



LUNDS UNIVERSITET  
Ekonomihögskolan

Företagsekonomiska institutionen

FEKH39

Kandidatkurs i Business and Data Analytics

HT 2022

# Drivande faktorer för individuell användning av big data system inom revisionsbolag

Kandidatuppsats

**Författare:**

Christoffer Yngwe

Jacob Svensson

Rasmus Thunberg

**Handledare:**

Magnus Johansson

# *Abstract*

**Title:** Drivande faktorer för individuell användning av big data system inom revisionsbolag

**Seminar date:** 11/1-2023

**Course:** FEKH39, Degree Project, Business Administration, Undergraduate level, 15 University Credit Points ECTS

**Authors:** Christoffer Yngwe, Jacob Svensson & Rasmus Thunberg

**Advisor:** Magnus Johansson

**Keywords:** Audit, big data systems, UTAUT-model, big data usage, general ledger systems.

**Research questions:** Which factors affect individual usage of big data systems for employees within the audit industry?

**Purpose:** Explaining and mapping driving factors for individual usage of big data systems for auditors.

**Methodology:** The study applies the UTAUT-model and performs a regression analysis in order to map the driving factors for individual usage of big data systems for auditors. Data is collected through a survey which was sent out to employees at four big auditing firms in Sweden. By using bootstrapping the study estimates the significance levels and coefficients for the factors affecting individual usage of big data systems.

**Theoretical perspectives:** The paper's theoretical framework is based on the UTAUT-model, which is adjusted accordingly to recent theories regarding big data usage and decision-making.

**Results:** The study identified *social influences*, *facilitating conditions* and *expected performance* as the driving factors for individual usage of big data systems within the auditing industry. All of the variables had a positive statistically significant relationship with the dependent variable, *actual usage*.

**Conclusions:** *Social influences*, *facilitating conditions* and *expected performance* had a significant positive relationship with *actual usage* of big data systems for auditors.

# Sammanfattning

**Titel:** Drivande faktorer för individuell användning av big data system inom revisionsbolag

**Seminariedatum:** 11/1-2023

**Kurs:** FEKH39, Examensarbete på kandidatnivå i Business & Data Analytics, 15 högskolepoäng

**Författare:** Christoffer Yngwe, Jacob Svensson & Rasmus Thunberg

**Handledare:** Magnus Johansson

**Nyckelord:** Revision, big data system, UTAUT-modellen, big data användning, huvudboksanalyssystem.

**Forskningsfråga:** Vilka faktorer påverkar individuell användning av big data system bland revisorer?

**Syfte:** Kartlägga och förklara drivande faktorer som påverkar individuell användning av big data system inom revisionsbranschen.

**Metod:** Studien tillämpar UTAUT-modellen och utför en regressionsanalys för att kartlägga drivande faktorer till individuell användning av big data system för revisorer. Data samlas in via en enkät som skickades ut till anställda på fyra revisionsföretag i Sverige. Genom användning av bootstrapping uppskattar studien signifikansnivån och koefficienterna för faktorer som påverkar individuell användningsgrad av big data system.

**Teoretiska perspektiv:** Uppsatsens teoretiska ramverk baseras på UTAUT-modellen som justeras utifrån ytterligare teorier om big data användning och beslutsfattande.

**Resultat:** Studien identifierade *socialt inflytande*, *underlättande förhållanden* och *förväntad prestation* som drivande faktorer för individuell användningsgrad av big data system inom revisionsbranschen. Samtliga variabler hade ett positivt statistiskt samband med den beroende variabeln *faktisk användning*.

**Slutsats:** *Socialt inflytande*, *underlättande förhållanden* och *förväntad prestation* hade positiva statistiskt signifikanta samband med *faktisk användning* av big data system för revisorer.

## *Förord*

Vi vill framföra ett tack till alla respondenter som tog sig tid att svara på vår enkät. Vi vill även rikta ett särskilt tack till de chefer på de olika medverkande revisionsföretagen för deras hjälp med att sprida och utforma vår enkät. Sist, men inte minst, vill vi rikta ett tack till vår handledare Magnus Johansson för hans handledning och värdefulla återkoppling under arbetets gång.

Januari 2023, Lund.

Christoffer, Jacob & Rasmus

<b>Lista över figurer</b>	<b>8</b>
<b>Lista över tabeller</b>	<b>9</b>
<b>Lista över förkortningar</b>	<b>10</b>
<b>1. Inledning</b>	<b>11</b>
<b>1.1 Bakgrund</b>	<b>11</b>
<b>1.2 Problematisering</b>	<b>13</b>
<b>1.3 Syfte och frågeställning</b>	<b>16</b>
<b>1.4 Avgränsning</b>	<b>16</b>
<b>2. Teori</b>	<b>18</b>
<b>2.1 Teknologiska acceptans-modeller</b>	<b>18</b>
<b>2.1.1 Self-efficiency Teori</b>	<b>18</b>
<b>2.1.2 TAM-teori</b>	<b>19</b>
<b>2.1.3 Organisationsteori</b>	<b>20</b>
<b>2.2 UTAUT-modellen</b>	<b>21</b>
<b>2.2.1 UTAUT-modellens oberoende variabler</b>	<b>22</b>
<b>2.2.1.1 Förväntad prestation</b>	<b>22</b>
<b>2.2.1.2 Förväntad ansträngning</b>	<b>23</b>
<b>2.2.1.3 Socialt inflytande</b>	<b>23</b>
<b>2.2.1.4 Underlättande förhållanden</b>	<b>23</b>
<b>2.2.2 UTAUT-modellens nyckelmoderatorer</b>	<b>24</b>
<b>2.2.2.1 Kön</b>	<b>24</b>
<b>2.2.2.2 Ålder</b>	<b>24</b>
<b>2.2.2.3 Erfarenhet</b>	<b>24</b>
<b>2.2.2.4 Frivillig användning</b>	<b>25</b>
<b>2.2.3 UTAUT-modellens mediator</b>	<b>25</b>
<b>2.2.4 UTAUT-modellens Beroende Variabel</b>	<b>25</b>
<b>2.3 Karaktärsdrag - Rationalist kontra icke-rationalist</b>	<b>26</b>
<b>2.4 Osäkerhet</b>	<b>27</b>
<b>2.5 Arbetets modifierade teoretiska referensram</b>	<b>28</b>
<b>2.5.1 Frivillig användning</b>	<b>29</b>
<b>2.5.2 Kön</b>	<b>29</b>

2.5.3	Ålder	29
2.5.4	Beteendemässig avsikt	30
2.5.5	Karaktärsdrag	30
2.5.6	Osäkerhet	30
2.6	Hypoteser	31
2.6.1	Förväntad prestation	31
2.6.2	Förväntad ansträngning	32
2.6.3	Socialt inflytande	34
2.6.4	Underlättande förhållanden	35
3.	Metod	36
3.1	Forskningsdesign	36
3.1.1	Kvantitativ deduktiv undersökning	36
3.1.2	Forskningsenkät	37
3.2	Datainsamling	38
3.2.1	Forskningspopulation	39
3.2.2	Stickprovsstorlek	39
3.3	Operationalisering	40
3.3.1	Beroende variabel	40
3.3.2	Oberoende variabler	41
3.4	Datakvalitet	42
3.4.1	Inre validitet	42
3.4.2	Yttre validitet	43
3.4.3	Reliabilitet	43
3.4.5	Täckningsfel	44
3.4.6	Bortfallsfel	44
3.5	Statistiska metodval, PLS-SEM	45
3.5.1	Ordinary Least Squares	45
3.5.2	Principal Component Analysis	46
3.5.3	Structural equation modeling	46
3.5.4	Partial Least Squares Structural Equation Modeling	47
3.5.5	PLS-SEM lämplighetsområden	48
3.5.6	Kritik mot PLS-SEM	48
3.5.7	Alternativa modeller	49

<b>3.5.8 Statistiska antaganden och tester</b>	<b>50</b>
<b>3.5.8.1 Intern konsistens</b>	<b>50</b>
<b>3.5.8.2 Konvergent validitet</b>	<b>50</b>
<b>3.5.8.3 Diskrimineringsvaliditet</b>	<b>51</b>
<b>3.5.8.4 Multikollinearitet</b>	<b>51</b>
<b>4. Resultat och analys</b>	<b>52</b>
<b>4.1 Bortfall</b>	<b>52</b>
<b>4.2 Deskriptiv statistik</b>	<b>53</b>
<b>4.3 Datakvalitet</b>	<b>57</b>
<b>4.3.1 Inre validitetstest - faktorladdningar</b>	<b>57</b>
<b>4.3.2 Intern indexkonsistens</b>	<b>58</b>
<b>4.3.3 Convergent validity - Average variance explained (AVE)</b>	<b>58</b>
<b>4.3.4 HTMT</b>	<b>59</b>
<b>4.3.5 Multikollinearitet - VIF-test</b>	<b>60</b>
<b>4.4 Hypotesprövning</b>	<b>61</b>
<b>5. Diskussion och slutsats</b>	<b>65</b>
<b>5.1 Diskussion</b>	<b>65</b>
<b>5.2 Slutsats</b>	<b>71</b>
<b>6. Förslag för framtida forskning</b>	<b>73</b>
<b>Litteraturlista</b>	<b>75</b>
<b>Appendix</b>	<b>80</b>
<b>Appendix A</b>	<b>80</b>
<b>Appendix B</b>	<b>81</b>
<b>Appendix C</b>	<b>82</b>
<b>Appendix D</b>	<b>83</b>

## Lista över figurer

**Figur 1** - Illustration av UTAUT-modellen - *Sida 22*

**Figur 2** - Förenklad illustration av Gressel, Pauleen & Taskin (2021) samband mellan karaktärsdrags påverkan på attityd till prestation och ansträngning av big data användning - *Sida 27*

**Figur 3** - Förenklad illustration av Gressel, Pauleen & Taskin (2021) samband mellan osäkerhet för datadrivna besluts påverkan på attityd till prestation och ansträngning av big data användning - *Sida 28*

**Figur 4** - Illustration av modifierad UTAUT-modell efter studiens hypoteser - *Sida 31*

**Figur 5** - Boxplot över distribution och extremvärden för faktisk användning – *Sida 54*

**Figur 6** - Illustration av korrelation mellan oberoende variabler och faktisk användning – *Sida 55*

**Figur 7** - Modellens resultat med moderatörer inkluderade - *Sida 61*

**Figur 8** - Modellens resultat med moderatörer exkluderade – *Sida 62*

**Figur 9** - Modellens resultat med moderatörer och extremvärden exkluderade – *Sida 63*

**Figur 10** - Stapeldiagram över likertsvarsfördelning för karaktärsdrag - *Sida 69*



# Lista över tabeller

**Tabell 1** - Beskrivande statistik över respondenternas egenskaper – *Sida 53*

**Tabell 2** - Beskrivande statistik över respondenternas uppskattade användning av big data system i andel av total arbetstid – *Sida 54*

**Tabell 3** - T-test på skillnader i medelvärden för faktisk systemanvändning mellan olika kontrollgrupper - *Sida 56*

**Tabell 4** - Indikatorers och latent variablers inre validitets- och reliabilitetstest - *Sida 57*

**Tabell 5** - HTMT-värden som visar ifall de latent variablerna och modererande variablernas korrelerar med varandra - *Sida 59*

**Tabell 6** - VIF-värden för multikollinearitet inom modellen – *Sida 60*

**Tabell 7** - Modellestimering med hypotesprövning – *Sida 64*

**Tabell 8** - Tabell över indikatorers standardavvikelse – *Sida 70, Appendix C - Sida 82*

**Tabell 9** - Fördelning av svar för indikatorer från respondenter, Appendix D - *Sida 83*

# Lista över förkortningar

**UTAUT** - Unified Theory of Acceptance and Use of Technology

**FP** - Förväntad Prestation

**FA** - Förväntad Ansträngning

**UF** - Underlättande Förhållanden

**SI** - Socialt Inflytande

**E** - Erfarenhet

**K** - Karaktärsdrag

**O** - Osäkerhet

**FKA** - Faktisk användning

**H** - Hypotes

**3Vs** - Volume, Velocity och Variety

**TAM** - Technology Acceptance Model

**DOI** - Diffusion of Innovation Theory

**TFF** - Task-Technology Fit

**RBV** - Resource Based View

**VIF** - Variance Inflation Factor

**AVE** - Average Variance Extracted

**PLS-SEM** - Partial Least Squares Structural Equation Modeling

**OLS** - Ordinary Least Squares

**CB-SEM** - Covariance-based Structural Equation Modeling.

**HTMT** - Heterotrait-Monotrait Ratio of Correlations

**EFA** - Exploratory Factor Analysis

**CFA** - Confirmatory Factor Analysis

**PCA** - Principal Component analysis

# 1. Inledning

---

*Detta kapitel ämnar förse läsaren med en djupare förståelse och bakgrund till studiens huvudämne, individuell big data användning inom revisionsbolag. Detta följs upp med en problematisering genom en beskrivning av forskningsändamålet som syftar till att motivera den forskningslucka som studien menar fylla. Slutligen presenteras uppsatsens syfte, avgränsning och frågeställning för att ge en bild om arbetets disposition.*

---

## 1.1 Bakgrund

Idag arbetar många företag med att försöka bli mer datadrivna för att kunna dra nytta av stora mängder data (Bean, 2022). Anledningen är att företag försöker tillskaffa sig konkurrensfördelar i form av välavvägt beslutsfattande, grundade i mer och bättre information (Bean, 2022). Problemet är att flera företag som investerat avsevärt i detta inte lyckas uppnå tillfredsställande grad av användning för att utveckla dessa konkurrensfördelar, på grund av olika, ofta oförutsedda, hinder (Bean, 2022). Det största hindret för att bli en datadriven organisation är inte tekniken, utan att få människorna i organisationen att bli mer datadrivna (Bean, 2022). Enligt Bean (2022) så ändras inte en organisations beteende över en natt. En bransch som har varit med om denna förändring och som nu behöver ställa om sina anställda är revisionsbranschen (Tschakert, Kokina, Kozlowski & Vasarhelyi, 2016). Branschen har börjat öka kraven på sina revisorer att bli mer datadrivna för att själva kunna utföra analysarbeten av big data i nya så kallade *huvudboksanalyssystem*, ett arbete som tidigare enbart utförts av data scientists och dylikt (Tschakert, Kokina, Kozlowski & Vasarhelyi, 2016)

Bean (2022) pekar på svårigheterna som företag upplever i att bli mer datadrivna, som ett resultat av att data idag skapas i rekordfart och av en stor uppsjö av källor. Dessa källor kan vara allt ifrån bilder, videos och meddelanden, till sensorer som ska behandlas och omformateras till värdefull data för att fatta beslut (Bean, 2022). Denna nya uppsjö av data och källor som data kan samlas in ifrån benämns som *big data* och skiljer sig från vanlig data utifrån tre dimensioner, så kallat de *3Vs*. 3Vs syftar till parametrarna *volume*, *velocity* och *variety* (McAfee & Brynjolfsson,

2012; Gressel, Pauleen & Taskin, 2020). Volume syftar till hur volymen av big data skiljer sig från traditionell data, där mängden data som klassificeras som big data betydligt överskrider traditionella datasets (McAfee & Brynjolfsson, 2012). Medan big data har terabyte som datamängd, har vanlig data megabyte eller gigabyte (McAfee & Brynjolfsson, 2012; Gressel, Pauleen & Taskin, 2020). Velocity avser hur snabbt data genereras. Till skillnad från traditionella datasets, som tar längre tid för bearbetning, sker bearbetning av big data näst intill i realtid (McAfee & Brynjolfsson, 2012; Gressel, Pauleen & Taskin, 2020). Big Data kan även separeras från traditionell data utifrån parametern variety; variation av data. Medan vanlig data enbart innefattar strukturerad numerisk data, består big data av vitt skilda format av data, både strukturerad och ostrukturerad (McAfee & Brynjolfsson, 2012; Gressel, Pauleen & Taskin, 2020).

Värdet av att implementera *big data system*, system som hanterar stora mängder data av olika format i realtid, grundar sig i big data systems förmåga att, i kombination med mänskligt omdöme, skapa ny kunskap (Gressel, Pauleen & Taskin, 2021). Denna process att skapa ny kunskap i beslutsfattande, genom tillämpning av big data benämns som *datadrivet beslutsfattande* och kan ha en betydande inverkan på en verksamhets lönsamhet (McAfee & Brynjolfsson, 2012). Detta samband påvisas även specifikt i fallet med revisorer, där Tschakert et al. (2016) pekar på att 61% av 2100 tillfrågade ekonomichefer hävdar att kunskap om hur data samlas in och analyseras är en förmåga som bedömdes positivt att besitta i rollen som revisor. Vikten av att samla in och analysera data stärks av McAfee & Brynjolfsson (2012), som påvisar ett positivt samband mellan lönsamhet och graden företag karakteriserar sig själva som datadrivna. Företag som tillhörde de 30% som främst använde datadrivet beslutsfattande var i genomsnitt fem procent mer produktiva och sex procent mer lönsamma än deras konkurrenter (McAfee & Brynjolfsson, 2012). Denna skillnad var signifikant och kvarstod även efter påverkan från arbetskraft, kapital, inköpta och investering i IT justerades för (McAfee & Brynjolfsson, 2012). Inverkan från big data kunde även observeras i företagets värdering på aktiemarknaden (McAfee & Brynjolfsson, 2012).

Tidigare studier visar att företag som tillämpar datadrivna beslut överlag är mer lönsamma (McAfee & Brynjolfsson, 2012). Därav finns det tydliga incitament för företag att använda sig av

big data system, varvid många företag försöker applicera det i sina verksamheter (Ross, Beath & Quaadgras, 2013). Det är dock få företag som lyckas med en implementering av big data system. Ross, Beath & Quaadgras (2013) redogör att majoriteten av företag inte ser tillräcklig avkastning från sina investeringar i big data system. Tabesh, Mousavidin & Hasan (2019) har dragit liknande slutsatser och hävdar att flera företag, oavsett ansträngningsnivå, misslyckas med att uppnå sina uppsatta mål med implementering av big data i sina processer. Vidare påvisar Baldwin (2015) att 79% av företag misslyckas med att integrera big data fullt ut och 67% av företag saknar väldefinierade kriterier att mäta hur lyckade deras big data projekt är.

## 1.2 Problematisering

---

Det vidsträckta problemet med implementering av big data system har lett till en betydande ökning av forskning inom området. Fokus i dessa studier har primärt legat på att identifiera och riva ner eventuella organisatoriska barriärer. Alharthi, Krotov & Bowman (2017) identifierar fem organisatoriska barriärer som kan orsaka misslyckad implementering av big data system; *beredskap av infrastruktur, komplexitet, bristande kompetens, integritet och kulturella barriärer*. Majoriteten av existerande forskning om big data system visar på att företag i första hand bör fokusera på att riva ner dessa organisatoriska barriärer genom investeringar i fungerande IT-infrastruktur, datahanteringssystem, kompetensförsörjning, personuppgiftshantering och generell organisationskultur (Alharthi, Krotov & Bowman, 2017). Även om tidigare beskrivna organisatoriska barriärer har överkommit, så utvinns inte värdet av implementering av big data system om inte den anställde, individen, tillämpar big data i sitt slutliga arbete i tillräcklig utsträckning.

Tidigare forskning som försökt kartlägga faktorer som påverkar individuell användning av big data system är relativt begränsad. Ännu mer sällsynt är forskning som fokuserat på big data system inom revision, då enbart två studier utförts om vad som påverkar revisionsbolags användning av big data system (Wongsim, Nilniyom, Sompong, Kaiwinit & Satchawatee, 2021; Deniswara et al., 2021). Den forskning som tidigare försökt kartlägga faktorer som påverkar individuell adoption av big data system har nästintill uteslutande utgått från *Unified Theory of Acceptance and Use of Technology*, förkortat UTAUT-modellen, som referensram (Venkatesh et

al., 2003). UTAUT-modellen är ursprungligen en modell som har sitt fundament i beteendevetenskapliga sammanhang för att undersöka faktorer som förklarar teknologisk acceptans (Venkatesh et al., 2003). Efterhand har modellen dock utvecklats till att användas i fler beteendekonomiska sammanhang i syfte att öka användningsgraden vid implementerandet av diverse system. Detta har inneburit att UTAUT-modellen sedermera har kunnat appliceras för att undersöka faktorer som påverkar individuell acceptansgrad av big data och numera blivit den modell som näst intill uteslutande använts för forskning inom ämnet (Deniswara, Kartono, Hamsal, Furinto, & Anthony, 2021; Cabrera-Sánchez & Villarejo-Ramos 2020; Brünink, 2016).

Den forskning som har utförts om individuell användning av big data har dessutom inte uteslutande inriktat sig på faktiskt användning, utan istället fokuserat huvudsakligen på förväntad användning (Deniswara et al., 2021; Brünink, 2016; Kwon, Lee, & Shin, 2014). Utöver detta är den forskning som undersökt faktorer till individuell big data användning tvetydig. Studier som utförts av Deniswara et al. (2021), Cabrera-Sánchez & Villarejo-Ramos (2020) och Brünink (2016) har dragit skilda slutsatser angående vilka faktorer som driver individuell användning av big data system, trots att samtliga studier baserats på samma ursprungliga modell.

Brünink (2016) undersöker individuell adoption av big data och introducerar en ny faktor i modellen; *förväntat företagsvärde*. Studien landar i att faktorerna *förväntad prestation*, *förväntat företagsvärde* och *socialt inflytande* har en positiv korrelation med användningen av big data system (Brünink, 2016). Avhandlingen hävdar även att det finns en modererande effekt från *erfarenhet* av dataanvändning och *datadrivet beslutsfattande* som har en negativ påverkan på individuell beteendemässig avsikt för slutgiltig big data användning.

Tidigare nämnda slutsats från Brünink (2016) går dock tvärtemot de slutsatser dragna av Deniswara et al. (2021), som menar på att varken *förväntad prestation*, *förväntad ansträngning*, *socialt inflytande* eller *underlättande förhållanden* har någon signifikant påverkan på individens beteendemässiga avsikt till användning av big data system. Den enda faktorn som enligt Deniswara et al. (2021) hade en signifikant påverkan på *intentionen att använda* big data var forskarnas egenintroducerade faktor *tillit*.

Även Cabrera-Sánchez & Villarejo-Ramos (2020) introducerar egna variabler till forskningsområdet. Studien undersöker användningen av big data system med introducerandet av tre nya faktorer: *motstånd till användning*, *uppfattad risk* och *alternativkostnad* (Cabrera-Sánchez & Villarejo-Ramos, 2020). Forskarnas slutsats i studien påvisade effekter på användning av big data system från *underlättande förhållanden*, *förväntad prestation* och *socialt inflytande*. Studien visade även att två av de nyintroducerade faktorerna; *alternativkostnad* och *motstånd till användning*, hade en statistiskt signifikant effekt på faktisk användning av big data system (Cabrera-Sánchez & Villarejo-Ramos, 2020).

Gressel (2019) undersökte användning av big data vid beslutsfattande, genom att göra en kvalitativ undersökning med beslutsfattare på olika nivåer. Gressel (2019) visar på ett samband mellan *karaktärsdrag*, samt grad av *osäkerhet* och slutligt användande av big data vid beslutsfattande. Utifrån resultat visat av Gressel (2019) integreras två nya moderatorer i detta arbetets modifierade UTAUT-modell; *karaktärsdrag* och grad av *osäkerhet*. Studien inkorporerar därmed ny kunskap inom området och testar nya potentiella samband i syfte att vidareutveckla förklaringsgraden UTAUT-modellen har för användning av big data system. Att använda karaktärsdrag har dessutom föreslagits av Brünink (2016) angående hur framtida forskning kan tillföra nya insikter. Denna studie identifierar en möjlighet att bidra med djupare förståelse för individuell användning av big data, genom att använda teori som inte fanns när Brünink (2016) genomförde sin undersökning.

Majoriteten av tidigare forskning som försökt kartlägga individuell användning av big data system har valt att använda variabeln *beteendemässig avsikt* som beroende variabel istället för *faktisk användning* (Deniswara et al., 2021; Brünink, 2016; Kwon, Lee, & Shin, 2014). Användandet av *beteendemässig avsikt* som beroende variabel motiveras i dessa studier av att de big data system som undersökts, vid den tidpunkt då studierna genomfördes, befunnit sig i början av sin implementationsfas inom organisationerna (Deniswara et al., 2021; Brünink, 2016). Tidigare forskning som tillämpat UTAUT-modellen i big data sammanhang har därmed inte uteslutande undersökt *faktisk användning* utan framförallt *beteendemässig avsikt*. Forskningsområdet, individuell adoption av big data system, är dessutom relativt nytt och ständigt föränderligt. Sedan tidigare studier genomfördes (Deniswara et al., 2021; Brünink, 2016;

Kwon, Lee, & Shin, 2014) har fler företag implementerat big data system i högre utsträckning. Som konsekvens av detta har denna studie större möjlighet att undersöka *faktisk användning*, och inte endast *beteendemässig avsikt* till att använda big data system. Därmed kan studien bidra med ny kunskap då UTAUT-modellen ställs i relation till faktisk användning för big data. Studien identifierar en forskningslucka baserat på introduktionen av två nya variabler, *karaktärsdrag* och *osäkerhet*, mätning av *faktiskt användning* istället för *beteendemässig avsikt* och avgränsningen till revisionsbranschen.

### 1.3 Syfte och frågeställning

---

Syftet med denna studie är att kartlägga och förklara drivande faktorer som påverkar individuell användning av big data system inom revisionsbranschen. Genom regressionsanalys ämnar studien även att estimerar individuella faktorerers inverkan på slutlig användning av big data system. Syftet mynnar ut i följande frågeställningar:

1. *Vilka faktorer påverkar individuell användning av big data system inom revisionsbranschen?*
2. *Vilka faktorer har störst inverkan på individuell användningsgrad inom revisionsbranschen?*

### 1.4 Avgränsning

---

Arbetet avgränsar sig till att testa den teoretiska referensramens påstådda anledningar till individuell användning av big data system. Det innebär att studien endast syftar till att pröva tidigare formulerade teorier kring anledningar till big data användning och inte till att försöka finna alla potentiella faktorer till revisorerers individuella användning av big data system.

Som ett resultat av studiens syfte; att undersöka faktorer som påverkar individuell användning av big data system inom revision, kommer urvalet till arbetets empiriska material bestå av revisorer som har möjlighet att arbeta med big data system. Studien har gjort en avgränsning i att enbart



studera större revisionsbyråer, då dessa bedöms ha tillräckligt stora kunder för att den data de hanterar ska kunna kategoriseras som big data. För att få en enhetlig spegling i användandet av systemen mellan de olika bolagen, har ytterligare en avgränsning gjorts i att enbart inkludera företagens *huvudboksanalyssystem*. Huvudboksanalyssystem syftar till system som analyserar huvudboken för att identifiera problem med bokföringen och öka kvalitén på arbetet. Då samtliga medverkande revisionsbyråer använder huvudboksanalyseringssystem som baseras på big data är resultatet mellan bolagen jämförbara.

För att undersöka huruvida individen väljer att tillämpa big data system i sitt arbete är det ett givet krav att big data system är tillgängliga för individen. Dessa big data system måste även vara tillgängliga genom företaget i fråga, då det inte är rimligt att förvänta sig att anställda tar eget initiativ att söka upp big data system externt. Slutligen måste individen ha en möjlighet att styra mängden av användningen av big data system i sitt arbete. Individen måste därför själv, i viss mån, kunna välja i vilken utsträckning hen använder big data system i sitt arbete.

Denna studie avgränsar sig mot revisionsbranschen och specifikt revisorers användning av big data system. Revisionsbranschen har genomgått en betydande förändring, då datadrivet beslutsfattande är ett krav och revisorer själva tillämpar big data system i deras arbete (Tschakert et al., 2016). Valet att undersöka revisionsbranschen grundar sig därav på det faktum att big data system redan är väletablerade inom industrin, i form av huvudboksanalyssystem. Då studien ämnar att undersöka faktisk användning av big data system, och inte endast förväntad användning, var ett krav att big data system redan användes inom företagen som studerades.

## 2. Teori

---

*Följande avsnitt beskriver de teoretiska modeller som arbetet utgått ifrån vid formulerandet av hypoteser. Eftersom arbetet ämnar undersöka förklaringsfaktorer till big data användning tar det teoretiska avsnittets första del ansats i att beskriva den forskning som redan är utförd inom ämnet teknologisk acceptans och big data användning för att sedermera kunna formulera big data relaterade hypoteser och en teoretisk referensram.*

### 2.1 Teknologiska acceptans-modeller

---

Den större andelen av tidigare genomförd forskning om faktorer som påverkar användningsgraden av big data system har fokuserat på organisatoriska barriärer. Även om organisatoriska barriärer har överkommit, så lyckas inte implementeringar av big data system om inte de anställda individuellt väljer att tillämpa big data systemen i sitt arbete. För att undersöka detta vidare så är det av vikt att initialt beskriva tidigare forskningsresultat utifrån olika teknologiska acceptansmodeller på individnivå, för att slutligen applicera passade teori på big data specifika system.

#### 2.1.1 Self-efficiency Teori

Bandura (1977) presenterar *self-efficiency* teorin som menar på att en individs inställning till något ges av ifall individen tror att hen klarar av att utföra det. Detta speglas i om en individ är benägen att ta till sig ny teknik eller inte, utifrån hur bra individen tror att den kommer lyckas med det. Utöver inställningen till hur bra en individ tror att de kan utföra något, påverkar även förväntad möda kontra förväntad avkastning huruvida en individ tar till sig något nytt (Payne, 1982). Payne (1982) lyfter här *cost-benefit paradigm* teorin som ska förklara hur en individ väger fördelarna kontra kostnaderna av olika alternativ som ges och om en individ väljer ett alternativ före ett annat, eller helt låter bli att välja något. Davis (1989) utvecklar resonemanget och kopplar det till att den förväntade kostnaden för att ta till sig en ny teknik kontra fördelarna

som förväntas ges av den nya tekniken påverkar ifall en individ väljer att ta till sig ny teknik eller inte. Davis (1989) beskriver att self-efficiency teorin är bra på att kartlägga önskvärda resultat i specifika situationer, men att den är något snäv och saknar förmågor i att kunna appliceras och mätas i mer generella situationer. Av den anledningen menar Davis (1989) att förbättra self-efficiency teorin genom att försöka uppnå en mer utvecklad teoretisk modell. Davis (1989) nya modell är applicerbar på fler generella situationer, inkluderar fler faktorer och ger ett resultat som är lättare att mäta.

### 2.1.2 TAM-teori

Utifrån tidigare nämnd forskning presenterar Davis (1989) *Technology Acceptans Model*, förkortat TAM-modellen, som är den första modellen vars syfte är att mer detaljerat försöka kartlägga användningen av teknologiska system på individnivå. Studien presenterar två variabler som utgör TAM-modellen: *uppfattad lätthet att använda* och *uppfattad användbarhet* (Davis, 1989). Trots att modellen var det första försöket i att förklara användningsgraden av ny teknologi används modellen fortfarande med vissa modifikationer i dagsläget. Exempelvis undersökte Soon, Lee & Boursier (2016) vad som ligger till grund för att individer adopterar nya teknologiska system, genom att slå samman TAM-modellen med den så kallade *DOI-modellen*, *Diffusion of Innovation Theory*, en annan modell inom området. Studien identifierar *uppfattad användbarhet* och *uppskattade fördelar* som de viktigaste faktorerna för om individen tillämpade big data system eller ej (Soon, Lee & Boursier, 2016). Ytterligare forskning som använt sig av TAM-modellen är Shahbaz, Zhai, Shahzad & Hu (2019), som försöker kartlägga användningen av teknologisk användningsgrad genom att kombinera TAM och *Task-Technology Fit*, TTF. Studien landar i att det ger en signifikant bättre förklaringsgrad än modellerna enskilt gör. Deras studie visade vidare på att *uppfattad användbarhet* och *uppfattad lätthet att använda* var avgörande för inställningen till användningen av big data system. Venkatesh, Smith, Morris, Davis & David (2003) utgick sedermera från TAM-modellen och TTF-modellen i syfte att försöka skapa en mer komplett och uppdaterad modell, för att än mer detaljerat förklara anledningar till individuell teknologisk acceptans. Venkatesh et al. (2003) kallar denna modell för *Unified Theory of Acceptance and Use of Technology*, förkortat UTAUT-modellen. UTAUT-modellen överträffar tidigare teoretiska modeller i förmåga att beskriva samband och andel förklarad varians till teknologisk acceptans (Venkatesh et al., 2003) och har av den

anledningen används i flertalet efterföljande studier som teoretisk referensram för att påvisa olika typer av teknologisk acceptans.

### 2.1.3 Organisationsteori

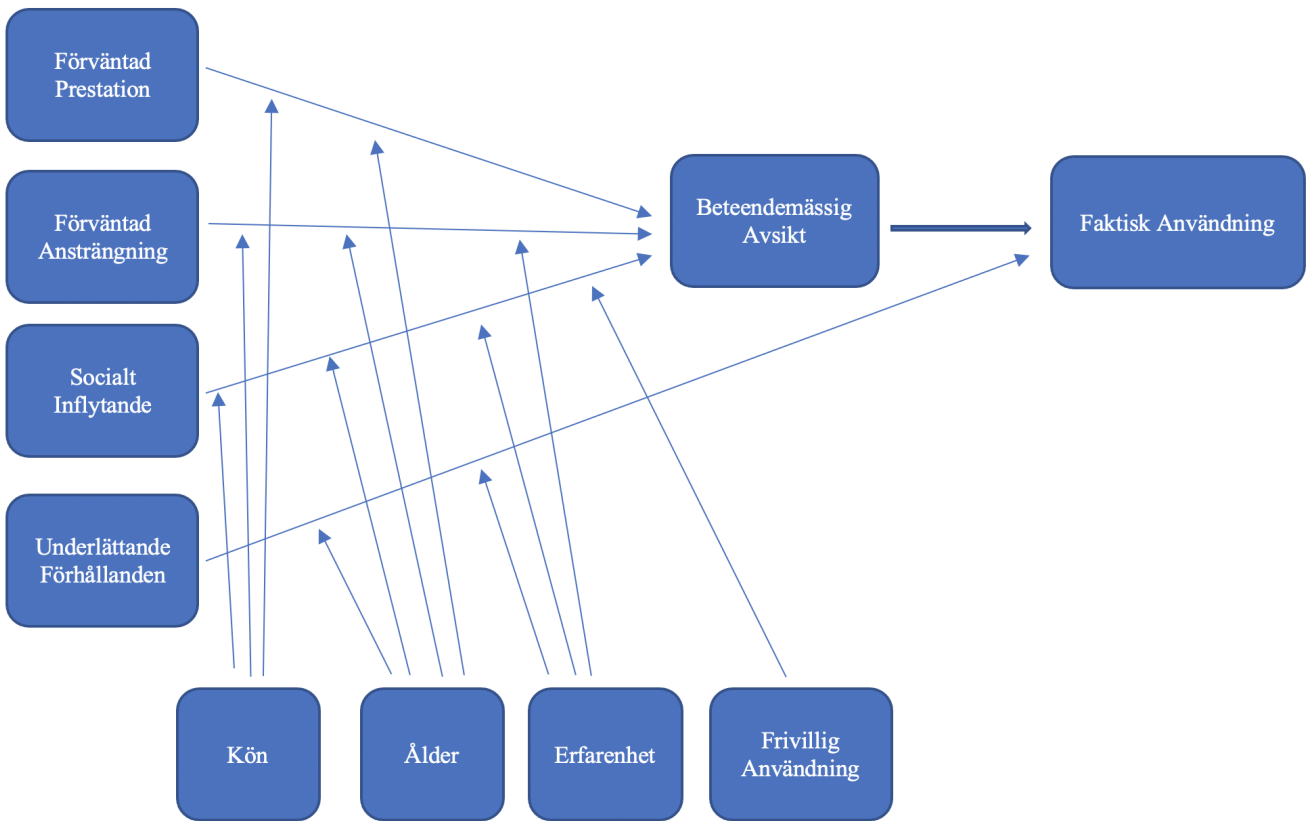
Utöver UTAUT-modellen så har ett fåtal andra studier försökt att genomföra liknande arbeten med egna referensramar som grund för hypotesprövning. Annan forskning som genomförts om individuell adoption av big data har bland annat försökt kartlägga samband mellan *förväntad prestation*, *underlättande förhållanden* och *faktisk intention* att använda big data analyssystem. Ett av dessa arbeten är studien av Kwon, Lee, & Shin (2014), som beskriver hur *förväntad prestation* bestäms av *graden av data-konsistens* tillsammans med *graden av komplett data*. Arbetet utgår från en teoretisk referensram med bas i en kombination av *resource based view* (RBV) och *isomorfism* för att försöka hypotiserar kring användandet av big data. Studien lyckas kartlägga ett samband mellan kvaliteten på data och hur stor den *förväntade prestationen* av att utnyttja big data analys tycks vara (Kwon, Lee, & Shin, 2014). Slutligen visar arbetet på att den *förväntade prestationen* och *underlättande förhållanden* har en direkt påverkan på individens intention att använda big data analys (Kwon, Lee, & Shin, 2014). Givet antagandet att *grad av data-konsistens* och *grad av komplett data* kan motsvara *förväntad ansträngning* av teknologi så innebär detta att Kwon, Lee, & Shin (2014) modellen starkt påminner om UTAUT-modellen. Skillnaden mellan UTAUT-modellen och modellen från Kwon, Lee, & Shin (2014) om isomorfism och RBV är att UTAUT-modellen är mer komplett och tar in fler oberoende variabler och moderatörer i sitt försök att kartlägga teknologisk acceptans. UTAUT-modellen anses därav vara mer fördelaktig i syfte att försöka påvisa fler förklaringsfaktorer till användningsgrad av big data system.

## 2.2 UTAUT-modellen

---

Majoriteten av den forskning i dagsläget som undersökt individuell användning av teknologiska system är baserad på UTAUT-modellen, *Unified Theory of Acceptance and Use of Technology*, (se Figur 1). Användandet av UTAUT-modellen motiveras utifrån att det är en sammanställning av flera tidigare modeller som avser att studera användning av ny teknologi (Venkatesh et al., 2003). Till skillnad från tidigare nämnda modeller, inkluderar UTAUT-modellen även ytterligare oberoende variabler och moderatorer, vilket tillåter fler dimensioner i analysen (Hair, Black, Babin, & Anderson, 2020). Motiveringen av UTAUT-modellen som fundamental modell i denna studie stärks ytterligare av att en av de två tidigare studierna om vad som påverkar individuell användning av big data system hos revisorer utgår från UTAUT-modellen (Deniswara et al., 2021). Utöver det har flera tidigare nämnda studier som undersökt individuell användning av big data system inom andra branscher (Brünink, 2016; Cabrera-Sánchez & Villarejo-Ramos, 2020), också har tillämpat modellen. Valet av UTAUT-modellen möjliggör därav att resultatet i denna studie blir jämförbart med tidigare forskning inom området. Utöver att resultatet ska vara jämförbart med tidigare studier anses modellen lämplig på grund av dess höga förklaringsgrad som är betydligt högre än alternativa modeller (Wu, Tao, & Yang, 2007). UTAUT-modellen bedöms även vara passande för statistisk analys, då tidigare forskning som använt sig av modellen har kodat de oberoende variablerna enligt en universell likertskala, vilket möjliggör för regressionsanalys (Hair et al., 2020). Genom användning av regressionsanalys kan oberoende variablers individuella och gemensamma inverkan på den beroende variabeln uppskattas. I denna studie innebär det att regressionsanalysen kan rangordna faktorer utifrån störst inverkan på individuell användning av big data system. Studiens syfte, att identifiera de drivande faktorerna för individuell användning av big data system, kan därmed uppnås genom tillämpning av regressionsanalys som möjliggörs av UTAUT-modellen.

## UTAUT-Modellen



Figur 1: Illustration av UTAUT-modellen.

### 2.2.1 UTAUT-modellens oberoende variabler

UTAUT-modellen från Venkatesh et al. (2003) identifierar fyra faktorer som direkt eller indirekt påverkar det faktiska användandet av teknologin i fråga (se Figur 1). Dessa oberoende faktorer är: *förväntad prestation*, *förväntad ansträngning*, *socialt inflytande* och *underlättande förhållanden* (Venkatesh et al., 2003).

#### 2.2.1.1 Förväntad prestation

En av de variabler som enligt Venkatesh et al. (2003) ska påverka acceptansnivån av teknologin är den förväntade prestationen av att använda den. Davis (1985) lyfter hur den viktigaste faktorn för att en individ ska adoptera en ny teknologi på arbetsplatsen är hurvida personen har en positiv eller negativ inställning till om teknologin kan vara användbar för individens egna

prestation. En negativ inställning till att teknologin skulle vara användbar för medarbetarna minskar därmed sannolikheten för att tekniken ska adopteras (Davis, 1985).

#### 2.2.1.2 Förväntad ansträngning

En annan variabel som ska påverka huruvida en individ väljer att acceptera en teknologi i sitt arbete eller ej är hur individen betraktar den *förväntade ansträngningen* av att lära sig och använda systemet (Venkatesh et al., 2003). Johnson & Payne (1985) presenterar *kostnad-fördelsmodellen*, som menar på att beslutsfattare gör en avvägning mellan strategin som de applicerar och ansträngningen det medför. Davis (1985) lyfter hur *kostnad-fördelsmodellen* även kan appliceras på teknologi, där en individ gör en avvägning och enbart är villig att ta till sig ny teknik om hen anser att fördelarna i tillräckligt stor grad överstiger ansträngningarna för att lära sig den nya tekniken.

#### 2.2.1.3 Socialt inflytande

*Socialt inflytande* menas enligt Venkatesh et al. (2003) kunna påverka den beteendemässiga avsikten att använda en teknologi. Detta samband påstås ligga i form av beteendemässiga förändringar för att försöka påverka andras uppfattning om dem själva (Venkatesh et al., 2003). Hausman & Stock (2003) beskriver hur *socialt inflytande* kan påverka inställningen till ny teknik. Som exempel lyfter de hur individer som påvisar en positiv attityd till användning av ny teknik influerar andra till en mer positiv inställning (Hausman & Stock, 2003). Gressel, Pauleen & Taskin (2021) drar liknande slutsatser, men kopplar även socialt inflytande till organisationskulturens påverkan på individen. Författarna menar på att en organisationskultur som är datadriven påverkar och uppmuntrar de anställda till att själva bli mer datadrivna genom socialt inflytande (Gressel, Pauleen & Taskin, 2021). Ett annat tillvägagångssätt som lyfts av Kulviwat, Bruner & Al-Shuridah (2009) för att påverka användning av teknologi med hjälp av *socialt inflytande* är att övertyga de tilltänkta att en relevant grupp redan har adopterat teknologin.

#### 2.2.1.4 Underlättande förhållanden

Venkatesh et al. (2003) skriver att *underlättande förhållanden* utgör i vilken grad medarbetare anser att det finns en organisatorisk och teknisk infrastruktur inom organisationen som

underlättar användandet av systemen. Underlättande kring användning av systemen innefattar enligt studien bland annat träning och vägledning inom teknologin (Venkatesh et al. 2003). Venkatesh et al. (2003) menar här på att *underlättande förhållanden* har en direkt påverkan på beteendet gällande i vilken grad teknologin slutligen används av individerna.

## 2.2.2 UTAUT-modellens nyckelmoderatorer

UTAUT-modellen presenterar fyra nyckelmoderatorer som ska influera hur stor påverkan tidigare nämnda variabler har på den indirekta eller direkta acceptansnivån av teknologin (se Figur 1).

### 2.2.2.1 Kön

Morris & Venkatesh (2000) menar på att de förväntade kraven för användning av teknologi är högre för kvinnor än för män. Studien hävdar att detta påverkar i vilken grad könen är beredda att ta till sig ny teknik. En annan faktor som kan påverka att *kön* har en inverkan är, enligt Bozionelos (1996), att kvinnor har dubbelt så hög nivå av teknikångest som män. Av den anledningen har *kön* tidigare motiverats att ingå i UTAUT-modellen som en påverkande moderator för acceptans av teknologi eller ej (Venkatesh et al., 2003).

### 2.2.2.2 Ålder

Att ålder ska ha en påverkan på teknologisk användning lyfts av Morris & Venkatesh (2000). Artikeln presenterar en tes om negativ korrelation mellan högre *ålder* och positiv inställning till användande av ny teknologi. Detta förklaras av att yngre individer har större sannolikhet för att ha utsatts för teknik under sin uppväxt (Morris & Venkatesh, 2000). En bieffekt av detta, som framhävs i forskningen, är att yngre människor i större grad känner sig mer bekväma med ny teknik gentemot deras äldre motsvarigheter (Morris & Venkatesh, 2000).

### 2.2.2.3 Erfarenhet

Morris & Venkatesh (2000) lyfter hur *erfarenhet* kan vara en möjlig faktor till i vilken utsträckning individen adopterar ny teknologi. Här lyfts det att tidigare träning och *erfarenhet* kan påverka huruvida en individ väljer att adoptera ny teknologi eller inte. Erfarenhetsmoderatorn undersöker därför hur stor *erfarenhet* en individ har av liknande



teknologier. Venkatesh et al. (2003) menar på att moderatorm ska ha en förstärkande eller försvagande effekt på hur stor påverkan de oberoende variablerna har på den faktiska användningen.

#### 2.2.2.4 Frivillig användning

Sambandet mellan socialt inflytande och beteendemässig avsikt ska enligt Venkatesh et al. (2003) påverkas av hur frivillig teknologin är att använda. Ramayah (2010) presenterar forskning på hur moderatorm, till vilken grad en individ frivilligt kan använda teknologin, påverkar korrelationen av *förväntad ansträngning*, *beteendemässig avsikt* och *faktiskt användande*. Därmed kan *frivillig användning* positivt påverka individens *faktiska användning* av ny teknologi.

#### 2.2.3 UTAUT-modellens mediator

Tidigare studier som anammat UTAUT-modellen i big data sammanhang har primärt applicerat modellen i implementationsstadiet för att försöka få till en så lyckad implementation och hög faktisk användningsgrad som möjligt (Brünink, 2016; Deniswara et al., 2021). Det innebär att modellen i huvudsak har använts i situationer där det inte funnits möjlighet att undersöka ett direkt samband med den faktiska användningsgraden. För att åtgärda detta dilemma har tidigare forskning som använt sig av UTAUT och inte kunnat studera faktisk användningsgrad, istället undersökt *beteendemässig avsikt* till att använda teknologin efter implementationsfasen. Detta innebär att variabeln mäter huruvida en individ har för avsikt att använda teknologin inom nästkommande månader. Tidigare användning av UTAUT-modellen i big data sammanhang har därav mätt *beteendemässig avsikt att använda*, istället för *faktisk användning* av big data system (Venkatesh et al., 2003; Brünink, 2016; Deniswara et al., 2021).

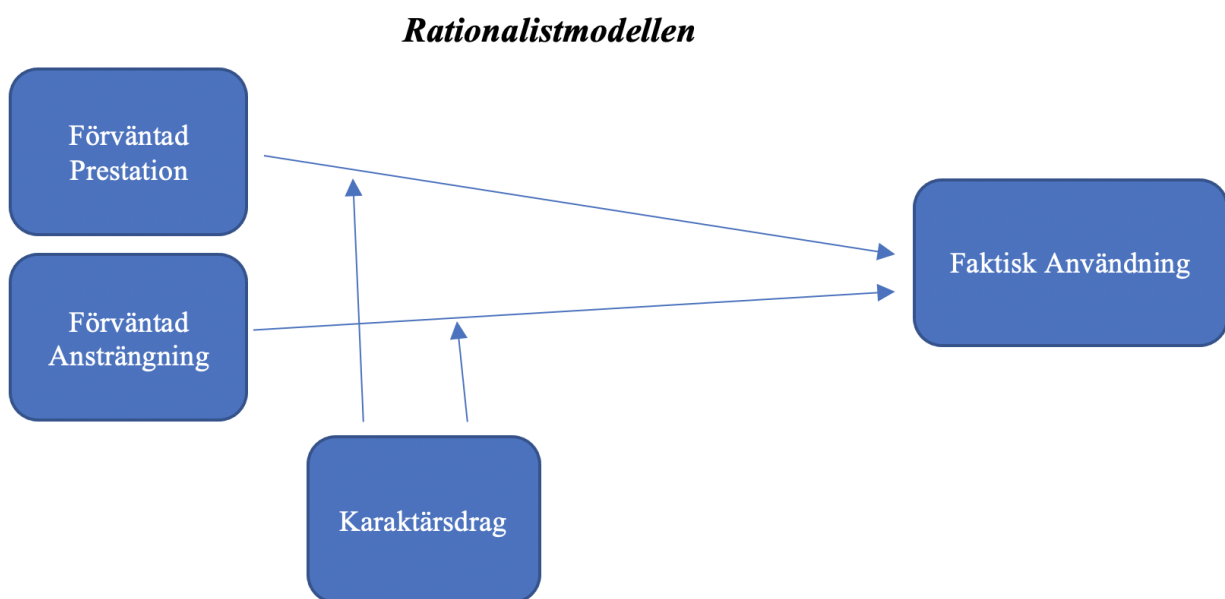
#### 2.2.4 UTAUT-modellens Beroende Variabel

Venkatesh et al. (2003) påvisar i sin studie signifikanta samband mellan samtliga tidigare beskrivna variablers påverkan på slutgiltig teknologisk acceptans dvs. *faktisk användning*. Det innebär att det ska finnas kausala samband mellan samtliga variabler och teknologisk slutgiltig acceptansnivå. Den beroende variabel som UTAUT-modellen undersöker är därav *faktisk användning* av undersökta teknologiska verktyg eller system.

## 2.3 Karaktärsdrag - Rationalist kontra icke-rationalist

---

Annan forskning som genomförts kring användandet av big data system undersöker andra faktorer till användning och framförallt attityd till nyttan av big data eller ej. Gressel, Pauleen & Taskin (2021) presenterar en skala av *karaktärsdrag* angående huruvida en person är *rationalistisk* eller inte. Ett rationalistiskt förhållningssätt karaktäriseras enligt Gressel (2019) av ett analytisk och kalkylerat tillvägagångssätt där individen skiljer fakta från sina egna känslor och förutfattade meningar. Gressel (2019) lyfter hur en rationalistisk person därav ser till siffror och data för att lösa sina uppgifter och undersöka om deras intuition kan stödjas eller dementeras. När en rationalistisk person ställs inför ett problem definierar och diagnostiserar de problemet, tar fram olika alternativa lösningar som de sedan väger mot varandra och tillslut landar de i den bästa lösningen (Gressel, Pauleen & Taskin, 2021). Motsatsen menar författaren är en *icke-rationalistisk* person. En icke-rationell person karaktäriseras enligt Gressel, Pauleen & Taskin (2020) av att de arbetar heuristiskt och med sin egna intuition. Tack vare att en icke-rationalistisk person arbetar heuristiskt är de snabba på att ta beslut, men det leder även till risken att missa ny information som uppkommit eller ignorerar tillgänglig information för att det dementeras deras uppfattade lösning. På grund av att en rationalistisk person har ett analytiskt och kalkylerande tillvägagångssätt så menar Gressel, Pauleen & Taskin (2021) på att de föredrar att använda sig av data för att fatta sina beslut. Samtidigt utgår en icke-rationalistisk person istället från sina egna erfarenheter och kunskap för att fatta beslut snabbt, något som inte är förenligt med databaserade beslut.



Figur 2: Förenklad illustration av Gressel, Pauleen & Taskin (2021) samband mellan karaktärsdrags påverkan på attityd till prestation och ansträngning av big data användning.

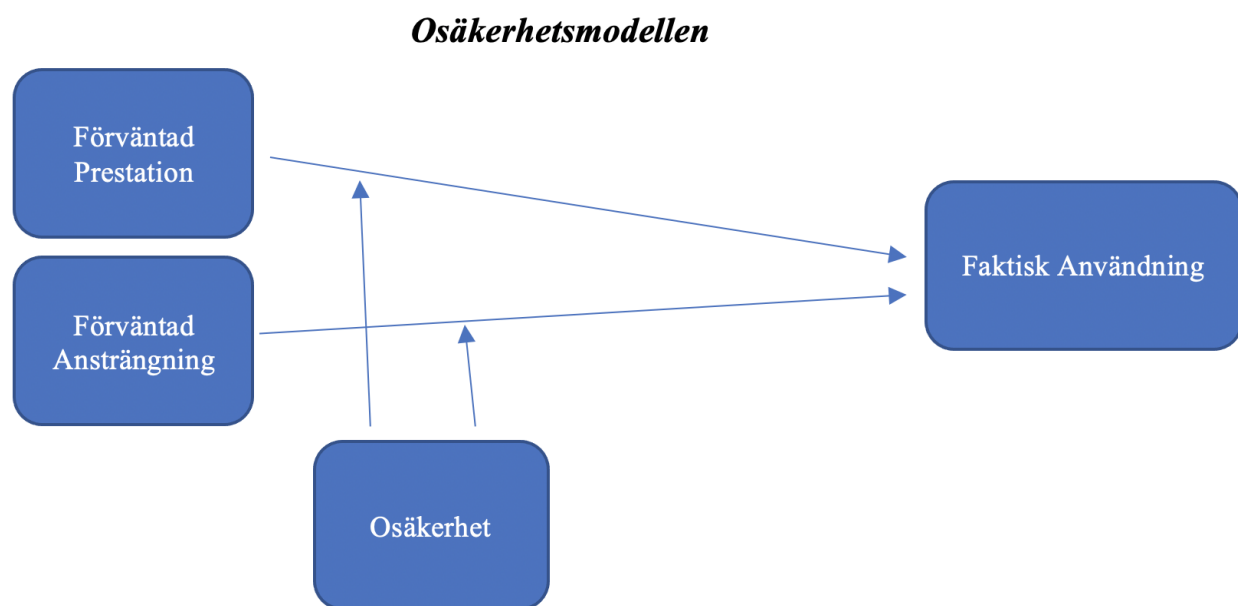
## 2.4 Osäkerhet

---

Gressel, Pauleen & Taskin, (2021) beskriver i sitt avsnitt om personlighetstyper ett samband i att personer som är osäkra på dataanvändning kommer, på grund av sin *osäkerhet*, använda data mindre i sitt beslutsfattande. Gressel, Pauleen & Taskin, (2021) lyfter hur denna osäkerhet i förlängningen leder till att individer tar färre databaserade beslut. Detta motiveras av att det finns en avsaknad av tillräckligt förtroende i datahanteringsförmågor som kan emotsäga individens intuition, något som leder till att individen primärt går på intuition och erfarenhet vid beslutsfattande. Som ett resultat löper individen risken att inte kunna se hela bilden (Gressel, Pauleen & Taskin, 2021). Därav argumenterar denna studie för att en individ som är mer säker inom dataanvändning förväntar en större prestationsökning och en lägre förväntad ansträngning vid användning av big data systemen i motsats till motsvarande, mer osäkra, datahanterare.

Gressel (2019) menar vidare även på att en organisation, genom att identifiera individer som är osäkra i big data användning, kan underlätta övergången till en mer big data-driven organisation.

Detta tack vare att enskilda individers säkerhet uppmuntrar andra att använda det, något som underlättar för andra individer att ta till sig big data och börja använda det (Gressel, 2019). Denna studie menar därav att en minskad osäkerhet i dataanvändning bland individer i organisationen leder till att förväntad ansträngning av att använda big data systemen minskar.



Figur 3: Förenklad illustration av Gressel, Pauleen & Taskin (2021) samband mellan osäkerhet för datadrivna besluts påverkan på attityd till prestation och ansträngning av big data användning.

## 2.5 Arbetets modifierade teoretiska referensram

---

Den teoretiska modell som tillämpas i denna studie är baserad på UTAUT-modellen med vissa modifierationer grundat i Gressel, Pauleen & Taskin (2021) tes om vad som kan påverka individuell användning av big data (se Figur 4). De modifierationer som gjorts baseras på faktorer som enligt Gressel (2019) ska påverka big data användning, men som inte tidigare använts i UTAUT-modellen. Det modifierade ramverket har även tagit bort nyckelmoderatorerna som tidigare big data specifika arbeten inte lyckats påvisa samband för.

### 2.5.1 Frivillig användning

Moderatorn *frivillig användning* tas bort som en nyckelmoderator i studiens modell. Resonemanget för detta grundar sig i studiens avgränsning, vilken ämnar att undersöka individer som har tillgång till, men inte är tvungna att konstant använda systemet. Frivillig användning är därmed ett kriterium som urvalsprocessen utgår ifrån och därmed bidrar nyckelmoderatorn inte med någon information inom modellen. Borttagning av redundanta variabler inom modellen är även önskvärt då det minskar multikollinearitet inom modellen och ökar sannolikheten att individuella variabler är signifikanta (Hair & Sarstedt, 2019).

### 2.5.2 Kön

UTAUT-modellen har sedan den först började användas i vetenskapliga sammanhang inkluderat *kön* som en *modererande variabel* för teknologisk acceptans. Att *kön* har en modererande effekt innebär att de olika könen har olika stark påverkan på sambanden mellan de oberoende och den beroende variabeln. Studier som genomförts i närtid pekar dock på att *kön* inte längre har en modererande effekt (Brünink, 2016; Cabrera-Sánchez & Villarejo-Ramos, 2020), i takt med att jämställdheten ökat i världen (Schwab Brende, Zahidi, Bekhouche, Guinault & Soo, 2013; Hausmann, Tyson & Zahidi, 2006). Därav resonerar denna studie att *kön* som moderator inte är försvarbar att inkludera i studien längre. *Kön* kommer dock fortfarande finnas som kontrollvariabel för att kunna se huruvida det är en stor snedvridning i könsfördelningen bland respondenter.

### 2.5.3 Ålder

På samma sätt som för *kön*, är forskning angående om *ålder* har en modererande effekt på inställningen till användning av ny teknik, över 20 år gammal. Nyare forskning pekar på att *ålder* inte längre har en signifikant modererande effekt på de oberoende faktorerna (Brünink, 2016). Brünink (2016) lyfter hur en möjlig förklaring till att ålder inte längre har en modererande effekt är den digitala revolution som har skett sedan UTAUT-modellen skapades 2003. Exempel på den digitala revolutionen är hur ökningen av antalet individer som har ett mobiltelefonabonnemang har mer än sexdubblats sedan UTAUT-modellen kom till (Statista, 2021). Något som talar för en

utbredd digitalisering och högre användningsgrad av ny teknik i alla åldrar. Därav väljer denna studie att exkludera *ålder* som en moderator i UTAUT-modellen.

#### 2.5.4 Beteendemässig avsikt

Tidigare studier inom användning av big data system har undersökt *beteendemässig avsikt* att använda big data system (Brünink, 2016; Cabrera-Sánchez & Villarejo-Ramos, 2020). Denna studie ämnar att särskilja sig från tidigare studier genom att istället studera *faktisk användning* av big data system. Detta motiveras främst av möjligheten till att nu kunna studera den faktiska användningen på grund av tillgången till företag som numera har implementerat big data system i sina organisationer. Till skillnad från tidigare studier tar arbetet sin ansats i att undersöka företag som inte befinner sig i implementationsfasen, utan redan har tillgängliga big data system som används i olika utsträckning, av olika medarbetare.

#### 2.5.5 Karaktärsdrag

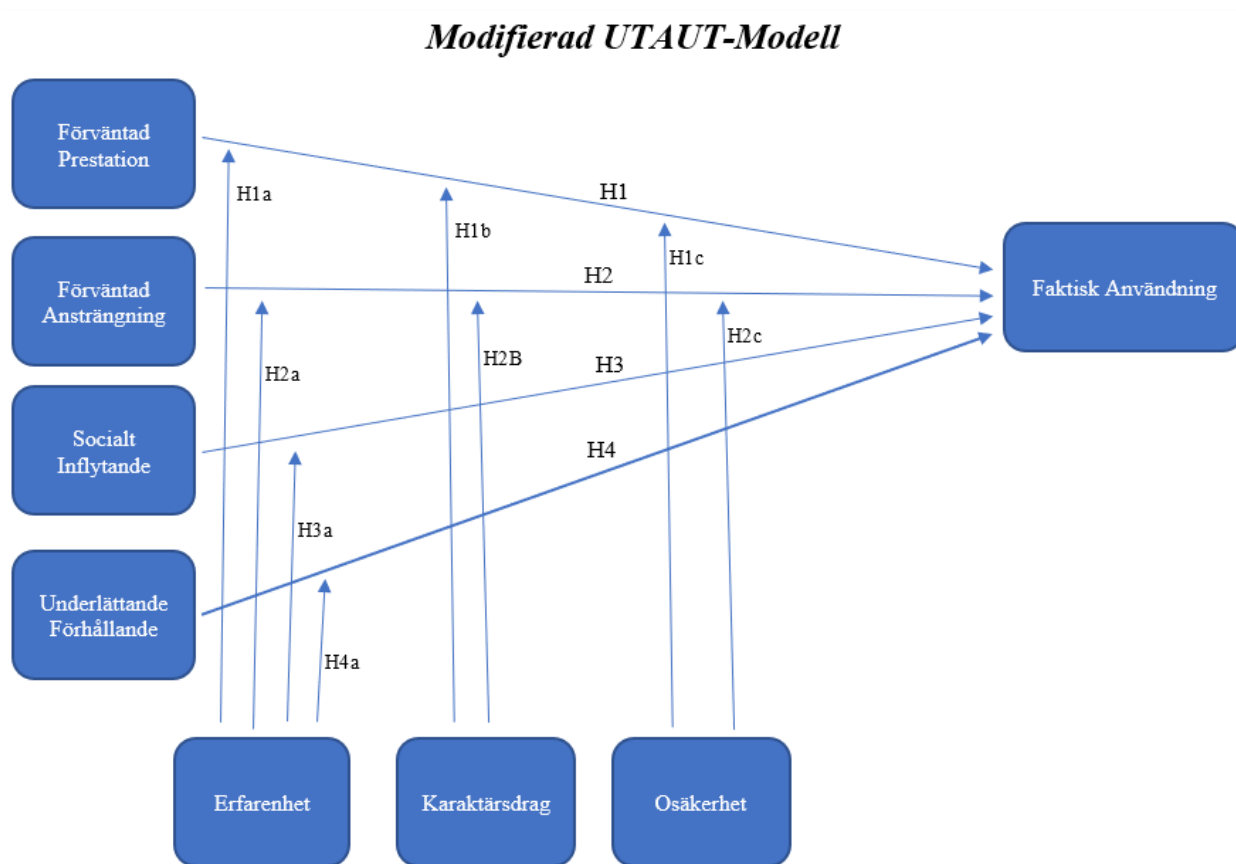
Med hänsyn till det samband som Gressel, Pauleen & Taskin (2021) beskriver kring *rationalist* kontra *icke-rationalist* som *karaktärsdrag* och hur det påverkar intresset av att använda data i sitt beslutsfattande, anser studien att *karaktärsdrag* kan addera kontext som modererande variabel till UTAUT-modellen. Därav väcker denna studie frågan huruvida en mer rationalistisk lagd person, i större utsträckning, tar till sig och börjar använda big data system eller inte. Studien lägger därför till moderatorm, *karaktärsdrag*, kring hur rationalistisk en individ är i UTAUT-modellen.

#### 2.5.6 Osäkerhet

Av den information som Gressel, Pauleen & Taskin (2021) presenterar om personlighetstyper så går det att se ett tydligt samband; att personer som är osäkra angående dataanvändning också kommer använda data i mindre utsträckning för att fatta beslut. Det går att resonera att detta också har en indirekt inverkan på den faktiska användningsgraden av big data system. Av den anledningen är det av intresse att även addera en variabel för *osäkerhet* av dataanvändnings påverkan på sambanden mellan *förväntad prestation* och *förväntad ansträngning* av big data system.

## 2.6 Hypoteser

Utifrån teori som beskrivits ovan har följande hypoteser ställts upp angående relationen mellan oberoende variabler, nyckelmoderatorer och den beroende variabeln.



Figur 4: Illustration av modifierad UTAUT-modell efter studiens hypoteser.

### 2.6.1 Förväntad prestation

Tidigare studier som tillämpat UTAUT-modellen har indikerat att det finns ett positivt samband mellan att använda en teknologi till en större utsträckning och om det finns skäl att tro att det kommer att förbättra respondenternas arbetsprestation (Venkatesh et al., 2003). Genom att följa denna logik är det möjligt att formulera följande hypotes:

**Hypotes 1:** Det finns ett positivt samband mellan *förväntad prestation* och *faktisk användning*.

De medarbetare som har tidigare erfarenhet av att använda big data är enligt Brünink (2016) mer öppna för användning av big data. Det innebär att de medarbetare som har mer tidigare erfarenhet av big data har lättare att förstå fördelarna med big data system och borde därför vara mer övertygade av den förväntade prestationsförbättringen. Detta innebär att medarbetare med tidigare erfarenhet kommer använda big data systemet till en större utsträckning än deras kollegor med mindre tidigare erfarenhet av big data.

**Hypotes 1a:** Sambandet mellan *förväntad prestation* och *faktisk användning* modereras av *erfarenhet*, detta på det vis att det blir en starkare effekt för de medarbetare som har tidigare erfarenhet av big data.

Gressel, Pauleen & Taskin (2021) skriver att det finns indikationer på att personer som är mer rationellt lagda sett till sina karaktärsdrag är mer benägna att ta till sig och arbeta mer med big data i förhållande till sina icke-rationalistiska motparter, som tenderar att inte lita på fördelarna med big data till samma utsträckning. Det är därför möjligt att formulera följande hypotes:

**Hypotes 1b:** Sambandet mellan *förväntad prestation* och *faktisk användning* modereras av *karaktärsdrag*, detta på det vis att det blir en starkare effekt för de medarbetare som är mer rationalistiska i förhållande till icke-rationalistiska.

Av det som Gressel, Pauleen & Taskin (2021) beskriver om osäkerhet kring dataanvändning som beslutsunderlag går det att formulera en hypotes om osäkerhets påverkan på den förväntade prestationen som:

**Hypotes 1c:** Sambandet mellan *förväntad prestation* och *faktisk användning* modereras av *osäkerhet*, detta på det vis att det blir en svagare effekt för de medarbetare som är osäkra på att fatta databaserade beslut.

## 2.6.2 Förväntad ansträngning

De studier som genomfört liknande studier har haft olika resultat angående sambandet mellan *förväntad ansträngning* och *faktisk användning* kan skiljas från slumpen. Brünink (2016)



hypotiserar att *förväntad ansträngning* har en negativ inverkan på *faktisk användning*. Ifall tekniken betraktas som krånglig och komplex så verkar det innebära lägre användningsnivåer av det faktiska systemet (Brünink, 2016). Varken Cabrera-Sánchez & Villarejo-Ramos (2020), Deniswara et al. (2021) eller Brünink (2016) lyckades bekräfta något signifikant samband mellan *förväntad ansträngning* och *intension att använda systemet*. På grund av variabelns centrala betydelse och byggsten i UTAUT-modellen bedöms dock variabeln ändå vara så pass betydande att den inkluderas i studiens hypotesprövning. Följande hypotes upprättas:

**Hypotes 2:** Det finns ett negativt samband mellan *förväntad ansträngning* och *faktisk användning* av big data system.

Tidigare studier föreslår att det finns ett samband mellan anställda som har liten erfarenhet av big data och deras uppskattning av hur komplext systemet är. Detta till den grad att anställda som har mindre erfarenhet av big data tenderar att underskatta komplexiteten av systemet och därmed ha en högre intension att faktiskt använda det. Studien antar därför att:

**Hypotes 2a:** Sambandet mellan *förväntad ansträngning* och *faktisk användning* modereras av *erfarenhet*, detta på det vis att det blir en starkare effekt för de medarbetare som har mindre tidigare erfarenhet av big data.

Ifrån den information som Gressel, Pauleen & Taskin (2021) presenterar om en rationalists syn på big data i förhållande till icke-rationalist, så går det att tyda att det ska finnas ett samband i att icke-rationalister har en tendens att överskatta den ansträngning som krävs för att fatta ett bra beslut. Detta i motsats till rationalister som istället tycks underskatta vad som krävs för att fatta ett perfekt beslut (Gressel, Pauleen & Taskin, 2020). Det innebär enligt detta synsätt att rationalister uppskattar den förväntade ansträngningen som lägre, medan icke-rationalister uppfattar den förväntade ansträngningen som högre. Detta borde även ge effekt i slutlig användningsgrad. Studien hypotiserar därför kring följande:

**Hypotes 2b:** Sambandet mellan *förväntad ansträngning* och *faktisk användning* modereras av *karaktärsdrag*, detta på det vis att det blir en svagare effekt för de medarbetare som är mer rationalistisk i förhållande till icke-rationalistisk.

Gressel, Pauleen & Taskin (2021) beskriver att osäkerhet på dataanvändning kommer att resultera i att den förväntade ansträngningen av big data användning stiger. Det innebär att osäkerhet av data indirekt borde ha en påverkan på den faktiska användningsgraden av big data system vilket gör det möjligt att formulera följande hypotes:

**Hypotes 2c:** Sambandet mellan *förväntad ansträngning* och *faktisk användning* modereras av *osäkerhet*, detta på det vis att det blir en starkare effekt för de som är osäkra på att använda data som beslutsunderlag.

### 2.6.3 Socialt inflytande

Tidigare studier föreslår att *socialt inflytande* påverkar den grad till hur respondenterna faktiskt använder en teknologi eller inte. Detta uppstår när respondenterna till en hög grad upplever att hög systemanvändning är högt ansett av kollegor och chefer (Alwahaishi & Snasel, 2013). Det innebär att om chefer utåt anser att big data systemen är användbara så kommer det ha en positiv inverkan på faktisk användning av big data system (Brünink, 2016). Följande hypotes ställs upp:

**Hypotes 3:** Det finns ett positivt samband mellan *socialt inflytande* och *faktisk användning*.

Brünink (2016) påvisar ett samband mellan erfarenhet av big data och känslighet för sociala påtryckningar till den grad att medarbetare med mer erfarenhet tycks vara mindre känsliga för sociala påtryckningar från kollegor och chefer än de anställda med mindre erfarenhet. Det bör dock sägas att Brünink (2016) påvisade ett negativt samband för variabeln innebärande att, till motsats av vad som hypotiserades kring, tycks vara så att de med mer erfarenhet av big data verkar påverkas mer av socialt inflytande än de med mindre erfarenhet av big data. Det är därför möjligt att anta att:

**Hypotes 3a:** Sambandet mellan *socialt inflytande* och *faktisk användning* modereras av *erfarenhet*, detta på det vis att det antingen blir en svagare eller starkare effekt för de medarbetare som har tidigare erfarenhet av big data.

## 2.6.4 Underlättande förhållanden

Med *underlättande förhållanden* menas saker som kan underlätta för individen att använda systemet, som stöttning och teknisk infrastruktur (Venkatesh et al., 2003). Cabrera-Sánchez & Villarejo-Ramos (2020) skriver att fler *underlättande förhållanden* för ett system leder till en större acceptans och användning av big data. Det innebär att det borde finnas ett samband mellan fler *underlättande förhållanden* och ökad *faktisk användning*.

**Hypotes 4:** Det finns ett positivt samband mellan *underlättande förhållanden* och *faktisk användning*.

Venkatesh et al. (2003) påvisar ett samband mellan *erfarenhet* av teknologi och underlättande förhållandens effekt på faktisk användning. Anställda med mer *erfarenhet* av liknande teknologi tycks ha lättare att själva hitta stöttning inom företaget och är därför inte lika beroende av formulerad stöttning från företaget. Anställda med mindre erfarenhet påvisar istället ett större behov av stöttning och underlättande förhållanden för att börja använda ett big data system (Venkatesh et al., 2003). Det går att anta att detta även är applicerbart i ett big data kontext. Trots att ingen tidigare studie tycks ha undersökt detta specifika område är det möjligt att formulera följande hypotes:

**Hypotes 4a:** Sambandet mellan *underlättande förhållanden* och *faktisk användning* modereras av *erfarenhet*, detta på det vis att det blir en starkare effekt för de medarbetare som har mindre tidigare erfarenhet av big data.

## 3. Metod

---

*Följande kapitel behandlar de metodiska val och källkritiska ageranden som arbetet vidtagit i syfte att få ett resultat med tillräckligt hög reliabilitet och validitet. För att kunna besvara de hypoteser som formulerats i tidigare avsnitt antar arbetet ett kvantitativt deduktivt arbetssätt. De kvantitativa metoder som studien använt sig av bygger på en enkätundersökning som datainsamlingsmetod för den teoretiska modellens olika variabler. Metodkapitlet avslutas med en genomgång av arbetets regressionsmetod, PLS-SEM och hur den tillämpas för att analysera sambanden i den teoretiska referensramen.*

### 3.1 Forskningsdesign

---

Syftet med arbetet var att undersöka huruvida det föreligger samband hos revisorer mellan olika oberoende variabler, som enligt teorin ska påverka individuell användningsgrad av ny teknologi och deras arbete med big data system. Ambitionen vid upprättandet av metoden var att arbetet dessutom skulle gynna forskningsmiljön med mer insikt i hur UTAUT-modellen kan vara applicerbar på big data specifika system inom revisionsbranschen. Detta skulle på samma sätt ge praktiker insikter som kan stimulera ökad användning av big datasystem i form av förslag på riktade insatser. Med detta som utgångspunkt var det därför av vikt att metoden skulle vara noggrant beskriven och kunna leda fram till att empiriska observationer kan verifiera eller förkasta studiens uppställda hypoteser.

#### 3.1.1 Kvantitativ deduktiv undersökning

Studien utgår från ett deduktivt arbetssätt vilket innebär att den teoretiska referensramen ligger till grund för arbetets hypoteser, vilka sedan testas och jämförs med empiriska observationer. Den teoretiska referensramen skapade tolv hypoteser som metoden designades efter. Kvantitativa undersökningar kommer oftast i form av case studier, databaser eller enkätundersökningar. En

enkätundersökning har överlag fler förmågor i replikerbarhet, generalisering och statistisk kraft i motsats till alternativa metoder (Bryman & Bell, 2017). Eftersom den teoretiska referensramen syftar till att teoretisera kring revisorers anledningar att anamma big data system, genomfördes enkätundersökningen på individer inom revisionsbranschen för att förklara deras användning av big data system.

### 3.1.2 Forskningsenkät

En enkätundersökning är en icke-experimentell anlagd frågeundersökning (Bryman & Bell, 2017). Enkätundersökningar används för att få en större förståelse för en grups ageranden eller perspektiv i förhållande till det ämne som undersöks (Bryman & Bell, 2017). Då studien syftade till att identifiera individuella drivkrafter till användning av big data system som skulle vara representativt för en hel yrkesgrupp, gjordes valet att studiens underlag skulle utgöras av en enkätstudie med så många yrkesverksamma revisorer som möjligt. Alternativa metoder hade kunnat ge mer djupgående svar angående orsaker till drivkrafter för individuell användning, men hade istället saknat förmågor i statistisk kraft för att kunna vara representativa för en hel grups ageranden och därmed saknat generaliserbarhet.

Frågorna i de utskickade enkäterna var utformade enligt en likertskala där respondenten instämde med ett påstående på en 1-5 skala. Likertskalan användes för att öka jämförbarheten med tidigare studier som studerat användningsgraden av big data system, då även tidigare studier tillämpat skalan (Brünink, 2016; Cabrera-Sánchez & Villarejo-Ramos 2020). Datanivån vid användning av likertskalan var på en intervallskala, vilket möjliggjorde multipel regression som var nödvändigt för att testa statistiska samband inom UTAUT-modellen. Likertskalor är generellt sätt en 1-5 eller 1-7 skala. I denna studie valdes en 1-5 skala för att förenkla svarsprocessen för respondenter och därmed minska arbetsbördan med förhoppningen att kunna öka svarsfrekvensen. Frågorna i enkäten var även slutna, vilket gjordes för att ytterligare förenkla svarsprocessen för respondenter och samtidigt underlätta arbetet med kodning av informationen. Utöver tidigare nämnda åtgärder skapades även separata enkäter för varje undersökt revisionsföretag. På så sätt

möjliggjordes identifieringen av potentiella systematiska felkällor i en undersökt grupp, exempelvis som att ett system i en grupp inte motsvarade systemet i de andra undersökta grupperna. Genom att göra individuella enkäter för varje enskilt företag var det även möjligt att namnge de system som individerna på företaget använde, istället för att använda ett allmänt begrepp för dem, som huvudboksanalyssystem eller big data system. Genom att enbart ändra namnet på huvudboksanalyssystemet mellan de olika enkäterna, skapades ett mycket litet utrymme för misstolkning av vad som efterfrågades och innebar en större jämförbarhet mellan enkäterna än om enkäterna hade haft större variation i sina frågor.

Utöver frågor för att kartlägga referensramens olika variabler innehöll enkäten frågor som ska förse läsarna med kontext för urvalet. Detta innebär frågor som undersöker ett antal bakgrundsvariabler för att ge information om eventuella snedvridningar i stickprovet och överrepresentationer av specifika grupper. De frågor som ger kontext till stickprovet i enkäterna är ålder, kön, antal år på företaget och yrkesroll.

## 3.2 Datainsamling

---

För att genomföra enkätundersökningen valdes 4 revisionsbolag ut. Efter att företag selekterats, påbörjades arbetet med enkäten där den gick igenom en iterativ process. Den iterativa processen utfördes genom dialog med systemansvariga på företagen, för att kartlägga vilket eller vilka system som användes på företagen och huruvida frågorna var applicerbara på företaget i fråga. Detta ledde till att enkätfrågorna successivt modifierades, för att ta bort utrymme för missuppfattningar och för att säkerställa att enkäten var applicerbar på varje företag. När frågorna ansågs vara färdiga, samlades mailadresser in till individer på tre av företagen genom LinkedIn. Kriteriet vid mailinsamlingen låg på att hitta individer som jobbade inom revision. För studiens fjärde företaget skickades enkäten istället ut till en personalchef som såg till att sprida enkäten till revisorer inom företaget. Givet att majoriteten av de kontaktade fanns på LinkedIn

uppkom en risk att urvalet blev snedvridet. Detta på grund av att LinkedIn är en social media plattform för jobb och karriär, något som är mer använt bland yngre, där enligt Statista (2022) över 60% av användarna är mellan 25-34 år.

### 3.2.1 Forskningspopulation

Studien undersöker nivåer av big data systems användning i revisionsbranschen, med bas i teori hos olika förklaringsvariabler som menas påverka hur hög användningsgraden av olika big data system är. För att undersöka teorins tillämplighet på big data system var därför det huvudsakliga kriteriet för arbetets forskningspopulation tillgång till ett sådant system i arbetet. Målet var att hitta respondenter med tillgång till big data system, men där deras primära arbetsuppgifter inte var att aktivt arbeta med big data system. Eftersom arbetets utgångspunkt var att öka användningsgraden av systemen, så var respondenternas användningsgrad av systemen tvungna att vara utifrån egen preferens. Hur respektive företag och avdelningar uppfyllde ovan beskrivna kriterier för att ingå i studien skulle inte uteslutande avgöras av författarna. För att undvika eventuella felbedömningar och minimera risker med bias i stickprovsurvalet valdes lämpliga respondentavdelningar ut i samråd med representanter från respektive företag.

### 3.2.2 Stickprovsstorlek

Den teoretiska UTAUT-modell som hypoteserna bygger på, teoretiserar kring en modell med ett flertal variabler för att förklara användningen av tekniska system (Venkatesh et al., 2003). För att statistiskt beskriva samband och korrelationer mellan flera oberoende variabler och en beroende variabel var multipel regression en lämplig statistisk metod (Hair et al., 2020). Att använda sig av multipel regression ställde dock en del krav på den data som skulle vara underlag för den statistiska analysen. Hair et al. (2020) beskriver bland annat att stickprovsstorleken är den ensamt mest inflytelserika faktorn för statistisk kraft, som går att kontrollera vid designandet av en undersökning. Hur stort stickprovet bör vara för att uppnå tillräckligt hög förklarad modellvarians råder det olika meningar om. Hair et. al. (2020) framhäver ett minimum kring antal observationer till minst 5 observationer per oberoende variabel, men att den önskvärda nivån beskrivs vara mellan 15 till 20 observationer per oberoende variabel. Det innebar, enligt

detta arbetes hypotesprövning med 4 oberoende variabler, att studien borde ha för avsikt att få mellan 60 och 75 observationer. Med tanke på att ett stort bortfall är att vänta vid enkätundersökningar (Hair et al., 2020) ställdes högre krav på studiens förmåga att nå ut till tillräckligt många respondenter inom urvalsgruppen och vidta möjliga åtgärder för att förebygga ett stort svarsbortfall.

### 3.3 Operationalisering

---

#### 3.3.1 Beroende variabel

Arbetet avsåg att studera big data system, som genom analys kan addera potentiellt värde till en organisation, genom utökad användningsgrad. Systemen som arbetet avsåg undersöka borde därför innehålla inslag av de 3Vs som i denna studie definerar big data (McAfee & Brynjolfsson, 2012). Det innebär att stor vikt lades vid bedömningen i huruvida de tekniska huvudboksanalyssystemen kvalificeras som big data system; vilket betydde att systemet uppfyllde tillräckliga egenskaper inom *variation*, *hastighet* och framförallt *volym*. Det innebar att systemen dels skulle innehålla inslag av stora mängder data där den undre gränsen sattes vid terabyte i förhållande till traditionella databassystem med megabyte och gigabyte (McAfee & Brynjolfsson, 2012). Systemen skulle även uppfylla kravet på att data fanns tillgänglig i realtid, i motsats till traditionella databassystem som har data tillgänglig för analys mer sällan; som på månads eller kvartalsbasis (McAfee & Brynjolfsson, 2012). Ett annat sätt som big data system skiljer sig från traditionella datasystem är variationen av den data som finns att tillgå i systemet. Utöver att innehålla inslag av både strukturerad, semi-strukturerad och ostrukturerad data har big data system förmågor att hantera data från en mängd olika källor simultant (TTG, 2022). Av den anledningen skulle systemen som var föremål för studien innehålla inslag av både strukturerad och ostrukturerad data från flera olika källor för att uppfylla kriteriet för variation (McAfee & Brynjolfsson, 2012). För att summera så påvisar big data system, till skillnad från traditionella data system, förmågor i att hantera utmaningar gällande skala, heterogenitet, komplexitet, brist på struktur, aktualitet och härkomsten på data (McAfee & Brynjolfsson, 2012; Brünink, 2016;



TTG, 2022; Jena, 2022). Varpå den huvudsakliga utgångspunkten vid operationaliseringen av den beroende variabeln var att huvudboksanalyssystemen som arbetet skulle analysera borde uppfylla dessa egenskaper för att kunna klassificeras som ett big data system.

Under kontakt med olika revisionsbolag för deltagande i undersökningen kom det fram att de huvudboksanalyssystem som arbetet undersöker många gånger används av olika medarbetare, för att rapportera in data som möjliggör framtida analys. Eftersom arbetets utgångspunkt ligger i funktionen att faktiskt hantera och analysera data var det av vikt att avgränsa den beroende variabeln till att bara undersöka de som använder systemet för bearbetning och analys av data, det vill säga output funktioner i systemen. De anställda som utnyttjar systemen för andra tänkbara ändamål som inte berörs av studiens syfte filtrerades därför bort, genom att formulera frågan för hur mycket respondenterna använde systemet för behandling och analys av data (se Appendix B). För att kontrollera att huvudboksanalyssystemen uppfyllde ovan beskrivna minimikriterier, fördes dialog med representanter för företagen. Efter att detta hade kontrollerats, genomfördes en pilotenkät med samma representanter för att kontrollera begripligheten i enkätfrågorna i aspekten att de var applicerbara på deras systemspecifika egenskaper.

### 3.3.2 Oberoende variabler

Vid designande av enkätfrågor finns det risker i att mäta fel på grund av otydligt formulerade eller ledande frågor. Därför bör studier i första hand försöka använda sig av tidigare arbetens redan granskade enkätfrågor (Bryman & Bell, 2017). Rapporten ämnade därför försöka använda sig av redan tidigare granskade frågekonstruktioner, vilket ger frågorna en högre tillförlitlighet och validitet i förmågan att mäta det som den teoretiska modellen menar på att mäta. Venkatesh et al. (2003) summerar de frågekonstruktioner som använts i tidigare arbeten med UTAUT-modellen. Av dessa selekterades de frågor som bedömdes applicerbara för arbetets ändamål (se Appendix A). Vidare utvecklades och omformulerades frågorna justerat för negationer och kontext, varpå de mest relevanta frågorna för arbetet selekterades som fundament till de slutgiltiga enkäterna. Frågorna som selekterats, omformulerats och använts för mätning av UTAUT-modellens variabler presenteras i Appendix B.

## 3.4 Datakvalitet

---

Målet med insamlingen av data under arbetet med enkätundersökningen var att få ett så representativt underlag som möjligt. Fokus låg därav på att få en så stor bredd i respondenternas attribut som möjligt sett till ålder, kön, antal år på företaget och huruvida en individ hade en strategisk position eller ej inom företaget. Vid arbete med olika typer av enkäter så finns det alltid risk att få ett partiskt resultat på grund av överrepresentation av enskilda grupper. För att hantera eventuella problem med detta sammanställdes en tabell som syftar till att ge en överblick i eventuella överrepresentationer av enskilda grupper (se Tabell 3). Vidare jämförs även resultaten från respektive grupper genom t-test på medelvärdeskillnader. Skulle det gå att identifiera signifikanta skillnader mellan olika gruppers resultat att använda big data system så kunde dessa dels påvisa tendenser till bias i arbetet, men även möjlighet till skillnader i olika grupper. Detta kan i sin tur fungera som grund för att ge förslag på mer riktade insatser för praktiker och framtida forskning, i syfte att öka användandet av big data system.

### 3.4.1 Inre validitet

Vid upprättande av kvantitativa studier är det av stor vikt att vara noggrann vid operationaliseringen av enkätfrågor för att mäta det som studien faktiskt syftar till att mäta (Hair & Sarstedt, 2019). Det innebär att en hög överensstämmelse mellan teoretiska begrepp och operationella begrepp är önskvärt. Fel i denna process kan innebära att studien mäter andra variabler än vad som beskrivs i arbetets teoretiska referensram. För att minimera riskerna med detta var det av vikt att utgå från redan tidigare formulerade enkätfrågor som använts av tidigare forskare med samma teoretiska referensram. De tidigare formulerade frågorna fanns dock bara att tillgå på engelska, något som skapade en risk för feltolkning i studiens egna enkät som utförs på svenska. Fördelarna med att utgå från redan formulerade enkätfrågorna är att dessa frågor redan genomgått en granskning i huruvida de operationella termerna överensstämmer med de teoretiska termerna eller ej.

### 3.4.2 Yttre validitet

För att hålla en hög kvalitet på undersökningen togs yttre validitet i beaktning. Bristningar i yttre validitet kan orsakas av att respondenter ljuger, överdriver eller minns fel. Yttre validitet är därmed svårare att mäta i förhållande till inre validitet i enkätsammanhang (Bryman & Bell, 2017). För att ta hänsyn till detta ställdes frågorna upp i en numerisk intervallskala med lika stort utrymme mellan varje steg för att minska eventuella risker med respondenter som överdriver eller kan vara subjektiva.

### 3.4.3 Reliabilitet

Vid genomförandet av enkätundersökningen vidtogs åtgärder för att uppnå en god reliabilitet. God reliabilitet uppstår när själva mätningen inte påverkas av den som har utfört mätningen. Arbetets resultat ska vara så signifikant att det kan vara skilt från slumpen. Detta innebär att studien ska gå att replikera och ge samma resultat. Ifall studien har lyckats med detta innebär det att det inte finns inneboende slumpmässiga fel, så kallad *common method bias*. Common method bias är enligt Podsakoff, MacKenzie, Jeong-Yeon & Podsakoff (2003) variansen som kan tillskrivas till mätmetoden, snarare än till konstruktionen som mätningarna representerar. Detta är ett metodfel som i mångt och mycket går att konstruera bort. Det första sättet är att göra tydliga formuleringar och akta sig för komplexa meningsuppbyggnader, något som studien såg till att undvika genom att utveckla frågorna iterativt tillsammans med potentiella respondenter. En annan typ av common method bias är misstaget att mäta alla variabler i samma enkät (Podsakoff et al., 2003). I dessa fall finns ett värde av att särskilja variablerna från olika källor. Studien gjorde därför en avvägning mellan antal förväntade svar kontra i vilken mån det kan antas att en särskiljning av olika variabler minskar common method bias. Studien landade här i att det värderades högre att få in fler svar istället för att skicka ut en eller några enskilda frågor som varje individ skall svara på. Vidare finns även andra dimensioner som påverkar huruvida en enkät är mer eller mindre benägen att ha inslag av common method bias. Det kan exempelvis vara om enkäten är strukturerad genomgående på samma sätt, ifall anonymitet garanteras i enkäten, om det är en tidsskillnad mellan när enkäten skickas ut till olika individer eller ifall det är socialt åtråvärt att svara positivt eller negativt på frågor (Podsakoff et al., 2003). Tidigare nämnda dimensioner som kan ge common method bias har alla varit i åtanke vid konstruktionen

av enkäten. Vidare har tidsskillnad tillgodosetts när enkäten skickats ut genom att skicka ut en påminnelse till alla som inte hade svarat på enkäten en vecka efter första utskicket och därpå en ytterligare påminnelse ännu en vecka senare. Detta resulterade i en skillnad i tid mellan olika respondenter. För att studien skulle vara replikerbar har arbetet lagt stor vikt i att vara transparenta i beskrivandet av operationaliseringsprocessen. Av samma anledning är därför alla enkätfrågor bifogade som appendix till rapporten (se Appendix B).

### 3.4.5 Täckningsfel

Täckningsfel innebär att det föreligger en skillnad i undersökt population och avsedd population (Bryman & Bell, 2017). Eftersom enkäten var frivillig att delta i måste försiktighet tas kring de slutsatser som går att dra utifrån enkätsvaren på grund av att de baserades på en frivillig grupps deltagande. Det är inte nödvändigtvis så att dessa bör anses som representativa för hela gruppen. Detta innebär att respondenter som svarade på enkäten i ett tidigt skede riskerar att göra det på andra grunder än de som besvarade enkäten efter upprepade påminnelser. För att hantera dessa risker genomfördes ett t-test för att jämföra om det fanns signifikanta skillnader i medelvärden hos gruppen som svarat på enkäten i ett tidigt skede och de som svarat efter enstaka eller upprepade påminnelser (se Tabell 3).

### 3.4.6 Bortfallsfel

Bortfallsfel är de element inom urvalsramen som tillhör målpopulationen, det vill säga de individer som studien ämnar att undersöka, som svar ej har kunnat utvinnas från. Det finns två huvudsakliga typer av bortfall; partiellt bortfall vilket innebär att en respondenten har svarat på vissa, men inte på alla frågor och individbortfall som syftar till när individen inte har svarat överhuvudtaget. Vid enkätundersökningar är bortfall överlag ett betydande problem som kan orsaka ett snedvridet resultat (Dahmström, 2011). För att motverka högt individbortfall konstruerades enkäten i syfte att framhäva tydlighet och minimera arbetsbördan för respondenten. För att öka tydligheten inom frågeformuläret var svarsalternativen för frågorna placerade på ett likformigt sätt. Majoriteten av frågorna tillämpar även en likertskala för att skapa en kontinuerlig svarsprocess som skulle minska respondenternas arbetsbörda. Frågorna

inom enkäten var även slutna, i syfte att förenkla svarsprocessen för respondenten (Dahmström, 2011). För att säkerställa att enkäten var förståelig för respondenterna utfördes en provundersökning med ett fåtal respondenter från respektive företag. Slutligen eliminerades risken för partiellt bortfall genom att enbart göra det möjligt för respondenten att skicka in frågeformuläret först när det var 100% genomfört.

Introduktionsbrev skickades ut i samband med frågeformuläret för att informera respondenterna om studiens syfte, forskarnas bakgrund och hur informationen används i studien. I introduktionsbrevet framhövs även att respondenternas resultat kommer vara anonyma, i syfte att öka svarsfrekvensen. Vid uteblivet svar skickades påminnelsebrev ut till respondenterna. Om respondenten inte svarade efter två påminnelsebrev bestod bortfallet.

## 3.5 Statistiska metodval, PLS-SEM

---

Studien kartlägger regressionssamband genom PLS-SEM, *Partial Least Square Structural Equation Model*. PLS-SEM är en statistisk regressionsmodell som baseras på latent variabler för att undersöka statistiska samband.

### 3.5.1 Ordinary Least Squares

Användningen av studiens regressionsmodell, *Partial Least Square Structural Equation Model*, PLS-SEM, som statistisk modell har på senare år blivit allt mer populär inom samhälls- och beteendevetenskap (Hair & Sarstedt, 2019). Modellen utgår från två väletablerade statistiska modeller, *Ordinary Least Square regression*, OLS och *Principal Component Analysis*, PCA (Hair, Risher, Sarstedt & Ringle, 2018).

OLS är den vanligaste formen av multipel och enkel regression. Tillämpning av OLS möjliggör undersökning av statistiska samband mellan oberoende variabler och den beroende variabeln. Dessa statistiska samband kan sedan utvärderas utifrån vikt, varpå oberoende variabler kan

rangordnas beroende på slutlig inverkan på den beroende variabeln. Statistiska sambands signifikansnivå kan även utvärderas för att bekräfta eller avfärda hypoteser (Hair et al., 2020).

### 3.5.2 Principal Component Analysis

Förutom OLS utgår arbetets valda regressionsmetod, PLS-SEM, även från PCA, *Principal Component Analysis*. PCA är en teknik inom *Exploratory Factor Analysis*, EFA. EFA syftar till att undersöka variabelernas underliggande strukturer i den statistiska modellen och till vilken grad det finns pålitlig intern korrelation mellan de olika variabelerna inom modellen (Hair et al., 2020). Detta kan sättas i relation till *Confirmatory Factor Analysis*, CFA, som istället syftar till att undersöka i vilken grad som variabelerna bekräftar forskarnas redan etablerade teoretiska ramverk (Hair et al., 2020). PCA är en EFA modell då den syftar till att undersöka hur det går att sammanställa flera variabler till färre, reducera antalet dimensioner i modellen, samtidigt som modellen behåller så stor förklaringsgrad som möjligt (Hair et al., 2020). Genom att sammanställa flera indikatorer för att beskriva en underliggande variabel kunde, så kallade *latenta variabler* tas fram som representation efter dimensionsreducering för arbetets oberoende variabler (Hair et al. 2020).

### 3.5.3 Structural equation modeling

Samlingen av statistiska regressionsmodeller som arbetets regressionsmetod, PLS-SEM, tillhör, kallas *Structural Equation Modeling* modeller. *Structural Equation Modeling*, SEM, är enligt Hair et al. (2020) en grupp av statistiska modeller som bygger på föregående två byggstenar, OLS och PCA. Hair et al. (2020) understryker vikten av att SEM modeller är ämnat för konfirmatorisk analys och därav syftar till att bekräfta teorier, något som kräver att teorin och alla förhållanden i modellen är tydligt formulerade i förväg. Detta för att undvika att teorin utvecklas för att bekräfta vad modellen visar. Författarna lyfter även hur orsakssamband måste ha etablerats i förväg (Hair et al., 2020). Då denna studies teoretiska ramverk baserades på redan etablerade teorier och orsakssamband inom området undveks därmed denna fallgrop.

### 3.5.4 Partial Least Squares Structural Equation Modeling

En SEM model som på senare år har ökat i användning inom samhäll- och beteendevetenskap är Partial Least Square Structural Equation Model, PLS-SEM (Hair & Sarstedt, 2019). Modellen liknar andra SEM modeller i att den består av två modeller, en yttre, så kallad mätmodell som syftar till att mäta hur de latent variablerna förklaras av *indikatorerna*, de faktiskt uppmätta variablerna, och en inre, så kallad strukturell modell som mäter hur de latent variablerna är associerade med varandra.

PLS-SEM inkorporerar dimensionsreducering genom PCA analys. Dimensionsreduceringen sker i denna studie genom att varje latent variabel, de oberoende variablerna och nyckelmoderatorerna, skapas baserat på vikter från tre individuella frågor, indikatorerna. Dessa tre individuella frågor var framtagna för att täcka alla dimensionerna av vad den latent variabeln syftade till att mäta (se Appendix B). Vikter som de olika indikatorerna fick gavs av hur bra varje indikator förklarade variansen i den latent variabeln (Hair et al., 2018). Utöver detta så beaktar PLS-SEM eventuella mätfel som uppkommer vid skapandet av latent variabler och korrigerar dessa, vilket skapar mer pålitliga statistiska uppskattningar (Hair et al. 2018). Genom att PLS-SEM gör dimensionsreduceringar minskade även risken för mätfel i den latent variabeln. Tillämpningen av PLS-SEM minskade därmed risken för att data som inkorporerades i modellen inte avsåg vad studien ämnade att undersöka, vilket ökade studiens reliabilitet. Alternativa statistiska metoder som regressionsanalys med OLS inkorporerar inte dimensionsreducering, vilket hade ökat risken att fel data använts i modellen.

PLS-SEM inkorporerar även *bootstrapping* inom modellen, vilket möjliggör uppskattning av signifikansnivåer mellan oberoende variabler och den beroende variabeln. Bootstrapping är särskilt användbart vid små stickprov eftersom det delar upp stickprovet i delstickprov för att testa hur väl det passar mot modellen. (Hair et al., 2018). Genom att upprepa denna process flera gånger, tillåts testningen av *diskrimineringsvaliditet*, vilket syftar till att bekräfta eller dementera om det finns en korrelation eller ej mellan de oberoende indikatorerna. Detta görs genom uppskattning av *heterotrait-monotrait ratio of correlations*, HTMT (Hair et al., 2018). HTMT uppskattar variansen varje indikator delar med modellen, kontra hur mycket den delar med andra

indikatorer, där hög samvarians med modellen, men låg samvarians med andra indikatorer är önskvärt (Hair et al., 2018). Hair et al. (2018) hävdar att 0.85 eller 0.9 är ett övre gränsvärde på samvarians som är beroende av hur lika varandra de latent variablerna som undersöks är. Bootstrapping används även senare när relationerna mellan de latent och den beroende variabeln ska undersökas (Hair et al., 2018), vilket möjliggjorde att arbetets hypotesprövningar var möjliga vid tillämpningen av PLS-SEM.

### 3.5.5 PLS-SEM lämplighetsområden

PLS-SEM används som statistisk metod för att undersöka att resultatet är väl anpassat till studien utifrån ett flertal faktorer. Vidare bör PLS-SEM användas om modellen som undersöks är komplex och inkluderar flera interna samband mellan variabler samt att forskarna vill kunna utföra hypotesprövningar (Hair et al., 2018). På grund av PLS-SEMs förmåga att behandla komplexa modeller och hypotestestning av teoretiska ramverk är den statistiska metoden användbar om målet med studien är att undersöka teoretiska förlängningar av etablerade teoretiska ramverk, så kallad explorativ forskning för teoriutveckling (Hair et al., 2018). Studien utgick från UTAUT-modellen, ett etablerat teoretiskt ramverk som är komplext och inkluderar flera interna samband mellan variabler. På grund av detta och det faktum att studien ämnade att utvidga modellen, genom att inkludera nya modererande variabler, var PLS-SEM ett lämpligt verktyg. Studien syfte; explorativ teoriutveckling av UTAUT-modellen i koppling till individuell adoption av big data system, var därav förenligt med PLS-SEM förmågor.

### 3.5.6 Kritik mot PLS-SEM

Kritik mot PLS-SEM tillskriver inte den ökade användningen av PLS-SEM till att det tillåter forskare att uträtta saker som inte är möjligt med andra modeller, utan menar istället att det kommit till på grund av att det saknas en grundläggande förståelse för PLS-SEM, dess antaganden och när den är lämplig att tillämpa (Ringle, Sarstedt & Straub, 2012). Ringle, Sarstedt & Straub (2012) hänvisar här till att de vanligaste argumenten till att tillämpa PLS-SEM som statistisk modell motiveras genom små urvalsstorlekar, icke-normalfördelad data eller för att forskarna ämnar använda *formativt uppmätta latent variabler*.



Ringle, Sarstedt & Straub (2012) lyfter hur ett vanligt argument för att tillämpa PLS-SEM över andra statistiska modeller är på grund av det förhållandevis lilla stickprovsstorleken som krävs, i jämförelse med andra modeller. Författarna menar här på att forskare använder detta som en ursäkt för att ha ett litet stickprov (Ringle, Sarstedt & Straub, 2012). Ringle, Sarstedt & Straub (2012) pekar på att en tredjedel av forskningen hade under 100 observationer och över fem procent uppfyllde inte minimikravet. Minimikravet ges av 10 multiplicerat med maximala antalet länkar som pekar på en inre eller yttre latent variabel för antal observationer, något som brukar benämnas *rule of ten* (Ringle, Sarstedt & Straub, 2012). Minimikravet inom denna studie uppgick enligt denna regel till 40 observationer, som med godo därmed uppfylldes.

### 3.5.7 Alternativa modeller

En alternativ SEM modell som används för att studera statistiska samband är CB-SEM, *Covariance-based structural equation modeling*. CB-SEM liknar PLS-SEM, då modellen tillämpar dimensionsreducering för att skapa latent variabler och sedan kartlägger statistiska samband mellan dessa latent variabler (Hair et al., 2020). Dock skiljer sig modellerna i deras primära tillämpningsområden. CB-SEMs primära användningsområden är i studier som ämnar att bekräfta existerande teori. PLS-SEM kan användas i syfte att bekräfta existerande teori, men kan även tillämpas för att utveckla teorier med explorativ avsikt. Då denna studie till stor grad var explorativ och ämnade att utveckla UTAUT-modellen med inkluderingen av två nya moderatorer lämpade sig PLS-SEM bättre för användning inom detta arbete.

PLS-SEM lämpar sig även bättre enligt Dash & Paul (2021) än alternativet, CB-SEM, då PLS-SEM är en *kompositbaserad modell*. En kompositbaserad modell innebär att modellen tar i beaktning att allt inte går att förklaras fullt ut av modellen i fråga, utan gör istället enbart en uppskattning av det teoretiska konceptet (Hair et al., 2020). Detta kan sättas i relation till CB-SEM som är en *common factor model*, som utgår från att allt kan mätas med indikatorerna i modellen och en felterm (Hair et al., 2017). Forskning har även visat att kompositbaserade modeller har en högre förklaringskraft med ett mindre stickprovsurval och är mer konsistenta i uppskattningen av parametrar i jämförelse med vad common factor modeller är (Hair et al., 2020). Då denna studie var baserad på ett mindre stickprovsurval lämpade sig PLS-SEM bättre för tillämpning.

På grund av studiens tidsrestriktioner, begränsades arbetets resultat av ett litet stickprovsurval. Till skillnad från andra möjliga statistiska metoder, som OLS och CB-SEM, kräver PLS-SEM inte ett lika stort stickprovsurval (Hair et al., 2018). Problem angående distribution, nämligen att residualer ej är normalfördelade, tenderar att uppkomma i fall med mindre stickprovsurval (Hair et al., 2018). Vid användning av PLS-SEM är detta dock inte ett problem, då modellen inte vilar på antagandet om normalfördelning (Hair et al., 2020). Därmed bedömdes modellen vara mer lämplig för användning av regressionsanalys.

### 3.5.8 Statistiska antaganden och tester

#### 3.5.8.1 Intern konsistens

Då studien tillämpar latent variabler som är en sammanställning av indikatorer var det av vikt att reliabiliteten av denna sammanställningen undersöktes, det vill säga att indikatorerna pålitligt mätte variationen av den latent variabeln. *Indikatorernas faktorladdningar*; korrelationen mellan indikator och respektive latent variabel, bör vara över 0,708 då detta visar att indikatorn förklarar mer än hälften av den latent variabeln variation (Hair et al., 2018). Inom PLS-SEM bör även *internal consistency reliability* undersökas, vilket är konsistensen av resultatet inom ett test. För att undersöka den interna konsistensen och reliabilitet tillämpades ett *composite reliability* test. Vid dessa tester är värden mellan 0,6 - 0,7 är acceptabla inom explorativ forskning och värden mellan 0,7 till 0,9 brukar bedömas som tillfredsställande till bra (Hair et al., 2018). Intern konsistens och reliabilitet testades även genom Cronbach alpha test. Cronbach alpha test är mer konservativ i dess uppskattning av intern konsistens jämfört med composite reliability test (Hair et al., 2018). Gränsvärdet för Cronbach alpha test som anses vara acceptabelt är över 0,7 (UCLA, n.d.).

#### 3.5.8.2 Konvergent validitet

Ringle, Sarstedt & Straub (2012) understryker vikten av att det finns en acceptabel nivå av mätningvaliditet innan de strukturella relationerna undersöks, något som författarna menar på ofta missas av de som tillämpar PLS-SEM. Studien syftade till att undersöka huruvida det fanns en acceptabel nivå av mätningvaliditet mellan indikatorernas faktorladdningar eller inte genom att undersöka de aspekter som Hair et al. (2020) presenterar. Först och främst lyfts vikten av att

räkna ut den *konvergenta validiteten* för PLS-SEM modellen. Den konvergenta validiteten beskriver graden varje enskild latent variabel förklarar variansen av sina indikatorer (Hair et al., 2018). Konvergenta validiteten mäts i studien genom AVE, *Average Variance Extracted*. Detta innebär att undersöka huruvida de olika latent variablerna delar varians, där ett AVE över 0,5 innebär att konvergent validitet existerar (Hair et al., 2020). För att säkerställa att denna studie uppfyllde dessa krav undersöktes den konvergenta validiteten genom att testa modellens AVE.

#### 3.5.8.3 Diskrimineringsvaliditet

Relationen mellan studiens olika latent variabler undersöktes även inom studien. Specifikt genomfördes undersökningar för *diskrimineringsvaliditet*, vilket är graden latent variabler inom modellen empiriskt skiljer sig från varandra inom modellen (Hair et al., 2018). För att undersöka huruvida de latent variablerna korrelerade med varandra examinerades modellens *HTMT-ratio* (Hair et al., 2018). Inom modeller där de latent variablerna är konceptuellt nära varandra anses 0,9 vara ett passande gränsvärde, om HTMT-ratio överstiger 0,9 är korrelationen mellan de latent variablerna för hög inom modellen (Hair et al., 2018). För latent variabler som konceptuellt skiljer sig åt anses 0,85 vara ett passande gränsvärde (Hair et al., 2018).

#### 3.5.8.4 Multikollinearitet

För att undersöka multikollinearitet inom modellen utfördes ett VIF-test med samtliga latent variabler. Multikollinearitet är problematiskt då inputs påverkar varandra och därmed minskar sannolikheten att individuella variabler inom modellen är signifikanta (Hair et al., 2020). Utöver minskad sannolikhet att individuella variabler är signifikanta kan multikollinearitet även medföra en betydande inverkan på modellens prediktiva förmåga (Hair et al., 2020). Studien genomförde därmed ett *Variance Inflation Factor* test, förkortat VIF-test, för att undersöka eventuell multikollinearitet. Det generellt föreslagna gränsvärdet för multikollinearitet är ett VIF-värde på 10 (Hair et al., 2020). Om testet överskrider 10 finns det betydande multikollinearitet inom modellen som kan påverka modellen och signifikansnivån på individuella variabler (Hair et al., 2020). Dock kan även VIF-värden under 10 fortfarande innebära betydande multikollinearitet som påverkar resultatet av modellen. Hair et al. (2020) utvecklar detta och menar på att ett VIF-värde över 5 pekar på en hög multikollinearitet, något som även det kan komma att påverka modellen.

## 4. Resultat och analys

---

*Följande avsnitt går igenom och analyserar undersökningens resultat. Inledningsvis presenteras en överblick av bortfall och deskriptiv data för att ge läsaren uppfattning och överblick om svarsfördelning i syfte att kunna belysa eventuella snedvridningar i arbetets stickprovsurval. Därefter går avsnittet igenom kvaliteten på data som enligt PLS-SEM regressionsmetod kan ha påverkan på metodens reliabilitet, inre validitet och förmåga att särskilja samband från slumpen. Avslutningsvis presenteras undersökningens resultat gällande sambandsstyrkor, signifikansnivåer och hypotesprövning för att bekräfta eller förkasta de hypoteser som redogjorts för i tidigare avsnitt.*

### 4.1 Bortfall

---

Enkäten mailades ut till 460 individer på tre företag. Efter första utskicket följde en påminnelse att svara på enkäten efter en vecka och sedan en slutgiltig påminnelse två veckor efter första utskicket. Utöver de 460 individer som enkäten mailades ut till, skickades enkäten även till personalansvarig på ett fjärde bolag som vidarebefodrade enkäten till sina anställda, något som ger ett mörkertal i bortfallet för det fjärde företaget. Utfallet blev 61 svar av de 460 individer som mailades, vilket gav en svarsgrad på 13,3%. Bortfallet uppgick därmed till 86,7% från tre av företagen.

## 4.2 Deskriptiv statistik

---

<b>Respondenternas egenskaper</b>			
<b>Mått</b>	<b>Värden</b>	<b>Frekvens</b>	<b>Andel i procent</b>
Kön	Män	49	57%
	Kvinnor	37	43%
Ålder	21-30	71	82,6%
	31-40	13	15,1%
	41-50	1	1,1%
	51-65	1	1,1%
Antal år på företaget	0	6	7%
	1	22	25,6%
	2	20	23,3%
	3	13	15,1%
	4	12	14%
	5	6	7%
	>5	7	8,1%
Personalansvar	Ja	28	32,6%
	Till viss del	15	17,4%
	Nej	43	50%
Företag	A	15	17,4%
	B	26	30,2%
	C	19	22,1%
	D	26	30,2%
<b>Total</b>		<b>86</b>	<b>100%</b>

Tabell 1: Beskrivande statistik över respondenternas egenskaper.

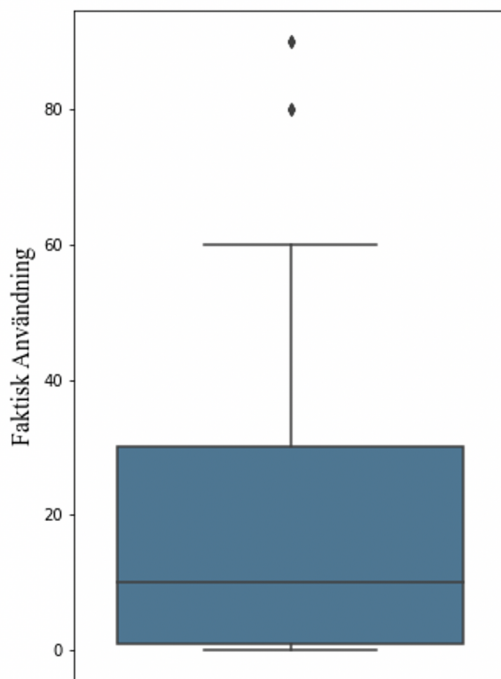
Tabell 1 illustrerar urvalets olika uppdelningar i syfte att kunna åskådliggöra eventuella snedvridningar i urvalets stickprov. Av vad som påvisas i Tabell 1 fanns det tecken på att urvalet kan vara snedvridet i form av en överrepresentation av unga respondenter samt respondenter som inte har arbetat så länge på företagen. I övrigt så verkar det ha varit en någorlunda jämn fördelning vad gäller både kön, position och företags-representation i stickprovet. Fördelning av respondenternas svarsfördelning för varje indikator presenteras i Appendix D. I Appendix C visas information om indikatorernas varians, standardavvikelse och de latent variablernas genomsnittliga standardavvikelse.

## Respondenternas big dataanvändning

Mått	Värden	Frekvens	Andel i procent
Uppskattad andel(%) av arbetstid i big data system	0	19	22,1%
	1-4	9	10,5%
	5-9	9	10,5%
	10-14	13	15,1%
	15-19	4	4,7%
	20-24	7	8,1%
	25-29	2	2,3%
	30-39	4	4,7%
	40-49	7	8,1%
	50-59	7	8,1%
	>60	5	5,8%
<b>Total</b>		<b>86</b>	<b>100%</b>

Tabell 2: Beskrivande statistik över respondenternas uppskattade användning av big data system i andel av total arbetstid.

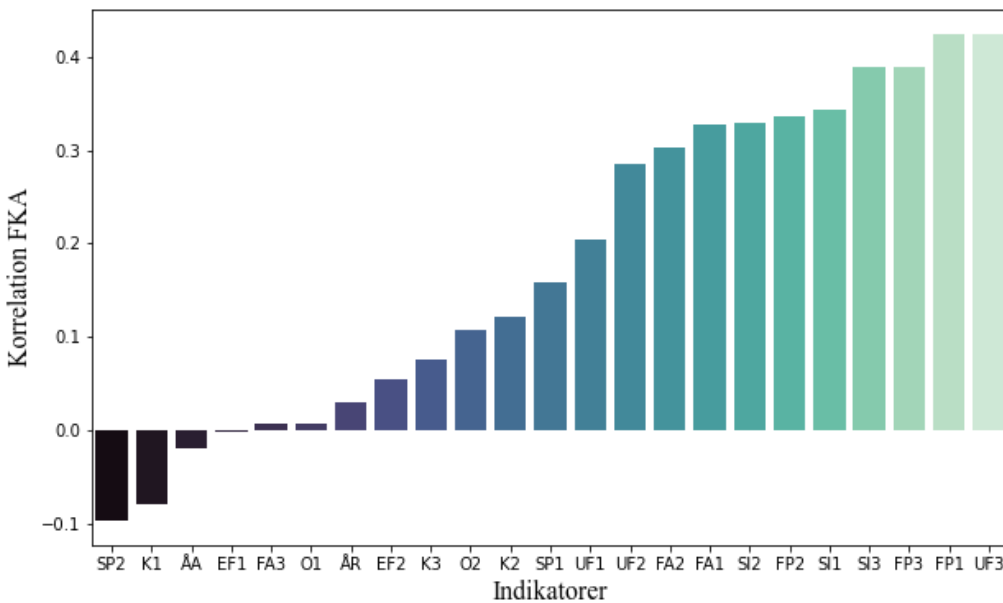
Tabell 2 redogör för frekvensen i hur respondenterna angav sin procentuella användning av undersökta big data system. Den största gruppen använde inte huvudboksanalyssystemen överhuvudtaget en vanlig månad medan resterande påvisar en ganska stor beteendemässig spridning i andel av arbetstid som spenderas i undersökta big data system.



Figur 5. Boxplot över distribution och extremvärden för Faktisk Användning

I Figur 5 kan distributionen av faktisk användning av big data system observeras. Majoriteten av respondenter använde big data system mindre än 20% av sin arbetstid och 75% av samtliga respondenter använde big data system under 30%. I urvalet kunde en positiv skevhet observeras mot högre användning av systemet. Inom urvalet kunde dessutom fyra extremvärden observeras, dessa utgjordes av individer som använde systemet mer än 60% av sin arbetstid.

### Korrelationsdiagram faktisk användning



Figur 6: Illustration av korrelation mellan oberoende variabler och faktisk användning.

I Figur 6 visas korrelationen mellan frågor och den beroende variabeln *faktisk användning*. Frågor kopplat till *förväntad prestation*, *FP*, hade stark positiv korrelation med *faktisk användning*, *FA*, då samtliga tre frågor (FP1, FP2 och FP3) hade över 0,3 i korrelation med den beroende variabeln. *Socialt inflytande*, *SI*, hade likväl som *förväntad prestation* stark korrelation med den beroende variabeln. Fråga nummer tre för *underlättande förhållanden*, *UF3*, var den fråga som hade starkast korrelation med *faktisk användning*, medan resterande två frågor, UF1 och UF2, hade en relativt svag korrelation. Av samtliga oberoende variabler hade *förväntad ansträngning* lägst korrelation med den beroende variabeln. För samtliga moderatörer inom modellen var korrelationen låg och det fanns inget tydligt mönster, då vissa frågor hade en

positiv korrelation, medan andra frågor hade en negativ korrelation. Liknande avsaknad av mönster kunde observeras för studiens kontrollvariabler.

<b>T-test</b>						
<b>Mått</b>	<b>Värden</b>	<b>N</b>	<b>Medelvärde</b>	<b>SD</b>	<b>t</b>	<b>p</b>
Kön	Män	49	21,16	25,67	1,246	0,19
	Kvinnor	37	15,14	16,50		
	Variation		6,028			
Ålder	< 28	49	18,29	18,73	0,135	0,898
	> = 28	37	18,95	26,53		
	Variation		0,66			
Antal år på företaget	0-3	48	19,31	21,80	-0,346	0,732
	>4	38	17,63	23,13		
	Variation		1,68			
Personalansvar	Ja	44	20,36	26,26	0,762	0,444
	Nej	42	16,60	17,28		
	Variation		3,67			
Svarat på utskick 1 eller 2	1	55	20,2	22,85	0,903	0,36
	2	31	15,68	21,28		
	Variation		4,52			

Tabell 3. T-test på skillnader i medelvärden för faktisk systemanvändning mellan olika kontrollgrupper.

Tabell 3 visar medelvärdesskillnader i stickprovets urvalsgrupper som kan åskådliggöra eventuella påverkan på resultatet, som i så fall behöver justeras för. Här testas bland annat om det finns någon skillnad i mellan de som svarat på enkäten vid första utskicket och den grupp som svarat först i ett senare skede efter andra den andra påminnelsen. Skulle det finnas en skillnad i detta innebär det att det föreligger ett så kallat non-response bias, vilket gör att resultatet inte skulle gå att betrakta som tillförlitligt. Av vad som går att tolka från informationen så finns det ingen signifikant skillnad i medelvärden för några av testpopulationernas beteendemässiga användning av big data system vilket innebär att stickprovets data godkänns för frånvaro av snedvridningar och bias.



## 4.3 Datakvalitet

### 4.3.1 Inre validitetstest - faktorladdningar

För att undersöka om indikatorerna förklarade variansen i modellens latenta variabler undersöktes faktorladdningarna, vilket redovisas i tabell 4. Hair et al. (2018) hävdar att faktorladdningar som understiger 0,708 bör tas bort, då dessa indikatorer förklarar mindre än hälften av variationen i den motsvarande latenta variabeln. I tabell 4 påvisas att totalt fyra indikatorer understeg detta gränsvärde; FA3, E1, K1 och K3. Som konsekvens av detta togs samtliga fyra indikatorer bort från modellen.

#### Reliabilitet och inre validitetstest

<b>Latent variabel</b>	<b>Påstående</b>	<b>Faktorladdning</b>	<b>AVE</b>	<b>Cronbachs alpha</b>	<b>Composite reliability</b>
<b>Förväntad prestation</b>			0,870	0,925	0,952
	FP1	0,933			
	FP2	0,945			
	FP3	0,920			
<b>Förväntad ansträngning</b>			0,898	0,887	0,946
	FA1	0,951			
	FA2	0,944			
	FA3	Borttagen 0,210			
<b>Socialt inflytande</b>			0,686	0,771	0,868
	SI1	0,858			
	SI2	0,778			
	SI3	0,847			
<b>Underlättande förhållanden</b>			0,629	0,729	0,835
	UF1	0,745			
	UF2	0,805			
	UF3	0,827			
<b>Erfarenhet</b>			1,00	1,00	1,00
	E1	Borttagen 0,609			
	E2	1,00			
<b>Karaktärsdrag</b>			1,00	1,00	1,00
	K1	Borttagen 0,317			
	K2	0,944			
	K3	Borttagen 0,431			
<b>Osäkerhet</b>			0,821	0,875	0,901
	O1	0,803			
	O2	0,999			

Tabell 4: Indikatorers och latenta variablers inre validitets- och reliabilitetstest.

### 4.3.2 Intern indexkonsistens

Den interna konsistensen och reliabiliteten av latent variabler visas i tabell 4. Ett composite reliability värde på 0,7 eller högre anses tillfredställande och 0,9 anses bra (Hair et al., 2018). Modellens variabler för *osäkerhet*, *förväntad prestation*, *socialt inflytande* och *underlättande förhållanden* låg alla inom detta önskvärda intervall. Två variabler, *karaktärsdrag* och *erfarenhet*, hade ett composite reliability värde på 1. Värden över 0,95 anser Hair et al. (2018) som problematiska, då det indikerar att indikatorerna är överflödiga. I detta fall var dock det höga resultatet en konsekvens av att de båda latent variablerna definierades av endast en indikator. De andra frågorna som utgjorde indikatorer för de två latent variablerna togs bort då faktorladdningarna för indikatorerna inte var signifikanta. Samtidigt är inte detta ett starkt försvar för de två latent variablernas reliabilitet. Då de två latent variablerna endast förklarades av en indikator förlitar sig studien på att de kvarvarande indikatorerna (K2 för *karaktärsdrag* och E2 för *erfarenhet*, se Appendix B) pålitligt mäter vad den latent variabeln ska representera. Som ett resultat av att de två latent variablerna endast förklarades av en indikator minskade reliabiliteten för dessa variabler inom modellen. Reliabiliteten av den latent variabeln *karaktärsdrag* kan därmed till viss del ifrågasättas. Liknande resultat visades återigen för de latent variablernas Cronbach's alpha tester. Samtliga latent variabler förutom *karaktärsdrag* och *erfarenhet* var inom ett acceptabelt intervall; då värdena var mellan 0,7-0,95. *Karaktärsdrag* och *erfarenhet* hade ett Cronbach's alpha värde på 1, som en konsekvens av att dessa variabler endast har en indikator.

### 4.3.3 Convergent validity - Average variance explained (AVE)

Convergent validity, hur varje enskild latent variabel förklarar variansen av sina faktorer, redogörs för i Tabell 4. Tabellen presenterar de latent variablernas AVE, Average Variance Extracted. Värden som överstiger 0,5 betraktas som acceptabla, då den latent variabeln förklarar mer än 50% av indikatorernas varians (Hair et al., 2018). Samtliga latent variablers AVE-värden översteg detta gränsvärde. AVE-värden för variablerna *karaktärsdrag* och *erfarenhet* uppgick till 1, som ett resultat av att de latent variablerna endast utgjordes av en indikator. Som en

konsekvens kan validiteten för dessa variabler ej anses vara hög. Resterande latenta variabler översteg alla gränsvärdet.

#### 4.3.4 HTMT

	FA	FKA	E	O	FP	K	SI	UF	ExFA	ExFP	ExSI	ExUF	OxFP	OxFA	KxFAK
FA															
FKA	0.352														
E	0.060	0.055													
O	0.136	0.065	0.472												
FP	0.924	0.427	0.037	0.051											
K	0.031	0.121	0.244	0.223	0.060										
SI	0.662	0.486	0.095	0.205	0.753	0.064									
UF	0.523	0.443	0.199	0.287	0.598	0.149	0.582								
ExUF	0.154	0.037	0.23	0.188	0.175	0.084	0.097	0.222							
ExFP	0.325	0.033	0.058	0.108	0.305	0.134	0.244	0.209	0.604						
KxFP	0.083	0.153	0.143	0.237	0.183	0.042	0.307	0.158	0.293	0.280					
OxFP	0.078	0.019	0.127	0.232	0.039	0.196	0.025	0.063	0.345	0.490	0.365				
ExFA	0.344	0.058	0.014	0.093	0.291	0.070	0.232	0.170	0.419	0.874	0.206	0.409			
ExSI	0.215	0.028	0.160	0.165	0.209	0.151	0.202	0.095	0.570	0.609	0.128	0.279	0.494		
OxFA	0.160	0.07	0.125	0.279	0.075	0.066	0.076	0.085	0.315	0.446	0.317	0.834	0.303	0.211	
KxFA	0.034	0.091	0.077	0.089	0.077	0.067	0.261	0.109	0.240	0.210	0.873	0.303	0.152	0.078	0.211

Tabell 5: HTMT-värden som visar ifall de latenta variablerna och modererande variablernas korrelerar med varandra.

HTMT-värden, som visar på hur de latenta variablerna korrelerar med varandra, kan ses i Tabell 5. De latenta värdena var konceptuellt inte nära varandra och därmed sattes gränsvärdet på 0,85 efter rekommendationer av Hair et al. (2018). Korrelationen mellan *förväntad prestation*, FP, och *förväntad ansträngning*, FA, uppgick till 0,924 och översteg därmed detta gränsvärde. HTMT-värdet på 0,924 visade i detta fall att de latenta variablerna var närliggande, trots att de teoretiskt sett inte borde vara det. Resterande variablers korrelation med varandra understeg gränsvärdet på 0,85.

#### 4.3.5 Multikollinearitet - VIF-test

<b>VIF-test latent variablerna</b>	
<b>Latent variabel</b>	<b>VIF-värde</b>
FA	4,768
E	1,541
O	1,909
<b>FP</b>	<b>5,213</b>
K	1,347
SI	2,067
UF	1,878

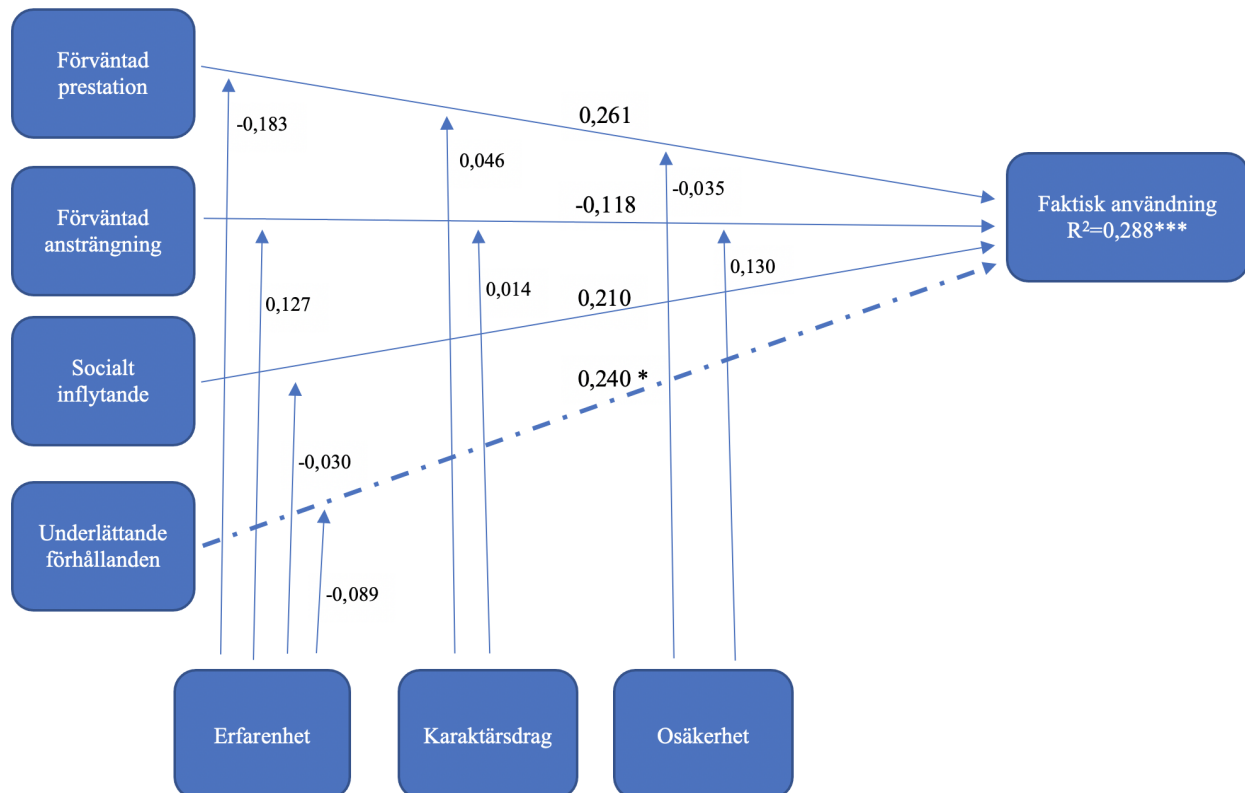
Tabell 6. VIF-värden för multikollinearitet inom modellen

VIF-värden som visar på eventuell multikollinearitet inom modellen presenteras i Tabell 6. Samtliga värden för de latent variablerna understeg VIF-gränsvärdet på 10. Därmed hade ingen av de latent variablerna ett såpass högt VIF-värde att de exkluderades ur modellen. Samtidigt visar Tabell 6 att *förväntad prestation*, FP, hade ett VIF-värde över 5, vilket innebär att den hade en betydande samvariation med andra oberoende variabler. Att ett VIF-värde är över 5 innebär att det måste tas i beaktning i studien då det kan komma att påverka den statistiska signifikansen hos de latent variablerna. Även *förväntad ansträngning*, FA, hade ett högt VIF-värde på 4,768 som nästan översteg 5. Den höga korrelationen mellan *förväntad ansträngning* och *förväntad prestation* som visades i HTMT-värden (se Tabell 5) är troligen förklaringen till att de båda variablerna visade relativt höga VIF-värden, då variablerna hade hög korrelation med varandra. VIF-värdena för *förväntad ansträngning* och *förväntad prestation* visade på betydande multikollinearitet inom modellen, vilket minskade sannolikheten att individuella effekter identifieras som statistiskt signifikanta. Sannolikheten för att dessa två variabler skulle blivit signifikanta minskade därmed.

## 4.4 Hypotesprövning

### Resultat av PLS-SEM sambandsprövning

#### *Modifierad UTAUT-Modell*

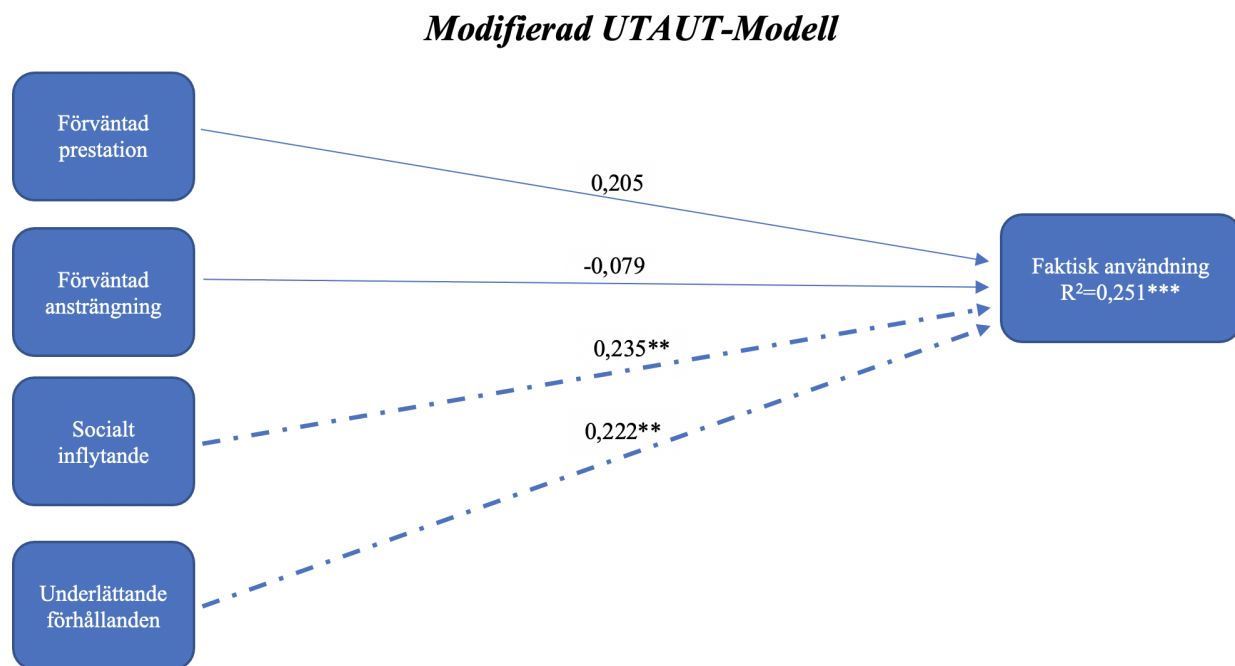


Figur 7: Modellens resultat med moderatorer inkluderade. Streckade linjer indikerar signifikanta riktningskoefficienter (\* $p < 0,1$ , \*\* $p < 0,05$ , \*\*\* $p < 0,01$ ); heldragna linjer indikerar icke signifikanta riktningskoefficienter.

De samband som gick att dra från studiens resultat redovisas i Figur 7. Modellen illustrerar riktningskoefficienternas påstådda styrka, moderatorernas påstådda påverkan,  $R^2$ -värde och huruvida de påstådda sambanden kan signifikant skiljas från slumpen till en 90, 95- kontra 99 procentig nivå. Av Figur 7 att döma så var det enda signifikanta sambandet som gick att påvisa *underlättande förhållanden* påverkan på *faktisk användning*, där sambandsstyrkan uppgick till 0,240 i hur starkt den kan predicera förändringar i den beroende variabeln. Resultaten indikerar därmed att det som mest påverkar huruvida en individ tar till sig big data systemet är *underlättande förhållanden*. Modellen visar också på en statistiskt signifikant förmåga att predicera förändringar i den beroende variabelns varians med ett  $R^2$ -värde på 0,288. Hair et al.

(2020) menar på att ett  $R^2$ -värde på 0,25 är svagt och ett  $R^2$ -värde på 0,5 är måttligt. Detta innebär att modellens förklaringskraft kring huruvida en individ faktiskt använder big data system inom revision kan anses svag. Författarna lyfter dock att vad som är ett starkt eller svagt  $R^2$ -värde inte är universellt bestämt, utan påverkas av vad som undersöks (Hair et al., 2020).

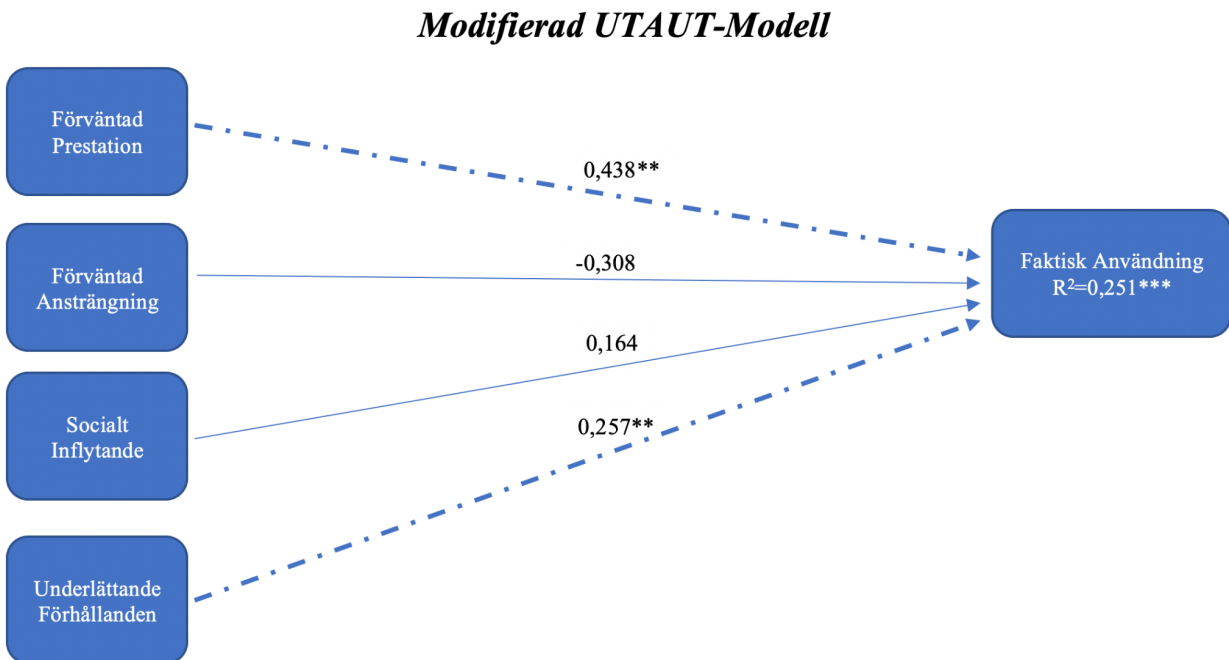
## Oberoende variabelers direkta sambandsstyrka



Figur 8: Modellens resultat med moderatörer exkluderade. Streckade linjer indikerar signifikanta riktningkoefficienter ( $*p < 0,1$ ,  $**p < 0,05$ ,  $***p < 0,01$ ); heldragna linjer indikerar icke signifikanta riktningkoefficienter.

Figur 8 illustrerar direkta samband mellan arbetets hypotetiska modell och beroende variabel utan inverkan av externa moderatörer. Testandet av dessa samband gjorde det möjligt att, förutom *underlättande förhållanden*, även kunna särskilja från slumpen att *socialt inflytande* hade en påverkan på *faktisk användning*. Det innebär att det fanns två positivt signifikanta samband mellan modellens oberoende variabler och modellens beroende variabler vid testningen för direkta samband. Det indikerar att förändring i dessa variabler bör resultera i förändring i den *faktiska användningen* av big data system. Vidare fanns det två andra icke-signifikanta samband mellan *förväntad prestation* och *förväntad ansträngning* som inte kunde konstateras ha någon påverkan på *faktisk användning* eftersom de inte gick att signifikant skilja från slumpen.

## Oberoende variabelers direkta sambandsstyrka, utan extremvärden



Figur 9: Modellens resultat med moderatörer och extremvärden exkluderade. Streckade linjer indikerar signifikanta riktningskoefficienter (\* $p < 0,1$ , \*\* $p < 0,05$ , \*\*\* $p < 0,01$ ); heldragna linjer indikerar icke signifikanta riktningskoefficienter.

Figur 9 illustrerar direkta samband mellan arbetets hypotetiska modell och beroende variabel utan inverkan av externa moderatörer. I Figur 9 har dessutom justeringar gjorts för eventuella extremvärden, för att studien inte skulle få ett potentiellt snedvridet resultat. Extremvärden i detta fall syftar till respondenter som använde systemet över 60% av sin arbetstid. Dessa extremvärden visas i Figur 5, som är en boxplot av respondenternas *faktisk användning*. Stickprovets extremvärden utgjordes av totalt fyra observationer, där två respondenter använde systemet 90% av tiden och de andra två respondenterna använde systemet 80% av tiden. Dessa extremvärden skilde sig drastiskt från medelvärdet som uppgick till 18,56% och medianen som uppgick till 10%. Testandet av dessa samband utan extremvärden gjorde det möjligt att, förutom *underlättande förhållanden*, dessutom kunna särskilja från slumpen att *förväntad prestation* hade en positiv påverkan på *faktisk användning*. *Förväntad prestation* var därmed signifikant inom modellen när extremvärden togs bort. Det innebär att, när justeringar för extremvärden gjordes, fanns det två positivt signifikanta samband mellan modellens oberoende variabler och modellens beroende variabel.

## Modellestimering

Hypotes	Riktningskoefficient	P-värde	Accepterad hypotes
H1: Förväntad prestation → Faktisk användning	0,205	0,308	Nej
H1a: Erfarenhet → Förväntad prestation	-0,183	0,583	Nej
H1b: Karaktärsdrag → Förväntad prestation	0,046	0,872	Nej
H1c: Osäkerhet → Förväntad prestation	-0,035	0,879	Nej
H2: Förväntad ansträngning → Faktisk användning	-0,079	0,659	Nej
H2a: Erfarenhet → Förväntad ansträngning	0,127	0,666	Nej
H2b: Karaktärsdrag → Förväntad ansträngning	0,014	0,959	Nej
H2c: Osäkerhet → Förväntad ansträngning	0,130	0,618	Nej
H3: Socialt inflytande → Faktisk användning	0,235	0,044	Ja
H3a: Erfarenhet → Socialt inflytande	-0,030	0,838	Nej
H4: Underlättande förhållanden → Faktisk användning	0,222	0,013	Ja
H4a: Erfarenhet → Underlättande förhållanden	-0,089	0,525	Nej

Tabell 7: Modellestimering med hypotesprövning.

Tabell 7 illustrerar sambandsstyrka och signifikansvärden för arbetets påstådda hypoteser. På grund av oförmåga i att kunna skilja sambanden från slumpen tvingades arbetet förkasta tio av de tolv på förhand uppställda hypoteserna. De hypoteser som accepteras är de direkta positiva sambanden för *socialt inflytande* och *underlättande förhållanden* som tidigare beskrivits i Figur 7 och Figur 8. Eftersom *förväntad prestation* endast fick signifikans vid exkluderandet av extremvärden kunde inte hypotesen accepteras, dock finns det tydliga indikationer att sambandet existerar givet att det är signifikant och har starkast samband av samtliga oberoende variabler vid borttagandet av extremvärden.



## 5. Diskussion och slutsats

---

*Följande kapitel går igenom och diskuterar de resultat som presenterats i föregående avsnitt. Avsnittet inleder med en diskussion om resultatet, hur det skiljer sig mot tidigare studiers resultat, vad det kan tänkas bero på och generella råd på åtgärder som studien menar att revisionsföretag kan vidta för att öka den organisatoriska användningsgraden av big data system. Slutligen redogörs det för de slutsatser som studien kan dra av resultatet och hur det bidrar till forskningsområdet. Slutsatserna tas upp både specifikt för revisionsbolag som vill få upp användningsgraden av big data system samt generella slutsatser som kan sättas i ett större perspektiv.*

### 5.1 Diskussion

Syftet med arbetet var att kartlägga drivande faktorer som påverkar individuell användning av big data system hos revisorer. Närmare bestämt menade arbetet att estimerar vilka variabler som kan påverka revisorers faktiska användningsgrad av big data system och till vilken utsträckning det gick att särskilja inbördes styrka i dessa variablers påverkan. Det är nu möjligt att konstatera att *underlättande förhållanden* och *socialt inflytande* tycks vara de faktorer som går att skilja från slumpen i skattningen av deras påverkan på undersökta revisorers anledning till användning av big data system. Inbördes verkar deras sambandsstyrka vara ungefär lika starka som förmåga att predicera faktisk användning av undersökta system. Intressant nog blir *förväntad prestation* signifikant först när justeringar gjorts för extremvärden. Sambandet blev dessutom starkare än vad båda de tidigare beskrivna sambanden varit i dess förmåga att predicera användning av big data systemen, med en sambandsstyrka på 0,438 i förhållande till tidigare beskrivna på cirka 0,225. Givet att sambandet tycks kunna särskiljas från slumpen först efter att extremvärden tagits bort talar för att fler studier behöver fokusera mer på att göra liknande undersökningar. Dessa skulle även kunna undersöka om det kan föreligga olika anledningar till användning av systemen hos olika grupper inom företagen. Möjligtvis kan de som använder systemen måttligt göra det på

grund av en förmodad prestationsökning medan de som använder systemen mycket gör det av anledningar mer förknippat till sociala och organisatoriska influenser.

Studiens resultat angående vilka oberoende variabler som är signifikanta skiljer sig från tidigare studier inom graden av individuell användning av big data system. Brünink, (2016) och Cabrera-Sánchez & Villarejo-Ramos (2020) får likt denna studie signifikanta resultat för *socialt inflytande* och *underlättande förhållanden*, men inte signifikans för *förväntad ansträngning*. Tidigare nämnda studier skiljer sig dock mot denna i att de även får *förväntad prestation* som signifikant (Brünink, 2016; Cabrera-Sánchez & Villarejo-Ramos, 2020). Dock får även denna studie signifikanta resultat för *förväntad prestation* när extremvärden tas bort. Denna studie undersöker även, likt Cabrera-Sánchez & Villarejo-Ramos (2020) *faktisk användning*. Detta skiljer studien från Brüninks (2016) studie som enbart undersöker intentionen att använda big data system, vilket även undersökningen av Deniswara et al. (2021) gör. Skillnaden mellan tidigare nämnda studier och Deniswara et al. (2021), som likt denna studie undersöker användningsnivåer inom revisionsbranschen, är dock att studien i fråga inte påvisar signifikans för någon av de originella variablerna.

Studien skiljer sig även branschmässigt mot tidigare studier. Cabrera-Sánchez & Villarejo-Ramos (2020) har ett tydligt fokus i att undersöka individer i ledningsposition på 199 serviceföretag och Brünink (2016) gör en casestudie där forskaren undersöker ett enskilt större företag inom en icke-specificerad industri. Deniswara et al. (2021) undersöker likt denna studie revisionsbolag, men landar i skilda slutsatser kring vilka variabler som är signifikanta. Skillnader i resultatet i denna studie kontra Deniswara et al. (2021) kan delvis bero på att denna studie undersökte revisionsbolag i Sverige, medan Deniswara et al. (2021) undersökte revisionsbolag i Indonesien. En möjlig förklaring är kulturella skillnader, som kan tänkas påverka i vilken grad socialt inflytande har en inverkan. Det kan även tänkas ha en effekt hur långt olika länder har kommit i sin utveckling av teknologisk infrastruktur.

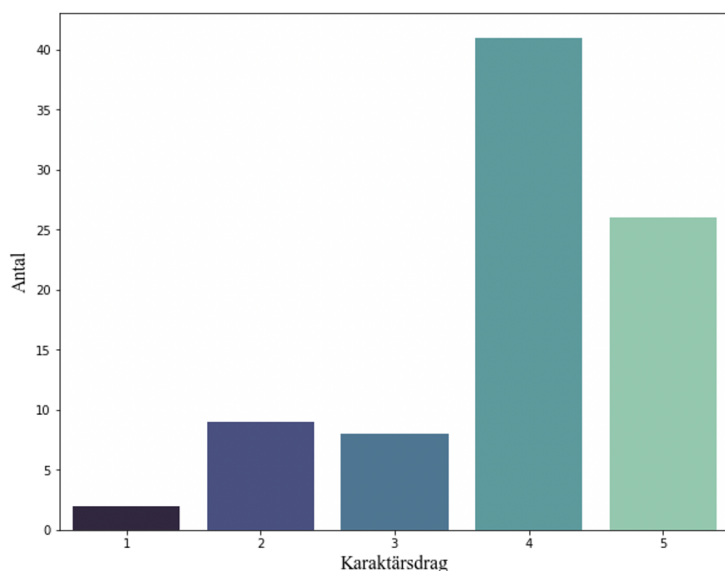
Arbetets skillnader i jämförelse mot Deniswara et al. (2021) studie kring påvisandet av vilka faktorer som är signifikanta kan ses som förvånande, givet att revision klassificeras som en homogen professionell yrkesgrupp. Professionsyrken, som revisorer, jurister och läkare, är ofta,

enligt Nanda & Narayandas (2021), lika sett till utbildning och professionsutövande och därav kan de antas vara en relativt homogen grupp. Speciellt förvånande blir det utifrån att denna studie får samma signifikanta faktorer som Brünink (2016) och Cabrera-Sánchez & Villarejo-Ramos (2020), som undersöker andra branscher som inte innefattar professionella yrkesgrupper. Då både Brünink (2016) och Cabrera-Sánchez & Villarejo-Ramos (2020) undersöker big data användning i olika europeiska länder kan det finnas en möjlighet att denna studies resultat går att applicera inom andra områden i Europa. Något som är mer troligt är dock att studien kan tänkas vara applicerbar på andra, homogena professionsyrken för närliggande branscher i Sverige. Exempelvis är revisorer en del av vad som brukar kategoriseras som professionella yrkesgrupper, varpå det finns en möjlighet att studiens resultat och slutsatser går att applicera inom diverse konsult- och juridikbyråer inom Sverige. Det är dock av vikt att akksamhet tas kring i vilken utsträckning studiens resultat går att applicera. Detta beror på att studiens  $R^2$ -värde var 0,288, något som ger modellen en betydligt lägre förklaringsgrad än vad Deniswara et al. (2021) modell med ett  $R^2$ -värde på 0,849 och Cabrera-Sánchez & Villarejo-Ramos (2020) med ett  $R^2$ -värde 0,483 uppnådde. Studiens låga  $R^2$ -värde i förhållande till tidigare studier kan tänkas ges av studiens stickprovsstorlek och heterogenitet.

Eftersom PLS-SEM modellen inte kunde visa att *förväntad prestation* och *förväntad ansträngning* var signifikant vid testning med extremvärden inkluderade, kunde ett statistiskt säkerställt samband ej påvisas. Detta resultat kan anses överraskande i bakgrund till tidigare forskningsresultat. Variablerna var även tydligt positivt korrelerade med den beroende variabeln *faktisk användning*, specifikt variabeln *förväntad prestation*, som hade starkast korrelation med *faktisk användning* (se Figur 6). Trots resultat från tidigare forskning och den positiva korrelationen förblir variablerna ej signifikanta inom modellen. En möjlig anledning till varför de två variablerna inte uppnår signifikans är på grund av multikollinearitet inom modellen. Utifrån informationen i VIF-testet har båda variablerna höga värden vilket påvisar multikollinearitet som minskar sannolikheten att effekterna från variablerna är signifikanta. Multikollineariteten orsakas troligen av att *förväntad prestation* och *förväntad ansträngning* är starkt korrelerade med varandra, där variablerna hade ett HTMT-värde på 0,924 (se Tabell 5).

Effekter av multikollinearitet behöver dock inte endast isoleras till dessa två variabler. Det är även möjligt att multikollinearitet kan förklara varför moderatorerna ej är signifikanta inom modellen, specifikt de två nya variablerna *karaktärsdrag* och *osäkerhet* som introducerades i studiens modifierade UTAUT-modell. Då *karaktärsdrag* och *osäkerhet* har en påstådd modererande effekt på sambanden mellan variablerna, *förväntad ansträngning och förväntad prestation*, och den beroende variabeln, *faktiskt användning*, påverkas moderatorerna även av multikollinearitet inom modellen. Därmed representerar höga VIF-värden för *förväntad ansträngning* och *förväntad prestation* även problematik angående multikollinearitet för moderatorerna *osäkerhet* och *karaktärsdrag*, vilket skulle kunna förklara varför moderatorerna ej blir signifikanta. Då modellen lider av viss multikollinearitet som kan påverka vilka variabler som är signifikanta inom modellen kan ett möjligt förbättringsområde för studien vara att utöka urvalsstorleken. Utökad urvalsstorlek kan reducera inverkan av multikollinearitet inom modellen, då större urvalsstorlek ökar sannolikheten för att individuella effekter blir signifikanta (Hair et al., 2020).

Större urvalsstorlek kan kompensera för eventuella effekter av multikollinearitet, men det finns tecken inom urvalet som pekar mot att *karaktärsdrag* och *osäkerhet* saknar betydande inverkan på *faktisk användning*, oavsett om multikollinearitet finns inom modellen eller ej. Korrelationen mellan moderatorerna och *faktisk användning* är låg för båda variablerna, då samtliga korrelationer mellan variablernas indikatorer och *faktiskt användning* är låg (se Figur 6). Utöver detta är vissa indikatorer negativt korrelerade och andra positivt korrelerade. Det finns inget tydligt mönster mellan moderatorerna och den beroende variabeln. Därmed är det inte troligt att ett statistiskt samband skulle kunna påvisas, även om modellen inte skulle lida av multikollinearitet.



Figur 10. Stapeldiagram över likertsvärsfördelning för karaktärsdrag

Den låga och tvetydiga korrelationen kan dock vara ett resultat av ett snedvridet urval. Figur 10 redovisar fördelningen av svar för variabeln karaktärsdrag (se Appendix D), med ett för de respondenter som klassas som starka icke-rationalister och fem för de respondenter som klassas som starka rationalister. En betydande majoritet av respondenter angav antingen fyra eller fem på likertskalan, vilket innebär att den överväldiga majoriteten klassificerade sig som rationalister. Detta snedvridna urval kan bero på studiens branschval, då revisorer som beslutstagare, enligt Nanda & Narayandas (2021), kan beskrivas som homogena. Resultatet i Figur 10 kan även vara en konsekvens av att frågorna uppfattas som ledande eller att respondenter ville framstå som datadrivna beslutsfattare. Resultatet från studien indikerar att *karaktärsdrag* och *osäkerhet* inte har en modererande effekt på *faktisk användning*, dock kan inte detta resultat säkerställas då fördelningen av respondenternas svar är snedvridet. Anledning varför *förväntad ansträngning* ej är signifikant inom modellen kan bero på studiens avgränsning mot revisionsbranschen. Under samtalen med representanter från de fyra företagen inom studien framgick det att inläring av big data systemen i någon form var obligatorisk. Då det är obligatoriskt att lära sig systemet är det rimligt att anta att samtliga respondenter hade kunskap att använda systemen. Ett betydande ansträngningsmoment vid användande av tekniska system är inlärningsprocessen, som i detta fall samtliga respondenter redan genomgått. Därmed är det troligt att variationen i *förväntad*

*ansträngning* är relativt liten mellan individer, då det största momentet som bidrar till ansträngning redan är överkommet. Om variationen är låg mellan individer minskar sannolikheten att variabeln är signifikant, då den individuella effekten av variabeln är lägre med mindre variation. Vid kontrollerandet av variationen i svarsmaterialet för *förväntad ansträngning* (Se Appendix C, Tabell 8) tydliggjordes variansen för respondenternas angivna *förväntade ansträngning*. Den genomsnittliga standardavvikelsen för *förväntad ansträngnings* indikatorer uppgick till 1,063, vilket var lägst av samtliga latenta variabler (se Appendix C, Tabell 8). *Förväntad ansträngning* hade därmed lägst varians av samtliga latenta variabler, vilket underbygger hypotesen om att låg varians kan ha påverkat signifikansnivån.

#### Standardavvikelse latenta variabler

<b>Latent variabel</b>	<b>Indikator</b>	<b>Standardavvikelse</b>	<b>Genomsnittlig standardavvikelse</b>
<b>Förväntad prestation</b>			1,180
	FP1	1,218	
	FP2	1,157	
	FP3	1,164	
<b>Förväntad ansträngning</b>			1,063
	FA1	1,095	
	FA2	1,079	
	FA3	1,016	
<b>Socialt inflytande</b>			1,190
	SI1	1,208	
	SI2	1,075	
	SI3	1,289	
<b>Underlättande förhållanden</b>			1,306
	UF1	1,367	
	UF2	1,324	
	UF3	1,226	

Tabell 8: Tabell på indikatorers standardavvikelse.

Studien ser vikten av att företag får en vägvisning i var de ska rikta sitt fokus, för att öka den generella användningsgraden hos medarbetare och utveckla de konkurrensfördelar som är nära förknippade med användning av big data system (McAfee & Brynjolfsson, 2012). Studiens resultat visade att *socialt inflytande*, *förväntad prestation* och *underlättande förhållanden* var signifikanta. *Socialt inflytande* var en av de faktorer som uppnådde signifikans vid prövandet av direkta samband. Studien visar därmed att sociala effekter har en inverkan på individens användningsgrad av big data system. Företag som ämnar att öka användningsgraden av sina big data system kan utnyttja detta faktum och främja socialt inflytande som driver ökad användning av big data system. Exempelvis kan företaget försöka etablera en organisationskultur som

förespråkar användandet av big data system för att lösa problem. Detta ligger i linje med forskning av Gressel, Pauleen & Taskin (2020) som menar att organisationer kan etablera en företagskultur och därmed inte blir beroende av individer för att bestå. Företaget kan även stötta förkämpar för big data användande inom företaget genom att utöka deras inflytande.

Studien ser även vikten av att företag förbättrar underlättande förhållandena för den enskilda individen i att kunna använda big datasystemen. Förslag på åtgärder är att fokusera resurser på hjälpfunktioner dedikerade till potentiella problem med systemen, som exempelvis IT-stöttning. Andra förslag på underlättande förhållanden som företagen kan vidta är att erbjuda interna utbildningar i systemen eller ha väl kommunicerade instruktionsmanualer till förfogande.

Vidare finns det anledning att uppmuntra företagen att lägga resurser på att förmedla de förväntade prestationsökningarna ännu tydligare till sina medarbetare, då detta är en av de tydligaste drivkrafterna som tidigare studier lyckats kartlägga. Även denna studie kunde påvisa att *förväntad prestation* blev signifikant vid borttagandet av moderatorerna och extremvärden i modellen. Det påstådda sambandet ska dessutom vara betydligt mycket starkare än vad exempelvis *socialt inflytande* var, innebärande att *förväntad prestation* har potential att påverka individernas beteende till stor grad om organisationen lyckas förmedla de förväntade prestationsökningarna till medarbetarna. Exempel på hur prestationsförbättring kan lyftas är att motivera för de anställda vad företaget ser för positiva fördelar med att använda systemet och vad de på sikt vill få ut av en ökad användningsgrad. Detta för att öka förståelsen för företagets motivering av införandet av systemen och få individer att se helhetsbilden av huvudboksanalyssystemens värdeskapande förmågor.

## 5.2 Slutsats

Denna studie undersökte de drivande faktorerna till individuell användning av big data system inom revisionsbranschen genom att tillämpa och konfigurera UTAUT-modellen. Studien visade att *socialt inflytande* och *underlättande förhållanden* var signifikanta inom modellen och hade ett positivt samband med den beroende variabeln, *faktisk användning*. När studien justerade för extremvärden i den beroende variabeln var även *förväntad prestation* signifikant och påvisade

det starkaste positiva sambanden av samtliga oberoende variabler. Den återstående oberoende variabeln, *förväntad ansträngning*, var inte signifikant i något fall. Därmed kan inget statistiskt samband mellan *förväntad ansträngning* och *faktisk användning* påvisas inom denna studie. Samtliga moderatorer inom modellen, både *erfarenhet* och studiens egna introducerade moderatorer *osäkerhet* och *karaktärsdrag*, var inte signifikanta inom modellen. Revisorer kan antas vara en homogen grupp, baserat på det empiriska materialet kring likheter i svar för *osäkerhet* och *karaktärsdrag*. Därmed är det möjligt att en annan bransch som påvisar mindre homogenitet hade kunnat påvisa signifikans för moderatorerna.

Revisionsföretag som ämnar att öka anställdas användningsgrad av big data system kan försöka skapa eller öka socialt inflytande som förespråkar big data system. Företag kan etablera en företagskultur som främjar användandet av big data system och datadrivet beslutsfattande. Utöver att påverka socialt inflytande, kan företag inom branschen förbättra sina underlättande förhållanden för användning av systemen. Detta kan uppnås genom att skapa stödfunktioner för användning av systemet, som teknisk support inom företaget. Slutligen kan företag utöka anställdas användning av big data system genom att belysa potentiella effektivitet och kvalitetsförbättringar som användning av big data system medför. Om anställda inom företag kan övertygas att användning av systemet medför individuella prestationsförbättringar är det troligt att användningsgraden inom företag ökar.

Genom att presentera vilka faktorer som är signifikanta, syftar studien till att tillföra att underlätta arbetet i att öka användningsgraden av big data system hos revisionsbolag. Studiens resultat lyckas även på så sätt bidra till att praktiker kan, vid tillvaratagande av rekommendationer, minska den totala arbetsbördan vid implementation och arbete med big data system.



## 6. Förslag för framtida forskning

---

*Följande kapitel går igenom de rekommendationer som arbetet vill ge till framtida forskning inom området baserat på den data som framkommit av studiens resultat och information som redogjorts för i tidigare avsnitt.*

---

Utifrån studiens resultat finns det anledning att ge förslag till framtida forskning i vad som kan tänkas vara intressant att försöka undersöka framöver. Arbetets resultat drabbades av stora utmaningar i form av frånvaro av signifikans. För att ta vidare de hypoteser som arbetet ställt upp så vore det intressant att genomföra ännu större tester på andra, inte revisionsexklusiva, branscher. Förhoppningsvis resulterar detta i ett större stickprov som har större möjlighet att signifikant skilja fler samband från slumpen. Detta skulle potentiellt också kunna ge mer information i om det finns branschspecifika egenskaper som kan göra att vissa branscher påverkas mer av vissa faktorer än andra. Vidare hade det varit intressant att fortsätta pröva UTAUT-modellen relevans genom adderandet av nya modererande, ej tidigare beprövade variabler. Studien ser här att det hade varit intressant att exempelvis gå in ännu djupare på personlighetstyper för att se om detta kan ha en påverkan på en individs beteendemässiga avsikt att använda systemen eller ej.

Modellen inom studien led av betydande multikollinearitet, som troligen orsakades av den höga korrelationen mellan variablerna *förväntad prestation* och *förväntad ansträngning*. Med bakgrund av detta och det faktum att samtliga tidigare studier som försökt kartlägga individuella drivkrafter till individuell användning av big data system har misslyckats med att få signifikans för variabeln *förväntad ansträngning*, finner arbetet tillräcklig grund i att föreslå en framtida modifiering i form av en borttagning av variabeln *förväntad ansträngning*. På grund av variabelns höga multikollinearitet med variabeln för *förväntad prestation* indikerar det att variablerna mäter samma sak och att framtida forskningsförsök kan bli mer framgångsrika i att uppnå signifikans genom frånvaron av multikollinearitet för dessa variabler. Det innebär dock inte att *förväntad ansträngning* saknar grund vid mätningar av beteendemässig förändring till ny

teknik men i samband med prediktionsmodeller för big data specifika system tycks faktorn sakna vetenskaplig grund för att framöver kunna ingå i den big data modifierade UTAUT-modellen.

Det bör framhållas att studiens resultat antas vara villkorat av den nu pågående problematiken som flera företag upplever med låga användningsnivåer av big data system. En studie som replikeras om ett par år riskerar att se andra samband beroende på hur de miljömässiga faktorerna har förändrats, så som generell användningsnivå inom branschen. Vidare är det inte heller säkert att de företag som ingår i studien bör anses som representativa för en hel bransch. Framtida studier, som har mer resurser och tid, kan även undersöka större stickprov, för att se huruvida *förväntad ansträngning* och de modererande variablerna fortsatt blir insignifikanta, givet att multikollineariteten minskar i ett större stickprov.

# Litteraturlista

---

Alharthi, A., Krotov, V. & Bowman, M. (2017). Addressing barriers to big data. *Business Horizons*, vol 60, pp. 285-292. Tillgänglig genom: LUSEMs bibliotek <https://www.lusem.lu.se/library> [Hämtad 2022-10-14]

Alwahaishi, S. & Snasel, V. (2013). Consumers' Acceptance and Use of Information and Communications Technology: A UTAUT and Flow Based Theoretical Model. *Journal of Technology Management and Innovation*, vol. 8, no. 2, pp. 61-73. Tillgänglig genom: LUSEMs bibliotek <https://www.lusem.lu.se/library> [Hämtad 2022-12-06]

Baldwin, H. (2015). *When Big Data Projects Go Wrong*. Forbes. Tillgänglig online: <https://www.forbes.com/sites/howardbaldwin/2015/01/22/when-big-data-projects-go-wrong/?sh=748bcd1f6427> [Hämtad 2022-10-17]

Bandura, A. (1977). Self-efficacy: Toward a Unifying Theory of Behavioral Change. *Psychological Review*, vol. 84, No. 2, pp. 191-215. Tillgänglig genom: LUSEMs bibliotek <https://www.lusem.lu.se/library> [Hämtad 2022-11-23]

Bean, R. (2022). Why Becoming a Data-Driven Organization Is So Hard. *Harvard Business Review Digital Articles*, pp. 1-6. Tillgänglig genom: LUSEMs bibliotek <https://www.lusem.lu.se/library> [Hämtad 2022-12-14]

Bozionelos, N. (1996). Psychology of Computer Use: XXXIX. Prevalence of Computer Anxiety in British Managers and Professionals. *Psychological Reports*, vol. 78, pp. 995-1002. Tillgänglig online: <https://journals.sagepub.com/doi/10.2466/pr0.1996.78.3.995> [Hämtad 2022-10-14]

Brünink, L. (2016). *Cross-functional big data integration: applying the UTAUT model*. Mastersuppsats, Department of Business Administration, University of Twente. Tillgänglig online: [https://essay.utwente.nl/71098/1/Brunink\\_MA\\_BMS.pdf](https://essay.utwente.nl/71098/1/Brunink_MA_BMS.pdf) [Hämtad 2022-10-17]

Bryman, A. & Bell, E. (2017). *Företagsekonomiska forskningsmetoder*. Malmö: Liber AB.

Cabrera-Sánchez, J-P. & Villarejo-Ramos, A. F. (2020). Acceptance and use of big data techniques in services companies, *Journal of Retailing and Consumer Services*, vol. 52. Tillgänglig online: <https://www.sciencedirect.com/science/article/pii/S0969698919302589> [Hämtad 2022-10-14]

Dahmström, K. (2011). *Från datainsamling till rapport - att göra en statistisk undersökning*. Polen: Studentlitteratur.

Dash, G. & Paul, J. (2021). CB-SEM vs PLS-SEM methods for research in social sciences and technology forecasting. *Technological Forecasting and Social Change*, vol. 173. Tillgänglig online: <https://www.sciencedirect.com/science/article/pii/S0040162521005254> [Hämtad 2022-12-02]

Davis, F. D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly*, vol. 13, no. 3, pp. 319-340. Tillgänglig genom: LUSEMs bibliotek <https://www.lusem.lu.se/library> [Hämtad 2022-10-14]

Deniswara, K., Kartono, R., Hamsal, M., Furinto, A. & Anthony, A. (2021). Analysis of the Auditor's Perspective on the Use of Big Data in Financial Statements: UTAUT Model Approach. Tillgänglig genom: LUSEMs bibliotek <https://www.lusem.lu.se/library> [Hämtad 2022-10-17]

Gressel, S., Pauleen, D. & Taskin, N. (2021). *Management Decision-Making, Big Data and Analytics*. London: SAGE Publications Ltd.

Gressel, S. (2019). *Management decision making in the age of big data: an exploration of the roles of analytics and human judgment*. Doktorsuppsats, Department of Management, Massey University, Tillgänglig online: <https://mro-ns.massey.ac.nz/bitstream/handle/10179/15999/GresselPhDThesis.pdf> [Hämtad 2022-11-17]

Hair, J., Black, W., Babin, B. & Anderson, R. (2020), *Multivariate Data Analysis*. Hampshire: Cengage Learning.

Hair, J. F., Hult, G. T. M., Ringle, C. M., Sarstedt, M., & Thiele, K. O. (2017). Mirror, Mirror on the Wall: A Comparative Evaluation of Composite-based Structural Equation Modeling Methods. *Journal of the Academy of Marketing Science (JAMS)*, 45(5), 616-632.

Hair, J., Risher, J., Sarstedt, M. & Ringle, C. (2018). When to use and how to report the results of PLS-SEM. *European Business Review*, vol. 31, no. 1. Tillgänglig online: [https://www.researchgate.net/publication/329443894\\_When\\_to\\_use\\_and\\_how\\_to\\_report\\_the\\_results\\_of\\_PLS-SEM](https://www.researchgate.net/publication/329443894_When_to_use_and_how_to_report_the_results_of_PLS-SEM)[Hämtad 2022-12-02]

Hair, J. F. & Sarstedt, M. (2019). Factors versus Composites: Guidelines for Choosing the Right Structural Equation Modeling Method. *Project Management Journal*, vol. 56. Tillgänglig online: <https://journals.sagepub.com/doi/epub/10.1177/8756972819882132> [Hämtad 2022-12-02]

- Hausman, A. & Stock, J. R. (2003). Adoption and implementation of technological innovations within long-term relationships. *Journal of Business Research*, vol. 56, pp. 681-686. Tillgänglig genom: LUSEMs bibliotek <https://www.lusem.lu.se/library> [Hämtad 2022-10-12]
- Hausmann, R., Tyson, L. D. & Zahidi, S. (2006). *The Global Gender Gap Report 2006*, World Economic Forum. Tillgänglig online: [https://www3.weforum.org/docs/WEF\\_GenderGap\\_Report\\_2006.pdf](https://www3.weforum.org/docs/WEF_GenderGap_Report_2006.pdf) [Hämtad 2022-10-13]
- Jena, S. (2022). *Difference between Traditional data and Big data*. Geeksforgeeks. Tillgänglig genom: <https://www.geeksforgeeks.org/difference-between-traditional-data-and-big-data/> [Hämtad 2022-11-7]
- Johnson, E. J. & Payne, J. W. (1985). Effort and Accuracy in Choice. *Management Science*, vol. 31, no. 4, pp. 395-414. Tillgänglig online: <https://pubsonline.informs.org/doi/abs/10.1287/mnsc.31.4.395> [Hämtad 2022-10-17]
- Kulviwat, S., Bruner, G. C. & Al-Shuridah, O. (2009). The role of social influence on adoption of high tech innovations: The moderating effect of public/private consumption, *Journal of Business Research*, vol. 62, pp. 706-712. Tillgänglig genom: LUSEMs bibliotek <https://www.lusem.lu.se/library> [Hämtad 2022-10-13]
- Kwon, O., Lee, N. & Shin, B. (2014). Data quality management, data usage experience and acquisition intention of big data analytics, *International Journal of Information Management*, vol 34, pp. 387-394. Tillgänglig online: [https://www-sciencedirect-com.ludwig.lub.lu.se/science/article/pii/S0268401214000127?ref=cra\\_js\\_challenge&fr=RR-1](https://www-sciencedirect-com.ludwig.lub.lu.se/science/article/pii/S0268401214000127?ref=cra_js_challenge&fr=RR-1) [Hämtad 2022-11-29]
- McAfee, A. & Brynjolfsson, E. (2012). Big Data: The Management Revolution, *Harvard Business Review*, vol. 90, no. 10. Tillgänglig online: <https://hbr.org/2012/10/big-data-the-management-revolution> [Hämtad 2022-10-11]
- Morris, M. G. & Venkatesh, V. (2000). Age Differences in Technology Adoption Decisions: Implications for a Changing Work Force, *Personnel psychology*, vol. 53, no. 2, pp. 375-403. Tillgänglig genom: LUSEMs bibliotek <https://www.lusem.lu.se/library> [Hämtad 2022-10-19]
- Nanda A. & Narayandas D. (2021). What Professional Service Firms Must Do to Thrive They have to tightly align their practices and clients. *Harvard Business Review*, vol. 99, no. 2, pp. 98. Tillgänglig genom: LUSEMs bibliotek <https://www.lusem.lu.se/library> [Hämtad 2022-10-19]

Payne, J. W. (1982). Contingent Decision Behavior *Psychological Bulletin*, vol. 92, no. 2, pp. 382-402. Tillgänglig genom: LUSEMs bibliotek <https://www.lusem.lu.se/library> [Hämtad 2022-11-23]

Podsakoff, P. M., MacKenzie, S. B., Jeong-Yeon, L. & Podsakoff, N. P. (2003). Common Method Biases in Behavioral Research: A Critical Review of the Literature and Recommended Remedies. *Journal of Applied Psychology*. Vol. 88, No. 5, 879–903

Ramayah, T. (2010). The Role of Voluntariness in Distance Education Students' Usage of a Course Website, *The Turkish Online Journal of Educational Technology*, vol. 9, no. 3, pp. 96-105. Tillgänglig genom: LUSEMs bibliotek <https://www.lusem.lu.se/library> [Hämtad 2022-10-13]

Ringle, C., Sarstedt, M. & Straub, D. (2012). An Assessment of the Use of Partial Least Squares Structural Equation Modeling in Marketing Research. *MIS Quarterly*, vol. 36, no. 1. Tillgänglig genom: LUSEMs bibliotek <https://www.lusem.lu.se/library> [Hämtad 2022-12-05]

Ross, J., Beath, C. & Quaadgras, A. (2013). You May Not Need Big Data After All. *Harvard Business Review*, vol. 91, no. 12, pp. 90-98. Tillgänglig genom: LUSEMs bibliotek <https://www.lusem.lu.se/library> [Hämtad 2022-10-7]

Schwab, K., Brende, R., Zahidi, S., Bekhouche, Y., Guinault, A. & Soo, A. (2013). *The Global Gender Gap Report 2013*, World Economic Forum. Tillgänglig online: [https://www3.weforum.org/docs/WEF\\_GenderGap\\_Report\\_2013.pdf](https://www3.weforum.org/docs/WEF_GenderGap_Report_2013.pdf) [Hämtad 2022-10-13]

Shahbaz, M., Zhai, L., Shahzad, F. & Hu Y. (2019). Investigating the adoption of big data analytics in healthcare: the moderating role of resistance to change. *Journal of Big Data*, vol. 6. Tillgänglig online: <https://journalofbigdata.springeropen.com/articles/10.1186/s40537-019-0170-y> [Hämtad 2022-10-12]

Statista. (2021). *Number of mobile (cellular) subscriptions worldwide from 1993 to 2021*. Tillgänglig online: <https://www.statista.com/statistics/262950/global-mobile-subscriptions-since-1993/> [Hämtad 2022-11-10]

Statista. (2022). *Distribution of LinkedIn users worldwide as of January 2022, by age group*. Tillgänglig online: <https://www.statista.com/statistics/273505/global-linkedin-age-group/> [Hämtad 2023-01-01]

Soon, K. W. K., Lee, C. A. & Boursier, P. (2016). *A study of the determinants affecting adoption of big data using integrated Technology Acceptance Model (TAM) and diffusion of innovation (DOI) in Malaysia*. Tillgänglig online:  
[https://www.researchgate.net/publication/304622794\\_A\\_study\\_of\\_the\\_determinants\\_affecting\\_a\\_doption\\_of\\_big\\_data\\_using\\_integrated\\_Technology\\_Acceptance\\_Model\\_TAM\\_and\\_diffusion\\_of\\_innovation\\_DOI\\_in\\_Malaysia](https://www.researchgate.net/publication/304622794_A_study_of_the_determinants_affecting_a_doption_of_big_data_using_integrated_Technology_Acceptance_Model_TAM_and_diffusion_of_innovation_DOI_in_Malaysia) [Hämtad 2022-10-17]

Tabesh, P., Mousavidin, E. & Hasani, S. (2019). Implementing big data strategies: A managerial perspective. *Business Horizons*, vol. 62, no. 3, pp. 347-358. Tillgänglig online:  
[https://www.researchgate.net/publication/331691114\\_Implementing\\_big\\_data\\_strategies\\_A\\_managerial\\_perspective](https://www.researchgate.net/publication/331691114_Implementing_big_data_strategies_A_managerial_perspective) [Hämtad 2022-10-5]

Tschakert, N., Kokina, J, Kozlowski, S. & Vasarhelyi, M. (2016). The next frontier in data analytics: Why CPAs and organizations need to learn to use advanced technology to predict and achieve outcomes. *Journal of Accountancy*. Tillgänglig Online:  
<https://www.journalofaccountancy.com/issues/2016/aug/data-analytics-skills.html> [Hämtad 2022-12-19]

TTG. (2022). Big Data vs. Traditional Data: What's the difference? Tillgänglig genom:  
<https://treehousetechgroup.com/big-data-vs-traditional-data-whats-the-difference/> [Hämtad 2022-11-7]

UCLA. (n.d.) *What does Cronbach's alpha mean?*. Tillgänglig online:  
<https://stats.oarc.ucla.edu/spss/faq/what-does-cronbachs-alpha-mean/> [Hämtad 2023-01-08]

Venkatesh, V., Smith, R. H., Morris, M. G., Davis, G. B. & David, F. D. (2003). User Acceptance of Information Technology: Toward a Unified View, *MIS Quarterly*, vol. 27, no. 3, pp. 425-478. Tillgänglig genom: LUSEMs bibliotek <https://www.lusem.lu.se/library> [Hämtad 2022-10-6]

Wongsim, M., Nilniyom, N., Sompong, A., Kaiwinit, S. & Satchawatee, N. (2021). *Factors influencing the adoption of big data analytics in accounting: A Study of Energy Companies*. *Journal of Renewable Energy for Community (J-REG)*.

Wu, Y., Tao, Y., & Yang, P. (2007). Using UTAUT to explore the behavior of 3G mobile communication users. *Proceedings of IEEE International Conference on Industrial Engineering and Engineering Management*, pp. 199-203. Tillgänglig genom: LUSEMs bibliotek <https://www.lusem.lu.se/library> [Hämtad 2022-11-8]

# Appendix

## Appendix A

### Appendix A

Frågekonstruktioner, före urval och anpassning av representanters input.

<b>Variabel</b>	<b>Påstående</b>	<b>Formulering</b>	<b>Källa</b>
<b>Förväntad prestation</b>			
FP1	Att använda systemet förbättrar kvaliteten på det arbete jag gör.		Moore and Benbasat, 1991
FP2	Att använda systemet förbättrar min produktivitet.		Moore and Benbasat, 1991
FP3	Om jag använder systemet kommer jag att öka kvaliteten på resultatet av mitt jobb.		Compeau et al. 1999
FP4	Om jag använder systemet kommer mina kollegor att uppfatta mig som kompetent.		Compeau et al. 1999
FP5	Att använda systemet skulle förbättra min arbetsprestation.		Davis et al., 1989
FP6	Jag skulle tycka att systemet är användbart i mitt jobb.		Davis et al., 1989
FP7	Användning av systemet kan avsevärt öka kvaliteten på resultatet av mitt jobb.		Thompson et al., 1991
<b>Förväntad ansträngning</b>			
FA1	Att lära sig hantera systemet skulle vara enkelt för mig.		Davis et al., 1989
FA2	Jag skulle tycka att det är enkelt att få systemet att göra det jag vill det ska göra.		Davis et al., 1989
FA3	Jag skulle tycka att systemet är flexibelt att interagera med.		Davis et al., 1989
FA4	Min interaktion med systemet är tydlig och begriplig.		Moore and Benbasat, 1991
FA5	Jag tror att det är lätt att få systemet att göra det jag vill att det ska göra.		Moore and Benbasat, 1991
FA6	Sammantaget tror jag att systemet är lätt att använda.		Moore and Benbasat, 1991
FA7	Att lära mig systemet skulle inte vara tidskrävande.		Venkatesh, Davis & Davis, 2003
<b>Socialt inflytande</b>			
SI1	Folk som påverkar mitt beteende tycker att jag ska använda systemet.		Davis et al., 1989
SI2	Jag använder systemet pga. andelen medarbetare som använder systemet.		Thompson et al., 1991
SI3	Generellt sett har organisationen stöttat användningen av systemet.		Thompson et al., 1991
<b>Underlättande förhållanden</b>			
UF1	Jag fick vägledning i valet av systemet.		Thompson et al., 1991
UF2	Specialiserade instruktioner angående systemet var tillgängligt för mig.		Thompson et al., 1991
UF3	En specifik person eller grupp är tillgänglig för mig för hjälp med systemproblem.		Thompson et al., 1991
UF4	Jag har den kunskap som krävs för att kunna använda systemet.		Ajzen, 1991
UF5	Systemet är kompatibelt med andra system som jag använder.		Ajzen, 1991
UF6	Jag har de resurser som krävs för att använda systemet.		Ajzen, 1991
<b>Erfarenhet</b>			
E1	Jag har hög erfarenhet i användning av dataanalys.		Egen tolkning av Brünink, 2016
<b>Faktisk användning</b>			
FA1	Vid arbete med (analys/företagsspecifikt synonym till analys) där användning av (företagsspecifikt namn på big data system) är möjligt, hur stor andel (%) av din arbetstid skulle du uppskatta att du använder (företagsspecifikt namn på big data system)?		Egen konstruerad fråga



# Appendix B

Appendix B

Frågekonstruktioner efter anpassning av representant input och undersökningens syfte

Variabel	Påstående	Formulering	Källa
<b>Förväntad prestation</b>			
FP1		Jag bedömer att användning av systemet ökar min produktivitet.	Moore and Benbasat, 1991
FP2		Jag bedömer att användning av systemet ökar min arbetsprestation.	Davis et al., 1989
FP3		Jag bedömer att användning av systemet ökar kvalitén på resultatet av mitt arbete.	Thompson et al., 1991
<b>Förväntad ansträngning</b>			
FA1		Jag bedömer att det är lätt att få systemet att utföra det jag vill.	Moore and Benbasat, 1991
FA2		I det stora hela bedömer jag att systemet är lätt att använda.	Moore and Benbasat, 1991
FA3		Jag bedömer att det är tidskrävande att lära sig systemet.	Venkatesh, Davis & Davis, 2003
<b>Socialt inflytande</b>			
SI1		Medarbetare tycker att jag ska använda systemet i större utsträckning.	Davis et al., 1989
SI2		Eftersom att andra medarbetare använder systemet, använder jag det i större utsträckning.	Thompson et al., 1991
SI3		Generellt sett har organisationen uppmuntrat mig att använda systemet i större utsträckning.	Thompson et al., 1991
<b>Underlättande förhållanden</b>			
UF1		Specialiserade instruktioner för användning av systemet är tillgängligt för mig.	Thompson et al., 1991
UF2		En specifik person eller grupp är tillgänglig för mig för hjälp med systemproblem.	Thompson et al., 1991
UF3		Jag har den kunskap som krävs för att kunna använda systemet i mitt arbete.	Ajzen, 1991
<b>Faktisk användning</b>			
FA1		Vid arbete med (analys/företagsspecifik synonym till analys) där användning av (företagsspecifikt namn på big data system) är möjligt, hur stor andel (%) av din arbetstid skulle du uppskatta att du använder (företagsspecifikt namn på big data system) för hantering och analys av data?	Egen konstruerad fråga
<b>Erfarenhet</b>			
E1		Jag har hög erfarenhet i användning av dataanalys.	Brünink, 2016
E2		Jag har stor erfarenhet i att hantera och analysera stora mängder data.	Gressel et al., 2020
<b>Karaktärsdrag</b>			
K1		När jag fattar beslut brukar jag gå på egen intuition och erfarenhet framför data.	Gressel et al., 2020
K2		Jag bedömer att mer data minskar osäkerheten kring hur effektivt ett beslut är.	Gressel et al., 2020
K3		Jag agerar hellre på data än intuition och erfarenhet.	Gressel et al., 2020
<b>Osäkerhet</b>			
O1		Jag kan bedöma när data är trovärdig och användbar för beslutsfattande.	Gressel et al., 2020
O2		Jag känner mig säker i att arbeta med data och ta beslut med data som underlag.	Gressel et al., 2020

## Appendix C

### Standardavvikelse latenta variabler

<b>Latent variabel</b>	<b>Indikator</b>	<b>Standardavvikelse</b>	<b>Genomsnittlig standardavvikelse</b>
<b>Förväntad prestation</b>			1,180
	FP1	1,218	
	FP2	1,157	
	FP3	1,164	
<b>Förväntad ansträngning</b>			1,063
	FA1	1,095	
	FA2	1,079	
	FA3	1,016	
<b>Socialt inflytande</b>			1,190
	SI1	1,208	
	SI2	1,075	
	SI3	1,289	
<b>Underlättande förhållanden</b>			1,306
	UF1	1,367	
	UF2	1,324	
	UF3	1,226	

Tabell 8: Tabell på indikatorers standardavvikelse.

## Appendix D

Indikatorer	Svar från respondenter				
	1	2	3	4	5
FP1	9	11	27	23	16
FP2	9	10	30	25	12
FP3	8	8	24	31	15
FA1	12	9	38	21	6
FA2	11	13	35	22	5
FA3	6	29	29	17	5
SI1	14	11	34	17	10
SI2	14	22	28	20	2
SI3	10	16	10	35	15
UF1	12	13	23	16	22
UF2	15	25	21	11	14
UF3	12	12	22	30	10
K1	3	14	24	35	10
K2	2	9	8	41	26
K3	2	10	33	27	14
EF1	3	15	30	25	13
EF2	3	10	31	31	11
O1	0	5	25	36	20
O2	0	3	23	42	18

Tabell 9: Fördelning av svar för indikatorer från respondenter