

KANDIDATUPPSATS I STATISTIK, 15HP

LUNDS UNIVERSITET

---

**Modelljämförelse för prediktioner av  
nettomigration**

---

*Författare:*  
Christina Anzén  
Sara Emilsson

*Handledare:*  
Jakob Bergman

HT 2018

## Sammanfattning

Migration är ett högaktuellt ämne i såväl samhällsdebatten som i den akademiska världen. Svårigheten att göra säkra prediktioner ökar osäkerheten i samhällsinstitutioners beslutsunderlag och påverkar därmed allt från politiska beslut till investeringar. Det finns ett uttalat behov av goda modeller för att minska osäkerheten och samhällskostnader. Utmaningarna i att åstadkomma säkra prediktioner driver därmed efterfrågan på nya ansatser och modeller. Detta arbetet jämför den väletablerade AR(1)-modellen med en interaktionsmodell, Gravitationsmodellen, för att prediktera den svenska nettomigrationen för år 2022. Dessa jämförs sedan med SCB:s prediktioner. I arbetet konstateras att Gravitationsmodellens prediktioner kraftigt överskrider övriga prediktioner. Prediktionen för AR(1)-modellen ligger mycket nära SCB:s alternativa antagande om högre migration, men överskrider SCB:s huvudprediktion betydligt. Arbetet förklarar skillnader i respektive prediktion med modellernas uppbyggnad, samt behandlar modellernas styrkor och svagheter gällande prediktioner.

## Abstract

The topic of migration causes great debate around the world. The difficulty in making accurate predictions affects society as a whole. There is an increased demand for accurate and reliable predictions both from the public and private sector, which has given rise to many new models, as well as a renaissance for older ones. The thesis compares the well-established AR(1) model with the Gravitational model to predict the Swedish net migration of the year 2022. The predictions are compared to the forecasts of Statistics Sweden. The study concludes that the Gravitational model's prediction greatly exceeds the others'. The prediction of the AR(1) model is in line with Statistics Sweden's alternative assumption of a higher migration. Still, the AR(1) prediction considerably exceeds the main forecast of Statistics Sweden. Finally, model structures, as well as results, are examined and evaluated through compared analysis.

# 1 Introduktion

Migration är en av vår samtids största politiska frågor. Frågan är på alla plan en universell beröringspunkt, från den lilla människan i en orange gummibåt till den stora världens globala gemenskap i FN. Få frågor väcker så mycket känslor och så mycket debatt som människors tendens till förflyttning, och de faktiska och upplevda kostnaderna förknippade med denna förflyttning. Välfärdsstater och deras ekonomier är i allt högre grad beroende av information kring migrationsflöden vilket driver efterfrågan av goda prediktioner. Utformningen av modeller har i takt med digitaliseringen av samhället och det stigande intresset för migrationsfrågor blivit ett allt mer omdiskuterat ämne. Statistikmyndigheten, SCB, är en myndighet med mycket gott förtroende bland svenskarna (SOM-Institutet, 2017). Trots det har SCB under senare år kritiserats öppet i etablerade medier för sina avvikande prediktioner (Gudmundson, 2017). Då SCB:s demografiska rapport *Sveriges framtida befolkning 2015–2060* publicerades i maj 2015 saknades höjd för den immigrationsökning som under hösten kom att beskrivas som en kris. Då prediktionernas tillförlitlighet ifrågasätts ges allt större utrymme till modeller vilka lägger tyngdvikt vid bakomliggande faktorer. Detta har resulterat i ett återupptäckande och en utveckling av äldre interaktionsmodeller, däribland Gravitationsmodellen. Syftet med detta arbete är att genom komparativ analys undersöka prediktioner av Sveriges nettomigration år 2022 gjorda utifrån AR(1)-modellen, respektive Gravitationsmodellen. I analysen behandlas skillnader i resultat och uppbyggnad, samt dess förhållande till SCB:s prediktioner för år 2022.

Arbetet fokuserar på den svenska nettomigrationen. Avgränsningen är gjord med hänsyn till frågeställningens omfattningen, samt kvalitet i tillgängligt datamaterial. Materialets kvalitet anses av stor betydelse då det är en förutsättning för prediktionernas tillförlitlighet. Vidare har prediktionerna en tidshorisont av fem år. Valet av tidsperiod är en avvägning mellan kortsiktiga prediktioners stora variation, och långsiktiga prediktioners förhållandevis låga variation (Wadensjö, 2012). Osäkerheten i prediktioner av nettomigration är mycket stor och tenderar att öka över tid. En femårig tidsperiod ger en prediktion relaterad till dagens geopolitiska förhållanden men med utrymme för att urskilja respektive modells långsiktiga utveckling. Vidare används femårsperioder flitigt i litteratur och praktiken vilket skapar goda förutsättningar för jämförelser mellan studier. Gällande val av modeller relaterar detta starkt till arbetets syfte. Syftet med arbetet är att jämföra en endogen modell och en modell med exogen ansats. Valet av AR(1)-modellen som endogen modell grundar sig del i modellens utformning. Med hänsyn uteslutande till tidigare observationer strävar AR(1)-modellen genom sin koefficient över tid mot sitt medelvärde. Modellen är isolerad från annan påverkan än observerade värden. Över tid kommer modellens prediktioner vara goda då den strävar mot sitt eget medelvärde; vilket också är dess styrka. Modellens svaghet gällande prediktioner ligger i att förutse oväntade, kraftiga förändringar. I en föränderlig period kommer det därför att förekomma större avvikelser från modellens medelvärde. Ett tydligt exempel på detta är flyktingvågen 2015 vilken inte kunde predikteras med AR(1)-modellen men däremot gav ett permanent avtryck i modellens skattade medelvärde. Modellen används av SCB då den på längre sikt ger en indikation om hur prediktioner av andra modeller förhåller sig till medelvärdet. Som exogen modell används Gravitationsmodellen. Gravitationsmodellen är en interaktionsmodell som används inom bland annat demografi, logistik och migrationsstudier för att förklara och skatta storleken av flöden mellan områden. Modellen tillåter variabler vilka i regel studeras i omvärldsanalyser och ger indikationer på kommande förändringar i flöden. Modellens styrka ligger därmed i att utifrån indikationer på förändring skatta plötsliga och kraftiga förändringar. Modellen möjliggör inkluderingen av variabler såsom HDI och säkerhetsläge, vilka kan förväntas

ändras innan nettoflödet av människor ändras. Dessa variabler fungerar därmed som indikatorer på kommande flödesförändringar. I synnerhet HDI är av intresse då variabeln används av SCB som en antydning om migration till följd av fattigdom. Modellens svaghet ligger i dess förklaringsgrad. Trots modellens popularitet förklarar den i regel en väldigt liten del av migrationsflöden (Poot et al. 2016, s. 3).

## 2 Metod

### 2.1 AR(1)

I en tidsserie används olika modeller för att förstå de underliggande faktorer som producerar den observerade datamängden. Resultaten används för att anpassa modellerna för att kunna prediktera framtida värden. Tidsserieanalys är studien av dessa modeller. En av de mest använda modellerna är den autoregressiva modellen,  $AR(p)$ , vilken används av SCB för att göra en stokastisk framskrivning av nettomigrationen (Statistiska Centralbyrån 2015, s. 202). En  $AR(1)$ -process är en regression med ett tidslagg och har ekvationen

$$Y_t = \phi Y_{t-1} + \varepsilon_t \quad (1)$$

där  $\varepsilon_t$  är oberoende  $Y_{t-1}$ , samt har varians  $\sigma^2$ . Det antas att  $E(Y_t) = 0$ . Om så inte är fallet kan medelvärdet subtraheras från  $Y_t$  i ekvation (1). För att bedöma om en modell är autoregressiv, samt för att få reda på vilket tidslagg den tillhör kan autokorrelationsfunktionen, ACF, samt partiella autokorrelationsfunktionen, PACF, användas.

Autokorrelationsfunktionen, ACF:en, visar korrelationen vid olika tidslagg  $k$  där den för en  $AR(1)$ -process ges av (Cryer & Chan, 2008, s. 110):

$$\rho_k = \phi^k, \quad \text{för } k = \pm 1, \pm 2, \pm 3, \dots$$

Värden på  $\phi$  nära  $\pm 1$  indikerar ett starkt linjärt beroende, samt värden nära 0 indikerar ett svagt beroende. När  $\phi = 0$  är  $Y_t$  och  $Y_{t+k}$  okorrelerade. Då  $|\phi| < 1$  avtar autokorrelationen exponentialartat när  $k$  ökar. Positiva korrelationer ges av när  $0 < \phi < 1$ , då när  $-1 < \phi < 0$  alternerar autokorrelationen mellan positiva och negativa värden.

Den partiella autokorrelationen används för att bestämma ordningen på  $AR(p)$ -processen. PACF:en är definierad vid tidslagg  $k$  såsom

$$\phi_{kk} = \text{Corr}(Y_t, Y_{t-k} \mid Y_{t-1}, \dots, Y_{t-k+1})$$

vilken mäter korrelationen mellan  $Y_t$  och  $Y_{t-k}$  betingat på  $Y_{t-1} \dots Y_{t-k+1}$  (Shumway & Stoffer, 2006, s. 107). Detta leder till  $\phi_{kk} = 0$  för  $k > p$  och en cut-off-effekt när tidslagget överskrider processens ordning.

Stationäritet är ett av de viktigaste antagandena inom stokastiska processer.  $AR(p)$ -processen är stationär om  $|\phi| < 1$ , varvid kovariansen mellan värden på givet avstånd i tid från varandra, inte beror på var de inträffar i tiden. Då

$$\text{Cov}(Y_t, Y_s) = \text{Cov}(Y_{t-k}, Y_{s-k}), \quad \text{för alla } t, s \text{ och } k$$

och genom att sätta  $k = s$  och därefter  $k = t$  erhålls

$$\gamma_{t,s} = \text{Cov}(Y_{t-s}, Y_0) = \text{Cov}(Y_0, Y_{s-t}) = \text{Cov}(Y_0, Y_{t-s}) = \gamma_{0,|t-s|}.$$

Kovariansen mellan  $Y_t$  och  $Y_s$  beror följaktligen på differensen  $|t - s|$  och ej de faktiska tiderna  $t$  samt  $s$ . Värden på  $Y$  en tidsenhet ifrån varandra har samma korrelation oberoende när i tiden de förekommer (Cryer & Chan, 2008, s. 16). Då processen är stationär, är  $\rho_{t,|t-k|}$  samma för alla värden av  $t$ .

## 2.2 Kroneckerprodukt

I konstruktionen av Gravitationsmodellens variabler används Kroneckerprodukter. Kroneckerprodukt är en operation bestående av två matriser av arbiträr storlek vilken resulterar i en blockmatris. Antag att  $\mathbf{A}$  är en  $m \times n$  matris och  $\mathbf{B}$  är en  $p \times q$  matris. Kroneckerprodukten,  $\otimes$ , är operationen till den  $mp \times nq$  blockmatris som uppkommer av multiplikationen av  $a_{ij}$  och  $b_{kl}$ .

$$\mathbf{A} \otimes \mathbf{B} = \begin{bmatrix} a_{11}b_{11} & a_{11}b_{12} & \cdots & a_{11}b_{1q} & \cdots & \cdots & a_{1n}b_{11} & a_{1n}b_{12} & \cdots & a_{1n}b_{1q} \\ a_{11}b_{21} & a_{11}b_{22} & \cdots & a_{11}b_{2q} & \cdots & \cdots & a_{1n}b_{21} & a_{1n}b_{22} & \cdots & a_{1n}b_{2q} \\ \vdots & \vdots & \ddots & \vdots & & & \vdots & \vdots & \ddots & \vdots \\ a_{11}b_{p1} & a_{11}b_{p2} & \cdots & a_{11}b_{pq} & \cdots & \cdots & a_{1n}b_{p1} & a_{1n}b_{p2} & \cdots & a_{1n}b_{pq} \\ \vdots & \vdots & & \vdots & \ddots & & \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots & & \ddots & \vdots & \vdots & & \vdots \\ a_{m1}b_{11} & a_{m1}b_{12} & \cdots & a_{m1}b_{1q} & \cdots & \cdots & a_{mn}b_{11} & a_{mn}b_{12} & \cdots & a_{mn}b_{1q} \\ a_{m1}b_{21} & a_{m1}b_{22} & \cdots & a_{m1}b_{2q} & \cdots & \cdots & a_{mn}b_{21} & a_{mn}b_{22} & \cdots & a_{mn}b_{2q} \\ \vdots & \vdots & \ddots & \vdots & & & \vdots & \vdots & \ddots & \vdots \\ a_{m1}b_{p1} & a_{m1}b_{p2} & \cdots & a_{m1}b_{pq} & \cdots & \cdots & a_{mn}b_{p1} & a_{mn}b_{p2} & \cdots & a_{mn}b_{pq} \end{bmatrix}$$

## 2.3 Gravitationsmodellen

Gravitationsmodellen bygger på Isac Newtons gravitationslag; kroppars rörelse mot varandra med kraft är proportionerlig till kropparnas massor och avståndet mellan dem. Modellen som vi känner den idag utvecklades av W. Alonso för beräkning av flöden inom geografi och demografi för att förklara migrationsflöden och benämns då som en spatial interaktionsmodell (Porell & Hua, 1981). Modellen används med andra ord till att förklara flöden mellan utgångsland,  $o$ , och destinationsland,  $d$ , som  $o \rightarrow d$ . Modellen beror av avstånd mellan utgångsland och destinationsland, samt förklarande variabler för landets egenskaper. I grundmodellen antas att avstånd som förklarande variabel eliminerar spartialt beroende bland  $o \rightarrow d$ -flöden med två regioner. Regionerna antas därmed oberoende. Grunden för antagandet diskuteras närmare i diskussionsavsnittet. Den beroende variabeln  $\mathbf{Y}$  är en  $n \times n$  matris av interregionalt flöde från  $n$  utgångsländer till  $n$  destinationsländer (LeSage & Pace, 2009, s. 212), där kolumnerna anger utgångsland och raderna destinationsland. Samtliga flöden ingår i ett slutet system vilket består av lika många utgångsländer som destinationsländer.

Tabell 1: Uppbyggnad av matris  $\mathbf{Y}$

Destinations-/Utgångsland	Utgångsland 1	Utgångsland 2	...	Utgångsland n
Destinationsland 1	$o_1 \rightarrow d_1$	$o_2 \rightarrow d_1$	...	$o_n \rightarrow d_1$
Destinationsland 2	$o_1 \rightarrow d_2$	$o_2 \rightarrow d_2$	...	$o_n \rightarrow d_2$
...	...	...	...	...
Destinationsland n	$o_1 \rightarrow d_n$	$o_2 \rightarrow d_n$	...	$o_n \rightarrow d_n$

Vidare transponeras utgångsmatrisen,  $\mathbf{Y}$ , för att generera en destinations-centrerad ordning, där destinationsland utgör kolumner och utgångsland rader. Därefter vektoriseras matrisen genom att stapla kolumner sekventiellt. Detta ger  $\mathbf{y} = \mathbf{vec}(\mathbf{Y}')$ , modellens beroende variabel (LeSage & Pace, 2009, s. 214). Då emigration anges i negativa värden är  $\mathbf{y}$  detsamma som nettomigrationen.

Modellen beror på en  $n \times k$  matris av förklarande variabler,  $\mathbf{X}$ , där  $k$  är antal egenskaper och  $n$  är antal utgångsländer. En av variablerna är ländernas populationsstorlek.

$$\mathbf{X} = \begin{bmatrix} \text{Population}_1 & \text{Variabel}_{12} & \dots & \text{Variabel}_{1k} \\ \text{Population}_2 & \text{Variabel}_{22} & \dots & \text{Variabel}_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ \text{Population}_n & \text{Variabel}_{n2} & \dots & \text{Variabel}_{nk} \end{bmatrix}$$

Destinationslandets egenskaper representeras av  $\mathbf{X}_d$  vilken är en  $N \times k$  matris där  $\mathbf{X}_d = \boldsymbol{\iota}_n \otimes \mathbf{X}$  och  $\boldsymbol{\iota}_n$  är en  $n \times 1$  vektor av ettor. Motsvarande procedur för utgångsland ger  $\mathbf{X}_o$ , där  $\mathbf{X}_o = \mathbf{X} \otimes \boldsymbol{\iota}_n$  (LeSage & Pace 2009, s. 214). Gravitationsmodellens förklarande variabler utgörs av  $\mathbf{X}_d$  och  $\mathbf{X}_o$ .

Tekniken för bildandet av  $\mathbf{X}_d$  samt  $\mathbf{X}_o$  illustreras nedan, där  $n = 3$  länder: Afghanistan, Frankrike och Sverige.

$$\mathbf{X}_d = \begin{bmatrix} \text{pop}_{Afg} & x_{Afg,2} & \dots & x_{Afg,k} \\ \text{pop}_{Frr} & x_{Frr,2} & \dots & x_{Frr,k} \\ \text{pop}_{Ssv} & x_{Ssv,2} & \dots & x_{Ssv,k} \\ \text{pop}_{Afg} & x_{Afg,2} & \dots & x_{Afg,k} \\ \text{pop}_{Frr} & x_{Frr,2} & \dots & x_{Frr,k} \\ \text{pop}_{Ssv} & x_{Ssv,2} & \dots & x_{Ssv,k} \\ \text{pop}_{Afg} & x_{Afg,2} & \dots & x_{Afg,k} \\ \text{pop}_{Frr} & x_{Frr,2} & \dots & x_{Frr,k} \\ \text{pop}_{Ssv} & x_{Ssv,2} & \dots & x_{Ssv,k} \end{bmatrix}$$

$$\mathbf{X}_o = \begin{bmatrix} \text{pop}_{Afg} & x_{Afg,2} & \dots & x_{Afg,k} \\ \text{pop}_{Afg} & x_{Afg,2} & \dots & x_{Afg,k} \\ \text{pop}_{Afg} & x_{Afg,2} & \dots & x_{Afg,k} \\ \text{pop}_{Frr} & x_{Frr,2} & \dots & x_{Frr,k} \\ \text{pop}_{Frr} & x_{Frr,2} & \dots & x_{Frr,k} \\ \text{pop}_{Frr} & x_{Frr,2} & \dots & x_{Frr,k} \\ \text{pop}_{Ssv} & x_{Ssv,2} & \dots & x_{Ssv,k} \\ \text{pop}_{Ssv} & x_{Ssv,2} & \dots & x_{Ssv,k} \\ \text{pop}_{Ssv} & x_{Ssv,2} & \dots & x_{Ssv,k} \end{bmatrix}$$

Variabeln  $\mathbf{G}$  betecknar avstånd utifrån en  $n \times n$  matris över kilometer mellan utgångsland och destinationsland. Därefter vektoriseras  $\mathbf{G}$ , vilket ger  $\mathbf{g} = \text{vec}(\mathbf{G})$ , en  $N \times 1$  vektor (LeSage & Pace 2009, s. 214).

Gravitationsmodellen ges slutligen i form av:

$$\mathbf{y} = \alpha \boldsymbol{\iota}_N + \mathbf{X}_d \boldsymbol{\beta}_d + \mathbf{X}_o \boldsymbol{\beta}_o + \gamma \mathbf{g} + \boldsymbol{\varepsilon} \quad (2)$$

där  $\mathbf{y}$  är migrationsflöden, koefficienten  $\alpha$  är den beroende variabelns intercept,  $\boldsymbol{\iota}_N$  är en  $N \times 1$  vektor av ettor,  $\boldsymbol{\beta}_d$  samt  $\boldsymbol{\beta}_o$  är  $k \times 1$  parametervektorer relaterade till destinationslandets respektive utgångslandets egenskaper, parametern  $\gamma$  visar på effekten av avstånd. Feltermen  $\boldsymbol{\varepsilon}$  är slumpmässiga störningar vars kovarians är  $\sigma^2 \mathbf{I}$ . Det är en  $N \times 1$  vektor med medelvärdet noll, konstant varians, samt har en kovarians av noll mellan dess störningar.

### 2.3.1 Prediktion för Gravitationsmodellen

Då  $\alpha$  samt  $\mathbf{g}$  är konstanter är det enbart matrisen  $\mathbf{X}$  som förändras över tiden. Vidare görs det framtida uppskattningar av de  $k$  variablerna i  $\mathbf{X}$ , vilken genererar matriserna  $\hat{\mathbf{X}}_d$  samt  $\hat{\mathbf{X}}_o$ . Ekvationen för prediktioner ett visst år är:

$$\hat{\mathbf{y}} = \hat{\alpha} \boldsymbol{\iota}_N + \hat{\mathbf{X}}_d \hat{\boldsymbol{\beta}}_d + \hat{\mathbf{X}}_o \hat{\boldsymbol{\beta}}_o + \hat{\gamma} \mathbf{g}. \quad (3)$$

En matris  $\mathbf{X}$  behövs därmed för vart år prediktionen görs. Parametrarna  $\alpha$ ,  $\boldsymbol{\beta}_d$ ,  $\boldsymbol{\beta}_o$ , samt  $\gamma$  estimeras genom minstakvadrat teknik från ekvation (2) samt används för att få fram  $\hat{\mathbf{y}}$ .

## 3 Data

### 3.1 SCB:s prediktioner

Statistiska Centralbyrån (2015) beskriver i sin rapport *Demografiska rapporter, Sveriges framtida befolkning 2015 – 2060* olika faktorer som påverkar migrationen. Huvudmodellen består av flertalet faktorer vilka kommer att redovisas i de nästkommande styckena. I rapporten beskrivs hur bland annat förmodade konflikter i världen, moderns födelseland, ålder, kön, arbets- samt studierelaterad migration påverkar migrationsflöden. Även den fria rörligheten i EU påverkar, då exempelvis immigrationen från Grekland ökade till följd av finanskrisen från 2007. I rapporten beskrivs hur den svenska ekonomin anses vara en av de mest robusta i EU, vilket leder till att Sverige är attraktivt land att emigrera till för individer födda inom EU.

Nettomigrationen definieras som antalet invandrare minus antalet utvandrare. I rapporten beskrivs att SCB har två alternativa antaganden för den predikterade nettomigrationen. Det första är ett antagande om en låg nettomigration där Sverige inte anses vara ett attraktivt land att emigrera till. Här antas en längre nivå av asylinvandring, vilket även medför en minskad anhöriginvandring. Även arbetskraftsinvandringen antas vara lägre då utgångsländer förutsätts ha en starkare ekonomi än i huvudprediktionen respektive den högre prediktionen. Det andra antagandet är en hög nettomigration. Asylinvandring antas här öka till följd av ett instabilt säkerhetsläge i omvärlden. Arbetskraftsinvandringen antas också som högre då ekonomiska omständigheter förutsätts som svårare för i utgångsländerna.

SCB:s prognoser baseras enbart på den folkbokförda befolkningen där individen har för avsikt att bosätta sig i det nya landet i minst ett år. Därmed exkluderas asylsökande som inte är folkbokförda samt visister från materialet. Antaganden för invandring skiljer sig beroende på födelsegrupp, där SCB delat in länder i grupperna: Sverige, Norden (utom Sverige), EU (utom Norden), övriga Europa, länder utanför Europa med hög utvecklingsnivå, länder utanför Europa med medel utvecklingsnivå samt länder utanför Europa med låg utvecklingsnivå. Utvecklingsnivån baseras på HDI enligt FN:s indelning. Vidare beskriver SCB att modellerna skiljer sig åt beroende på vilken region människor migrerar till eller från.

Detta utgör SCB:s huvudmodell för prediktion av migration. En av SCB:s kompletterande modeller beskrivs i rapporten vara en AR(1) modell:

$$Migr_t = a_1 Migr_{t-1} + \mu(1 - a_1) + e_t$$

där  $\mu$  anpassas så att framskrivningen av migrationen blir konsistent med huvudalternativet (Statistiska Centralbyrån, 2015). För att fånga upp osäkerheten i prognosen görs ett lågt, respektive högt antagande. I tabell 2 visas SCB:s prediktionsskattningar för år 2022 (Statistiska Centralbyrån, 2015).

Tabell 2: SCB:s prognos för år 2022

Skattningar	Immigration	Emigration	Nettomigration
Aktuell skattning	127929	67377	60552
Hög skattning	160484	78630	81854
Låg skattning	66162	44451	22161

### 3.2 Migrationsdata

Datamaterialet för AR(1)-modellen är hämtat från SCB:s statistikdatabas. Materialet har sammanställts utifrån Folk- och Bostadsräkningarna av SCB utifrån uppgifter rörande in- och utvandring (Statistiska Centralbyrån, 2017a). Materialet rapporterades mellan åren 1960–2017. Årlig rapportering finns från 1980 vilket det därför görs till första året i tidsserien för detta arbete.

För gravitationsmodellen används 2017 års data (Statistiska Centralbyrån, 2017b). Materialet består av antal invandringar och utvandringar efter ut- och inflyttningsland för 203 ”länder”. Vid bearbetning exkluderas icke längre existerande stater från materialet; Jugoslavien, samt Serbien-Montenegro. Migration till och från ”okänt land” exkluderas. Sverige inkluderades. Territorier utan självstyre, samt av FN ej erkända stater vilka inkluderades i materialet är Brittiska Jungfruöarna, Taiwan, Kosovo, Palestina och Vatikanstaten. Materialet består efter bearbetning av 201 länder.

### 3.3 Internationell befolkningsstatistik

Information om internationell befolkningsstatistik delas upp i två variabler: befolkningsstorleken i land  $i$  år 2017, respektive framskrivning av befolkningsstorlek i land  $i$  år 2022. Datamaterialet för befolkningsstorlek 2017 är hämtat från Världsbankens databas (World Bank Group, 2018b). Data tas in för de länder som angivits i datamaterialet från SCB. Kvaliteten på populationsberäkningarna varierar mellan länder. Användning av folkbokföringsregister ligger till grund för beräkningar i vissa länder, medan andra



länders befolkningar anges som skattningar av tidigare folkräkningar. Skillnader i kvalitet diskuteras inte närmare i detta arbete. Befolkningsmängd anges i tusental innan den skrivs om till ental. Framskrivning av befolkningsstorlek 2022 hämtas även den från världsbankens databas (World Bank Group, 2018a). Populationstillväxten antas vara exponentiell, ett antagande som inte diskuteras närmare i detta arbete men som givetvis har betydelse för antalet individer i beräkningarna. Befolkningsstorlek för land  $i$  anges i ental.

### 3.4 HDI

Human Development Index (HDI) är ett mått som används för att beskriva den mänskliga utvecklingen i ett land och baseras på tre kriterier som sedan vägs ihop till ett index mellan 0 och 1. Kriterierna är förväntad livslängd vid födseln, utbildningsnivå samt bruttonationalinkomst per capita.

HDI-ranking hämtas från UNDP, FN:s utvecklingsprogram (United Nations Development Programme, 2018). Rankingen kategoriseras av FN i fyra kategorier från högt till lågt med "very high HDI" som högst och "Low HDI" som lägst. SCB har modifierat rankingen och omarbetat den till tre kategorier (Statistiska Centralbyrån, 2005). Index för år 2017 har i detta arbete justerats efter denna princip genom att slå samman de två övre kategorierna, "Very high HDI" och "High HDI". HDI kodas om till dummyvariabler, där kategori "Low HDI" behandlas som referens. Valet av införandet av dummyvariabler grundar sig i att det exakta värdet för HDI för ett land är av mindre intresse då det är kategoriseringen av landet och de implikationer som följer med den kategoriseringen som är av intresse. Kategoriseringen underlättar därmed jämförelser och vidare analys. Resonemanget ligger i linje med SCB:s val av variabelkodning.

### 3.5 Avstånd

Information om avstånd mellan Stockholm och världens övriga huvudstäder har inhämtats. Avståndet anges i kilometer för den rakaste vägen mellan huvudstäderna efter att ha omarbetats från mil (Flygtid.se, 2018). Inmatning sker manuellt i en  $n \times n$  matris. Kolumn  $i$  anger avstånd mellan Sverige och land  $i$ , på samma sätt anger rad  $i$  ett avstånd mellan Sverige och land  $i$ . Övriga element i matrisen sätts till noll. Notera att avståndet Stockholm–Stockholm är noll.

### 3.6 Utrikesdepartementets avrådan

Utrikesdepartementet, UD, avråder från resor till ett land eller delar av detta land då säkerhetssituationen bedöms som "mycket farlig, oberäknelig eller snabbt föränderlig" (Regeringskansliet, 2015). UD:s avrådan omfattar endast ett fåtal gånger ett land i sin helhet då säkerhetsläget kan variera stort inom landet. Förekomsten på listan indikerar dock en instabilitet vilken med troligen kan ligga till grund för emigration från landet. Därför inkluderas samtliga 48 länder som UD listat (i Appendix återfinns listan med samtliga länder). Information om de länder och områden UD avråder ifrån inhämtas och sammanställs manuellt (Regeringskansliet, 2018). Avrådan kodas till en dummyvariabel, med icke-avrådan som referens.

### 3.7 Region

Valet att inkludera "Region" möjliggör att jämföra grupper av länder med liknande geografisk position. Följaktligen går det att urskilja potentiella skillnader för olika regioner för variablerna "HDI" och "UD avråder". Detta kan vara av betydelse i konfliktdrabbade områden då flöden från grannländer kan relateras till säkerhetsläget i regionen. Länder fördelas mellan nio regioner skapade för detta arbete. Överlappning mellan regioner förekommer ej då fördelning sker enligt principen en region per land. Sammanställning sker manuellt. Region kodas till dummyvariabler, med Norden som referens. Regionerna är:

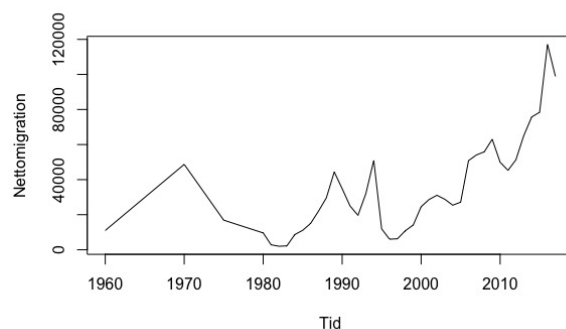
1. Norden
2. EU, exklusive de nordiska länderna
3. Övriga europeiska länder, exklusive de nordiska länderna samt EU-länder
4. Afrika
5. Asien
6. Mellanöstern
7. Sydamerika
8. Nordamerika
9. Oceanien

## 4 Beräkningar

Inmatning och sammanställning av datamaterialet görs i datahanteringsprogrammet Excel. Modeller och samtliga beräkningar görs i R-studio. Valet av program grundar sig i dess utformning och anpassning till arbetsuppgift.

### 4.1 Sveriges nettomigration över tid

Figur 1 illustrerar nettomigrationen till Sverige år 1960–2017, där det ser ut att vara en ökning över tid. Notera att data mellan åren 1960–1980 rapporterats in vart femte år. Från år 1980 rapporteras data in på årsbasis.

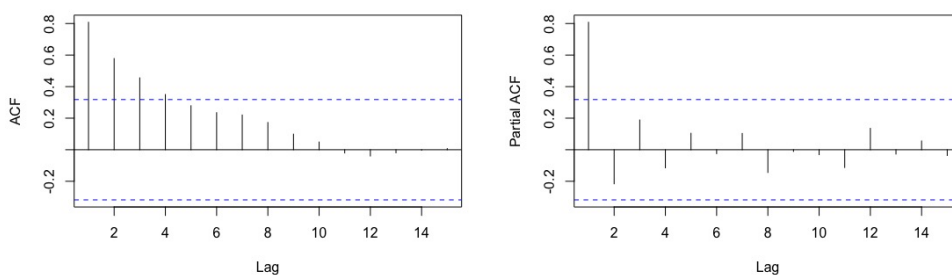


Figur 1: Migration över tid

## 4.2 AR(1)

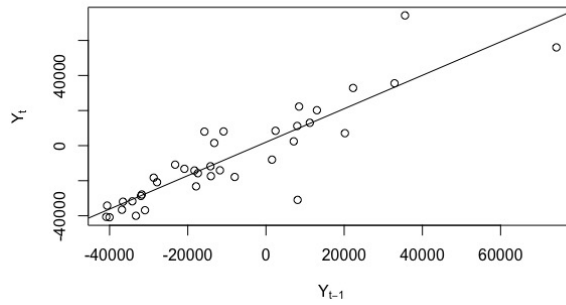
Då en AR(1)-process har antagandet att  $\mu = 0$  subtraheras observationernas medelvärde från den ursprungliga tidsserien, innan en modellen anpassas och värdet  $\hat{\phi} = 0.9277$  erhålls. För att processen ska vara stationär krävs att  $|\phi| < 1$ . Då värdet på  $\hat{\phi}$  är nära ett utförs ett Dickey-Fuller test som undersöker om det finns en enhetsrot ( $\rho = 1$ ), d.v.s. om processen är icke-stationär. Ett p-värde på 0.8635 erhålls och nollhypotesen att processen är icke-stationär går inte att förkasta. Vidare är residualerna för nettomigration i AR(1)-modellen normalfördelade utifrån grafisk analys.

Vidare undersöks ACF och PACF. Den avtagande ACF:en (den vänstra grafen i Figur 2) tyder på att processen är autoregressiv, medan PACF:en (den högra grafen i Figur 2) visar att processen har ordning 1. Dessa två grafer antyder att processen är en AR(1) och beslutet tas att inte differentiera modellen då antagandet om stationaritet anses vara uppfyllt.



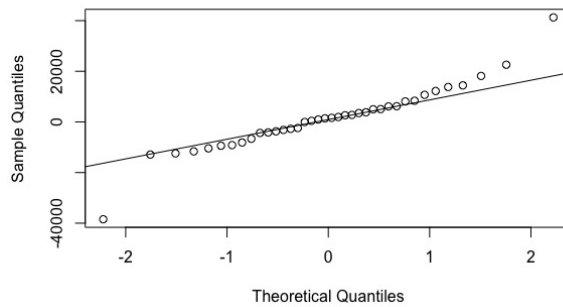
Figur 2: ACF samt PACF

Figur 3 visar en graf av  $Y_t$  mot  $Y_{t-1}$ . Då observationerna är samlade runt den räta linjen verkar det rimligt att anta att ett linjärt samband mellan  $Y_t$  och  $Y_{t-1}$  existerar.



Figur 3:  $Y_t$  mot  $Y_{t-1}$

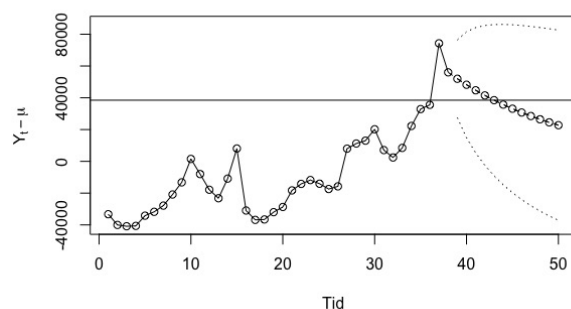
QQ-plotten i Figur 4 visar på normalfördelade residualer för nettomigrationen. Residualanalysen indikerar att det råder oberoende mellan residualerna, vilket är en premis för AR(1)-processens normalfördelningsantagande.



Figur 4: QQ-plot över residualer för AR(1)-modellen

### 4.3 Prediktion AR(1)

Predikterade värden för AR(1) modellen illustreras i figur 5. Den horisontella axeln anger tid ( $t$ ), där  $t = 0$  motsvarar år 1980 och  $t = 37$  motsvarar år 2017. Prediktionsintervall illustreras genom de streckade linjerna. Värdet för nettomigrationen år 2022 visas med en horisontell linje när  $Y = 38487.18$ . Prediktionen för nettomigrationen år 2022 blir då  $38487.18 + \mu = 81340.89 \approx 81000$ .



Figur 5: Prediktion för AR(1)

#### 4.4 Gravitationsmodell

Gravitationsmodellen sätts upp enligt ekvation (2). I matrisen  $\mathbf{X}$  finns 12 variabler totalt: "Population", om "UD avråder", "Region" uppdelat på 8 dummyvariabler, samt "HDI" uppdelat på två dummyvariabler. Region Norden, samt lågt HDI är referensnivåer. Parametrarna skattas med minstakvadratteknik.

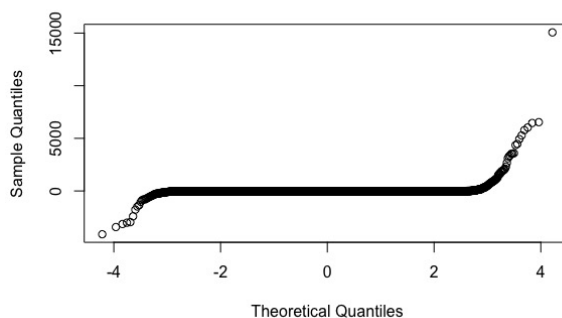
Tabell 3: Parameterskattningar, medelfel samt p-värden för skattad Graviationsmodell

Koefficient	Skattning	Standardavvikelse	p-värde
Intercept	83.87	7.683	$< 2 \cdot 10^{-16}$ ***
Xd.Population	$1.565 \cdot 10^{-8}$	$4.787 \cdot 10^{-9}$	0.00108 **
Xd.UD avråder	3.434	1.680	0.04099 *
Xd.EU	68.92	5.109	$< 2 \cdot 10^{-16}$ ***
Xd.Övriga Europa	66.46	5.224	$< 2 \cdot 10^{-16}$ ***
Xd.Afrika	59.22	5.187	$< 2 \cdot 10^{-16}$ ***
Xd.Asien	62.15	5.150	$< 2 \cdot 10^{-16}$ ***
Xd.Mellanöstern	72.62	5.533	$< 2 \cdot 10^{-16}$ ***
Xd.Sydamerika	61.25	5.458	$< 2 \cdot 10^{-16}$ ***
Xd.Nordamerika	61.85	5.122	$< 2 \cdot 10^{-16}$ ***
Xd.Oceanien	61.05	5.403	$< 2 \cdot 10^{-16}$ ***
Xd.HDI hög	-3.770	2.163	0.08138
Xd.HDI medel	-2.670	2.124	0.20863
Xo.Population	$-5.153 \cdot 10^{-9}$	$4.787 \cdot 10^{-9}$	0.28164
Xo.UD avråder	0.1526	1.680	0.92765
Xo.EU	-147.0	5.109	$< 2 \cdot 10^{-16}$ ***
Xo.Övriga Europa	-145.5	5.224	$< 2 \cdot 10^{-16}$ ***
Xo.Afrika	-144.9	5.187	$< 2 \cdot 10^{-16}$ ***
Xo.Asien	-144.7	5.150	$< 2 \cdot 10^{-16}$ ***
Xo.Mellanöstern	-145.2	5.533	$< 2 \cdot 10^{-16}$ ***
Xo.Sydamerika	-145.0	5.459	$< 2 \cdot 10^{-16}$ ***
Xo.Nordamerika	-145.3	5.122	$< 2 \cdot 10^{-16}$ ***
Xo.Oceanien	-145.5	5.403	$< 2 \cdot 10^{-16}$ ***
Xo.HDI hög	-0.4052	2.163	0.85139
Xo.HDI medel	-0.07358	2.124	0.97236
g.Avstånd	0.01324	$9.687 \cdot 10^4$	$< 2 \cdot 10^{-16}$ ***

Tabell 3 visar parameterskattningar och medelfel för den skattade gravitationsmodellen. Majoriteten av variablerna är signifikanta med ett p-värde mindre än 0.001 för både destinationslandets egenskaper, utgångslandets egenskaper, samt avståndet mellan utgångsland och destinationsland. Även då variablerna "Avstånd" respektive "nettomigration" standardiserats genom logarimtering, förblir resultatet oförändrat sett till signifikans. Destinationslandets populationsstorlek är signifikant vid 0.01. UD:s avrådan är signifikant för destinationslandet vid 0.05. Icke signifikanta variabler är de dummyvariabler relaterade till HDI. Den variabel bland de icke-signifikanta som ligger närmast signifikans är destinationslandets HDI med ett p-värde av 0.08138. Modellen har en förklaringsgrad  $R^2$  på enbart 4 %.

För att undersöka om det finns skillnader mellan olika regioner i medelvärde på variablerna "UD avråder" och "Immigration" används Tukeys HSD på nivån fem procent. Med hänsyn till variabeln "UD avråder" finns det en skillnad i medelvärde mellan regioner i åtta fall. Den vanligast förekommande regionen är Mellanöstern, vilken förekommer i fem fall. Testet genomförs även för variabel "Invandring" där Mellanöstern skiljer sig från alla regioner utom Norden och EU. Invandringen är i dessa fall högre, med undantag för Norden och EU.

Se Figur 6. Residualerna för nettomigrationen är inte normalfördelade då merparten av observationerna ligger kring noll. Det påverkar dock inte metodvalet då punktskattningen av nettomigration inte kräver normalfördelade residualer. Följaktligen krävs inte heller en normalfördelning av datamaterialet för prediktion av nettomigrationen.

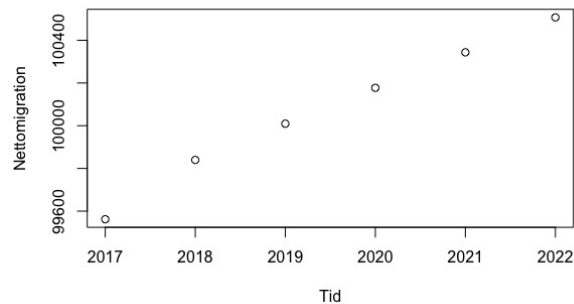


Figur 6: Residualanalys av Gravitationsmodellen

## 4.5 Prediktion för Gravitationsmodellen

Parameterskattningarna  $\hat{\alpha}$ ,  $\hat{\beta}_d$ ,  $\hat{\beta}_o$ , samt  $\hat{\gamma}$  estimeras med minstakvadratteknik. Av de 12 variablerna i  $\mathbf{X}$  är det enbart variabeln population som antas förändras över tid. Världsbankens populationsskattningar för år 2022 ersätter 2017 års population i  $\mathbf{X}$ . Då prediktionen enbart avser ett kort tidsperspektiv antas HDI inte förändras nämnvärt och hålls därmed konstant. Ekvation (3) används för att bestämma  $\hat{y}_{2022}$  vilken summeras för att erhålla den totala nettomigrationen till Sverige. Gravitationsmodellen ger en predikerad nettomigration på  $100507.1 \approx 101000$  år 2022.

Gravitationsmodellens prediktioner för år 2018 till 2022 illustreras i Figur 7 samt den faktiska nettomigrationen för år 2017. Nettomigrationen år 2022 predikteras till 101000.



Figur 7: Prediktioner av nettomigration för Gravitationsmodellen

## 5 Analys och diskussion

### 5.1 AR(1)

Inledningsvis anpassades en autoregressionsmodell till datamaterialet, varpå autokorrelationsfunktionen och den partiella autokorrelationsfunktionen undersöktes. Den exponentiellt avtagande ACF:en tyder på att en autoregressiv modell. PACF:en med ett signifikant tidslag och en tydlig cut-off-effekt visar på ordning 1. Anpassningen till datamaterialet indikerar också stationaritet. I Figur 1 är  $\mu = 42900$  och det kan urskiljas att den faktiska nettomigrationen har en uppåtgående trend över tid. Detta tolkas som en indikation på att AR(1)-processen i datamaterialet 1980 befann sig i ett globalt minimum och 2017 ett globalt maximum, där variationerna under perioden 1980–2017 kan tolkas som lokala minimum och lokala maximum. Tolkningen om globalt minimum 1980 kan möjligen stödjas av tidigare data från 1960–1975. Nettomigrationen det första registrerade året i datamaterialet, år 1960, var runt 11000 och ett decenium senare var den nästan 49000, vilket ligger över det skattade medelvärdet. År 1975 minskade nettomigrationen till nästan 17000. De tidigare observationerna alternerar mellan att ligga över samt under  $\hat{\mu}$ . Då tidigare värden historiskt sett också varierat runt medelvärdet kan detta också styrka stationaritetsantagandet. Vidare visar det höga värdet av  $\hat{\phi}$  en stark återkoppling och förklarar varför ökningen av nettomigration är stor med en lång uppåtgående trend.

Som tidigare nämnts är AR-modeller mycket användbara inom tidsserieanalys då den använder tidigare värden för att förklara framtida värden samt att göra prediktioner. Dock ger chocker permanenta förändringar i processen, vilken kan tänkas var fallet år 2015 när migrationen till Sverige ökade kraftigt under en kort period. Då AR-processen enbart tar tidigare värden i beaktning fångar den inte upp samhällsförändringar som kan påverka hur migrationen förändras.

### 5.2 Gravitationsmodellen

De variabler vilka i modellen förklarar nettomigrationen är främst de vilka relaterar till region. Migrationsflöden i länder tillhörande samma region är mer lika än länder tillhörande olika regioner. Relaterat till avstånd kan detta ses som en av modellens styrkor; länder inom samma region med ungefärligt samma avstånd har liknande flöden.

Tolkningen av Tukeys test är att länder i Mellanöstern har större sannolikhet att finnas med bland de länder UD avråder ifrån att resa till, än länder i jämförda regioner. Testet genomfördes även för variabel "Invandring" där Mellanöstern skiljde sig från alla regioner utom Norden och EU där invandringen i dessa fall är högre, med undantag för Norden och EU. Detta indikerar att Mellanösterns säkerhetsläge i högre utsträckning än andra regioners påverkar emigrationen från regionen. Pearsons korrelation visade tydligt på en samvariation. Detta indikerar att Mellanösterns säkerhetsläge i högre utsträckning än andra regioners påverkar emigrationen från regionen.

Populationen i destinationslandet är signifikant, i motsats till populationsstorleken i utgångslandet. Vid immigration till Sverige är den svenska populationsstorleken av betydelse, medan utgångslandets populationsstorlek inte är det. Detta kan anses vara en brist i modellen då teorin bygger på att just populationsstorlek är avgörande för flödesstorlek. Resultatet kan tolkas som att populationsstorleken i utgångslandet är mindre betydelsefull för immigrationen till Sverige. Med andra ord sätter utgångslandets populationsstorlek den maximala storleken av emigrationsflödet. Det innebär dock inte att populationens storlek avgör flödets storlek.

Gällande UD:s avrådan skiljde sig signifikansen mellan destinationsland och utgångsland. Att UD:s avrådan är signifikant i destinationsland kan tolkas som att människor i högre utsträckning väljer att emigrera till länder med ett högre eller likvärdigt säkerhetsläge som det land de lämnar. Människor verkar i regel inte emigrera till länder med ett sämre säkerhetsläge. Säkerhetsläget i utgångslandet är däremot mindre relevant då anledningarna till emigrationen skiljer sig mellan länder. Att "UD avråder" inte är en signifikant variabel kan tolkas som att immigrationen till Sverige är så pass diversifierad i sin sammansättning, med hänsyn till denna variabel, att immigrationen inte entydigt kan förklaras med osäkerhet i omvärlden.

Variabeln "Avstånd" har hög signifikans. Detta kontradikterar modellens anstats: att kortare avstånd ska generera större flöde. Resultatet tolkas som att ju längre avstånd mellan utgångsland och destinationsland, desto högre nettomigration mellan länder. Sett till den ovan diskuterade betydelsen av region är resultatet att förstå utifrån denna aspekt.

### 5.3 Komparativ analys

Genom Gravitationsmodellens flödesberäkning av nettomigrationen predikteras 2022 års nettomigration till 101000. Motsvarande prediktion för AR(1)-modellen ger en nettomigration av 81000. Differensen uppgår ganska precist till 20000. Differensens storlek kan till stor del förklaras av modellernas utformning, vilka kommer att driva respektive prediktion i två olika riktningar.

AR(1)-modellen strävar genom sin koefficient efter att återgå till sitt medelvärde. Då koefficienten  $\phi$  skattas till 0.9277, är tillväxten mindre än 1. Följaktligen minskar prediktionerna fram till den punkt då AR(1)-processen nått sitt medelvärde. Med andra ord anger koefficienten tillväxttakten som i detta fall ger en nedgående prediktionstrend.

Gravitationsmodellens prediktion baseras på skattningarna av ländernas populations-tillväxt. Detta har två viktiga implikationer för prediktioner med Gravitationsmodellen. Det första har att göra med modellens  $\mathbf{X}_d$  och  $\mathbf{X}_o$ . Modellen bygger, som tidigare nämnts, på teorin att avstånd och populationsstorlek avgör flöden mellan områden. Då ländernas populationsstorlekar ökar, bör teoretiskt sett flödet mellan populationerna tillta. Med andra ord förväntas  $\hat{\beta}\mathbf{X}_i$  öka där  $i = d, o$ . Det bidrar både till en ökning i immigration och en ökning i emigration. Givet att förhållandet mellan dem förblir oförändrat,



ökar nettomigrationen. Det andra är en mer direkt effekt av världens generella populationsökning. Modellens X-matriser ger givetvis ökade variabelvärden till följd av större populationer i länderna och betakoefficienter större än 0, något som genererar en prediktion av nettomigrationen vilken är större än 2017 års nettomigration.

Med andra ord avgörs differensen av att AR(1)-processens koefficient är mindre än 1 och Gravitationsmodellens koefficienter större än 0. Ju längre tidsperiod, desto större differens kan förväntas mellan modellernas prediktioner.

SCB:s huvudmodell predikterar att nettoflödet för migration uppgår till 60552 individer år 2022. AR(1)-modellen ligger närmast med en skattning av 81341 individer. AR(1)-modellen ligger mycket nära SCB:s alternativa antagande om en hög immigration, vilken uppgår till 81854 individer. Prediktionen kan delvis förklaras med 2015 års förhöjning av modellens medelvärde, vilket med stor sannolikhet är avgörande för att prediktionen överskrider SCB:s huvudprediktion men ligger så nära dess alternativa antagande. Gravitationsmodellens prediktion uppgår som tidigare nämnts till hela 101000 individer. Skattningen är betydligt högre än både SCB:s huvudprediktion och det alternativa antagandet om högre migration. Förklaringen ligger delvis som tidigare nämnts i modellens antagande att populationen ökar, allt annat lika. Delvis kan den höga prediktionen förklaras med den höga nettomigrationen för 2017. I relation till övriga prediktioner måste Gravitationsmodellens prediktion anses ha ett lågt tillitsvärde. Vilken av de två modellerna som givit bäst prediktion är i skrivande stund omöjligt att avgöra. Men skulle uppsatsförfattarna satsa sina slantar är AR(1)-modellens prediktion favorit.

## 6 Slutsatser

Båda modellernas prediktioner ligger över SCB:s huvudantagande. Den modell som är närmast SCB:s huvudantagande är AR(1)-modellen vilken även låg oerhört nära SCB:s alternativa antagande om högre immigration. Vad som ligger till grund för modellernas så olika prediktioner är deras ansatser. Valet av modeller grundar sig i modellernas fundamentala olikhet. AR(1)-modellen karaktäriseras av dess endogena uppbyggnad, medan Gravitationsmodellen av exogenitet. Modellernas prediktion av nettomigrationen 2022 skiljde sig betydande åt med en differens av 20000 individer. Differensen är på sätt och vis inte helt oväntad utifrån modellernas så vitt skilda ansatser. Pondera att Gravitationsmodellen jämförts med en annan spatial interaktionsmodell, då hade ansatsen, uppbyggnaden samt variablerna varit ungefär de samma. Följaktligen kan mindre differens i prediktion förväntas mellan dessa två. Differensens storlek visar även på vikten av att använda olika modeller som varandras komplement. Modellernas olika ansatser tillför olika information. AR(1)-modellen bör kanske även ses som ett verktyg för att kontrollera andra modellers prediktioner, snarare än en faktisk modell för prediktion av nettomigration. Det gör den inte mindre betydelsefull men tillskriver den ännu en roll i prediktioner. Redan idag används AR(1)-modellen på detta vis av bland andra SCB (Statistiska Centralbyrån, 2015). Den kan även komplettera andra modeller genom att inkluderas i dessa.

Gravitationsmodellens största svaghet är som tidigare nämnts dess låga förklaringsgrad. En modell med mindre än 4 procent av variationen förklarad må anses vara högst osäker i sina prediktioner. Modellen lider av grundantagandet att det råder oberoende mellan utgångsländer och destinationsländer. Genom att förlänga modellen till spatialt autoregressiv filtreras det spatiala beroendet mellan utgångsländer och destinationsländer (LeSage & Pace, 2009, s. 215–216).

$$(\mathbf{I}_N - \rho_d \mathbf{W}_d)(\mathbf{I}_N - \rho_o \mathbf{W}_o)\mathbf{y} = \alpha \boldsymbol{\nu}_N + \mathbf{X}_d \boldsymbol{\beta}_d + \mathbf{X}_o \boldsymbol{\beta}_o + \gamma \mathbf{g} + \mathbf{e} \quad (4)$$

Modellen visar genom variabeln  $\mathbf{W}_i$ , den spatials anslutningen mellan  $n$  regioner, tendensen hos flöden i ett land att ha inverkan på flöden i ett grannland, samt de flöden från ett utgångslands grannländer till destinationslandets grannländer (LeSage & Pace, 2009, s. 218–220). Annorlunda uttryckt visar det tendensen hos individer i Pakistan att emigrera till Sverige då individer från Afghanistan emigrerar till Sverige, respektive där emigration från Afghanistan till Sverige påverkar emigrationen från Pakistan till Norge. För att öka precisionen i flödesberäkningarna kan den utvecklade formen av gravitationsmodellen kompletteras med en Maximum-likelihood-skattning (LeSage & Pace, 2009, s. 218–220). Som nämnts i stycket ovan skulle även en inkludering av autoregression kunna vara till gagn för modellen. Utvecklingen av modellen begränsades med hänseende till arbetets omfattning.

Förklaringsgraden skulle kunna förbättras genom förfining av modellens variabler. Säkerhetslägets klassificering skulle kunna bli mer beskrivande. Då variabeln "UD avråder" kodats som en dummy-variabel, inkluderas samtliga former av avrådan. Därmed särskiljs inte länder med olika klassificeringar vilka kan ha stor skillnad i innebörd. Ett tydligt exempel är Sydsudan och Turkiet; Sydsudan är ett land med akut livsmedelsbrist och pågående inbördeskrig där svenskar av UD uppmanas lämna landet; Turkiet är ett populärt turistmål med avrådan från resor till områden direkt gränsande till den Syriska gränsen. Ett lands förekomst på UD:s lista speglar inte i sig säkerhetsläget för den genomsnittlige invånaren i landet. Genom att koda om säkerhetsläget till UD:s tre nivåer av avrådan kan större skillnader fångas upp. Vidare fattas variabler vilka kan indikera mer precist vilken typ av migration det rör sig om. Att introducera en variabel för arbetskraftsinvandring skulle möjliggöra att prediktioner i högre utsträckning skulle kunna relateras till länders ekonomiska läge och konjunktur. En annan variabel är den för asylinvandring. Båda former av immigration kan analyseras genom AR(1) processen för att möjligen visa på skillnader i flödena för dessa. Detta skulle möjligen kunna öka prediktionernas precision ytterligare.

Prediktioner med denna möjliggör utifrån denna modell underlag för samhällsanalyser och politiska beslut gällande resursfördelning. Givet att användningen av den utvecklade modellen bidrar till bättre prediktioner, har det ett stort ekonomiskt och samhälleligt värde. Därtill finns givetvis den enskildes perspektiv. Bakom varje individ finns en person. Med möjlighet att planera och budgetera för variationer i nettomigrationen kan myndigheters handläggningstider förkortas. Värde av en människas tid är betydande i en etableringsprocess och bör därför behandlas med största varsamhet. Ju kortare tid en person uppehålls från att fortsätta med sin livsplan; vidare till arbete, studier eller annan sysselsättning, desto bättre.

## 7 Referenser

- Cryer, J. D., Chan, K. 2008. *Time Series Analysis With Applications in R*. 2:a upplagan. New York: Springer.
- Flygtid.se. 2018. *Lista över världens huvudstäder med avstånd från Stockholm, Sverige*. <https://flygtid.se/varldens-huvudstader> (Hämtad 2018-10-01)
- Gudmundson, P. 2017. Så underskattas invandringen - år efter år. *SvD*. 7 februari. <https://www.svd.se/sa-underskattas-invandringen--ar-efter-ar> (Hämtad 2018-12-21)
- LeSage, J., Pace, R. K. 2009. *Introduction to Spatial Econometrics*. Boca Raton: Taylor & Francis Group.
- Porell, F.W., Hua, C-I. 1981. *An econometric procedure for estimation of a generalized systemic gravity model under incomplete information about the system* Regional Science and Urban Economics. 11(4), s.585-606 <https://www.sciencedirect.com/science/article/pii/0166046281900405>(Hämtad: 2019-01-21)
- Poot, J., Alimi, O., Cameron, M. P., Maré, D. C. 2016. *The Gravity Model of Migration: The Successful Comeback of an Ageing Superstar in Regional Science*. Discussion paper series. IZA DP No.10329. Bonn: Institute for the Study of Labour. <http://ftp.iza.org/dp10329.pdf>(Hämtad: 2019-01-03)
- Regeringskansliet. 2015. *Vad innebär en avrådan*. <https://www.regeringen.se/uds-reseinformation/vad-innebar-avradan/> (Hämtad 2018-11-07)
- Regeringskansliet. 2018. *UD avråder*. <https://www.regeringen.se/uds-reseinformation/ud-avrader/> (Hämtad 2018-11-07)
- Shumway, R.H., Stoffer, D. S. 2006. *Time Series Analysis and Its Applications With R Examples*. 2:a upplagan. New York: Springer.
- SOM-Institutet. 2017. *SOM-Institutet*. [https://som.gu.se/digitalAssets/1644/1644259\\_26-k--nedom-om-och-f--rtroende-f--r-scb-2016.pdf](https://som.gu.se/digitalAssets/1644/1644259_26-k--nedom-om-och-f--rtroende-f--r-scb-2016.pdf) (Hämtad 2019-01-21)
- Statistiska Centralbyrån. 2005. *SCB:s modell för befolkningsframskrivningar*. Bakgrundsmaterial om demografi, barn och familj 2005:1. [https://www.scb.se/statistik/\\_publikationer/be0401\\_2005a01\\_br\\_be52op0501.pdf](https://www.scb.se/statistik/_publikationer/be0401_2005a01_br_be52op0501.pdf)
- Statistiska Centralbyrån. 2015. *Sveriges framtida befolkning 2015-2060*. Demografiska Rapporter, rapport 2015:2. [https://www.scb.se/contentassets/2403a2a8ed644d0b8cf378e831359264/be0401\\_2015i60\\_br\\_be51br1502.pdf](https://www.scb.se/contentassets/2403a2a8ed644d0b8cf378e831359264/be0401_2015i60_br_be51br1502.pdf)
- Statistiska Centralbyrån. 2017a. *folk- och bostadsräkningarna (FoB)*. <http://www.scb.se/Statistik/BE/BE0101/2013A02F/be0101tab8samdrag.xls>
- Statistiska Centralbyrån. 2017b. *Invandringar och utvandringar efter ut-/inflyttningsland och kön*. [http://www.statistikdatabasen.scb.se/pxweb/sv/ssd/START\\_\\_BE\\_\\_BE0101\\_\\_BE0101J/ImmiEmiFlytt/?rxid=b8106016-94d7-4add-bde3-3e7d90f526ce](http://www.statistikdatabasen.scb.se/pxweb/sv/ssd/START__BE__BE0101__BE0101J/ImmiEmiFlytt/?rxid=b8106016-94d7-4add-bde3-3e7d90f526ce)
- United Nations Development Programme. 2018. *Human Development Index and its components*. Human Development Reports. <http://hdr.undp.org/en/composite/HDI> (Hämtad 2018-10-04)
- Wadensjö, E. 2012. *Framtidens migration*. Underlagsrapport/Statsrådsberedningen:5. Stockholm: Statsrådsberedningen <https://www.samfak.gu.se/digitalAssets/1441/>

1441788\_underlagsrapport\_5\_framtidskommissionen\_w.pdf

World Bank Group. 2018a. *Population estimates and projections*. <http://databank.worldbank.org/data/source/population-estimates-and-projections/preview/>on (Hämtad 2018-10-21)

World Bank Group. 2018b. *Population, total*. <https://data.worldbank.org/indicator/SP.POP.TOTL> (Hämtad 2018-10-01)

## Appendix

Lista över länder för vilka UD utfärdat avrådan:

Afghanistan	Georgien	Libyen	Ryska Federationen
Algeriet	Haiti	Malaysia	Saudiarabien
Azerbajdzan	Indien	Mali	Somalia
Burkina Fazo	Indonesien	Mauretanien	Sudan
Burundi	Irak	Moçambique	Sydsudan
Centralafrikanska Republiken	Iran	Nicaragua	Syrien
Colombia	Israel	Niger	Tchad
Demokratiska Republiken Kongo	Jemen	Nigeria	Thailand
Ecuador	Jordanien	Nordkorea	Tunisien
Egypten	Kamerun	Pakistan	Turkiet
Elfenbenskusten	Kenya	Palestina	Ukraina
Filippinerna	Libanon	Republiken Kongo	Venezuela