečią svorių funkciją, kuri procesą aprašytų tiksliau, nors jai ir nebegaliojotų tiriami apribojimai. Tegul ši tikroji funkcija ir alternuojanti  $b(\theta, l) = (-1)^l f(\theta)$ , čia  $f$  - apribota svorių funkcija) (žr. 6(a) pav.). Tuomet akivaizdu, kad įprasta MIDAS svorių funkcija nesugebės teisingai įvertinti modelio, kurio ženklai yra alternuojantys. Iš tikro, tokiu atveju funkcinė forma atmetama, o vertinant prognozės paklaida paprasčiausiai auga didėjant imties dydžiui. Bet aprašius svorių funkciją, kuri atsižvelgtų į neigiamas reikšmes (taip pat alternuojanti) rezultatai rodo (žr. 6(b) pav.), kad žinant tikrąją svorių funkciją, kad ir kokia ji būtų, galima įvertinti tokį MIDAS modelį, kurio prognozė būtų geresnė už Lasso bet kokiam imties dydžiui iš tiriamų.

### 3.3 Modelių palyginimas, kai svoriai aprašomi Gompertz polinomu

Gompertz polinomu generuojami analogiški procesai, kaip ir Almon polinomo atveju. Pirmojo proceso ankstinių pasiskirstymas glodus (žr. 14(a) pav.), antrojo - su pradiniu nuliniu svoriu (žr. 15(a) pav.), trečiojo - su neigiama reikšme (žr. 16(a) pav.). Visais trimis atvejais MIDAS prognozė tiksliausia (žr. 14-16(b) pav.).

Kitiems atvejams pridedama daugiau lūžių ir neigiamų svorių (žr. 17(a), 18(a) pav.). Rezultatai panašūs į ankstesniojo skyrelio - augant imties dydžiui Lasso pradeda prognozuoti tiksliau nei MIDAS. Tačiau bendrai šio polinomo atveju rezultatai, tiksliau, skirtumas tarp prognozių tikslumo, net tokie akivaizdūs lyginant su Almon polinomo atveju. Skirtumas tarp prognozių paklaidos mažesnis, o MIDAS paklaidai padidėjus, ji išauga ne tiek daug, palyginus su praeito skyrelio rezultatais. Tačiau tai nekeičia fakto, kad prognozės tikslumu ir čia Lasso aplenkia MIDAS augant imties dydžiui.

Galiausiai, tariama, kad net jei funkcija netenkina apribojimų ir yra alternuojanti, bet žinoma (kaip ir Almon polinomo atveju), galima ją gana tiksliai aprašyti ir vertinti MIDAS regresiją. Gauti rezultatai (žr. 19 pav.) patvirtina šį teiginį. MIDAS prognozės paklaida gauta mažesnė už LASSO bet kokiam tirtam imties dydžiui.

### 3.4 Modelių palyginimas, kai svoriai aprašomi Beta polinomu

Beta polinomas, literatūroje naudojamas taip pat dažnai, nors kiek rečiau nei Almon. Šiame skyrelyje ir vėl taikoma panaši procedūra. Pirmojo generuojamo proceso ankstinių pasiskirstymas glodus (žr. 20(a) pav.), antrojo - su pradiniu nuliniu svoriu (žr. 21(a) pav.). Šiais atvejais, kaip ir ankstesniuose skyreliuose, MIDAS prognozės paklaida mažiausia visiems tiriamiems imties dydžiams (žr. 20(b), 21(b) pav.).

Kaip ir ankstesniais atvejais, pridėjus "lūžių" bei neigiamų reikšmių, MIDAS prognozė tampa prastesnė nei Lasso augant imties dydžiui. Beta polinomo atveju skirtumas gaunamas mažiausias ir prognozės paklaidos apsilenkia esant didesnei imčiai (žr. 22-23 pav.). Tačiau rezultatai panašūs. Išskyrus penktąjį procesą, kurio atveju MIDAS prognozė visuomet prastesnė už tiek U-MIDAS, tiek Lasso, nes įvertiniai įvertinami nepakankamai tiksliai (žr. 24 pav.)

Galiausiai, ir vėl generuojamas procesas su alternuojančia svorių funkcija. Tariant, kad tikroji modelio struktūra žinoma, galima gana tiksliai aprašyti procesą ir vertinti su apribotu MIDAS, kurio prognozė tiksliausia bet kokiam tiriamam imties dydžiui (žr. 25 pav.).

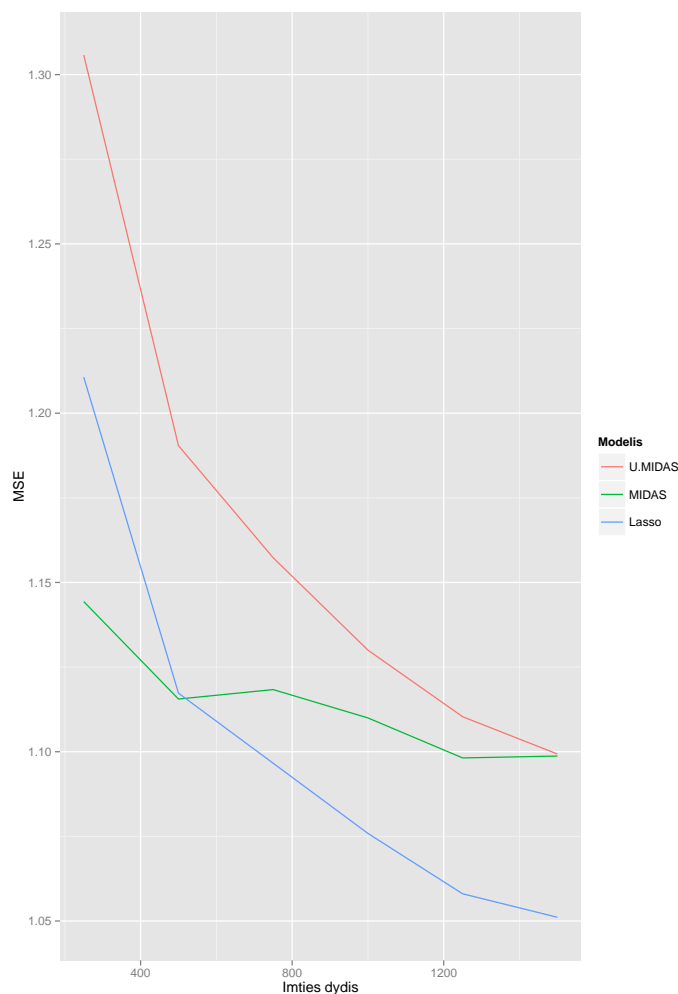
### 3.5 Palyginimas kelių kintamųjų atveju

Vieno kintamojo atveju situacija, kai tikrasis ankstinių pasiskirstymas aprašomas glodžia funkcija, tenkinančia apribojimus, duoda gana akivaizdžius rezultatus - modelis geras su tinkamais apribojimais, todėl prognozės tikslumas lenkia kitus. Tačiau įdomesnis atvejis, kai tikroji funkcija neatitinka apribojimų. Būtent tokia situacija susimuluota šiame skyrelyje kelių kintamųjų atveju. Modelis su  $i = 1, 2, 3$  aprašomas kaip

$$Y_t = \alpha + \varphi t + \sum_{i=1}^3 \beta_i \sum_{q=0}^{l_i} f_q^i X_{tm_i-q}^i + \varepsilon_t,$$

čia nagrinėjamas trijų kintamųjų atvejis, kai du iš jų aprašomi neglodžiu ankstinių pasiskirstymu, bet modeliuojami pagal ankstesnius pavyzdžius (konkrečiai, 3 ir 5 procesai iš 3.2 skyrelio) taip, kad funkcinės formos adekvatumo testai jų neatmestų. Be to, visi kintamieji skirtingų dažnių ( $m_1 = 4$ ,  $m_2 = 12$ ,  $m_3 = 260$ ).

Gauti rezultatai taip pat atitinka anksčiau pristatytus: MIDAS prognozė augant imties dydžiui tampa netokia tiksli lyginant su Lasso (žr. 13 pav.). Taigi, didesnis kintamųjų skaičius situacijos nekeičia. O remiantis Uematsu ir Tanaka (2015), dar didesnis kintamųjų skaičius suteiktų daugiau pranašumo Lasso tipo regresijai.



7 pav.: MSE 3-jų kintamųjų atveju

### 3.6 Rezultatai

Taigi, iš atlikto tyrimo matoma, kad MIDAS prognozė yra geriausia iš visų trijų modelių (MIDAS, U-MIDAS, Lasso) tik dviem atvejais: kai tikroji ankstinių pasiskirstymą aprašanti funkcija tenkina reikalaujamus apribojimus arba tuomet, kai tikroji funkcija yra žinoma ir ją galima tiksliai aprašyti ignoruojant apribojimus. Pastaruoju atveju tirti apribojimai netenka prasmės.

Darant prielaidą, kad Sims (1973) teiginys, jog glodus ir tik teigiamas sąryšis realybėje mažai tikėtinas, yra teisingas, tuomet pirmojo atvejo tikimybė - taip pat nedidelė. Be to, tikrajam ankstinių pasiskirstymui esant neglodžiam arba tokiam, kurio svoriai įgyja neigiamas reikšmes, apribotas MIDAS modelis gali būti priimtinas (funkcinės formos adekvatumo testo atžvilgiu), nors jo rezultatai (nei įvertiniai, nei prognozė) nėra geriausi.

Taigi, nežinant tikrosios ankstinių pasiskirstymą aprašančios funkcijos, kas yra įprasta dirbant su realiais duomenimis, naudingiau yra vertinti modelį, kuris turi dimensijos mažinimo savybių, tačiau neapriboja funkcinės modelio formos. Tai yra, U-MIDAS su Lasso. Tokiu atveju tiesiog apsisaugoma nuo galimų neteisingų svorių. Be to, prisiminus Uematsu ir Tana-

ka (2015) išvadą apie tai, kad MIDAS prognozuoja prasčiau turint didelį kintamųjų skaičių, galima ją papildyti, kad prie aptartų realių ir tikėtinų sąlygų, MIDAS prognozė prastėja net nereikalaujant didelio kintamųjų skaičiaus. Rezultatai matosi net vieno kintamojo atveju. Nėgana to, akivaizdu ir tai, kad Lasso visuomet tiksliau prognozuoja nei paprasta U-MIDAS regresija, todėl matoma beveik nuolatinė kintamųjų atrinkimo (dimensijos mažinimo) nauda neapribotam modeliui.

Tačiau verta atsižvelgti ir į tai, kad MIDAS prognozė ima prastėti (paklaida didesnė nei Lasso, o rečiau ir ima augti) turint didelę imtį ir netgi ne visais atvejais. Praktikoje imtys dažnai būna mažesnės, todėl konkrečiais atvejais MIDAS prognozė vistiek bus tikslesnė už Lasso net jei apribojimai ne visiškai atitinka realybę. Nors prognozės atžvilgiu tokia situacija palanki MIDAS, tačiau tokiu atveju norint pritaikyti modelį ateities prognozei visuomet išlieka neužtikrintumas, jog parinktas modelis nebeveiks.

Kitą vertus, jei tikroji ankstinių pasiskirstymo funkcija yra žinoma, kad ir kokia ji būtų, tuomet MIDAS modelis gali būti aprašomas tikslia svorių funkcija nepaisant apribojimų ir tokia regresija prognozuos geriau nei U-MIDAS ar Lasso.

Bendrai galima daryti išvadą, kad 2-ame skyriuje įvardintas tvirtinimas nėra klaidingas.

## 4 Išvados

Šio tyrimo tikslas buvo ištirti, kaip veikia korektiškas MIDAS modelis prognozės atžvilgiu lyginant su Lasso modeliui, nors tikri kintamųjų sąryšiai gali netenkinti įprastų glodumo ar reikšmių teigiamumo apribojimų.

Taigi, išsiaiškinta, kad MIDAS prognozė visuomet yra tikslesnė už U-MIDAS bei U-MIDAS su Lasso dviem atvejais: kai tikroji ankstinių pasiskirstymą aprašanti funkcija tenkina reikalaujamus apribojimus arba tuomet, kai tikroji funkcija yra žinoma ir ją galima tiksliai aprašyti.

Pirmas atvejis, remiantis Sims (1971, 1973) prielaidomis, realybėje yra mažai tikėtinas, o ir ekonominė teorija negali pasiūlyti stiprių argumentų jam paremti. Taigi, norint išvengti neteisingo įvertinių reikšmingumo ar ženklų, ar tiesiog norint tikslesnės prognozės, naudingiau vertinti neapribotą U-MIDAS modelį su Lasso. Nėgana to, atsižvelgus į Uemat-su ir Tanaka (2015) rezultatus apie Lasso tipo regresijų pranašumą turint didelį kintamųjų skaičių, galima papildyti jų išvadą ir tuo, kad prie aptartų realių ir tikėtinų sąlygų Lasso pranašumas aptinkamas net vieno kintamojo atveju.

Kitą vertus, išsiaiškinta, kad apribotas MIDAS ima prognozuoti prasčiau nei U-MIDAS su Lasso tuomet, kai imties dydis gana didelis. Mažesnės imties atveju MIDAS beveik visuomet gaunamas pranašesnis. Tačiau net ir tada lieka neužtikrintumas, kad modelio gerumas neišliks toks pat ar net pablogės, jei modelis nebus sudaromas iš naujo kiekvienam imties padidėjimui.

Antruoju atveju tikroji ankstinių pasiskirstymą aprašanti funkcija yra žinoma, todėl aki-

vaizdu, kad ir MIDAS regresija galima pakankamai tiksliai šį sąryšį aprašyti nepaisant svarių funkcijos apribojimų. Tokiu atveju MIDAS veikai geriausiai iš visų lyginamų modelių. Tačiau ir vėl kyla problema, kad tokia situacija nėra dažnai sutinkama praktikoje, ypač prognozuojant makroekonominius rodiklius.



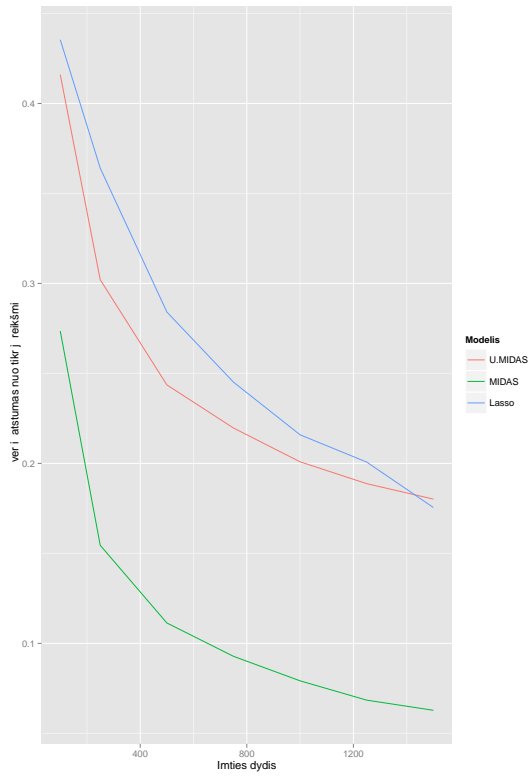
## 5 Literatūra

- [1] M. P. Clements, A. B. Galvão, *Macroeconomic Forecasting with Mixed Frequency Data: Forecasting US Output Growth*, Department of Economics, London, Working Paper Nr. 616, 2007
- [2] E. Ghysels, P. Santa-Clara, R. Valkanov, *The MIDAS Touch: Mixed Data Sampling Regression Models*, UNC ir UCLA, 2002
- [3] E. Ghysels, P. Santa-Clara, R. Valkanov, *The MIDAS Touch: Mixed Data Sampling Regression Models*, 2004
- [4] E. Ghysels, V. Kvedaras, V. Zemlys *Mixed Frequency Data Sampling Regression Models: The R Package midasr*
- [5] J. Geweke *Temporal Aggregation in the Multiple Regression Model*, *Econometrica*, Vol. 46, Nr. 3, 1978 (p.643 - 661)
- [6] C. Marsilli, *Variable Selection in Predictive MIDAS Models*, Banque de France, Working Paper Nr. 520, 2014
- [7] Christopher A. Sims, *Discrete Approximations to Continuous Time Distributed Lags in Econometrics*, *Econometrica*, Vol. 39, Nr. 3, 1971 (p.545 - 563)
- [8] Christopher A. Sims, *Distributed Lags*, Center for Economic Research, Discussion Paper Nr. 28, , 1973
- [9] Y. Uematsu, S. Tanaka, *Macroeconomic Forecasting and Variable Selection with a Very Large Number of Predictors: A Penalized Regression Approach*, 2015
- [10] U. Tibshirani *Regression Shrinkage and Selection via the Lasso*, *Journal of Royal Statistical Society*, Vol. 58, 1996 (p.267 - 288)

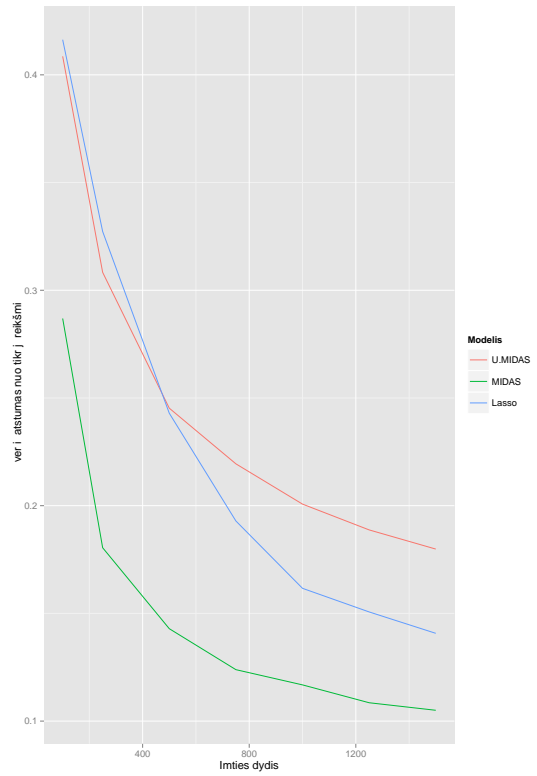
# A Priedas

## A.1 Almon polinomo atvejis. Įverčių atstumas nuo tikrųjų parametrų reikšmių

Kitų polinomų atveju rezultatai analogiški, todėl čia nepateikiami.

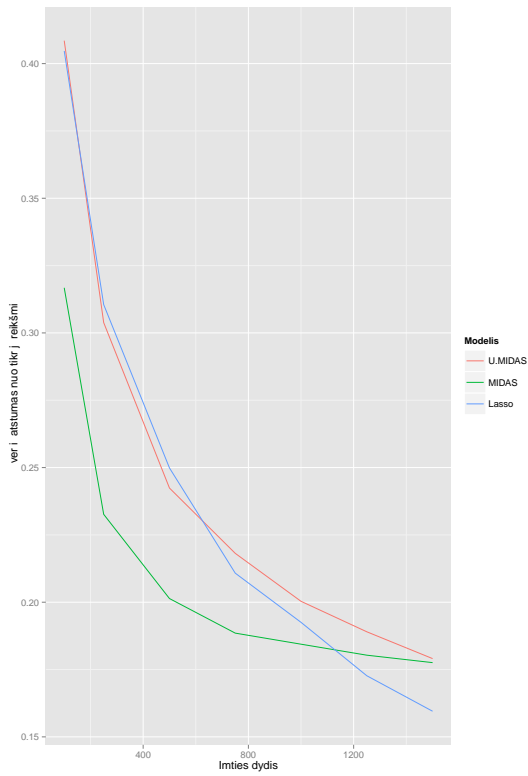


(a) 1-asis procesas

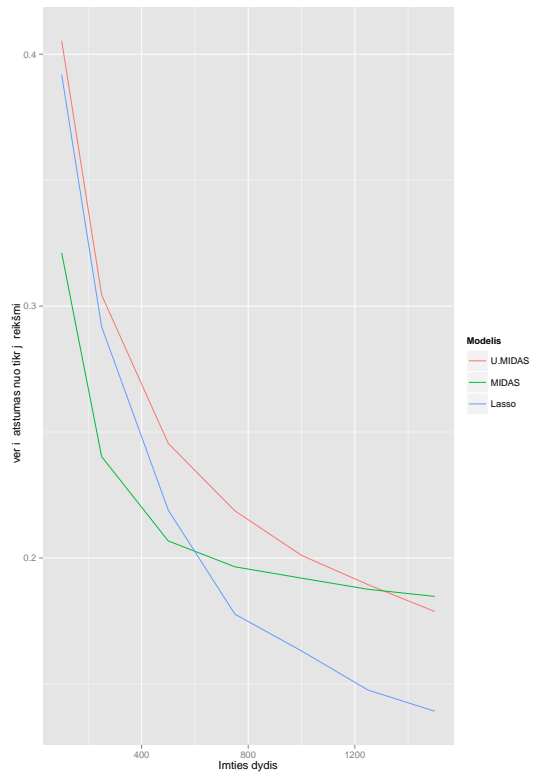


(b) 2-asis procesas

8 pav.: Vidutiniai kvadratiniai įverčių nuokrypiai nuo tikrųjų reikšmių

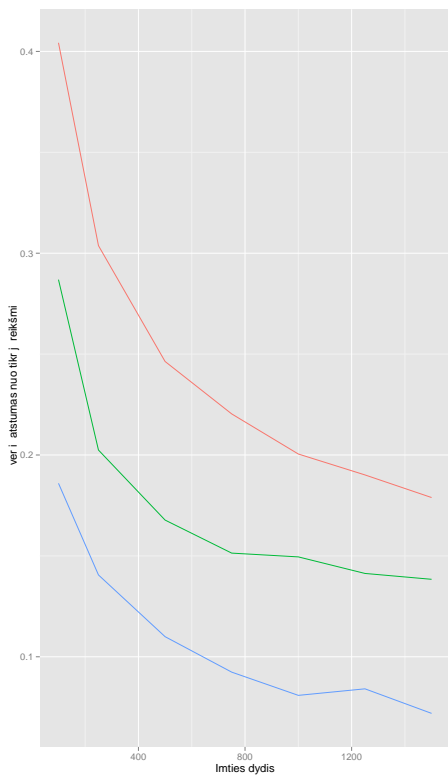


(a) 3-asis procesas

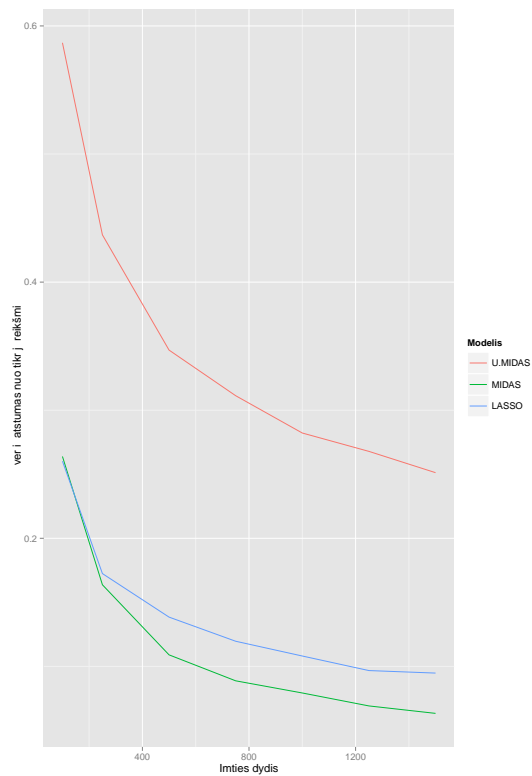


(b) 4-asis procesas

9 pav.: Vidutiniai kvadratiniai įverčių nuokrypiai nuo tikrųjų reikšmių



(a) 5-asis procesas

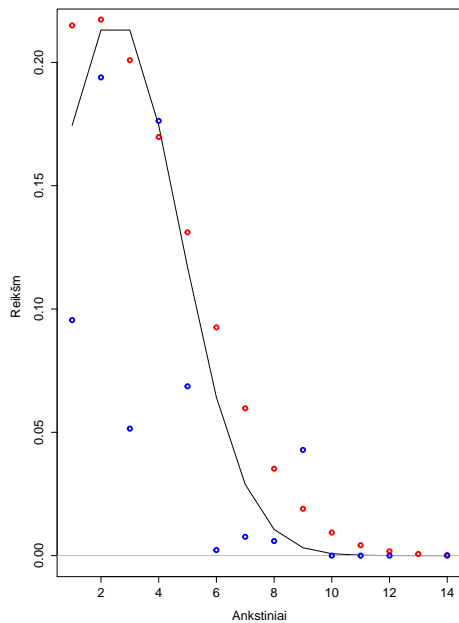


(b) Alternuojanti svorių f-ija

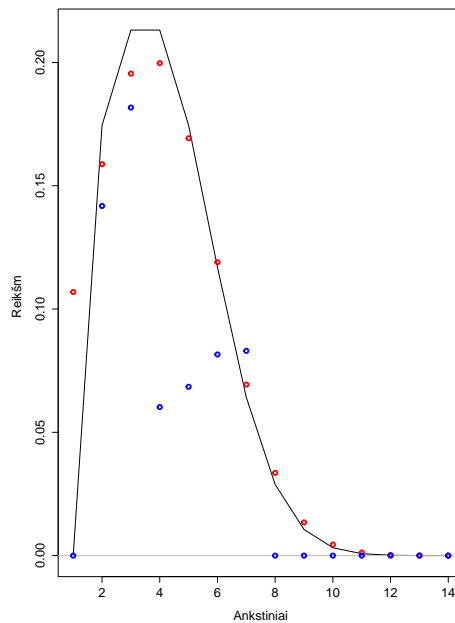
10 pav.: Vidutiniai kvadratiniai įverčių nuokrypiai nuo tikrųjų reikšmių

## A.2 Almon polinomo atvejis. Konkretus įverčių palyginimas

Čia pateikiami MIDAS ir Lasso įverčiai, lyginant su tikrosiomis reikšmėmis. Grafikai yra konkretiems atvejams su  $N = 200$ .

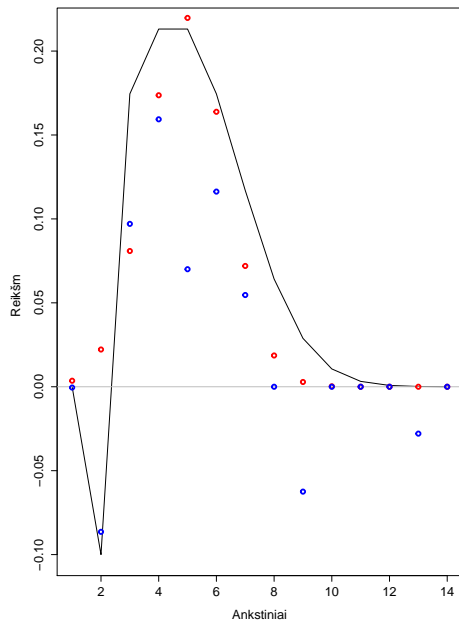


(a) 1-asis procesas

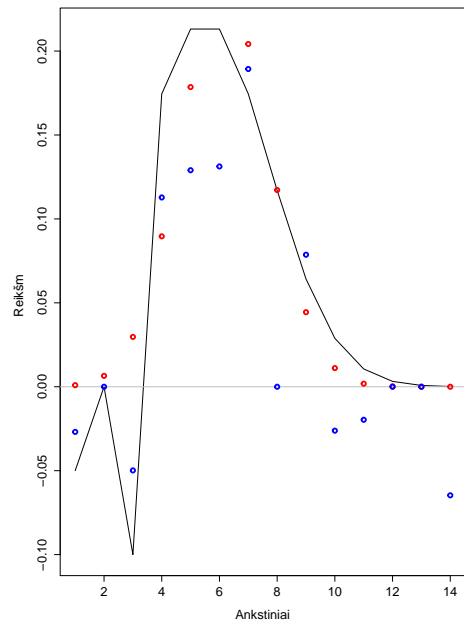


(b) 2-asis procesas

11 pav.: Įverčiai lyginant su tikraisiais: juoda - tikrosios reikšmės, raudona - MIDAS, mėlyna - Lasso

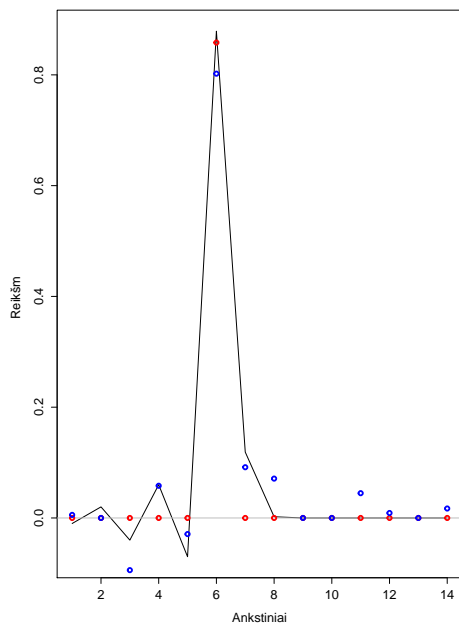


(a) 1-axis procesas

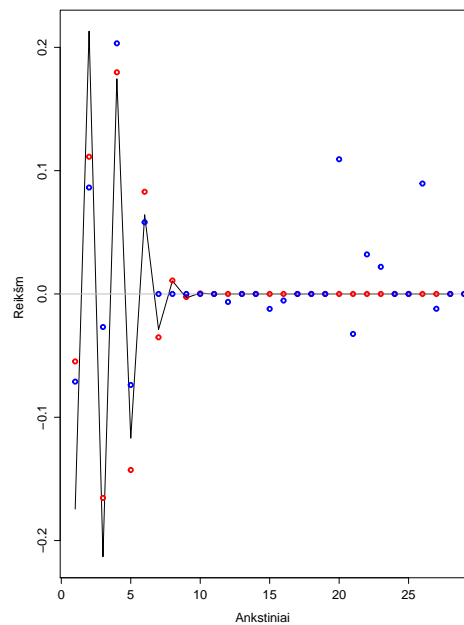


(b) 2-axis procesas

12 pav.: Įverčiai lyginant su tikraisiais: juoda - tikrosios reikšmės, raudona - MIDAS, mėlyna - Lasso



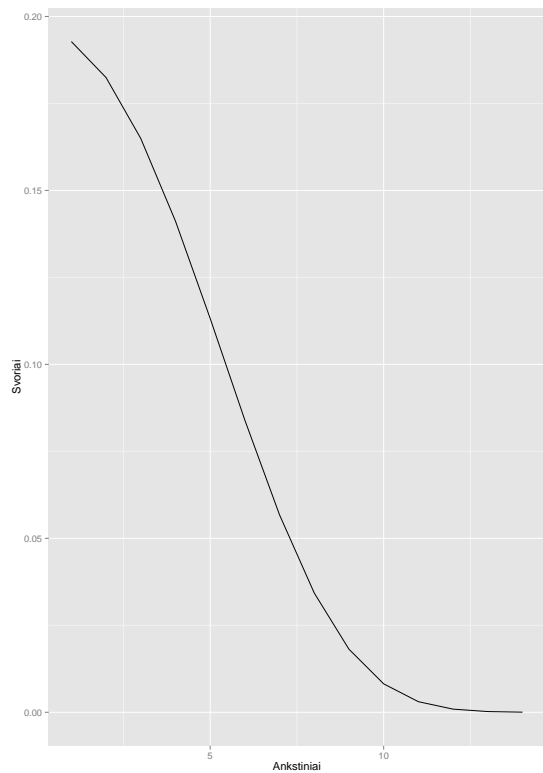
(a) 1-axis procesas



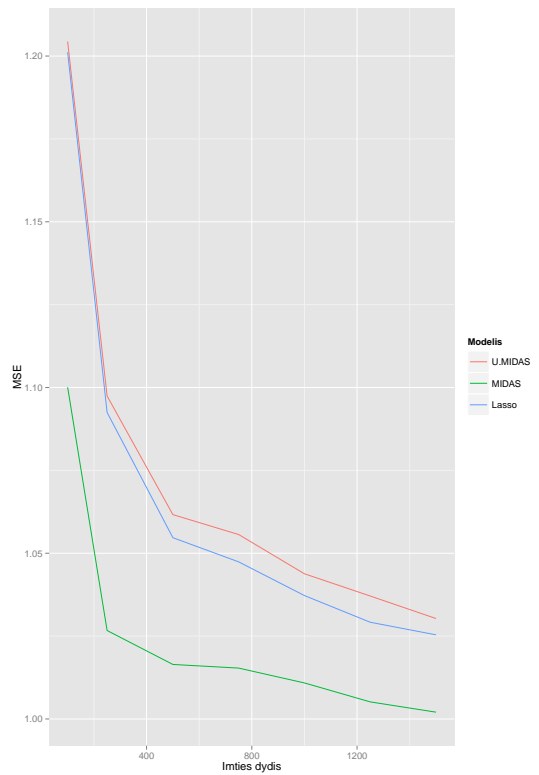
(b) 2-axis procesas

13 pav.: Įverčiai lyginant su tikraisiais: juoda - tikrosios reikšmės, raudona - MIDAS, mėlyna - Lasso

### A.3 Gompertz polinomo atvejis. MSE rezultatai

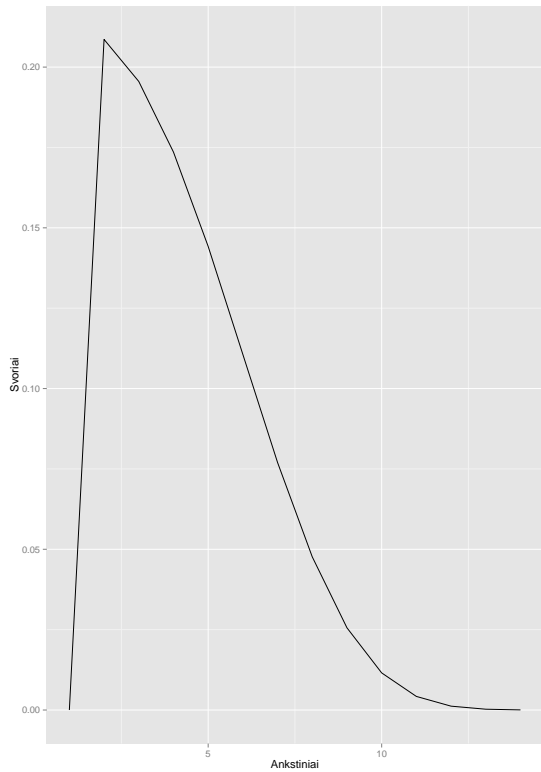


(a) Svoriai

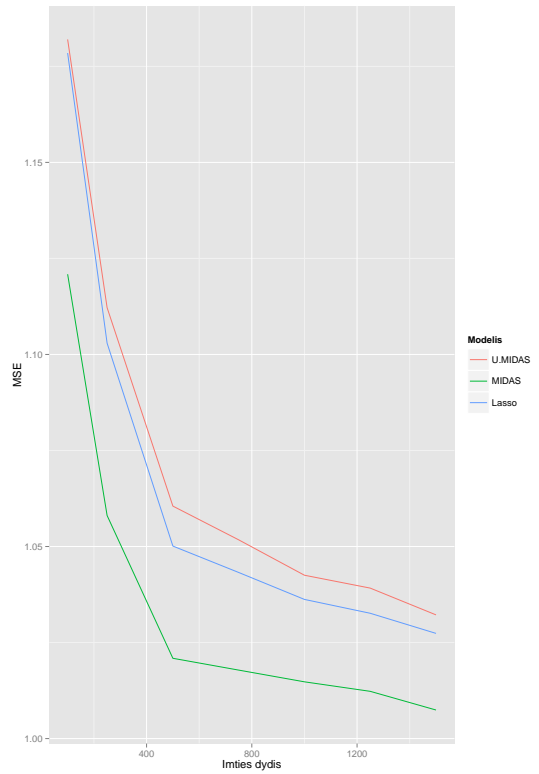


(b) MSE

14 pav.: 1-ojo proceso svorių funkcija ir gauta prognozės paklaida MSE (raudona - U-MIDAS, žalia - MIDAS, mėlyna - Lasso)

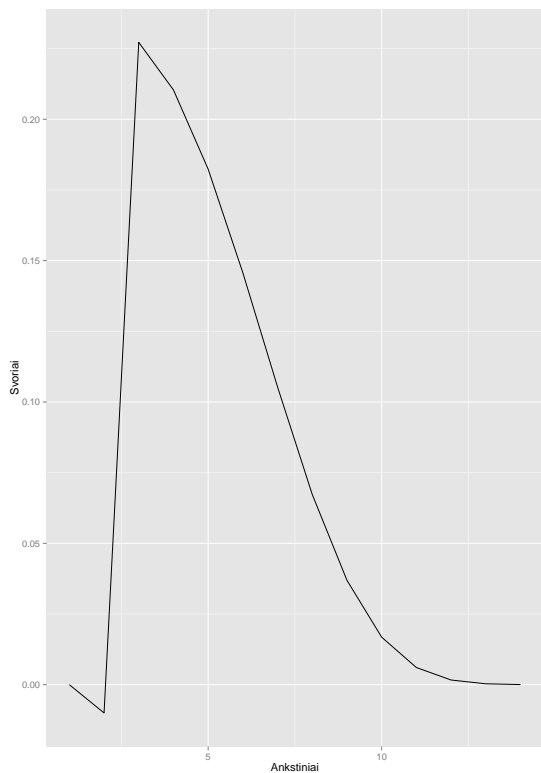


(a) Svoriai

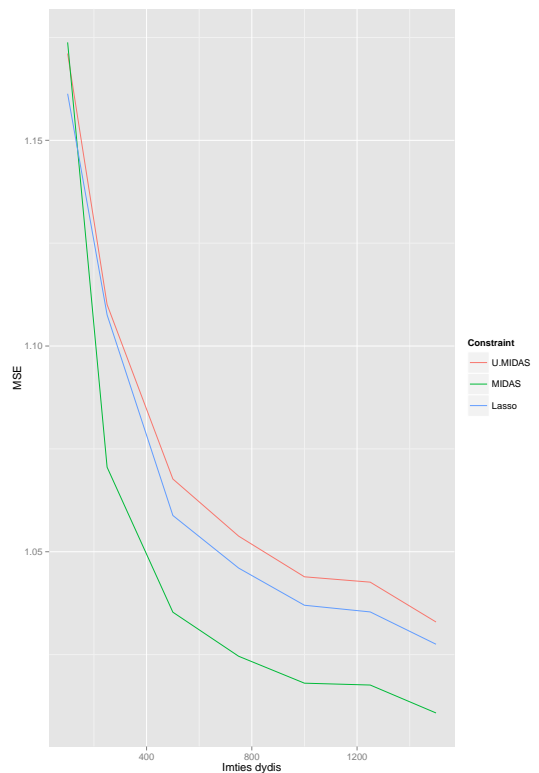


(b) MSE

15 pav.: 2-ojo proceso svorių funkcija ir gauta prognozės paklaida MSE (raudona - U-MIDAS, žalia - MIDAS, mėlyna - Lasso)

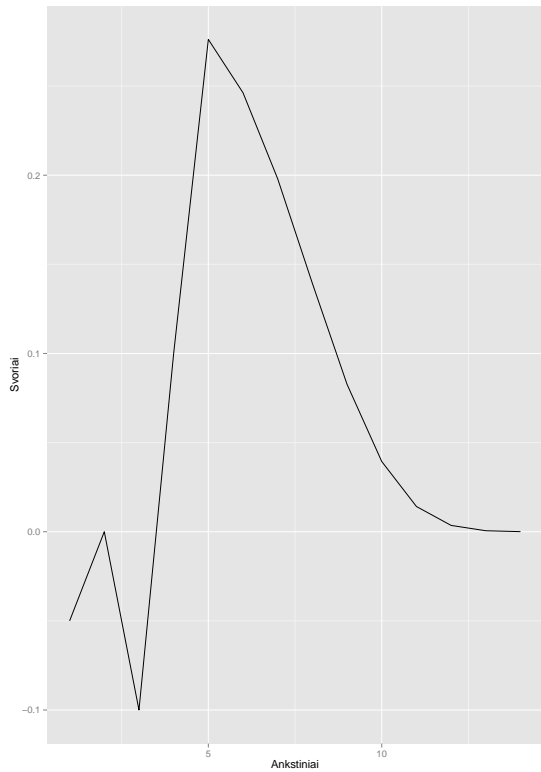


(a) Svoriai

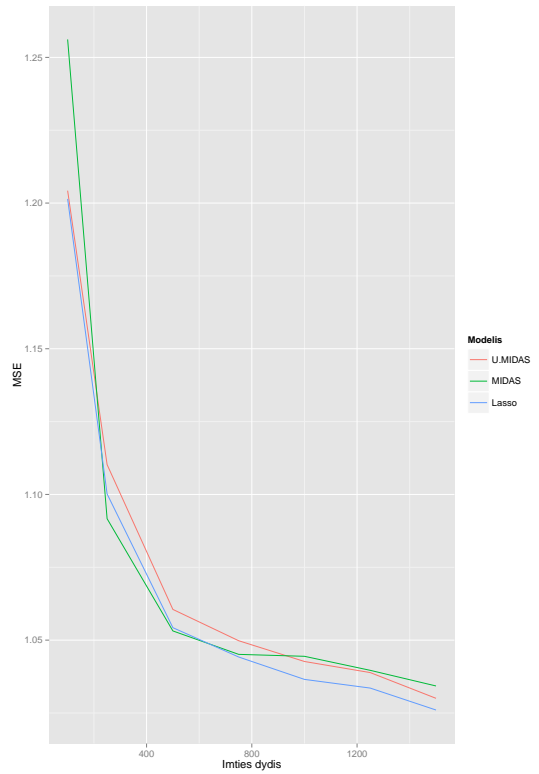


(b) MSE

16 pav.: 3-iojo proceso svorių funkcija ir gauta prognozės paklaida MSE (raudona - U-MIDAS, žalia - MIDAS, mėlyna - Lasso)

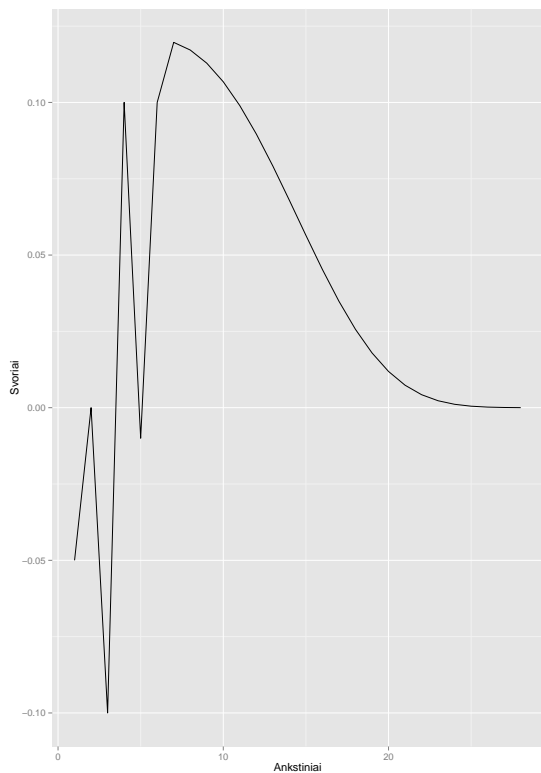


(a) Svoriai

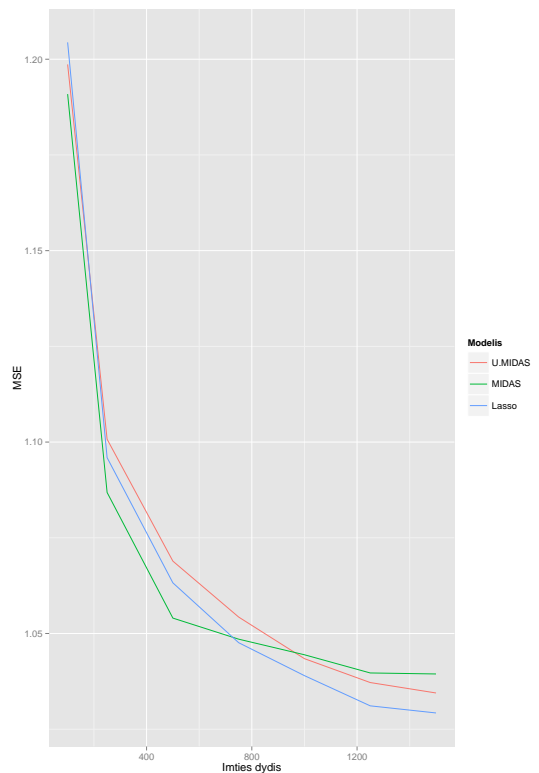


(b) MSE

17 pav.: 4-ojo proceso svorių funkcija ir gauta prognozės paklaida MSE (raudona - U-MIDAS, žalia - MIDAS, mėlyna - Lasso)



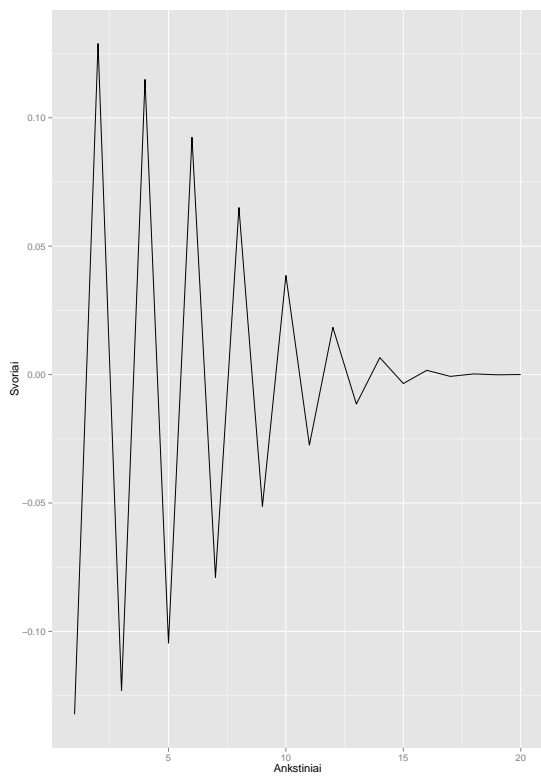
(a) Svoriai



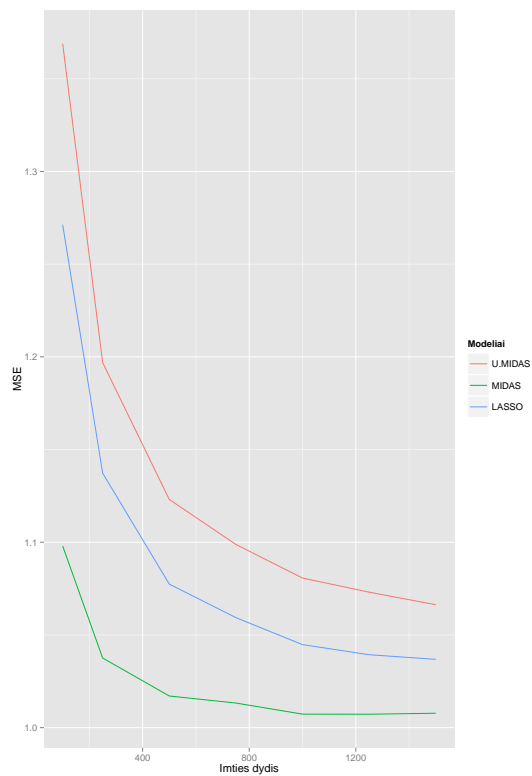
(b) MSE

18 pav.: 5-ojo proceso svorių funkcija ir gauta prognozės paklaida MSE (raudona - U-MIDAS, žalia - MIDAS, mėlyna - Lasso)





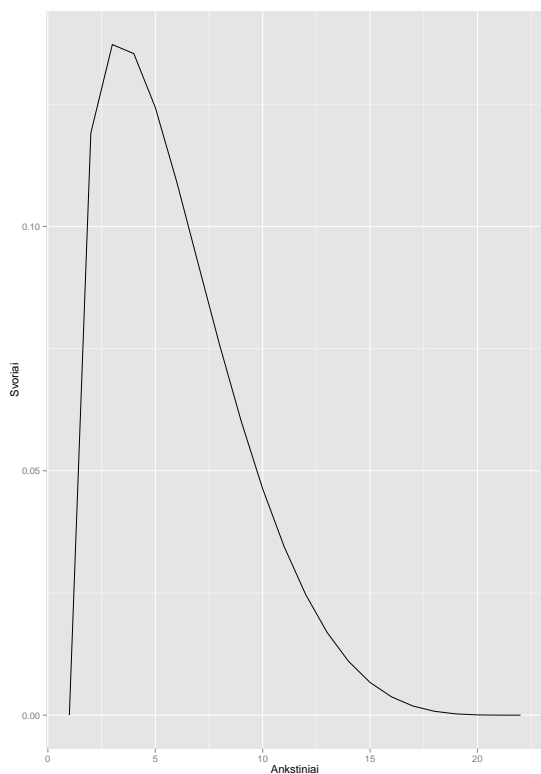
(a) Svoriai



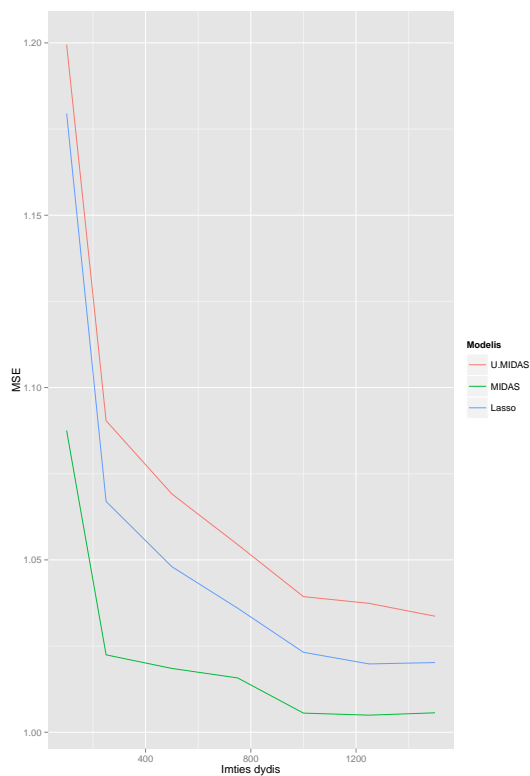
(b) MSE

19 pav.: Alternuojanti svorių funkcija ir gauta prognozės paklaida MSE (raudona - U-MIDAS, žalia - MIDAS, mėlyna - Lasso)

## A.4 Beta polinomo atvejis. MSE rezultatai

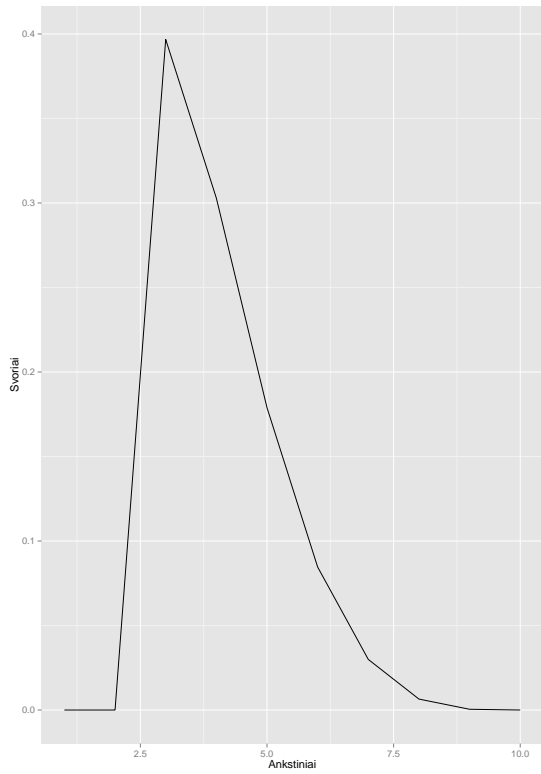


(a) Svoriai

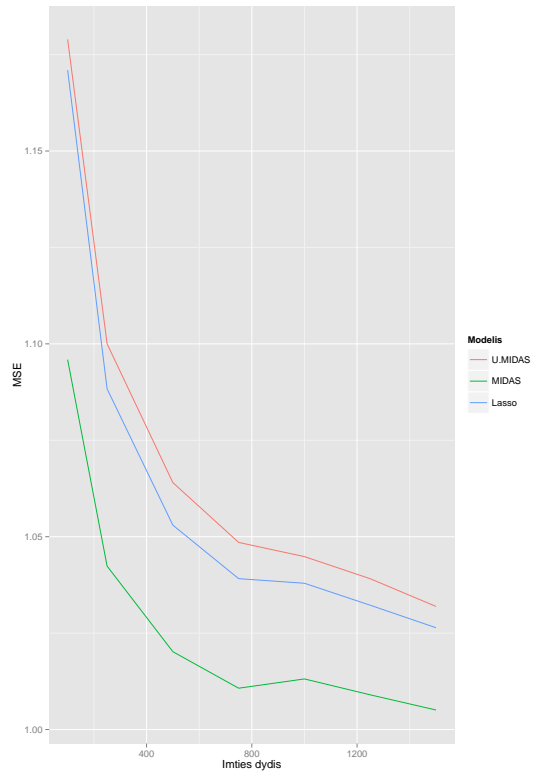


(b) MSE

20 pav.: 1-ojo proceso svorių funkcija ir gauta prognozės paklaida MSE (raudona - U-MIDAS, žalia - MIDAS, mėlyna - Lasso)

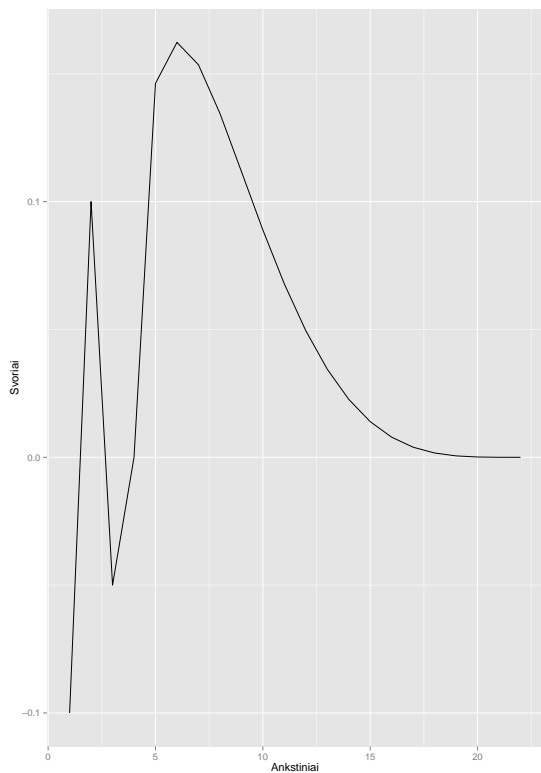


(a) Svoriai

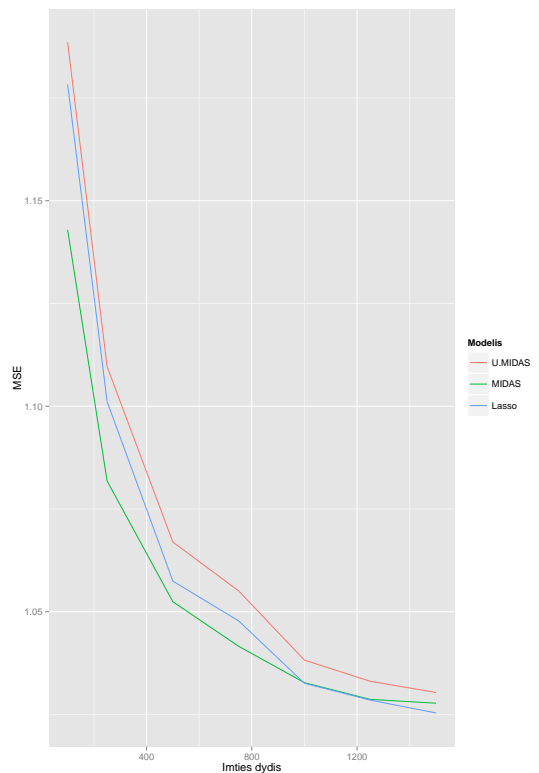


(b) MSE

21 pav.: 2-ojo proceso svorių funkcija ir gauta prognozės paklaida MSE (raudona - U-MIDAS, žalia - MIDAS, mėlyna - Lasso)

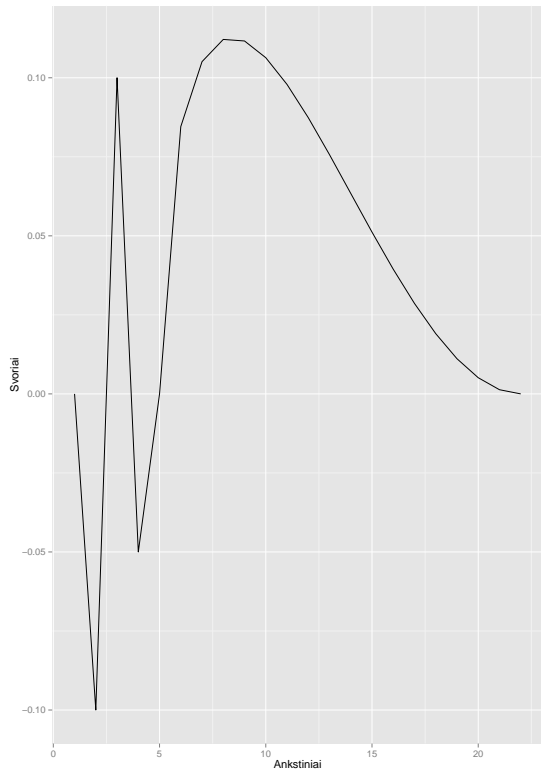


(a) Svoriai

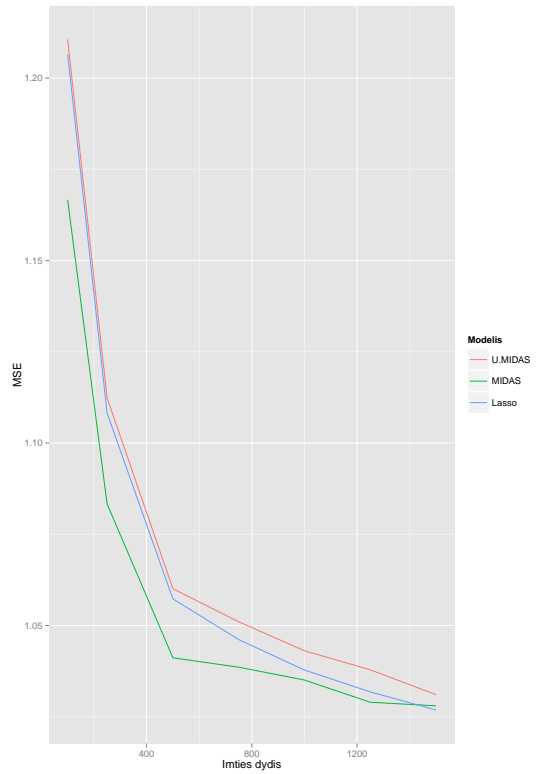


(b) MSE

22 pav.: 3-iojo proceso svorių funkcija ir gauta prognozės paklaida MSE (raudona - U-MIDAS, žalia - MIDAS, mėlyna - Lasso)

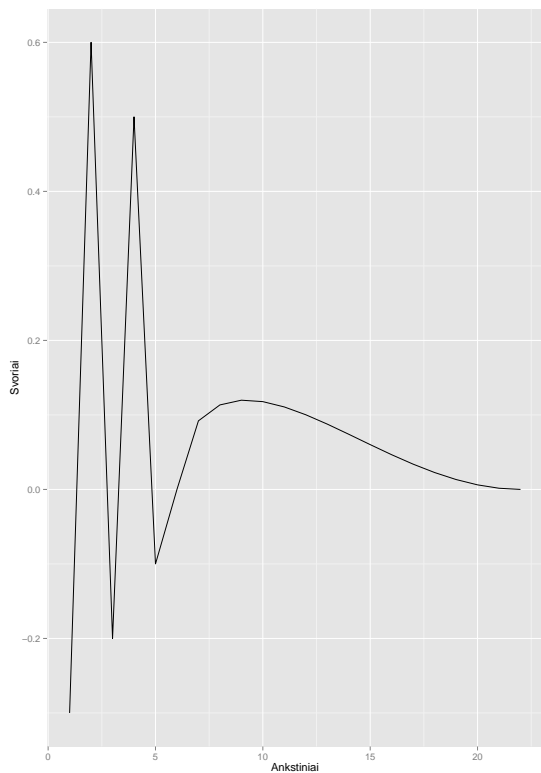


(a) Svoriai

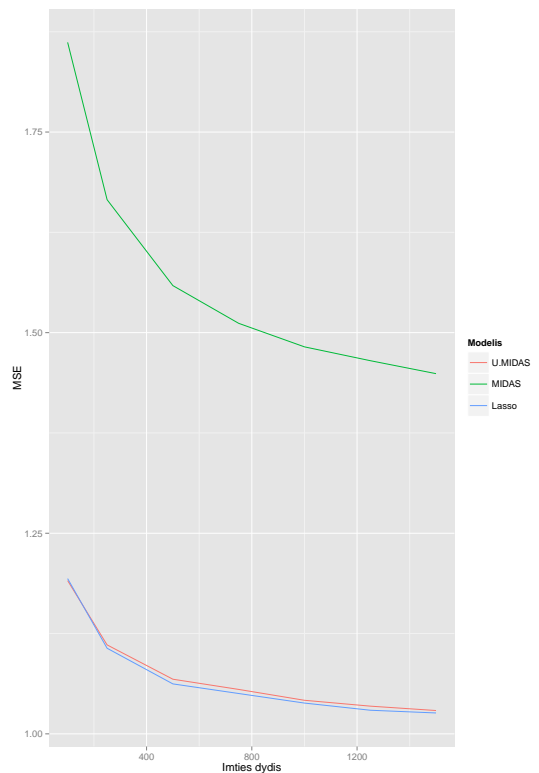


(b) MSE

23 pav.: 4-ojo proceso svorių funkcija ir gauta prognozės paklaida MSE (raudona - U-MIDAS, žalia - MIDAS, mėlyna - Lasso)

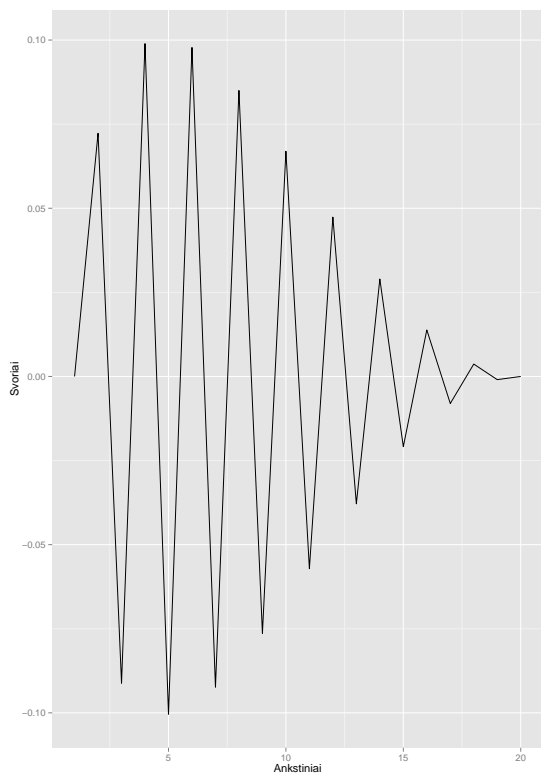


(a) Svoriai

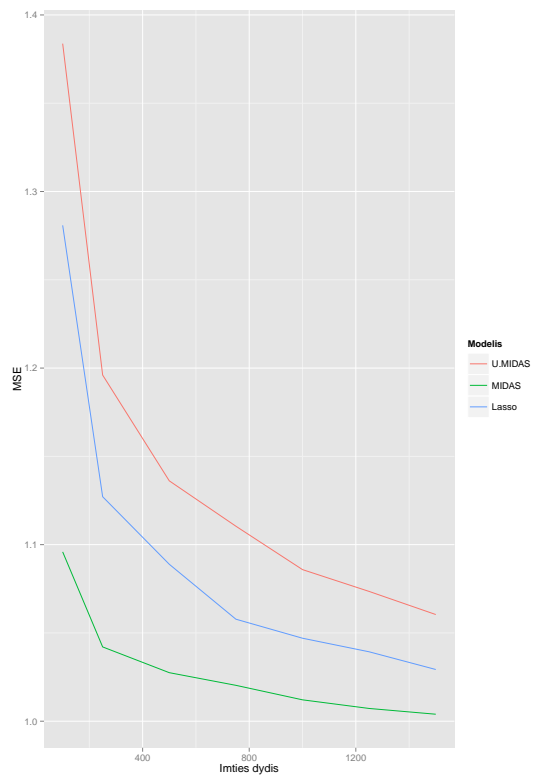


(b) MSE

24 pav.: 5-ojo proceso svorių funkcija ir gauta prognozės paklaida MSE (raudona - U-MIDAS, žalia - MIDAS, mėlyna - Lasso)



(a) Svoriai



(b) MSE

25 pav.: Alternuojanti svorių funkcija ir gauta prognozės paklaida MSE (raudona - U-MIDAS, žalia - MIDAS, mėlyna - Lasso)