

Pimeän työn määrä

Mittaamisselvitys elinkeinonharjoittajien
verotarkastustietojen perusteella

Selvitys 5/2022

Julkaisun nimi:

Pimeän työn määrä – Mittaamisselvitys elinkeinonharjoittajien verotarkastustietojen perusteella

Tekijät:

Asiantuntija Riitta Nummila sekä analyytikot Kirsi Ristola ja Lasse Winter

Diaarinumero: VH/880/02.98.00.00/2022 / 2021_086

Julkaisija: Verohallinto, Harmaan talouden selvitysyksikkö

Julkaistu: 22.11.2022

Julkaisutapa: Sähköinen (PDF)

Julkisuus: Julkinen

Lisätietoja medialle:

Johtaja Janne Marttinen, puh. 029 512 6066

Apulaisjohtaja Marko Niemelä, puh. 029 512 6070

Muut yhteydenotot htsy@vero.fi

Julkaisualustat harmaa-talous-rikollisuus.fi ja harmaatalous.fi (vain luvite-
tuille viranomaisille)

Velvoitteidenhoitoselvityksen voi pyytää ilmiöselvityksessä kuvatun ryhmän perusteella. (Laki Harmaan talouden selvitysyksiköstä (1207/2010) 5:3 §)

Harmaan talouden selvitysyksikön selvityksissä esitetyt näkemykset ja tulokintakannanotot ovat yksikön omia, eivätkä ne sido Verohallintoa tai muita viranomaisia.

Tiivistelmä

Harmaan talouden selvitysyksikkö on julkaissut samanaikaisesti neljä ilmiöselvitystä pimeästä työstä ja pimeästä palkasta. Pimeä työ aiheuttaa verovajetta monella tapaa ja on yksi osa-alue harmaan talouden kokonaisuudesta. Tässä selvityksessä tarkastellaan elinkeinonharjoittajien pimeää työtä ja pimeää palkanmaksua verotarkastustietojen avulla. Data-analyysi suoritetaan usealla eri menetelmällä ja se kattaa vuodet 2018–2020. Elinkeinonharjoittajilla pimeä työ aiheuttaa verovajetta silloin, kun elinkeinotoiminnan tuloja salataan, kuluja liioitellaan ja maksetaan palkansaajille palkkoja pimeästi.

Elinkeinonharjoittajat työllistävät lähtökohtaisesti itsensä ja toiminta on usein pienimuotoista. Suomessa on noin 220 000 elinkeinonharjoittajaksi rekisteröitynyttä henkilöä. Noin 10 prosenttia näistä ilmoittaa palkanmaksua. Elinkeinonharjoittajan verotuksen ominaispiirre on se, että yritystoiminnasta saatavat tulot verotetaan elinkeinonharjoittajalla itsellään ansio- ja pääomatulona. Veroja välttääkseen elinkeinonharjoittajat saattavat jättää osan tuloista ilmoittamatta tai ilmoittaa liikaa kuluja ja vähennyksiä, jotta verotettava yritystulo olisi mahdollisimman pieni.

Elinkeinonharjoittajien pimeän työn ja pimeästi maksettujen palkkojen määrän arvioidaan olevan vuodessa noin 250–450 miljoonaa euroa. Tämä vastaa noin 8–14 prosenttia vuosittain verotuksessa vahvistetuista elinkeinotoiminnan tulosten yhteismäärästä. Valtaosa (90 %) arviosta muodostuu elinkeinonharjoittajien liian vähäisenä ilmoittamista elinkeinotoiminnan tuloista ja perusteettomista vähennyksistä. Työntekijöille pimeästi palkat muodostavat vain pienen osan kokonaisarviosta (10 %). Selvityksen mukaan tuloja salaamalla, kuluja liioittelemalla ja pimeästi maksetuilla palkoilla aiheutetaan vuosittain noin 75–140 miljoonan euron verovaje. Yrittäjän eläkevakuutusmaksujen mahdollista vajetta ei tässä selvityksessä arvioida.

Selvityksen mukaan kohonnut riski pimeään työhön on erityisesti elinkeinonharjoittajalla, jonka toiminta on melko pientä ja jonka veronalaiset tulot muodostuvat pääosin elinkeinotoiminnasta. Työntekijöille maksettavien pimeiden palkkojen riski on puolestaan kohonnut laajamittaisempaa elinkeinotoimintaa harjoittavilla.

Pimeän työn riskin arvioidaan olevan kohonnut erityisesti ravitsemistoiminnan, maaliikenteen, talonrakentamisen, majoituksen sekä vuokraus- ja leasingtoiminnan toimialoilla.

Euroopan Komission pimeän työn arvioon verrattuna Suomessa elinkeinonharjoittajien pimeä työ näyttää Harmaan talouden selvitysyksikön arvion mukaan olevan melko maltillista. Pimeän työn torjuminen on tärkeää erityisesti veronmenetysten vähentämisen näkökulmasta. Vähäinen pimeän työn määrä turvaa myös edellytykset reilulle kilpailulle.

Avainsanat: Mittaaminen; Pimeä työ; Verovaje; Verotarkastus

Sisällys

1	Johdanto ja selvityksen tarkoitus	3
1.1	Pimeän työn ja palkan määritelmät sekä selvityksen rajaukset	3
1.2	Selvityksen tavoite ja tutkimuskysymykset	4
1.3	Verovajeen laskentaperusteet	5
1.3.1	Verovaje työntekijöiden pimeistä palkoista	5
1.3.2	Verovaje yrittäjän pimeästä työstä	6
2	Euroopan komission arvio itsensä työllistäjien pimeästä työstä	8
3	Elinkeinonharjoittajat Suomessa	10
3.1	Elinkeinonharjoittajien lukumäärä ja rekisteröintiasema	10
3.2	Toiminnan laajuus	11
3.3	Elinkeinonharjoittajien verotettavat tulot	11
3.4	Toimialat	13
3.5	Päätoimista ja sivutoimista elinkeinotoimintaa harjoittavat	14
3.6	Elinkeinonharjoittajat palkanmaksajina	16
3.7	Elinkeinonharjoittajan verotus	17
4	Elinkeinonharjoittajan verotettavan tulon yleisimmät virhetilanteet	19
4.1	Pimeä eli puuttuva myynti	19
4.2	Perusteettomat kulut	19
4.3	Yksityiskäytön tulouttaminen veroilmoituksella	20
4.4	Lisävähennykset ja muut kulujen oikaisut veroilmoituksella	21
4.5	Työntekijälle maksettu pimeä palkka	21
5	Data-analyysissä käytetyt verovalvonnan aineistot	23
5.1	Elinkeinonharjoittajiin kohdistuneet verotarkastukset	23
5.2	Pimeää työtä koskevat tarkastushavainnot	24
5.3	Säännönmukaisen valvonnan aineistot	25
6	Elinkeinonharjoittajien pimeän työn määrä eri menetelmillä arvioituna	26
6.1	CEM-kaltaistaminen	26
6.2	Asiakaskohtainen odottama	27
6.2.1	Menetelmä ja aineistot	27
6.2.2	Pimeän työn määrä	28
6.2.3	Asiakaskohtaisen odottaman tarkempia tuloksia	28
6.3	Profilointi	30
6.3.1	Estimaatit	30
6.3.2	Toimialat	31
6.3.3	Koko	32
7	Tulosten yhteenveto sekä verojen ja muiden maksujen menetykset	34
7.1	Pimeän työn määrä sekä verojen ja muiden maksujen menetykset	34
7.1.1	Pimeän työn määrä	34
7.1.2	Verojen ja sosiaalivakuutusmaksujen menetykset	34
7.2	Tyypillinen pimeä elinkeinonharjoittaja	35
7.2.1	Pimeää työtä tekevät elinkeinonharjoittajat liikevaihtoluokittain	35
7.2.2	Pimeää työtä tekevät elinkeinonharjoittajat toimialan mukaan	36

7.2.3	Päätoimisten ja sivutoimisten elinkeinonharjoittajien pimeä työ	37
7.3	Estimaattien laskentaan liittyvistä epävarmuuksista	38
8	Johtopäätökset	40
9	Lähteet	41
10	Luettelo kuvioista ja taulukoista	42
11	Liitteet	43
	Liite 1. CEM-kaltaistaminen	1
	Liite 2. Yrityskohtaisen odottaman laskenta	3
	Liite 3. Profilointi	6
	Liite 4. Ansiotulon verovajeen laskeminen	9

1 Johdanto ja selvityksen tarkoitus

Tämä selvitys on osa Harmaan talouden selvitysyksikön laatimaa pimeän työn ja palkan julkaisukokonaisuutta. Harmaan talouden selvitysyksikössä on laadittu neljä ilmiöselvitystä¹, joissa on tarkasteltu pimeää palkkaa ja pimeää työtä usealla eri menetelmällä. Tämä selvitys tuottaa tietoa siitä, kuinka yleistä ja laajamittaista pimeä työ on elinkeinonharjoittajien keskuudessa ja kuinka paljon verotuloja tämän takia menetetään. Sen sijaan tarkoituksena ei ole tuottaa erillisiä riskiyritysten joukkoja operatiivista toimintaa varten.

Pimeää työtä käsitteleviä tutkimuksia on esitelty saman aikaisesti julkaistussa osakeyhtiöiden pimeän työn ilmiöselvityksessä (*Pimeän työn määrä – Mittaamisselvitys osakeyhtiöiden verotarkastustietojen perusteella*). Tässä selvityksessä ei esitellä uudelleen niiden tuloksia. Luvussa kaksi on esitelty lyhyesti Euroopan komission raportin tuloksia itsensä työllistäjien osalta, sillä tutkimuksen osaa ei ole käsitelty muissa Harmaan talouden selvitysyksikön pimeän työn selvityksissä. Tässä elinkeinonharjoittajien pimeän palkan selvityksessä teoreettisen viitekehyksen muodostaa yleiskuvaus elinkeinonharjoittajista Suomessa sekä lyhyet kuvaukset elinkeinonharjoittajien verotuksesta ja elinkeinonharjoittajien verotuksessa yleisimmin havaituista virheistä.

1.1 Pimeän työn ja palkan määritelmät sekä selvityksen rajaukset

Elinkeinonharjoittajalla tarkoitetaan henkilöä, joka on rekisteröity elinkeinonharjoittajaksi ja hänellä on y-tunnus. Elinkeinonharjoittaja toimii omista nimissään ja vastaa itse omista taloudellisista sitoumuksistaan. Y-tunnuksettomia kevytyrittäjiä ja muita itsensä työllistäjiä tämä selvitys ei käsittele.

Suurin osa elinkeinonharjoittajista työllistää vain itsensä. Elinkeinonharjoittaja ei voi maksaa itselleen palkkaa, vaan hänen verotettavaa tuloaan on elinkeinotoiminnasta saatu voitto. Näistä syistä ongelmaksi nousee, kuinka mitata elinkeinonharjoittajan pimeää työtä tai sen arvoa.

Tässä selvityksessä mitataan elinkeinonharjoittajien tekemän pimeän työn arvoa ja sen heille tuottamaa rahallista hyötyä vältettyjen verojen muodossa. Pimeän työn määrän mittaaminen esimerkiksi tehtyinä työtunteina olisi ollut haastavaa verotusaineistojen perusteella. Lisäksi tarkoitus on nimenomaan ollut mitata enemmän pimeän työn arvoa kuin määrää tunteina ja johtaa pimeän työn estimaateista edelleen menetettyjen tuloverojen arvio. Selvityksessä käytetty pimeän työn mittaaminen soveltuu näihin käyttötarkoituksiin hyvin.

Pimeän työn arvoa mitataan tässä selvityksessä elinkeinotoiminnan verotettavan tulon avulla, joka jää pimeän työn seurauksena liian pieneksi. Tällaisia verotettavan tulon

¹ Selvitykset on julkaistu 22.11.2022. Pimeän työn ilmiöselvitykset ovat: *Pimeän työn määrä – Mittaamisselvitys osakeyhtiöiden verotarkastustietojen perusteella 4/2022*, *Pimeän työn määrä – mittaamisselvitys elinkeinonharjoittajien verotarkastustietojen perusteella 5/2022*, *Pimeän työn mittaaminen profilointi menetelmällä - Osakeyhtiöiden pimeän työn määrä 06/2022* ja *Pimeän työn mittaaminen LIM-menetelmällä – Yksityisen sektorin palkan-saajien pimeän työn määrä 7/2022*. Lisäksi pimeän työn selvityksistä on julkaistu kooste: *Pimeä työ Suomessa – kooste selvityksistä 2022, 8/2022*.

virhetilanteita ovat: elinkeinotoiminnan tuotot, joita ei ilmoiteta veroilmoituksella, perusteettomasti elinkeinotoiminnan verotuksessa vähennyksiksi ilmoitetut kulut² sekä veroilmoituksella tehdyt perusteettomat kuluoikaisut ja tulouttamatta jäänyt yksityiskäyttö.

Pieni osa elinkeinonharjoittajista käyttää ulkopuolista työvoimaa. Tällöin pimeällä työllä tarkoitetaan myös työntekijöille maksettuja pimeitä palkkoja eli palkkoja, joista työnantaja ei ole asianmukaisesti hoitanut lakisääteisiä vero- ja sosiaalivakuutusmaksuvelvoitteitaan. Työntekijöille maksetut pimeät palkat jäävät sivurooliin tässä selvityksessä, koska elinkeinonharjoittajat työllistävät vain harvoin muita itsensä lisäksi.

Laajasti ajateltuna pimeänä työnä voidaan pitää myös tilannetta, jossa työnantaja teettää työntekijällä työtä, mutta maksaa siitä alle yleis- tai normaalisitovan työehtosopimuksen mukaista palkkaa. Tätä kutsutaan usein alipalkaksi. Työnantaja hyötyy menettelystä, koska työ tulee tehtyä pienin kustannuksin, kun osa palkasta jätetään maksamatta. Näiden väärinkäytöstilanteiden ja niissä menetettyjen verojen ja lakisääteisten maksujen määrän arvioiminen on hankalaa. Tapausten selvittäminen edellyttää usein monien viranomaisten yhteistyötä, tapauskohtaista tarkastelua ja työrikosten tutkintaa (työsyryntä, kiskonnantapainen työsyryntä, jopa ihmiskauppa). Tämä selvitys ei koske alipalkkaustilanteita. Harmaan talouden selvitysyksikkö on laatinut ulkomaalaisten alipalkkausta harmaan talouden ilmiönä käsittelevän selvityksen vuonna 2020. Selvitys on luettavissa harmaa-talous-rikollisuus.fi-sivustolla³.

1.2 Selvityksen tavoite ja tutkimuskysymykset

Selvityksen tavoitteena oli arvioida elinkeinonharjoittajien pimeää työtä ja siitä aiheutuvia verojen ja sosiaalivakuutusmaksujen menetyksiä vuosina 2018–2020. Tavoitteen saavuttamiseksi asetettiin seuraavat tutkimuskysymykset:

- Kuinka paljon Suomessa on elinkeinonharjoittajia, kuinka laajaa elinkeinotoiminta on ja millä toimialoilla elinkeinonharjoittajat toimivat?
- Kuinka paljon tuloja elinkeinonharjoittajat saavat ja miten tulot jakautuvat eri lähteittäin?
- Millaisia ovat elinkeinonharjoittajien pimeään työhön liittyvät virhetilanteet?
- Kuinka suuri on elinkeinonharjoittajien pimeiden palkkojen määrä?
- Kuinka suuri on pimeiden palkkojen määrä suhteessa elinkeinotoiminnan tulokseen?
- Kuinka suuri on elinkeinonharjoittajan osuus pimeistä palkoista? Entä palkansaajien osuus?
- Millä toimialoilla pimeän työn riski on suurin?
- Miten pimeiden palkkojen estimaatit jakautuvat kokoluokittelussa (liikevaihto)?
- Miten suuri osa pimeästä työstä tapahtuu päätoimisessa ja/tai sivutoimisessa toiminnassa?

² Esimerkiksi kuluihin sisällytetyt yksityistalouden menot tai muut menot, jotka eivät ole täytä elinkeinotulon verottamisesta annettuun lain mukaista vähennyskelpoisten kulujen määritelmää. EVL 7 §:n mukaan vähennyskelpoisia ovat elinkeinotoiminnassa tulon hankkimisesta tai säilyttämisestä johtuneet menot ja menetykset.

³ Harmaa talous & talousrikollisuus / Laajuus / Harmaan talouden selvityksiä / Ulkomaalaisten alipalkkaustapauksista oikeuteen vain jäävuoren huippu. Selvitys 05/2020 julkaistu 20.10.2020 Ulkomaalaisten alipalkkaus harmaan talouden ilmiönä Suomessa – oikeustapausanalyysi.

Ennen data-analyysijä muodostettiin yleiskuva Suomessa toimivista elinkeinonharjoittajista, kuten lukumääristä, toimialoista ja liiketoiminnan laajuudesta (luku 3). Yksi keskeinen elinkeinonharjoittajia jakava piirre on elinkeinotoiminnan päätoimisuus tai sivutoimisuus. Kaikki elinkeinonharjoittajat eivät ole päätoimisia yrittäjiä, vaan osalla pääasiallinen tulonlähde on esimerkiksi palkkatyö. Elinkeinoharjoittajien pimeään työhön liittyviä verotuksen virhetilanteita teorian tasolla kuvataan luvussa neljä.

Data-analyysien aineistoja esitellään luvussa viisi. Pimeää työtä tarkastellaan elinkeinonharjoittajien verotarkastuksilla tehtyjen havaintojen ja verojen maksuunpanojen avulla. Aineistoa on täydennetty soveltuvin osin myös elinkeinonharjoittajien säännönmukaisessa verotuksessa tehdyillä verotustoimenpiteillä. Aineistojen aikarajaus käsitti vuodet 2018–2020 ja sen taustalla on Verohallinnon tietojärjestelmä uudistus, jossa henkilöiden tuloverotus otettiin uudessa järjestelmässä käyttöön verovuodesta 2018 alkaen.

Varsinaiset data-analyysit ja niiden tuloksia on esitelty luvuissa kuusi ja seitsemän. Käytetyt analyysimenetelmät olivat CEM-kaltaistaminen, asiakaskohtaista riskisyyttä ennustava koneoppimismenetelmä sekä Harmaan talouden selvitysyksikössä kehitetty profilointimenetelmä.

1.3 Verovajeen laskentaperusteet

Selvityksen varsinainen data-analyysi tuotti estimaatteja pimeiden palkkojen määrästä. Arviot menetetyistä veroista ja palkkojen sosiaalivakuutusmaksuista johdettiin näistä estimaateista. Seuraavaksi on kuvattu verovajeen laskentaperusteet erikseen työntekijöiden pimeistä palkoista ja yrittäjän pimeästä työstä.

1.3.1 Verovaje työntekijöiden pimeistä palkoista

Työntekijöille maksettujen pimeiden palkkojen verovaje laskettiin samalla prosenttikertoimella kuin selvityksessä *Pimeän työn määrä – Mittaamisselvitys osakeyhtiöiden verotarkastustietojen perusteella*. Seuraavaksi on kuvattu lyhyesti, mihin tietoihin pimeiden palkkojen kertomien laskenta perustuu.

Ennakonpidätykset

Selvityksessä käytettiin ennakonpidätysten määrän laskemiseen kaavamaista 25 prosenttia. Taustalla on korkeimman oikeuden (KKO) ratkaisu⁴, jossa KKO on ottanut kantaa vältetyn veron määrän arviointiin törkeän veropetoksen rikosasiassa. Kyse oli pimeästä palkanmaksusta, ja päätöksen perusteluissa on todettu aiempien oikeusasteiden laskentaperusteen (25 prosenttia) olleen soveltuva ennakonpidätyksen määräksi.

⁴ KKO 2017:75.

Palkan sivukulut

Sosiaalivakuutusmaksujen⁵ määristä säädetään erikseen kutakin maksua koskevassa laissa tai niitä täydentävissä asetuksissa⁶. Sosiaalivakuutusmaksujen menetysten laskennassa käytettiin sosiaali- ja terveysministeriön vuodelle 2021 vahvistamia prosentteja (Sosiaali- ja terveysministeriö). Seuraavassa taulukossa on esitetty kooste palkan sivukulujen osuudesta⁷, joiden määrä oli yhteensä noin 28 prosenttia.

Taulukko 1. Palkan sivukulut ja niiden suuruus vuonna 2021.

Sosiaalivakuutusmaksut	Sivukulun suuruus (prosenttiosuus bruttopalkasta)
TyEL (työnantajan osuus)	16,95 %
TyEL (työntekijän osuus)	7,15 %
Työnantajan työttömyysvakuutusmaksu	0,50 %
Työntekijän työttömyysvakuutusmaksu	1,40 %
Työnantajan sairausvakuutusmaksu	1,53 %
Työnantajan tapaturmavakuutusmaksu	0,70 %
Yhteensä	28 %

Lähde: Sosiaali- ja terveysministeriö 2021.

1.3.2 Verovaje yrittäjän pimeästä työstä

Elinkeinotoiminnan verotettava tulo verotetaan kokonaisuudessaan yrittäjän (ja puolison) henkilökohtaisena tulona. Tässä selvityksessä laskettiin vain pimeän työn aiheuttamat tuloverojen (ansio- ja pääomatulo) menetykset. Sitä verovajeen määrää, joka aiheutuu muista elinkeinonharjoittajan veroista⁸ ja lakisääteisistä maksuista⁹ ei ole tässä selvityksessä arvoitu.

Tuloverotus (ansio- ja pääomatulojen vero)

Suomessa tulovero jakautuu yksityishenkilöillä ansiotuloista ja pääomatuloista perittäviin tuloveroihin. Vuoden 2019¹⁰ verotustietojen mukaan keskimäärin 94 prosenttia elinkeinonharjoittajien veronalaisesta yritystulosta verotettiin ansiotuloina ja kuusi prosenttia pääomatuloina.

Pääomatuloista veroa¹¹ maksetaan 30 tai 34 prosentin mukaan riippuen pääomatulojen vuotuisesta yhteismäärästä. Verotustietojen mukaan suurimmalla osalla elinkeinonharjoittajia vuotuiset pääomatulot jäivät alle 30 000 euron – toisin sanoen valtaosa maksaa pääomatuloista veroa 30 prosentin mukaisesti.

⁵ Työntekijän työeläkevakuutusmaksu (TyEL-maksu), työntekijän työttömyysvakuutusmaksu (TVR-maksu), työnantajan eläkevakuutusmaksu, työnantajan työttömyysvakuutusmaksu, työnantajan sairausvakuutusmaksu (savamaksu).

⁶ Työntekijän eläkelain 153 §, Laki työttömyysetuuskien rahoituksesta 18 §, Laki työnantajan sairausvakuutusmaksusta, Valtioneuvoston asetus sairausvakuutusmaksujen maksuprosenteista vuonna 2021.

⁷ Oletukset perustuvat ajatukseen ”keskimääräisestä” pimeän palkansaajasta. Tosiasialliset sivukulujen maksujen määrät vaihtelevat eri ikäryhmien kesken. Laskennassa käytetään oletusta, että pimeän palkansaaja on täysi-ikäinen, alle 53-vuotias ja työskennellyt sellaisen yrityksen palveluksessa, jonka vuoden 2020 aikana maksamat palkat ovat olleet alle 2 125 500 euroa.

⁸ Esim. arvonnisäverot salatuissa tuloissa tai perusteettomissa vähennyksissä.

⁹ Yrittäjän eläkevakuutusmaksu (YEL).

¹⁰ Verovajeen laskennan tietoja on tarkasteltu vuoden 2019 verotustiedoista.

¹¹ Tuloverolaki (31.12.1992/1535) 124 § 2 momentti.

Ansiotulot verotetaan valtionverotuksessa progressiivisen asteikon mukaan, eli tulojen noustessa myös veroprosentti nousee. Progressiivinen veroprosentti huomioitiin verovajeen laskennassa seuraavasti: ensin arvioitiin ansiotulojen osuus elinkeinonharjoittajien pimeän työn estimaatista. Seuraavaksi ansiotulojen osuus estimaatista jaettiin luokkiin liikevaihdon sekä päätoimisuuden tai sivutoimisuuden mukaan. Eri luokkiin kuuluville elinkeinonharjoittajille laskettiin keskimääräiset ansiotulot,¹² johon lisättiin arvioitu keskimääräinen pimeä työ. Näin muodostettiin ansiotulojen määrä, jonka mukaan veroprosentti määräytyi. Veroprosentilla kerrottiin pimeän työn estimaatit luokittain ja siten saatiin arvio menetetyistä ansiotuloeroista. Verovajeen laskenta on kuvattu tarkemmin liitteessä 4 ja sen tulokset on esitetty raportin luvussa 7.1.2.

Yrittäjän eläkevakuutusmaksu (YEL)

Tässä selvityksessä ei arvioitu YEL-vakuuttamisen puutteista aiheutuvaa maksujen vajetta, koska selvityksessä ei ollut käytettävissä YEL-vakuuttamisen yrittäjäkohtaisia tietoja. Jokaisen yrittäjän tulee vakuuttaa itsensä, jos yrittäjän työtulot ylittävät vakuuttamisen alarajan¹³. YEL-maksuja maksetaan yrittäjän ilmoittaman ja vakuutusyhtiön vahvistaman työtulon perusteella sosiaali- ja terveysministeriö vuosittain vahvistaman prosentin¹⁴ mukaisesti. Nykyinen eläkejärjestelmä sallii melko laajan valinnanvapauden yrittäjälle myös yli- ja alivakuuttamiseen, koska vakuutusmaksut perustuvat työtuloon¹⁵ eikä yritystoiminnasta tosiasiallisesti saatuun tuloihin. Selvityksen laatimishetkellä on ollut vireillä muun muassa yrittäjän eläkelain muuttaminen¹⁶, jossa on tarkoitus kehittää myös YEL-työtulon määrittämistä ja tarkistamista.

¹² Keskimääräiset ansiotulot laskettiin profiloitimenetelmän valitsemille riskitapauksille. Vertailun vuoksi keskimääräiset ansiotulot laskettiin myös koko perusjoukon elinkeinonharjoittajille. Riskitapausten ansiotulot olivat keskimäärin matalammat kuin koko elinkeinonharjoittajien perusjoukossa.

¹³ Yrittäjän työtulo tarkoittaa yrittäjän työpanoksen rahallista arvoa. Vuonna 2022 YEL-vakuuttamisen työtulon minimiraja on 8 261,71 euroa ja yläraja 187 625 euroa.

¹⁴ YEL-maksun suuruudet (2022) ovat alle 53-vuotiaille 24,10 %, 53–62-vuotiaille 25,60 % ja yli 62-vuotiaille yrittäjille 24,10 % vahvistetun työtulon määrästä.

¹⁵ Minimityötulo (8 261,71 € v.2022) on tarkoitettu vain osa-aikaisen yrittäjän työtuloksi. Yrittäjän tulisi arvioida työtulonsa vuosittain ja hakea tarvittaessa muutosta vahvistettuun työtuloon eli YEL-maksun perusteeseen. Takautuvasti muutosta ei voi hakea.

¹⁶ Hallituksen esitys HE 102/2022 vp.

2 Euroopan komission arvio itsensä työllistäjien pimeästä työstä

Harmaan talouden selvitysyksikön selvityksessä *"Pimeän työn määrä – Mittaamisselvitys osakeyhtiöiden verotarkastustietojen perusteella"*¹⁷ on esitetty kootusti pimeää työtä ja pimeää palkkaa koskevien tutkimusten tuloksia. Tulokset koskivat pääosin palkansaajien pimeää työtä. Sen sijaan elinkeinonharjoittajien pimeän työn määrää ja laajuutta on tutkittu vain vähän.

Euroopan komissio (myöhemmin komissio) on julkaissut vuonna 2017 tutkimuksen¹⁸ pimeästä työstä. Julkaisussa on esitetty arvioita ilmoittamattoman työn esiintyvyydestä EU-alueella ja sen jakautumisesta työntekijöiden pimeään palkkaan, itsensä työllistävien pimeään työhön ja kotitalouksien pimeään työhön. Arviot on esitetty EU-alueella keskimäärin ja EU-maittain. Arviot on laadittu käyttäen LIM-menetelmää (The Labour Input Method). Aineistona LIM-menetelmässä on käytetty muun muassa eri EU-maiden työllisyystutkimuksia työn tarjonnasta ja kysynnästä. Kysynnän ja tarjonnan eroavuuksiin perustuen on muodostettu arviot ilmoittamattoman työn määrästä. Harmaan talouden selvitysyksikkö on laatinut oman arvion työntekijöiden pimeän työn esiintyvyydestä ja määrästä Suomessa käyttäen LIM-menetelmää. Selvitys on nimeltään *Pimeän työn mittaaminen LIM-menetelmällä – Yksityisen sektorin palkansaajien pimeän työn määrä*¹⁹.

Komission pimeän työn määritelmä on laaja ja vastaa pääosin myös OECD:n pimeän työn määritelmää. Tiivistetysti pimeän työn määritelmään kuuluu kaikki laillisesta toiminnasta maksetut suoritukset, joita ei ole ilmoitettu viranomaisille²⁰.

Komission julkaisussa on esitetty arvioita pimeän työn esiintyvyydestä ja määrästä verrattuna työpanokseen. Euromääräisiä arvioita se ei sisällä. Julkaisussa on esitetty, että eri EU-maiden välillä on huomattavan suuria eroja pimeän työn esiintyvyydessä ja määrässä. EU-maiden välillä on myös huomattavia eroja muun muassa bruttokansantuotteessa, työttömyysprosentissa, luottamuksessa viranomaisiin ja esimerkiksi korruption esiintyvyydessä. Näillä on arvioitu olevan vaikutusta pimeän työn esiintyvyyteen ja määrään. Itsensä työllistävien osuus kaikesta pimeästä työstä on EU-alueella keskimäärin noin kolmasosa²¹ (Williams, 2017, p. 1). Julkaisun mukaan pimeän työn määrä korreloi ostovoimalla korjatun bruttokansantuotteen kanssa. Mitä suurempi on bruttokansantuote, sitä vähemmän pimeää työtä esiintyy (Williams, 2017, p. 19). Eri maista kerätyt tiedot perustuvat samoihin kriteereihin, mutta tietojen luotettavuus vaihtelee maiden välillä.

¹⁷ Selvitys 4/2022, julkaistu 22.11.2022.

¹⁸ An evaluation of the scale of undeclared work in the European Union and its structural determinants: estimates using the Labour Input Method, 2017.

¹⁹ Selvitys 07/2022, julkaistu 22.11.2022.

²⁰ Undeclared work is defined as "any paid activities that are lawful as regards their nature, but are not declared to the public authorities, taking into account the differences in the regulatory systems of the Member States" (European Commission, 2007: 1).

²¹ Euroopan komission arvion mukaan pimeä työ jakautuu seuraavasti: työsuhteessa tehty pimeä työ 61,8 %, itsensä työllistävien pimeä työ 37,3 % ja kotitalouksien pimeä työ 0,3 %.

Komission julkaisussa esitetyn arvion mukaan Suomessa pimeä työ oli yleisempää itsensä työllistävillä kuin työsuhteessa ja kotitalouksissa tehdyssä työssä²². Itsensä työllistäneillä pimeän työn osuus arvioitiin olevan 45,3 prosenttia työpanoksesta (total labour input) (Williams, 2017, p. 15). Suomessa jopa kaksi kolmasosaa pimeästä työstä liittyi itsensä työllistämiseen ja kolmasosa työsuhteessa tehtyyn työhön. Kotitalouksien pimeän työn osuudeksi oli arvioitu vain hyvin vähäinen määrä (Williams, 2017, p. 17). Komission julkaisusta ei yksiselitteisesti ole tulkittavissa mitä kaikkia yritysmuotoja ja toimijoita itsensä työllistäville²³ tarkoitetaan. Selvää kuitenkin on, että esimerkiksi elinkeinonharjoittajat lukeutuvat itsensä työllistävien joukkoon.

Seuraavassa luvussa on kuvattu elinkeinonharjoittajien toimintaa ja laajuutta Verohallinnon rekisteritietojen ja elinkeinonharjoittajien antamien veroilmoitustietojen avulla. Kuvauksen tarkoituksena on luoda yleiskuva siitä, minkälaista ja kuinka laajaa yritystoimintaa elinkeinonharjoittajilla on Suomessa.

²² Itsensä työllistäjien pimeän työn arvioitiin olevan muuta pimeää työtä yleisempää myös seuraavissa maissa: Kypros, Tanska, Kroatia, Irlanti, Liettua, Alankomaat, Romania ja Slovakia.

²³ Definition: Self-employed persons are persons who are the sole or joint owners of the unincorporated enterprises in which they work, excluding those unincorporated enterprises that are classified as quasi-corporations (European Commission/Eurostat/Ramon/Metadata, luettu 16.9.2022).

3 Elinkeinonharjoittajat Suomessa

Tässä luvussa on muodostettu yleiskuva Suomessa toimivista elinkeinonharjoittajista Verohallinnon tietojen perusteella. Yksityinen elinkeinonharjoittaja²⁴ on henkilö, joka harjoittaa liiketoimintaa itsenäisesti omaan lukuunsa, omissa nimissään, omalla y-tunnuksellaan. Elinkeinonharjoittaja vastaa kaikista yritystoimintansa velvoitteista koko henkilökohtaisella omaisuudellaan. Elinkeinonharjoittajien joukko on moninainen. Monella palkkatyössä käyvällä henkilöllä on sivutulojen hankkimista varten oma y-tunnus. Tällainen elinkeinotoiminta voi olla hyvinkin pientä. Toisaalta myös päätoimisesti²⁵ yritystoimintaa harjoittavalla elinkeinonharjoittajilla voi olla muita tulonlähteitä, kuten palkkatöitä, maataloustoimintaa tai muuta elinkeinotoiminnan ulkopuolista toimintaa.

3.1 Elinkeinonharjoittajien lukumäärä ja rekisteröintiasema

Suomessa on ollut vuosittain noin 220 000 y-tunnuksellista elinkeinonharjoittajaa selvityksen tarkastelujaksolla 2018–2020. Nämä y-tunnukset, jotka olivat rekisteröityneinä elinkeinonharjoittajiksi vähintään yhden päivän vuoden aikana ja joiden tuli antaa elinkeinotoiminnan veroilmoitus (lomake 5), muodostivat myös tämän selvityksen data-analyyseihin niin sanotun **perusjoukon**. Perusjoukkoa käytettiin eri menetelmiä soveltavissa data-analyyseissä.

Noin 65 prosenttia elinkeinonharjoittajista kuului arvonlisäverovelvollisten rekisteriin, valtaosa oli ilmoittautunut liiketoiminnasta arvonlisäverovelvollisten rekisteriin. Moni vähäistä liiketoimintaa harjoittava oli hakeutunut vapaaehtoisesti²⁶ arvonlisäverovelvollisten rekisteriin, vaikka vähäisen toiminnan 15 000 euron raja ei ollut ylittynyt. Lähes kaikki elinkeinonharjoittajat ovat hakeutuneet ennakkoperintärekisteriin (yli 90 prosenttia). Palkkoja maksavaksi työnantajaksi oli rekisteröitynyt noin 4 % kaikista elinkeinonharjoittajista. Vain noin 3 prosenttia elinkeinonharjoittajista ei kuulunut mihinkään Verohallinnon rekisteriin²⁷.

Taulukko 2. Elinkeinonharjoittajat Verohallinnon rekistereissä vuosina 2018–2020.

Vuosi	Elinkeinonharjoittajat (lkm)	Ennakkoperintärekisterissä (lkm)	Alv-rekisterissä* (lkm)	Työnantajarekisterissä (lkm)	Ei rekistereissä (lkm)
2018	213 200	198 600	140 000	9 400	7 100
2019	222 300	206 700	147 000	8 800	7 800
2020	222 300	208 300	153 100	8 100	6 600

*) Kaikki arvonlisäverovelvollisuus asemat (yleisimpiä ovat arvonlisäverovelvolliset liiketoiminnasta, -alkutuotannosta ja -kiinteistönkäyttöoikeuden luovutuksesta hakeutuneet).

Lähde: Verohallinto 2/2022.

²⁴ Puhekielessä elinkeinonharjoittajasta käytetään usein sanaa toiminimi tai toiminimiyrittäjä.

²⁵ Päätoimisella yrittäjyydellä tarkoitetaan yksinkertaistettuna sitä, että suurin osa henkilön veronalaisista tuloista on peräisin elinkeinotoiminnasta. Sivutoimisuudella tarkoitetaan tilannetta, jossa henkilön veronalaiset tulot tulevat pääosin muualta kuin elinkeinotoiminnasta.

²⁶ Yritys voi rekisteröityä vapaaehtoisesti alv-rekisteriin, vaikka niin sanottu vähäisen toiminnan liikevaihtoraja ei ylittyisi. Tällöin kaikista arvonlisäverolain mukaan verollisista myynneistä tulee suorittaa arvonlisäveroa, mutta vastaavasti elinkeinotoimintaan kohdistuvien ostojen sisältämät arvonlisäverot ovat vähennyskelpoisia.

²⁷ Nämä ovat rekisteröity elinkeinonharjoittajiksi eli heillä on y-tunnus, mutta he ole eivät ole merkittynä arvonlisäverovelvollisten rekisteriin, ennakkoperintärekisteriin tai palkkoja maksavien työnantajien rekisteriin.

3.2 Toiminnan laajuus

Kaikkien elinkeinonharjoittajien yhteenlaskettu veroilmoituksella ilmoitettu liikevaihto oli hie-
man alle 11 miljardia euroa vuodessa tarkastelujaksolla 2018–2020. Liikevaihtojen hajonta
oli huomattavaa. Liikevaihtojen keskiarvo oli noin 50 000 euroa ja mediaani noin 10 000 eu-
roa. Esimerkiksi neljäsosa elinkeinonharjoittajista ei ilmoittanut²⁸ veroilmoituksella lainkaan
liikevaihtoa ja vain 10 prosentilla ilmoitettu liikevaihto ylitti 100 000 euroa. Taulukossa 3 on
esitetty elinkeinonharjoittajien lukumääriä eri liikevaihtoluokkiin jaettuna.

Taulukko 3. Elinkeinonharjoittajien lukumäärät liikevaihtoluokissa, vuodet 2018–2020.

Vuosi	Ei liikevaihtoa	1 € - 10 000 €	10 000 € - 50 000 €	50 000 € - 100 000 €	yli 100 000 €
2018	62 600	51 200	54 900	25 800	18 700
2019	62 600	56 600	57 500	27 200	18 500
2020	56 800	62 600	59 700	26 400	16 900

Lähde: Verohallinto 2/2022.

3.3 Elinkeinonharjoittajien verotettavat tulot

Elinkeinotoiminnan voittoa ilmoittivat vuosittain noin 60 prosenttia ja tappiota noin 15 prosent-
tia elinkeinonharjoittajista. Neljännes (25 %) ei ilmoittanut lainkaan veronalaista elinkeinotoi-
minnan tuloa eikä kuluja. Elinkeinotoiminnan verotettava tulos oli usein huomattavan alhai-
nen verrattaessa esimerkiksi palkansaajien keskimääräisiin palkkoihin²⁹. Elinkeinotoimintaa
ilmoittaneiden voitollisen tuloksen keskiarvo oli noin 22 000 euroa ja mediaani 12 600 euroa.
Hajonta elinkeinotoiminnan tuloksissa oli huomattavaa.

Toisaalta osalla elinkeinonharjoittajista oli myös muita tulonlähteitä. Noin puolet heistä oli toi-
sen yrityksen palveluksessa työskenteleviä palkansaajia, ja palkkatulojen keskiarvo oli noin
30 000 euroa. Eläkkeitä ja etuuksia sai noin puolet elinkeinonharjoittajista. Osa sai tuloja
myös maa- ja metsätaloudesta. Lähes 40 000 elinkeinonharjoittajaa sai tulonsa yksinomaan
elinkeinotoiminnasta.³⁰

²⁸ Taulukossa 3 on esitetty lukumäärät perustuen asiakkaan veroilmoituksella ilmoittamiin tietoihin. Lopullisen ve-
rotuksen tietoja ei ole huomioitu.

²⁹ Tilastokeskuksen kokeellisen Tulorekisterin palkat ja palkkiot -tilaston mukaan tammikuussa 2022 palkka- ja
palkkiotulojen mediaani oli 2 968 euroa. Alin neljännes palkka- ja palkkiotuloja saaneista sai 2 084 euroa tai vä-
hemmän. Ylin neljännes palkka- ja palkkiotuloja saaneista sai 4 100 euroa tai enemmän (julkaistu 8.3.2022
www.stat.fi).

³⁰ Vuonna 2019 noin 38 000 elinkeinonharjoittajaa sai elinkeinotoiminnan lisäksi muita tuloja enintään 100 euroa
per tulolaji – eli palkat, eläkkeet ja veronalaiset etuudet tai muut pääomatulot.

Taulukko 4. Elinkeinonharjoittajien tulot, vuosi 2019.

	Lkm	Keskiarvo (€)	Mediaani (€)
Elinkeinotoiminnan tulos tai tappio			
Elinkeinotoiminnan [voitollinen] tulos	134 600	22 400	12 600
Elinkeinotoiminnan tappio	29 900	-3 800	-1 400
Muut tulot			
Palkat	115 200	30 500	27 400
Eläkkeet ja veronalaiset etuudet	97 100	11 900	7 600
Muut pääomatulot *)	59 300	6 600	1 400
Muut tulot yhteensä	178 300	28 300	21 400
Maa- ja metsätaloustoiminta			
Maataloustoimintaa**)	8 300	--	--
Metsätaloustoimintaa**)	13 100	--	--

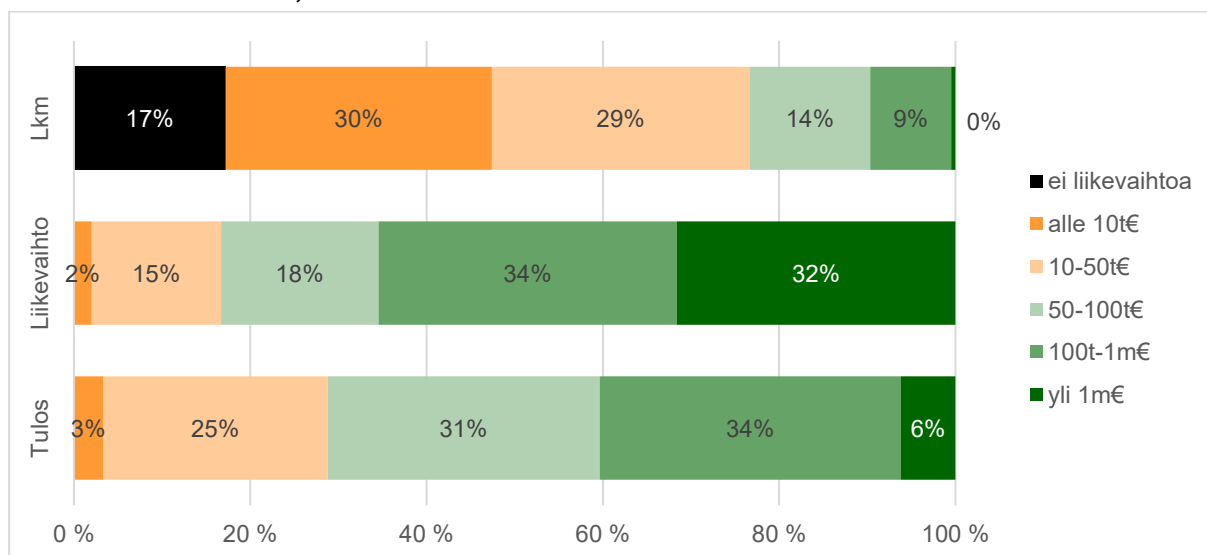
*) Huomioitu yli 100 € pääomatulot vuonna 2019.

***) Elinkeinonharjoittajat, joiden tulee antaa elinkeinotoiminnan veroilmoituksen lisäksi myös erillinen maatalouden ja/tai metsätalouden veroilmoitus.

Lähde: Verohallinto 2–5/2022.

Kuviossa 1 on esitetty yhteenveto elinkeinonharjoittajien lukumäärästä, liikevaihdosta ja tuloksesta³¹. Nämä elinkeinonharjoittajien ominaisuudet on summattu ja jaettu liikevaihdon mukaisiin kokoluokkiin. Lukumäärien tarkastelusta havaitaan, että elinkeinonharjoittajista yli kolmasosa ilmoittaa hyvin vähän tai ei lainkaan liikevaihtoa. Kun tarkastelukulmaksi vaihtuu elinkeinonharjoittajien liikevaihto tai tulos, tämä lukumääräisesti suuri joukko miltei katoaa kuvasta. Liikevaihdolla ja tuloksella mitattuna suurin ryhmä on elinkeinonharjoittajat, joiden liikevaihto on yli 50 000 euroa, mutta lukumäärästä näiden osuus on vain alle neljäsosa.

Myöhemmin analyysissä pimeän työn estimaatteja verrataan tulokseen (luku 6) ja esitetään myös estimaattien jakautuminen liikevaihdon mukaisiin kokoluokkiin (kuvio 8).

Kuvio 1. Elinkeinonharjoittajien lukumäärän, liikevaihdon ja tuloksen suhteellinen jakautuminen liikevaihtoluokkiin, vuosi 2019.

Lähde: Verohallinto 2–5/2022.

³¹ Tarkasteluun valittiin elinkeinonharjoittajien vuoden 2019 verotustietojen mukaiset elinkeinotoiminnan tulokset ilman elinkeinotoiminnan tappioita, koska voitollisista tuloksista maksetaan veroja.

3.4 Toimialat

Taulukossa 4 (seuraavalla sivulla) on esitetty elinkeinonharjoittajien 15 yleisintä toimialaa³². Yleisin toimiala oli muut henkilökohtaiset palvelut. Lähes joka kuudes elinkeinonharjoittaja toimi kyseisellä alalla. Toimialan elinkeinonharjoittajien liikevaihdot olivat usein hyvin pieniä, keskimäärin noin 25 000 euroa. Toiseksi suurin toimiala oli terveystalot³³, johon kuului joka kymmenes elinkeinonharjoittaja. Terveystaloihin liittyvä erityispiirre on se, että osa toimialan elinkeinonharjoittajista tuottaa arvonnäköverottomia terveystalot.

Kolmanneksi yleisin toimiala oli vähittäiskauppa. Huomattava toimialakohtainen erityispiirre liittyy apteekkeihin³⁴. Lääkelain säädöksistä johtuen apteekin yritysmuoto on toiminimi ja apteekkari on toiminimiyrittäjä. Apteekit tilittävät valtiolle vuosittain huomattavat verot arvonnäköverojen, tuloverojen ja erityisesti apteekiverojen muodossa. (Apteekkariliitto, 2021) Apteekkien liikevaihdot olivat huomattavan suuret verrattuna muiden elinkeinonharjoittajien liikevaihtoihin.³⁵ Apteekkien lisäksi vähittäiskaupassa oli myös suurta liikevaihtoa ilmoittavia supermarketteja³⁶. Toisaalta monien erikoismyymälöiden liikevaihdot olivat hyvin pieniä: keskimäärin vähittäiskauppaa harjoittavien elinkeinonharjoittajien liikevaihto (ilman apteekkeja ja supermarketteja) oli noin 80 000 euroa³⁷ vuodessa.

Yleisimpiin toimialoihin lukeutui myös maaliikenne, johon kuuluu henkilö- ja tavaraliikenne. Taksiliikenneluvan omaavia elinkeinonharjoittajia oli yli 8 000 eli toimialasta lähes puolet³⁸. Huomion arvoista oli myös, että ravitsemistoiminta on keskikokoinen toimiala (15. yleisin) ja alalla oli noin 4 700 elinkeinonharjoittajaa.

³² TOL2008 luokituksen 2-numerotasolla.

³³ Toimialaan terveystalot (86) kuuluvat tutkimus- ja hoitopalvelut, joita tuotetaan lyhyt- ja pitkäaikaishoitoa antavissa yleis- ja erikoissairaaloissa, sairaskodeissa, kuntoutuslaitoksissa ja muissa terveydenhuollon laitospalvelusyksiköissä. Myös yleis- ja erikoislääketieteen alojen vastaanotto ja hoito, yleis- ja erikoishammaslääkäripalvelut sekä terveydenhuollon muiden laillistettujen ammattihenkilöiden tuottamat terveystalot kuuluvat tähän toimialaan.

³⁴ Toimiala 47730 Apteekit.

³⁵ Apteekkeja on elinkeinonharjoittajissa noin 640 kpl ja niiden yhteenlaskettu liikevaihto on reilut 2 miljardia euroa vuodessa. Apteekkien liikevaihdot kattavat noin 16 % kaikkien elinkeinonharjoittajien yhteenlasketusta liikevaihdosta.

³⁶ Toimiala 47111 Isot supermarketit (yli 1000 m²).

³⁷ Toimialan keskimääräinen liikevaihto on laskettu elinkeinonharjoittajien vuoden 2019 veroilmoituksilla ilmoitetuista liikevaihtoista. Apteekkien ja supermarkettien veroilmoituksilla ilmoitetut liikevaihdot olivat keskimäärin 3,7 miljoonaa euroa vuonna 2019.

³⁸ Vuoden 2020 lopussa taksiliikenneluvan haltijoiden määrä oli 12 225, jotka olivat yritysmuodoltaan pääosin yksityisiä elinkeinonharjoittajia (67 %). Joukkoliikenne luvan haltijoista (2020 lopussa 1 063 kpl) 18 % oli elinkeinonharjoittajia ja tavaraliikenneluvan (2020 lopussa 11 634 kpl) haltijoista 25 %.

Taulukko 5. Elinkeinonharjoittajien yleisimmät toimialat vuonna 2019.

Toimiala	Lukumäärä*	Osuus, %
Kaikki toimialat	222 300	100 %
96 Muut henkilökohtaiset palvelut	27 200	12 %
86 Terveyspalvelut	25 900	12 %
47 Vähittäiskauppa	15 500	7 %
43 Erikoistunut rakennustoiminta	13 000	6 %
74 Muut erikoistuneet palvelut liike-elämälle	11 500	5 %
41 Talonrakentaminen	10 200	5 %
49 Maaliikenne	10 000	4 %
90 Kulttuuri ja viihdetoiminta	9 700	4 %
81 Kiinteistön- ja maisemanhoito	7 900	4 %
85 Koulutus	7 200	3 %
70 Pääkonttorien toiminta; liikkeenjohdon konsultointi	6 400	3 %
45 Moottoriajoneuvojen tukku- ja vähittäiskauppa sekä korjaus	6 200	3 %
71 Arkkitehti- ja insinööripalvelut; tekninen testaus ja analysointi	5 100	2 %
93 Rahapeli- ja vedonlyöntipalvelut	5 100	2 %
56 Ravitsemistoiminta	4 700	2 %

(*) Lukumäärät on pyöristetty lähimpään sataan.

Lähde: Verohallinto 2/2022.

3.5 Päätoimista ja sivutoimista elinkeinotoimintaa harjoittavat

Jako päätoimista ja sivutoimista elinkeinotoimintaa harjoittaviin on tehty tarkastellen sitä, mistä tulonlähteestä³⁹ henkilön verotettavat tulot muodostuvat, ja sen lisäksi sitä, kuinka laajaa yritystoimintaa on ollut. Määrittelyllä pyritään kuvaamaan, onko elinkeinotoiminta luonteeltaan päätoimista tai muuten laajamittaista vaiko sivutoimista tai hyvin vähäistä.

Päätoimisella elinkeinotoiminnalla tarkoitetaan tässä selvityksessä sitä, että yli puolet henkilön verotettavista ansio- ja pääomatuloista muodostuu elinkeinotoiminnan tuloksesta. Lisäksi, jos elinkeinotoiminnan liikevaihto on ollut yli 50 000 euroa, on elinkeinotoiminta katsottu päätoimiseksi muiden tulojen määrästä riippumatta. Mikäli liikevaihto on ollut alle 50 000 euroa, mutta yli puolet henkilön veronalaisista tuloista on tullut elinkeinotoiminnasta, on tämä tulkittu päätoimiseksi elinkeinotoiminnaksi.

Jos henkilöllä ei ollut lainkaan tuloja, mutta hänen odotettiin antavan elinkeinotoiminnan veroilmoitus, hänet katsottiin päätoimiseksi elinkeinonharjoittajaksi. Tällaisia täysin tulottomia elinkeinonharjoittajia oli noin 3 700 vuonna 2019. Vastaavasti, jos tuloista alle puolet oli peräisin elinkeinotoiminnasta ja liikevaihto oli alle 50 000 euroa, katsottiin elinkeinotoiminnan olevan sivutoimista.

Edellä esitettyjen määritelmien mukaisesti hieman alle puolet elinkeinonharjoittajista luokiteltiin päätoimiseksi (108 100 kpl) ja hieman yli puolet sivutoimiksi (114 200 kpl) vuonna 2019.

Taulukosta 6 näemme, kuinka päätoimiseksi ja sivutoimiseksi luokiteltujen elinkeinonharjoittajien tulorakenteet erosivat toisistaan: niillä päätoimisilla, joiden elinkeinotoiminta oli voitollista, keskimääräinen tulos oli noin 33 000 euroa vuonna 2019. Tämä oli heidän suurin yksittäinen

³⁹ Tässä yhteydessä tulonlähteellä tarkoitetaan sitä, mistä henkilön verotettavat tulot ovat peräisin: elinkeinotoiminnasta, saaduista palkoista, muista pääomatuloista.

tulonlähteensä. Sivutoimisilla voitollisen tuloksen keskiarvo oli vain 5 500 euroa, mutta keskimääräiset palkkatulot olivat lähes 37 000 euroa.

Taulukko 6. Päätoimisten ja sivutoimisten elinkeinoharjoittajien tulot, vuosi 2019.

Elinkeinoharjoittajien tulot	Päätoimiset			Sivutoimiset		
	Lkm	Keskiarvo (€)	Mediaani (€)	Lkm	Keskiarvo (€)	Mediaani (€)
Elinkeinotoiminnan tulos tai tappio						
Elinkeinotoiminnan [voitollinen] tulos	83 100	32 800	23 800	51 500	5 500	3 000
Elinkeinotoiminnan tappio *)	6 800	-7 400	-2 900	23 100	-2 600	-1 100
Muut tulot						
Palkat	26 700	10 600	3 500	88 500	36 500	33 400
Eläkkeet ja veronalaiset etuudet	38 200	7 000	3 600	58 900	15 100	11 800
Muut pääomatulot **)	24 500	6 200	1 400	34 700	6 900	1 500
Muut tulot yhteensä	65 300	10 800	4 700	113 600	38 400	32 600

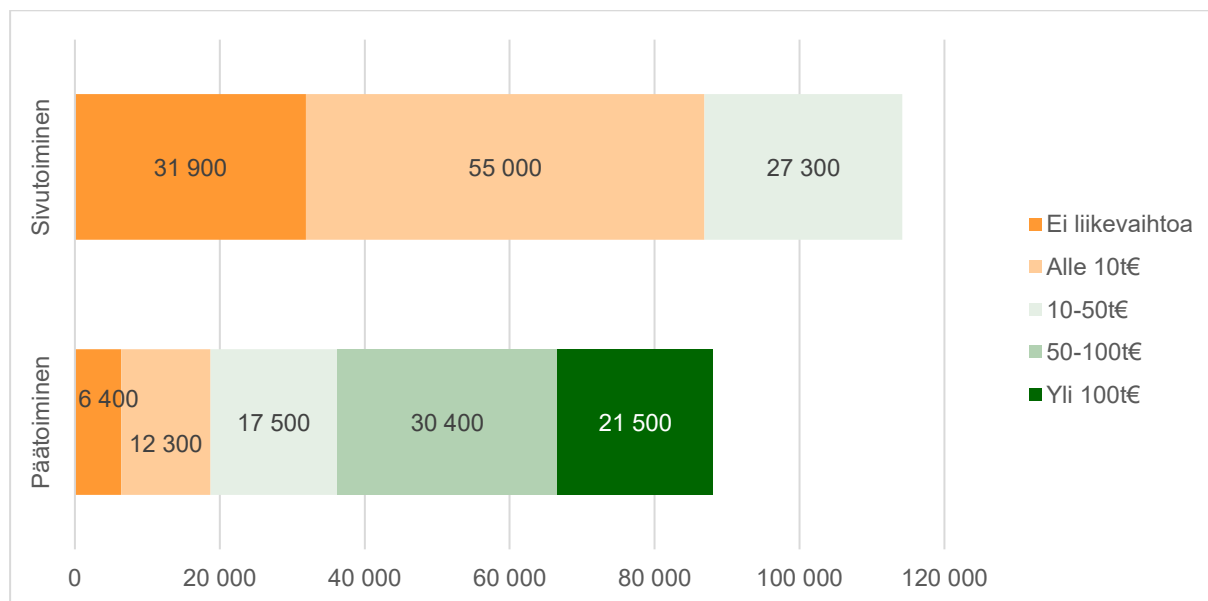
*) Jos tulos on tappiollinen, elinkeinotoiminnan tuloja ei ole. Jos liikevaihto on kuitenkin ollut yli 50 000 euroa tai jos muita tuloja ei ole ollut, niin elinkeinotoiminta on määritelty päätoimiseksi.

***) Huomioitu yli 100 € pääomatulot vuonna 2019.

Lähde: Verohallinto 2–6/2022.

Seuraavassa kuviossa 2 päätoimiset ja sivutoimiset elinkeinoharjoittajat on jaettu samoihin liikevaihto luokkiin kuin edellä kohdassa 3.1 *Rekisteröintiasema ja toiminnan laajuus*.

Kuvio 2. Päätoimisten ja sivutoimisten elinkeinoharjoittajien lukumäärät liikevaihtoluokittain vuonna 2019.



Lähde: Verohallinto 2–6/2022.

Taulukossa 7 on esitetty elinkeinoharjoittajien yleisimmät toimialat (samat toimialat kuin taulukossa 4). Näistä maaliikenteessä, ravitsemistoiminnassa ja talonrakentamisessa selvä enemmistö arvioitiin olevan päätoimisia elinkeinoharjoittajia.

Sen sijaan koulutuksessa, arkkitehti- ja insinööripalveluissa sekä liikkeenjohdon konsultoinnissa selvä enemmistö arvioitiin olevan sivutoimisia elinkeinoharjoittajia.

Taulukko 7. Päätoimisten ja sivutoimisten elinkeinonharjoittajien suurimmat toimialat vuonna 2019.

Toimiala	Yhteensä	Päätoimiset		Sivutoimiset	
		Lkm	Osuus, %	Lkm	Osuus, %
Kaikki toimialat	222 331	108 100	49 %	114 200	51 %
96 Muut henkilökohtaiset palvelut	27 166	17 187	63 %	9 802	36 %
86 Terveyspalvelut	25 863	11 950	46 %	14 071	54 %
47 Vähittäiskauppa	15 470	5 939	38 %	9 245	60 %
43 Erikoistunut rakennustoiminta	13 032	7 422	57 %	5 464	42 %
74 Muut erikoistuneet palvelut liike-elämälle	11 530	6 653	58 %	6 853	59 %
41 Talonrakentaminen	10 242	6 653	65 %	3 618	35 %
49 Maaliikenne	9 970	7 476	75 %	2 362	24 %
90 Kulttuuri ja viihdetoiminta	9 658	3 711	38 %	5 998	62 %
81 Kiinteistön- ja maisemanhoito	7 870	4 235	54 %	3 580	45 %
85 Koulutus	7 217	1 972	27 %	5 228	72 %
70 Pääkonttorien toiminta; liikkeenjohdon konsultointi	6 390	2 041	32 %	4 412	69 %
45 Moottoriajoneuvojen tukku- ja vähittäiskauppa ...	6 160	3 607	59 %	2 503	41 %
71 Arkkitehti- ja insinööripalvelut; ...	5 130	1 388	27 %	3 701	72 %
93 Rahapeli- ja vedonlyöntipalvelut	5 076	1 854	37 %	3 140	62 %
56 Ravitsemistoiminta	4 699	3 179	68 %	1 485	32 %

Lähde: Verohallinto 2–6/2022.

Jakoa päätoimiseen ja sivutoimiseen elinkeinotoimintaan on noudatettu profiloitimenetelmällä suoritettussa data-analyysissä, jonka tuloksia kuvataan luvussa 6.3.

3.6 Elinkeinonharjoittajat palkanmaksajina

Elinkeinonharjoittajista noin 9 000 eli neljä prosenttia oli merkittynä säännöllisesti palkkoja maksavien työnantajien rekisteriin. Lisäksi noin 13 000 elinkeinonharjoittajaa ilmoitti vuosittain satunnaista palkanmaksua⁴⁰. Yhteensä siis noin 10 prosenttia elinkeinonharjoittajista oli säännöllisesti tai satunnaisesti palkkoja maksavia työnantajia. Elinkeinonharjoittajat työllistivät osa- tai kokonaikaisesti Suomessa vuosittain noin 70 000–80 000 henkilöä ja ilmoittivat maksettuja palkkoja noin 700 miljoonaa euroa.

Elinkeinonharjoittajat, jotka työllistivät useimmin palkansaajia toimivat maaliikenteen, vähittäiskaupan, ravitsemistoiminnan, erikoistuneen rakennustoiminnan ja talonrakentamisen aloilla. Lisäksi huomion arvoista oli, että palkansaajia työllistävien elinkeinonharjoittajien toiminnan laajuus oli moninkertaista verrattuna niihin elinkeinonharjoittajiin, joilla ei ollut palkattuja työntekijöitä. Palkkoja maksavien elinkeinonharjoittajien liikevaihtojen keskiarvo on yli 300 000 euroa.

Tulorekisteriin ilmoitettujen palkkojen keskiarvosta voitiin päätellä, että elinkeinonharjoittajat työllistivät keskimäärin 1–2 työntekijää tai maksoivat palkkaa esimerkiksi kausittaisesti palkatulle lisätyövoimalle osan vuotta. Työnantajarekisterissä olleiden elinkeinonharjoittajien maksama mediaanipalkka oli noin 20 000 euroa vuodessa ja satunnaisesti palkkoja maksaneiden maksama mediaanipalkka vain noin 5 000 euroa vuodessa.

⁴⁰ Yritys on niin sanotusti satunnaisesti palkkoja maksava, jos sillä on palveluksessaan vain 1 vakituinen työntekijä tai sillä on palveluksessaan 1–5 työntekijää, joiden työsuhte ei kestä koko kalenterivuotta.

Taulukko 8. Työnantajarekisterissä olevat ja satunnaisesti palkkoja maksavat elinkeinonharjoittajat: elinkeinonharjoittajien lukumäärät sekä palkkojen mediaanit ja keskiarvot vuosina 2018–2020.

Vuosi	Työnantajarekisterissä*			Satunnaisesti palkkoja maksaneet**		
	Lkm	Palkkojen mediaani (€)	Palkkojen keskiarvo (€)	Lkm	palkkojen mediaani (€)	palkkojen keskiarvo (€)
2018	9 400	24 700	66 500	14 400	5 800	13 100
2019	8 800	23 800	68 400	13 500	5 500	13 400
2020	8 000	19 800	67 600	12 000	5 000	12 800

*) Säännöllisesti palkkoja maksavien työntekijöiden rekisteriin merkityt elinkeinonharjoittajat.

***) Elinkeinonharjoittajat, jotka olivat ilmoittaneet maksettuja palkkoja tulorekisteriin, mutta eivät olleet säännöllisesti palkkoja työntekijöiden rekisterissä.

Lähde: Verohallinto 2–5/2022.

3.7 Elinkeinonharjoittajan verotus

Elinkeinonharjoittajan verotettava tulo lasketaan tulolähteittäin. Tulolähteitä ovat elinkeinotoiminnan tulolähde, maataloustoiminnan tulolähde ja henkilökohtainen tulolähde eli muun toiminnan tulolähde. Verotettavaa tuloa laskettaessa tulot ja menot kohdistetaan oikeaan tulolähteeseen, myös tappiot vahvistetaan tulolähteittäin. Kullekin tulolähteelle on oma veroilmoituksensa. Elinkeinotoiminnan tulolähteen tuotot ja kulut ilmoitetaan 5-lomakkeella, ja ne verotetaan elinkeinotulon verottamisesta annetun lain mukaisesti. Elinkeinotoiminnan veroilmoitus täytyy palauttaa aina, myös vaikka toimintaa ei olisi ollut. Lisäksi jokaisen elinkeinonharjoittajan tulee vuosittain tarkistaa, täydentää ja korjata oma henkilökohtainen veroilmoitus eli esitäytetty veroilmoitus.

Kirjanpitolain mukaan luonnollinen henkilö on kirjanpitovelvollinen harjoittamastaan liike- ja ammattitoiminnasta⁴¹. Maatalouden ja kalastuksen harjoittajaa tämä velvollisuus ei koske. Kirjanpitolaissa on määritelty rajat, jolloin liikkeen- ja ammatinharjoittaja on velvollinen pitämään kahdenkertaista kirjanpitoa⁴². On paljon toimialoja, joilla toiminta on luvanvaraista. Luvanvaraista toimintaa harjoittavilla velvollisuus kahdenkertaiseen kirjanpitoon voi olla määritelty muualla laissa. Myös luvanvaraista toimintaa valvova viranomainen voi edellyttää kahdenkertaista kirjanpitoa. Hieman vajaa puolet elinkeinonharjoittajista ilmoittaa veroilmoituksellaan pitävänsä elinkeinotoiminnastaan kahdenkertaista kirjanpitoa⁴³.

Elinkeinotoiminnan veroilmoituksella ilmoitetut tuotot ja kulut muodostavat elinkeinonharjoittajan elinkeinotoiminnan tuloksen. Voitollisesta tuloksesta vähennetään aikaisempien vuosien vahvistetut tappiot ennen verovuoden tuloksen jakamista verotettavaksi yritystuloksi. Yritystulosta vähennetään myös yrittäjävähennys⁴⁴. Verotettava yritystulo jaetaan sen jälkeen

⁴¹ Kirjanpitolaki (31.12.1997/1336) 1a §.

⁴² Kirjanpitolaki 1a § 2 momentin mukaan kahdenkertaista kirjanpitoa ei tarvitse pitää, jos sekä päätyneenä, että sitä edeltäneellä tilikaudella on täyttynyt enintään yksi seuraavista edellytyksistä 1) taseen loppusumma ylittää 100 000 euroa; 2) liikevaihto tai sitä vastaava tuotto ylittää 200 000 euroa; tai 3) palveluksessa on keskimäärin yli kolme henkilöä.

⁴³ Vuonna 2019 veroilmoituksen (5-lomakkeen) antaneista elinkeinonharjoittajista noin 49 % ilmoitti pitävänsä kahdenkertaista kirjanpitoa.

⁴⁴ Verohallinto tekee 5 %:n yrittäjävähennyksen verovuoden elinkeinotoiminnan tuloksesta. Yrittäjävähennyksen saa vain elinkeinotoiminnan voitosta eikä se voi olla suurempi kuin voitto. Yrittäjävähennys ei siis kasvata verovuoden tappiota.

elinkeinonharjoittajan ansio- ja pääomatuloksi⁴⁵. Jos elinkeinonharjoittajan puoliso työskentelee yrityksen hyväksi, yrityksestä saatu tulo jaetaan puolisoitten kesken verotettavaksi. Pääomatulo-osuus jaetaan verotuksessa puolisoitten kesken sen mukaisesti, kuinka suuren osuuden kumpikin yrityksen nettovarallisuudesta omistaa. Ansiotulo-osuus jaetaan puolisoille yrityksen hyväksi tehdyn työn osuuden mukaisesti. Tieto puolison omistamasta nettovarallisuudesta ja työskentelystä ilmoitetaan vuosittain veroilmoituksella prosentteina. (Verohallinto, 2021)

Elinkeinonharjoittajien pääasiallinen verovalvonta tapahtuu vuosittain säännönmukaisen verotuksen toimittamisen yhteydessä. Kaikki veroilmoitukset valvotaan teknisesti, jonka myötä osa siirtyy virkailijäkäsittelyyn. Siinä virkailija tarkistaa asiakkaan Verohallinnolle antaman veroilmoituksen ja tarvittaessa pyytää lisätietoja. Säännönmukaisen verotuksen toimittaminen perustuu Verohallinnolle lähetettyihin tietoihin, eikä valvonta ole yhtä kokonaisvaltaista kuin esimerkiksi verotarkastuksilla. Verotarkastuksilla käytetään laajasti vertailutietoja ja tarkistetaan verovelvollisen kirjanpito ja/tai muistiinpanot.

Verovalvonnassa korjataan vuosittain paljon huolimattomuuteen tai veroilmoituksen puutteelliseen täyttämiseen perustuvia virheitä. Tällaisia virheitä ei voida laskea tämän selvityksen mukaiseksi pimeäksi työksi tai palkaksi, vaikka nekin aiheuttavat veronmenetyksiä, jos ne jäävät korjaamatta. Huomattava osa säännönmukaisesta verovalvonnasta kohdistuu elinkeinonharjoittajiin, jotka eivät ole antaneet veroilmoitusta ajallaan. Myös arvioverotukset ovat melko yleisiä toimenpiteitä säännönmukaisessa verotuksessa.

Jos elinkeinonharjoittaja jättää veroilmoituksellaan ilmoittamatta tuloja tai tulouttamatta esimerkiksi yksityiskäyttöä, tai toisaalta ilmoittaa kuluissaan aiheettomia ostoja tai tekee perusteettomia lisävähennyksiä, muodostuu elinkeinonharjoittajan verotettava tulo liian pieneksi. Tällaisessa tilanteessa elinkeinonharjoittaja saa perusteetonta etua verotuksessa. Näitä tilanteita on kuvattu tarkemmin seuraavassa luvussa 4.

⁴⁵ Pääsäännön mukaan 20 % edellisen vuoden nettovarallisuutta vastaavasta tulosta on elinkeinonharjoittajan pääomatuloa ja loput ansiotuloa. Pääsäännöstä poiketakseen elinkeinonharjoittaja voi vaatia veroilmoituksella, että vain 10 % tulosta katsotaan pääomatuloksi tai että koko tulos tulee verottaa ansiotulona.

4 Elinkeinonharjoittajan verotettavan tulon yleisimmät virhetilanteet

Tässä luvussa kuvataan elinkeinonharjoittajien yleisimpiä verotarkastuksilla havaittuja virheitä. Palkansaajille maksetut pimeät palkat kuuluvat luonnollisesti selvittävän käsitteen piiriin. Elinkeinonharjoittajan pimeällä työllä tarkoitetaan tässä selvityksessä kaikkia niitä tekoja ja toimia, joilla elinkeinonharjoittajan elinkeinotoiminnan verotettava tulo jää tosiasiallista määrää alhaisemmaksi. Seuraavissa alaluvuissa on kuvattu tunnistettuja elinkeinonharjoittajien pimeän työn ja palkan tekojen ilmenemismuotoja.

4.1 Pimeä eli puuttuva myynti

Pimeällä myynnillä tarkoitetaan tässä selvityksessä sellaisia elinkeinotoiminnasta saatuja tuloja, joita ei ilmoiteta elinkeinotoiminnan veroilmoituksella tuottoina. Pimeän työn käsitteeseen ei tässä selvityksessä katsota niitä tuloja, jotka elinkeinonharjoittaja ilmoittaa elinkeinotoiminnan sijaan esimerkiksi henkilökohtaisella tai maatalouden veroilmoituksella. Tuloja ei ole yritetty salata, jos ne on ilmoitettu väärällä ilmoituksella. Pimeää eli puuttuvaa myyntiä on perinteisesti myyntitulot, jotka jätetään kokonaan kirjanpidon ulkopuolelle ja ilmoittamatta veroilmoituksella.

Elinkeinotoiminnan salattuja myyntituloja havaitaan usein verotarkastuksilla. Säännönmukaisessa verotuksessa elinkeinonharjoittajien verotettavan tulon lisäykset aiheutuvat usein veroilmoituksen ja oma-aloitteisten verojen ilmoituksen välisestä eroavuudesta sekä veroilmoituksen ja sivullisilmoittajilta saatujen tietojen eroista. Huomattava osa tulon lisäyksistä aiheutuu siitä, ettei veroilmoitusta anneta lainkaan ja verotus toimitetaan arvioimalla⁴⁶.

Jos elinkeinonharjoittaja jättää ilmoittamatta tuloja veroilmoituksella, muodostuu hänen verotettavat tulot⁴⁷ liian pieniksi ja tuloverotuksessa veroja tulee maksettavaksi liian vähän. Usein salatut myyntitulot jäävät elinkeinonharjoittajan omaan käyttöön⁴⁸, ja verot tuloista jäävät maksamatta. Pimeä myynti eli myyntitulojen salaaminen on elinkeinonharjoittajan pimeän työn ilmenemismuodoista usein se, joka aiheuttaa merkittävää verojen⁴⁹ vajetta.

4.2 Perusteettomat kulut

Elinkeinotoiminnan tuloista vähennyskelpoisia⁵⁰ ovat elinkeinotoiminnan tulojen hankkimisesta tai säilyttämisestä johtuneet menot ja menetykset. Perusteettomilla kuluilla tarkoitetaan tässä selvityksessä sellaisia kuluja, jotka ovat pienentämässä elinkeinotoiminnan tulosta,

⁴⁶ Verotusmenettelystä annetun (18.12.1995/1558) 27 § 1 momentin mukaan verotus on toimitettava arvioimalla, jos veroilmoitusta ei ole annettu tai sitä ei voida panna oikaistunakaan verotuksen perusteeksi.

⁴⁷ Ansio- ja pääomatulot.

⁴⁸ Salatuilla myyntituloilla saatetaan maksaa myös pimeitä palkkoja. Jos elinkeinonharjoittaja työllistää vain itsensä, salatut myyntitulot jäävät lähtökohtaisesti elinkeinonharjoittajan käyttöön.

⁴⁹ Elinkeinonharjoittajan tuloverotuksen ansio- ja pääomatuloista maksettavat verot. Lisäksi jos myynnit jätetään ilmoittamatta myös arvonlisäveroa koskevilla ilmoituksilla, jää myös arvonlisäveroja maksatta. Arvonlisäverojen määrää ei tässä selvityksessä arvioida.

⁵⁰ EVL 7 §, Laki elinkeinotulon verottamisesta (24.6.1968/360).

vaikka ne eivät kohdistu elinkeinotoiminnan tulon hankkimiseen. Tällaisia kuluja ovat esimerkiksi yksityistalouden menot. Esimerkiksi väärässä tulolähteessä⁵¹ ilmoitettuja kuluja ei tässä selvityksessä pidetä perusteettomina kuluina, jos ne ovat aidosti vähennyskelpoisia toisessa tulolähteessä.

Jos elinkeinotoiminnan tulosta pienentämässä on esimerkiksi yksityistalouden menoja tai tekaistuja kulutositteita, jää elinkeinotoiminnan tulos liian pieneksi. Liian pienen tuloksen vuoksi tuloveroja jää kertymättä, joten siksi perusteettomat kulut on katsottu kuuluvan selvityksen pimeän työn piiriin.

4.3 Yksityiskäytön tulouttaminen veroilmoituksella

Elinkeinonharjoittajan tulee tehdä tuloutus yksityiskäytöstä veroilmoituksella, jos elinkeinotoiminnan varallisuutta⁵² käytetään omiin tarpeisiin. Usein yksityiskäytön tuloutusta tulee ilmoittaa esimerkiksi auton yksityiskäytöstä. Jos ajoneuvon kulut sisältyvät elinkeinotoiminnan kuluihin eikä auton yksityiskäytön tuloutusta tehdä, jää yksityiskulutuksen menoja rasittamaan elinkeinotoiminnan tulosta.

Yksityiskäytön tuloutuksia tulee tehdä myös silloin, kun elinkeinonharjoittaja ottaa omaan käyttöön tavaroita tai palveluita elinkeinotoiminnastaan. Elinkeinonharjoittajan veroilmoituksella on erilliset kohdat yksityiskäyttöjen tuloutuksille: auton käytöstä, tavaroiden yksityiskäytöstä⁵³ sekä muusta yksityiskäytöstä⁵⁴ (Verohallinto, 2021). Mikäli elinkeinonharjoittaja ei ilmoita yksityiskäytön tuloutuksia oikaisuna veroilmoituksellaan, jää elinkeinotoiminnan tulosta rasittamaan yksityistalouden kuluja ja siten myös tuloverotuksessa veroja kertymättä liian pienenä esitetyn tuloksen vuoksi.

Aineiston tarkastelun yhteydessä havaittiin, että noin 14 000 elinkeinonharjoittajaa eli 7,5 prosenttia ilmoitti vuodelta 2019 annetulla veroilmoituksella yksityiskäytön tuloutuksia. Ajoneuvoneuvon käytön tuloutuksia oli näistä puolet. Tietyille toimialoille⁵⁵ on ominaista, että elinkeinotoiminnassa tarvitaan autoa. Jos elinkeinotoiminnan kuluissa on ajoneuvon kaikki kulut, tulee veroilmoituksella tulouttaa yksityiskäytön osuus (näitä ovat esimerkiksi matkat työpaikalle, harrastuksiin, ruokakauppaan, lomamatkoille). Jos puolestaan elinkeinotoiminnassa käytetään yksityistalouden autoa ja auton kuluja ei ole vähennetty elinkeinotoiminnan

⁵¹ Tuloverotuksessa on kolme tulolähdettä: elinkeinotoiminnan tulolähde, maatalouden tulolähde sekä henkilökohtainen tulolähde, josta käytetään myös nimitystä muun toiminnan tulolähde.

⁵² Esimerkiksi ajoneuvo kuuluu elinkeinotoiminnan kalustoon, kun elinkeinotoiminnan ajojen määrä on yli puolet verovuonna ajettujen kilometrien kokonaismäärästä. Ajoneuvon kulut voivat olla vähennetty elinkeinotoiminnan kuluina ja yksityiskäytön osuus tuloutetaan veroilmoituksella.

⁵³ Yksityistalouden käyttöön otetuista tavaroista on tuloutettava alkuperäisten hankintamenojen määrä ilman arvonlisäveroa.

⁵⁴ Elinkeinotoimintaan kuuluvan puhelimen, kiinteistön, loma-asunnon tai veneen käytöstä. Jos elinkeinonharjoittaja käyttää elinkeinotoimintaan kuuluvaa kiinteistöä yksityistarkoituksiin, tuloutetaan veroilmoituksella yksityiskäyttöä vastaava osuus sekä kuluista että kiinteistön poistoista.

⁵⁵ Esimerkiksi henkilökuljetus, tukkukäynnit ravitsemusosalalla, tarvikkeiden ja työkalujen kuljetukset rakennusala jne.

kuluina, on elinkeinonharjoittajalla mahdollisuus tehdä autonkuluista lisävähennys veroilmoituksella⁵⁶.

Yksityiskäytön tuloutuspuutteet kuuluvat selvityksessä tarkoitetun pimeän työn piiriin, koska ne vaikuttavat elinkeinotoiminnan tulokseen ja siten elinkeinonharjoittajan tuloverojen määrään.

4.4 Lisävähennykset ja muut kulujen oikaisut veroilmoituksella

Elinkeinonharjoittaja voi tehdä veroilmoituksella lisävähennyksiä, jos yksityistalouden omaisuutta⁵⁷ käytetään elinkeinotoiminnassa. Lisäksi elinkeinonharjoittaja voi vähentää tilapäisistä työmatkoista aiheutuneita kuluja veroilmoituksella⁵⁸. Jos vähennyksiä ilmoitetaan todellisuutta suurempana tai aiheuttomasti, saa yrittäjä perusteetonta etua tuloverotuksessa liian pienen tuloksen muodossa ja veroja jää kertymättä. Perusteettomat lisävähennykset ja muut kulujen oikaistu kuuluvat selvityksessä pimeään työn piiriin.

Aineiston tarkastelun yhteydessä havaittiin, että noin 51 000 elinkeinonharjoittajaa eli 27 prosenttia ilmoitti vuodelta 2019 annetulla veroilmoituksella lisävähennyksiä. Kaikista lisävähennyksistä ajoneuvojen vähennyksiä oli noin kaksi kolmasosaa.

Tietyille toimialoille on ominaista, että elinkeinotoiminnassa tarvitaan autoa. Jos auto on elinkeinotoiminnassa välttämätön, eikä veroilmoituksella ilmoiteta ajoneuvon yksityiskäytön tuloutusta tai lisävähennystä ajoneuvon käytöstä, tarkoittaa se käytännössä, että elinkeinonharjoittajalla tulisi olla toinen auto yksinomaan elinkeinotoiminnan ajoihin ja toinen auto yksinomaan yksityisajoihin (myös matkat työpaikalle ovat yksityisajoja). Ajoneuvon yksityiskäytön tuloutuksia ja ajoneuvojen käytön lisävähennyksiä ilmoitetaan veroilmoituksilla melko harvoin.

4.5 Työntekijälle maksettu pimeä palkka

Elinkeinonharjoittaja voi maksaa työntekijöilleen palkkaa, kuten muissakin yritysmuodoissa. Tässä selvityksessä pimeällä palkalla tarkoitetaan elinkeinonharjoittajan työntekijöille maksamia palkkoja, joista ennakonpidätykset ja työnantajavelvoitteet jätetään ilmoittamatta ja siten myös maksamatta. Pimeän palkan määritelmään sisältyy lisäksi sellaiset työntekijöille maksetut verovapaat kustannusten korvaukset⁵⁹, joissa verovapauden edellytykset eivät täyty.

⁵⁶ Elinkeinonharjoittajan veroilmoituksen kohta 2.3 Lisävähennykset (kuvaus lisävähennyksistä raportin kohdassa 4.4.)

⁵⁷ Lisävähennyksen voi tehdä yksityistalouden autosta, kiinteistöstä ja huoneistosta, jota käytetään osittain elinkeinotoiminnan käyttöön.

⁵⁸ EVL 55§, Laki elinkeinotulon verottamisesta (24.6.1968/360). Elinkeinonharjoittaja voi vähentää tilapäisistä työmatkoista aiheutuvia kuluja. Lisävähennysten enimmäismäärä on Verohallinnon verovapaista matkakustannusten korvauksista antaman päätöksen mukaisten päivärahajen ja kilometrikorvausten enimmäismäärien ja todellisten kirjanpitoon tai muistiinpanoihin kirjattujen kustannusten erotus.

⁵⁹ Esimerkiksi päivärahat ja kilometrikorvaukset. Verohallinto vahvistaa vuosittain verosta vapaaksi katsottavien matkakustannusten korvausten perusteet ja määrät. Matkakustannuksia ovat verovelvolliselle Suomessa tai ulkomailla tehdystä työmatkasta aiheutuneet kustannukset.

Myös palkansaajien tuloiksi ilmoittamattomat tai aliarvostetut luontoisedut⁶⁰ kuuluvat pimeän palkan määritelmään.

Palkkoja maksava elinkeinonharjoittaja hakeutuu palkkoja maksavien työnantajien rekisteriin ja ilmoittaa maksamansa palkat (ja luontoisedut sekä kustannustenkorvaukset) tulorekisteriin. Lisäksi palkanmaksaja pidättää työntekijälle maksettavasta palkasta ennakonpidätyksen ja työntekijän osuuden TyEL-maksuista ja tilittää palkkaan liittyvät ennakonpidätykset sekä muut lakisääteiset työnantaja velvoitteet valtiolle. Mikäli palkkoja tai rahanarvoisia etuuksia ei ilmoiteta asianmukaisesti, jää veroja ja lakisääteisiä maksuja kertymättä.

⁶⁰ Luontoisedulla tarkoitetaan työnantajan muuna kuin rahana työstä suorittamaa vastiketta palkansaajalle. Luontoisetu on työnantajan järjestämä ja kustantama hyödyke (tavara tai palvelu), jonka työnantaja luovuttaa palkansaajan käyttöön. Tuloverolain 64 §:n 1 momentin mukaan työnantajalta saatu luontoisetu on veronalaista ansiotuloa ja se arvioidaan käypään arvoon. Verohallinto määrää vuosittain luontoisetujen käypien arvojen laskentaperusteet.

5 Data-analyysissä käytetyt verovalvonnan aineistot

Elinkeinonharjoittajien pimeän työn muotoja esiteltiin edellä luvussa 4. Käytännössä havain- toja pimeästä työstä saadaan kahdesta eri verovalvonnan prosessista, verotarkastuksista ja säännönmukaisesta verovalvonnasta. Säännönmukainen valvonta kattaa kaikki elinkeinon- harjoittajat, joista osa valvotaan tarkemmin. Verotarkastuksia tehdään melko vähän, mutta niissä perehdytään syvällisesti valvottavan yrityksen toimintaan.

Valvontaprosessit tuottavat aineistoja, joita selvityksen data-analyyseissä hyödynnettiin. Kes- keisessä roolissa olivat erityisesti verotarkastusten tuottamat aineistot, mutta yksi menetel- mistä hyödynsi myös säännönmukaisen verovalvonnan tuottamia havaintoja.

5.1 Elinkeinonharjoittajiin kohdistuneet verotarkastukset

Kaikkiaan 614 elinkeinonharjoittajaa on verotarkastettu vuosina 2018–2020⁶¹, mikä tarkoit- taa, että tarkastuskattavuus oli noin 0,24 prosenttia kyseisellä ajanjaksolla. Esimerkiksi osa- keyhtiöitä tarkastettiin selvästi enemmän, ja niissä tarkastuskattavuus on suurempi.⁶²

Kaikki elinkeinonharjoittajien verotarkastukset olivat riskiperusteisia, toisin sanoen satunnais- otantaan perustuvia verotarkastuksia ei ollut käytettävissä tätä selvitystä varten.⁶³ Riskiperus- teisista tarkastuksista saatu aineisto on vinoutunut, esimerkiksi verotarkastuksia on koh- dennettu enemmän suuriin elinkeinonharjoittajiin kuin pieniin. Verotarkastetuista noin 40 pro- sentilla liikevaihto oli yli 100 000 euroa, kun koko perusjoukosta vain noin 10 prosentilla liike- vaihto ylittää tämän summan. Lisäksi kolmannes verotarkastetuista ja koko perusjoukosta vain noin yksi prosentti kuului työnantajarekisteriin. Luvussa 7 on käsitelty lähemmin vinoutu- neen aineiston vaikutusta data-analyysin ja mitä ongelmien torjumiseksi on tehty.

Elinkeinonharjoittajien verotarkastukset keskittyivät melko harvoille toimialoille. Nyt tarkastel- tavalla jaksolla Verohallinnolla on ollut käynnissä julkisuudessakin esitellä olleita valvonta- hankkeita muun muassa taksi-, ravintola-, kampaamo- ja kauneudenhoitoalojen yrityksiin. Taulukossa 9 on esitetty tarkastusten lukumääriä sekä prosentiosuuksia siitä, kuinka monta elinkeinonharjoittajaa eri toimialoilta tarkastettiin vuosina 2018–2020. Eniten tarkastettiin ra- vitsemistoiminnan, maaliikenteen ja vähittäiskaupan elinkeinonharjoittajia.

⁶¹ Verotarkastus kohdistui verovuosiin 2018–2020.

⁶² Selvityksessä *Pimeän työn määrä – Mittaamisselvitys osakeyhtiöiden verotarkastustietojen perusteella* käytet- tiin aineistona osakeyhtiöiden EPL ja EVL-verolajeihin kohdistettuja verotarkastuksia. Tarkastuskattavuus oli noin viisinkertainen verrattuna elinkeinonharjoittajiin eli 1,24 prosenttia osakeyhtiöiden perusjoukosta vuosina 2014– 2018.

⁶³ Vuosia 2018–2020 koskevia satunnaisesti valittuja verotarkastuksia tehtiin kaiken kaikkiaan hyvin vähän.

Taulukko 9. Toimialat, joissa elinkeinonharjoittajia on verotarkastettu eniten. Tarkastettujen osuus kaikista toimialan elinkeinonharjoittajista vuosina 2018–2020.

Toimiala	Tarkastetut, lkm	Tarkastettujen osuus perusjoukosta, %
56 Ravitsemistoiminta	132	2,2 %
49 Maaliikenne	108	0,9 %
47 Vähittäiskauppa	101	0,5 %
45 Moottoriajoneuvojen kauppa	60	0,8 %
96 Muut henkilökohtaiset palvelut	29	0,1 %
43 Erikoistunut rakentamistoiminta	24	0,2 %
41 Talonrakentaminen	19	0,2 %
46 Tukkukauppa	18	0,4 %
62 Ohjelmistot, konsultointi ja siihen liittyvä toiminta	11	0,1 %
74 Muut erikoistuneet palvelut liike-elämälle	10	0,2 %

Lähde: Verohallinto 2/2022.

5.2 Pimeää työtä koskevat tarkastushavainnot

Selvityksen määritelmään mukaiseen pimeään työhön viittaavia virheitä havaittiin noin 55 prosentissa elinkeinonharjoittajien verotarkastuksista. Näitä on eritelty lähemmin taulukossa 10. Pääsääntöisesti havainnot liittyivät kirjanpidosta puuttuvaan myyntiin tai muuhun veronalaiseen tuloon. Perusteettomia vähennyksiä ja työntekijöille maksettuja pimeitä palkkoja havaittiin melko vähän. Samalla elinkeinonharjoittajalla on voinut olla samanaikaisesti useita eri väärinilmoittamisen muotoja.

Taulukko 10. Verotarkastuksissa havaittujen virheiden lukumäärät vuosina 2018–2020.

Vuosi	Verotarkastukset, lkm	Havaitut virheet, lkm	Havaitut virheet, %	Puuttuva myynti, lkm	Perusteettomat vähennykset, lkm	Palkansaajien** pimeät palkat, lkm
Yht.*	1 107	604	55 %	516	114	69
2018	434	258	59 %	226	31	40
2019	428	229	54 %	190	50	23
2020	245	117	48 %	100	33	6

*) Yhteensä rivillä on laskettu tehdyt tarkastukset ja havaitut virheet per y-tunnus ja vuosi. (Esimerkiksi, jos y-tunnuksen tarkastus kattaa kolme vuotta, kasvaa verotarkastusten lukumäärä 3:lla. Jos siinä havaitaan puuttuvaa myyntiä kahtena vuonna, kasvaa puuttuvan myynnin summa 2:lla.)

***) Työntekijälle maksettu pimeä palkka, joka on Suomessa yleisesti tai rajoitetusti verovelvollinen.

Lähde: Verohallinto 2/2022.

Verotarkastuksilla todetut virheet olivat huomattavan suuria, kun niitä verrattiin kaikkien elinkeinonharjoittajien keskimääräiseen liikevaihtoon. Tämä luonnollisesti johtuu myös siitä, että verotarkastettujen elinkeinonharjoittajien toiminta oli usein keskimääräistä laajamittaisempaa, jolloin myös havaitut virheet ovat voineet olla suurempia. Puuttuvan myynnin virheet voivat olla liikevaihtoon nähden suuria, jos myyntejä ei ole ilmoitettu kuin hyvin vähän tai ei lainkaan.

Taulukko 11. Verotarkastuksilla havaittujen virheiden suuruudet vuosina 2018–2020.

Verotarkastuksella havaittu virhe	Havaitut virheet, lkm	Havaitun virheen mediaani, €	Havaitun virheen keskiarvo, €
Puuttuva myynti	516	25 000	44 000
Perusteettomat vähennykset	114	12 000	37 000
Palkansaajien pimeät palkat	69	24 000	50 000

Lähde: Verohallinto 2/2022.

5.3 Säännönmukaisen valvonnan aineistot

Elinkeinonharjoittajien verovalvontaa tehdään myös niin sanotussa säännönmukaisessa valvonnassa, joka tapahtuu vuosikohtaisesti ja on edellytys asiakkaan verotuksen toimittamiseksi.⁶⁴ Säännönmukainen valvonta on lukumääräisesti laajamittaista, mutta myös paljon kevyempi toimenpide kuin verotarkastus, ja siinä tehdään paljon asiakkaan huolimattomuudesta ja osaamattomuudesta johtuvia korjauksia. Säännönmukaisen valvonnan ensisijainen tarkoitus on varmistaa oikeasisältöinen verotus. Säännönmukainen valvonta voidaan jakaa kahteen osaan, joista toinen on riskiperusteinen valvonta ja toinen laadunvarmistus. Jälkimmäisessä korjataan ilmiselviä virheitä, jotka eivät johdu tietoisesta pyrkimyksestä verojen välttämiseen.

Datan näkökulmasta oli mahdotonta päätellä luotettavasti isolle joukolle, johtuiko verotukseen tehty korjaus aidosti harmaan talouden toiminnasta vai oliko verotusta korjattu esimerkiksi vain huolimattomuusvirheen takia. Havaittujen virheiden sijaan osa tämän selvityksen analyysimenetelmistä hyödynsi säännönmukaisessa valvonnassa moitteettomiksi todettuja tapauksia. Tällä tarkoitetaan niitä elinkeinonharjoittajia, joita oli säännönmukaisessa verotuksessa valvottu tarkemmin ja pyydetty lisäselvitystä. Lisäselvityksen jälkeen elinkeinonharjoittajan oli todettu toimineen oikein. Näissä tapauksissa oli olemassa melko suuri varmuus siitä, että kyseiset elinkeinonharjoittajat olivat toimineet verosäännösten mukaisesti.

Riskiperusteisesti valvottuja elinkeinonharjoittajia, joille oli lähetetty selvityspyyntö tai jotka oli arvioverotettu, oli vuosina 2018–2020 noin 3 500–5 000. Valvonnassa havaittiin virheitä noin 18–25 prosentilla. Tässä selvityksessä hyödynnettiin sitä osaa noin 75–80 % riskiperusteisesti valvotuista, joilta ei ollut löydetty tietojen väärin ilmoittamiseen viittaavaa asiakkaan kuulemisen jälkeen.

⁶⁴ Säännönmukaisessa valvonnassa jokaiseen elinkeinonharjoittajaan kohdistetaan ensiksi koneellinen tarkistus. Elinkeinonharjoittajat, joiden tiedoissa on koneellisen tarkastuksen perusteella epäselvyyksiä, nostetaan virkailijoiden työjonoille tarkastettavaksi. Tarvittaessa virkailija pyytää asiakkaalta täydentäviä tietoja ja lisäselvityksiä.

6 Elinkeinonharjoittajien pimeän työn määrä eri menetelmillä arvioituna

Tässä luvussa esitellään elinkeinonharjoittajien pimeän työn kokonaismäärää arvioivia tuloksia sekä käytettyjä data-analyysimenetelmiä. Kokonaisarvio laskettiin kolmella eri menetelmällä, jotka ovat CEM-kaltaistaminen, asiakaskohtaista riskisyyttä ennustava koneoppimismenetelmä sekä Harmaan talouden selvitysyksikössä kehitetty profilointimenetelmä. Näistä CEM-kaltaistaminen sekä profilointimenetelmä hyödynsivät pelkästään verotarkastusten tuottamia tietoja elinkeinonharjoittajien pimeästä työstä, mutta koneoppimismenetelmässä täydennettiin verotarkastusaineistoja säännönmukaisen valvonnan tiedoilla.

Kokonaisarvio elinkeinonharjoittajien pimeän työn määrästä on estimaatti, joka sisältää puuttuvan myynnin, perusteettomien vähennysten ja pimeiden palkkojen euromääräisen arvon. CEM-kaltaistaminen ja koneoppimismenetelmä tuottivat kokonaisestimaatteja, jotka sisältävät kaikki edellä mainitut pimeän työn muodot. Profilointimenetelmällä laskettiin erikseen kolme eri estimaattia, jotka olivat päätoimisen elinkeinotoiminnan puuttuva myynti ja perusteeton vähennys, päätoimisen elinkeinotoiminnan pimeät palkat sekä sivutoimisen elinkeinotoiminnan puuttuva myynti ja perusteeton vähennys.

Tässä selvityksessä elinkeinonharjoittajien pimeän työn määräksi on arvioitu vuosittain noin 250–450 miljoonaa euroa ajalla 2018–2020. Suurin kokonaisestimaatti oli lähes kaksinkertainen verrattuna pienimpään estimaattiin. Tuloksen vaihteluväliä voitiin kuitenkin pitää siedettävänä, etenkin kun huomioidaan käytetyn aineiston epävarmuudet (luku 7.1.2).

Kun katsotaan tuloksia tarkemmalla jaolla, eri menetelmillä päädyttiin hieman erilaisiin päätelmiin elinkeinonharjoittajien pimeän työn jakautumisesta alajoukkoihin, esimerkiksi eri toimialoille.

Vaikka selvityksessä esitetään vuosittaisia arvioita elinkeinonharjoittajien pimeän työn määrästä, ei tehtyjen analyysien tarkoituksena varsinaisesti ollut arvioida sen trendiä. Tarkoituksena ei siis ollut ollut arvioida, mihin suuntaan elinkeinonharjoittajien pimeä työ Suomessa on kehittynyt. Sen sijaan pyrkimyksenä oli saada arvio pimeän työn suuruusluokasta.

6.1 CEM-kaltaistaminen

Coarsened Exact Matching eli CEM on yksinkertainen ja tehokas kaltaistamisen menetelmä, jolla voidaan tasoittaa verotarkastusaineiston valikointiharhaa. CEM-menetelmässä jokaiselle tarkastetulle elinkeinonharjoittajalle pyrittiin löytämään vastinpari tai pareja tarkastamattomien elinkeinonharjoittajien joukosta, jotka vastasivat valittujen ominaisuuksien⁶⁵ suhteen

⁶⁵ Kaltaistamiseen käytettävät muuttujat tuotettiin pääkomponenttianalyysillä pienentämällä aineiston dimensioonalisuus kolmeentoista. Tällä tavalla pyrittiin hyödyntämään kaikki numeeriset aineiston muuttujat ja samalla vältettiin tekemästä subjektiivisia valintoja siitä, mitä muuttujia kaltaistamiseen tulisi käyttää. Kategorisia muuttujia ei kaltaistuksessa huomioitu. Dimensioiden määrä 13 valittiin sillä perusteella, että 13 ensimmäistä pääkomponenttia selittivät yli 50 % aineiston varianssista.

täysin tarkastettua elinkeinonharjoittajaa.⁶⁶ Verotarkastettujen elinkeinonharjoittajien tarkastusten tulokset kopioitiin heille löydetyille vastinpareille sellaisenaan. Ilman paria jääneiden tarkastamattomien elinkeinonharjoittajien pimeän työn määräksi tuli nolla euroa. Menetelmän kannalta ei ollut merkityksellistä, kuinka iso prosentti tarkastetuista yrityksistä oli harmaan talouden toimijoita, joten menetelmässä ei hyödynnetty säännönmukaisen valvonnan havainnottoja.

Menetelmän avulla laskettiin elinkeinonharjoittajien pimeän työn määrän olleen vuosien 2018–2020 valvontatietojen mukaan noin 250–400 miljoonaa euroa vuodessa (taulukko 12). Tämä vastasi noin 8–12 prosenttia elinkeinonharjoittajien vuosittain verotuksessa vahvistetuista elinkeinotoiminnan tulosten yhteismäärästä⁶⁷.

Taulukko 12. Arvio elinkeinonharjoittajien pimeästä työstä CEM-menetelmällä 2018–2020 (miljoonaa euroa).

Vuosi	Estimaatti	Osuus tuloksesta
2018	242	7,6 %
2019	351	10,6 %
2020	384	12,1 %

Lähde: Verohallinto 2–5/2022.

6.2 Asiakaskohtainen odottama

6.2.1 Menetelmä ja aineistot

Asiakaskohtaisella odottamalla tarkoitetaan kullekin elinkeinonharjoittajalle algoritmilähtöisesti estimoitua pimeän työn euromääräistä odotusarvoa. Menetelmä arvioi kunkin toiminimen riskisyyttä verotarkastustietojen ja muiden verotustietojen avulla.

Niin sanottu harjoitusaineisto muodostettiin kaikista verotarkastetuista sekä säännönmukaisessa valvonnassa moitteettomiksi todetuista elinkeinonharjoittajista. Harjoitusaineistossa valvonnan lopputulos luokiteltiin kahteen luokkaan "valkoisiin" (moitteettomat) ja "harmaisiin" (ne, joilta oli löytynyt virheitä). Tähän aineistoon opetettiin valvonnan lopputulosta ennustava algoritmi; se tapahtui toiminimen erilaisia ominaisuuksia kuten liikevaihtoa hyödyntämällä. Kun tarkastettujen toiminimien joukossa oli opittu suhde ennustavien muuttujien sekä tarkastuksen lopputuloksen välille, opitulla mallilla voitiin päätellä pimeän työn esiintyvyyttä myös tarkastamattomien toiminimien joukossa, tiettyjen reunaehtojen vallitessa. Jokaiselle elinkeinonharjoittajalle pääteltiin kaksi arvoa kahdella eri mallilla: kuinka todennäköisesti elinkeinonharjoittaja tekee pimeää työtä sekä pimeän työn euromääräinen suuruus, jos riski toteutuu. Kertomalla asiakaskohtainen pimeän työn todennäköisyys pimeän työn euromäärällä, voitiin jokaiselle elinkeinonharjoittajalle päätellä pimeän työn odottama.

⁶⁶ Jokainen valittu numeerinen ominaisuus jaetaan vaihtelevaan määrään kategorisia luokkia sen saaman arvon mukaan. Esimerkiksi ominaisuus liikevaihto jaetaan kuuteen luokkaan suuruuden mukaan. Tarkastamaton yritys on vastinpari tarkastetulle, jos sen jokainen valittu ominaisuus osuu samaan luokkaan kuin tarkastetun yrityksen ominaisuudet.

⁶⁷ Pimeän työn määrää verrattiin elinkeinonharjoittajien elinkeinotoiminnan tuloksien summaan, koska niiden perustella määräytyy myös elinkeinonharjoittajien verotettava tulo. Tappiollisia tuloksia ei otettu tähän vertailuun mukaan, koska tappiollisesta tuloksesta ei muodostu verotettavaa tuloa.

Esimerkiksi, jos asiakkaalle X arvioitiin 30 prosentin todennäköisyys pimeiden palkkojen maksamiseen ja mallin arvioima pimeän työn määrä oli 20 000 euroa, tuli asiakkaan odottamaksi tällöin 6 000 euroa. Kun jokaiselle yritykselle oli päätelty odottama, voitiin muodostaa arvio pimeiden palkkojen kokonaismäärästä kunakin vuonna.

Pimeän työn todennäköisyyttä ennustettiin niin kutsutulla ensemble⁶⁸-mallilla, joka koostuu joukosta neuroverkko- ja XGBoost-malleja. Lopullinen ennuste saatiin yhdistelemällä usean eri mallin tekemä ennuste. Sen avulla voitiin kohtuullisella tarkkuudella ennustaa, kuinka todennäköisesti elinkeinonharjoittaja teki pimeää työtä, kun tunnettiin toiminimen ominaispiirteet. Pimeän työn kokoa arvioiva regressiomalli rakennettiin XGBoost-algoritmilla.

6.2.2 Pimeän työn määrä

Koneoppimismallilla laskettuna elinkeinonharjoittajien pimeän työn määräksi arvioitiin vuosittain noin 400–450 miljoonaa euroa, mikä vastasi noin 13–14 prosenttia elinkeinonharjoittajien vuosittain verotuksessa vahvistetuista elinkeinotoiminnan tulosten yhteismäärästä.

Taulukko 13. Arvio elinkeinonharjoittajien pimeästä työstä koneoppimismenetelmällä 2018–2020 (miljoonaa euroa).

Vuosi	Estimaatti	Osuus tuloksesta
2018	451	14,1 %
2019	440	13,3 %
2020	407	12,8 %

Lähde: Verohallinto 2–5/2022.

6.2.3 Asiakaskohtaisen odottaman tarkempia tuloksia

Koska asiakaskohtaisen odottaman laskennassa nimenmukaisesti pääteltiin kullekin elinkeinonharjoittajalle henkilökohtaisen pimeän työn odottama⁶⁹, voitiin tuloksia tarkastella myös kaltaistusmenetelmiä tarkemmalla tasolla, esimerkiksi toimialoittain tai liikevaihdon mukaan. Alla avataan hieman menetelmän perusteella tehtyjä tarkempia arvioita pimeän työn esiintymisestä elinkeinonharjoittajien joukossa vuonna 2019.

Kuviossa 3 elinkeinonharjoittajien pimeän työn odottamat on summattu liikevaihtoluokittain – toisin sanoen kuva esittää kokonaisestimaatin jakautumisen liikevaihtoluokkiin vuonna 2019. Pimeän työn odottama ei kuvaa pelkästään pimeän työn todennäköisyyttä, vaan se huomioi myös odotetun virheen koon. Jos pimeä työ on yleistä⁷⁰ jossain luokassa, niin se kasvattaa pylvään kokoa. Toisaalta, jos ennustetut virheet ovat suuria, niin myös se kasvattaa pylvään kokoa.

Mallin mukaan eniten pimeää työtä tehtiin elinkeinonharjoittajien joukossa, joiden liikevaihto on 10 000–50 000 euroa. Tähän liikevaihtoluokkaan kuuluvat elinkeinonharjoittajat tekivät

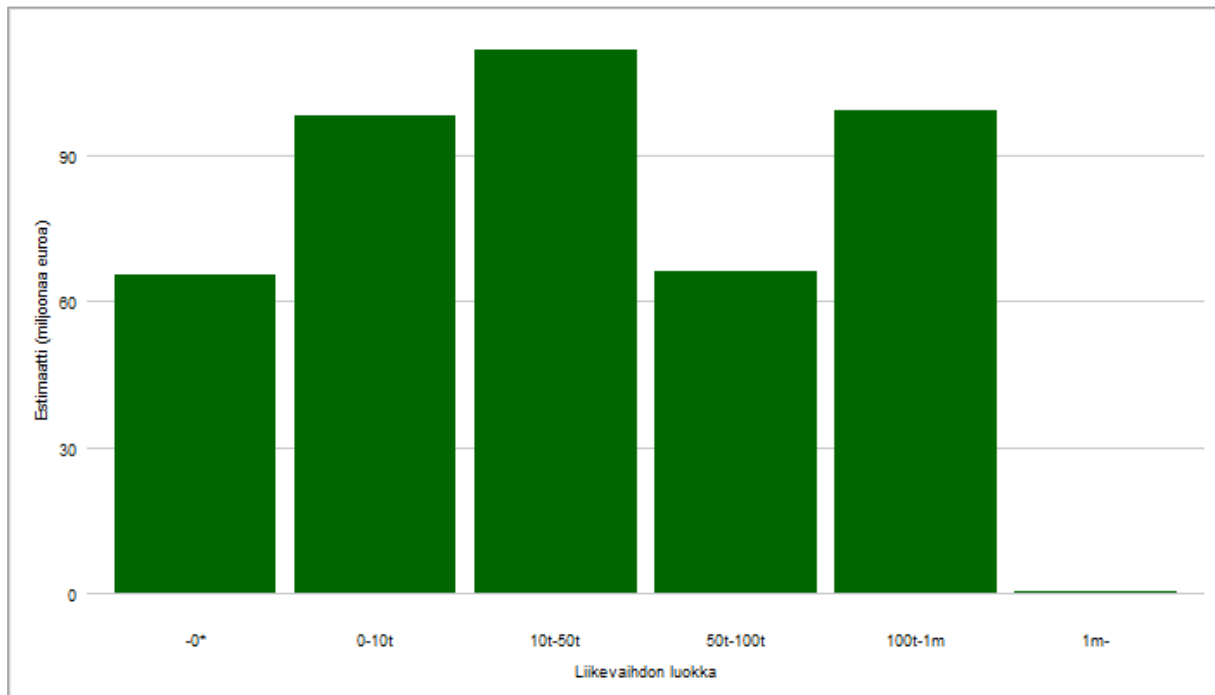
⁶⁸ Ensemble- eli yhdistelmämallit koostuvat joukosta eri malleja ja algoritmeja, joiden tekemät itsenäiset ennusteet yhdistetään lopulta lopulliseksi ennusteeksi. Itse asiassa myös XGBoost on yhdistelmämalli, koska se yhdistelee monta pientä päätöspuuta niin sanotulla boosting-tekniikalla yhdeksi lopulliseksi ennusteeksi.

⁶⁹ Pimeän työn odottama on pimeän työn todennäköisyys kertaa ennustettu euromäärä. Odottama lasketaan jokaiselle elinkeinonharjoittajalle erikseen.

⁷⁰ Yleisyyteen voi vaikuttaa kaksi asiaa: toisaalta luokan koko, eli elinkeinonharjoittajien määrä, ja toisaalta virheen todennäköisyys.

pimeää työtä noin 110 miljoonan euron edestä. Myös liikevaihtoluokissa yli 100 000 euroa ja alle 10 000 euroa pimeän työn ennuste oli lähes yhtä suuri.

Kuvio 3. Pimeän työn estimaatti liikevaihtoluokittain vuonna 2019.



*) Luokkaan -0 kuuluvat myös ne elinkeinonharjoittajat, joilta liikevaihtotieto puuttuu. Liikevaihtotietoa ei ole, jos sitä ei ilmoiteta 5-lomakkeella.

Lähde: Verohallinto 2–5/2022.

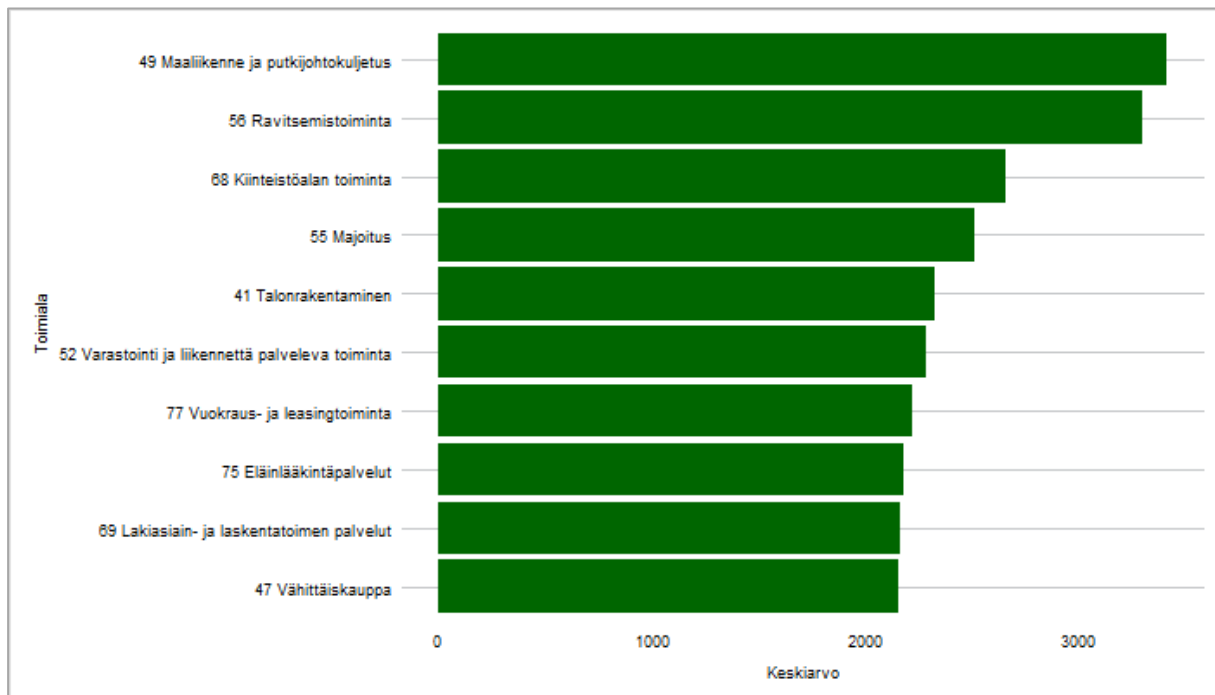
Kuviossa 4 tarkastellaan keskimääräisiä odottamia⁷¹ toimialoittain. Kaksi toimialaa, maaliikenne ja ravitsemistoiminta, nousivat keskimääräisen odottaman tarkastelussa muita suuremmiksi. Näillä toimialoilla keskimääräinen pimeän työn odottama oli yli 3 000 euroa⁷².

Kuviosta 3 ei ole suoraan luettavissa pimeän työn keskimääräisiä odottamia kokoluokittain. Erikseen voidaan kuitenkin todeta, että mallin mukaan keskimääräinen pimeän työn odottama oli sitä suurempi, mitä suurempi liikevaihto elinkeinonharjoittajalla oli.

⁷¹ Keskimääräinen odottama = odottamien summa luokassa jaettuna elinkeinonharjoittajien lukumäärällä luokassa.

⁷² Pimeän työn keskimääräinen odottama voidaan tulkita seuraavasti: jos kaikkiin toimialan elinkeinonharjoittajiin suoritettaisiin verotarkastus, olisi salattujen tulojen, perusteettomien vähennysten ja työntekijöille maksettujen pimeiden palkkojen määrä kaikille toimijoille jaettuna keskimäärin 3 000 euroa / elinkeinonharjoittaja. Keskimääräisen odottaman tarkastelussa on huomattava, että malli ei odota kaikkien elinkeinonharjoittajien tehneen pimeää työtä. Valtaosa elinkeinonharjoittajista toimii oikein ja pieni osa vilpillisesti. Ilmoitettu tulos on vain toimialan keskiarvo.

Kuvio 4. Elinkeinoharjoittajien pimeän työn odottaman keskiarvo toimialoittain vuonna 2019 koneoppimismallin mukaan.



Lähde: Verohallinto 2–5/2022.

6.3 Profilointi

Harmaan talouden selvitysyksikössä on kehitetty profilointimenetelmä, jonka avulla voidaan laskea yrityksille riskipisteitä erilaisten veronvälttämistapojen suhteen. Pisteytyksessä harmaan talouden toimijoiksi arvioitujen elinkeinoharjoittajien (vertailujoukko) ominaisuuksia verrataan muihin yrityksiin (perusjoukko). Ne perusjoukon elinkeinoharjoittajat, joilla on samankaltainen ominaisuusprofiili kuin vertailujoukon elinkeinoharjoittajilla, saavat korkeita riskipisteitä ja muodostavat niin sanotun riskiyritysten joukon. Riskiyritysten avulla muodostetaan lukumääräarvioita väärin toimineista elinkeinoharjoittajista ja edelleen euromääräisiä arvioita pimeän työn määrästä. Menetelmää on kuvattu tämän selvityksen liitteessä 3.

Riskiyritysten joukolle voidaan hakea erilaisia yrityskohtaisia tietoja. Tämä mahdollistaa tulosten monipuolisen tarkastelun, esimerkiksi muodostamalla riskiyritysten toimiala- tai kokojakaumia ja vertaamalla näitä perusjoukon vastaaviin jakaumiin.

6.3.1 Estimaatit

Myös profilointimenetelmällä laskettu estimaatti käsitti vuodet 2018–2020, ja se on kolmen vuoden keskiarvo. Tulos jaettiin myös päätoimisten ja sivutoimisten yrittäjien kesken.

Pimeän työn kokonaisestimaatti oli noin 360 miljoonaa euroa vuodessa (taulukko 14). Päätoimisten osuus pimeän työn kokonaisestimaatista oli noin kolme neljäsosaa ja sivutoimisten

osuus noin yksi neljäsosa.⁷³ Päätoimisten elinkeinonharjoittajien estimaattiin sisältyy vajaat 40 miljoonaa euroa työntekijöille maksettuja pimeitä palkkoja. Niiden osuus kokonaisestimaa- tista oli noin 10 prosenttia.

Pimeän työn riski oli tehdyn analyysin perusteella todennäköisempi päätoimisten (10,2 %) kuin sivutoimisten (6,4 %) elinkeinonharjoittajien joukossa. Palkansaajien pimeän työn toden- näköisyys oli alhainen (2,0 %), koska elinkeinonharjoittajat työllistävät tyypillisesti vain it- sensä. Lukumäärinä nämä prosenttiosuudet vastaavat yhteensä noin 18 000 elinkeinonhar- joittajaa vuosittain.

Toisaalta sivutoimisilla elinkeinonharjoittajilla pimeän työn estimaatin osuus elinkeinotoimin- nan tuloksesta oli selvästi suurempi (27 %) kuin päätoimisilla (yhteensä noin 10 %). Tätä se- littää se, että toiminta on sivutoimisilla keskimäärin vähäistä, mutta tarkastuksissa havaitut virheet ovat melko suuria.⁷⁴

Taulukko 14. Arvio elinkeinonharjoittajien pimeästä työstä profiloitimenetelmällä, keskimäärin vuodessa ajalla 2018–2020 (miljoonaa euroa).

Selite	Estimaatti, milj. €	Osuus tuloksesta, %	Virheiden todennäköisyys perusjoukossa, %
Yhteensä	362	11 %	8–9 % (*)
Puuttuva tulo ja perusteettomat vähennykset, päätoimiset	239	8,2 %	10,2 %
Pimeät palkat, päätoimiset	37	1,3 %	2,0 %
Puuttuva tulo ja perusteettomat vähennykset, sivutoimiset	86	27,2 %	6,4 %

(*) Todennäköisyyden alaraja on laskettu olettaen, että kaikilla pimeitä palkkoja maksaneilla on myös puuttuvaa tuloa tai perusteettomia vähennyksiä. Yläraja on laskettu olettaen, että yhdelläkään pimeitä palkkoja maksaneella ei ole puuttuvaa tuloa tai perusteettomia vähennyksiä.

Lähde: Verohallinto 2–6/2022.

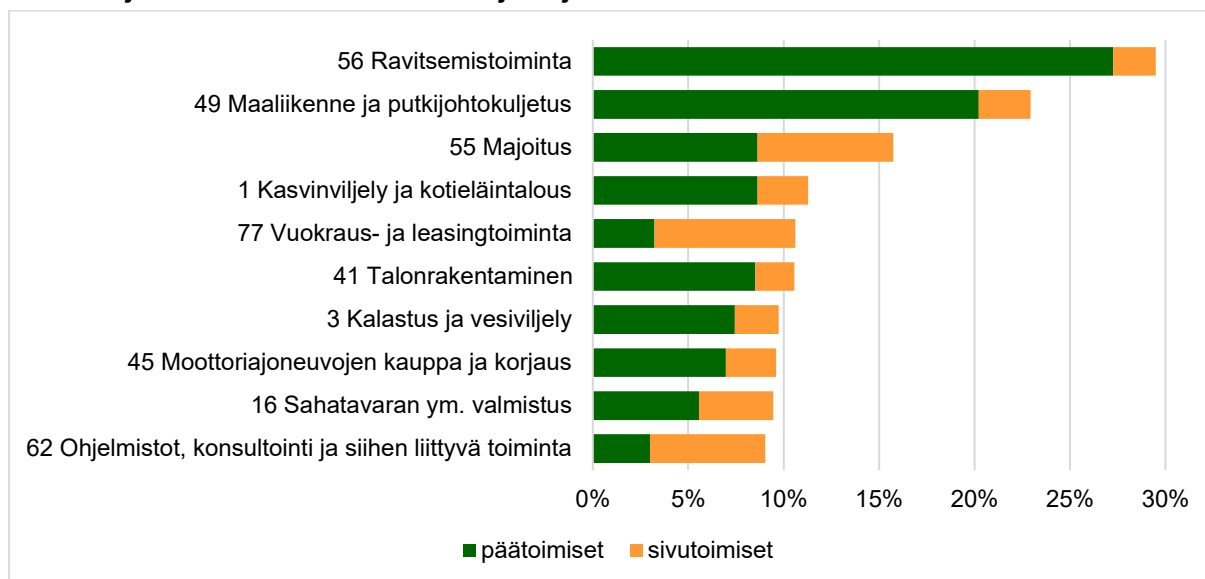
6.3.2 Toimialat

Toimialatarkastelussa ravitsemistoiminnan elinkeinonharjoittajista lähes 30 prosenttia ja maaliikenteen elinkeinonharjoittajista yli 20 prosenttia valikoitui riskiyritysten joukkoon. Li- säksi riskiyritysten osuus oli yli 10 prosenttia majoituksessa, maataloudessa, vuokraus- ja leasingtoiminnassa sekä talonrakentamisessa.

⁷³ Laskennan jako ei täysin noudata päätoimisten ja sivutoimisten määrittelyä, sillä sivutoimisten estimaatti sisäl- tää myös ne päätoimiset, joiden liikevaihto oli alle 5 000 euroa vuodessa. Tähän ratkaisuun päädyttiin, koska hy- vin pienen volyymin toiminta poikkeaa rajusti muista päätoimisista ja muistuttaa toiminnan laajuuden suhteen enemmän sivutoimisiksi määriteltyjä.

⁷⁴ Aiemmin taulukossa 11 esitettiin virheiden keskiarvot ja mediaanit kaikille verotarkastetuille elinkeinonharjoitta- jille yhteensä. Kun tarkastetut jaettiin päätoimisten ja sivutoimisten alaryhmiin, saatiin seuraavat mediaanit: pää- toimisten puuttuva tulo ja perusteettomat vähennykset 25 700 €, palkansaajien pimeä palkka 28 300 €, sivutoimis- ten puuttuva tulo ja perusteettomat vähennykset 12 200 €.

Kuvio 5. Riskiyritysten osuus kaikista toimialan elinkeinonharjoittajista vuonna 2019. Jako päätoimisiin ja sivutoimisiin elinkeinonharjoittajiin.



Lähde: Verohallinto 2–6/2022.

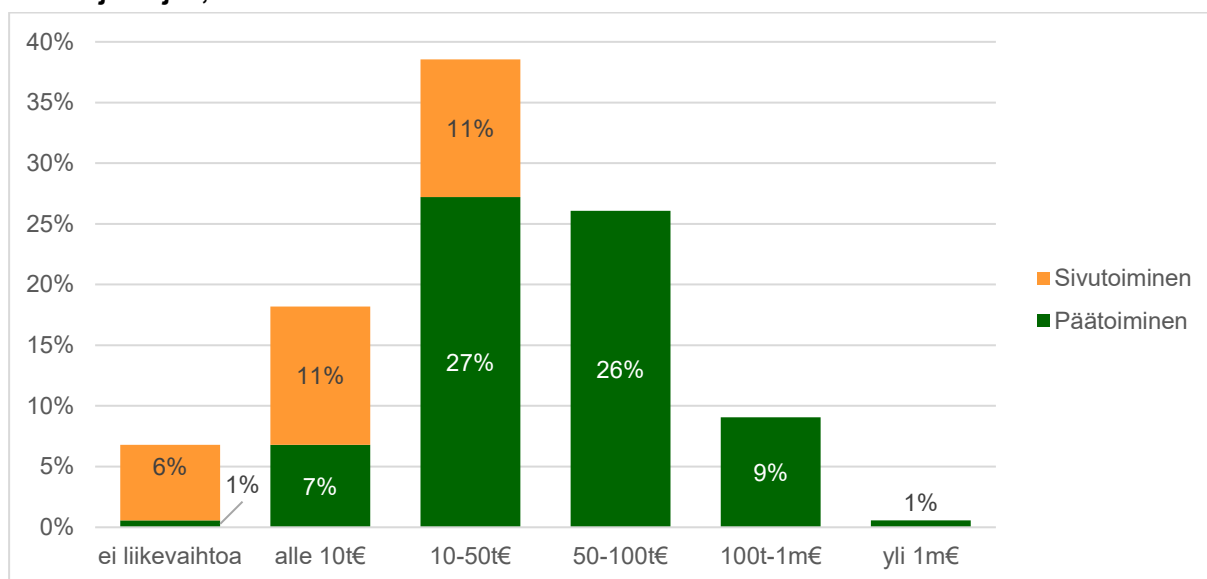
6.3.3 Koko

Kuviossa 6 on esitetty riskiyritysten kokojakauma liikevaihdon mukaan. Jokainen liikevaihtoluokka on jaettu edelleen päätoimisten ja sivutoimisten kesken. Kuviossa 7 on vastaava jakauma koko perusjoukolla.

Kuviosta 6 nähdään, että riskiyrityksistä yli puolet oli päätoimisia elinkeinonharjoittajia, joiden liikevaihto oli 10 000–100 000 euroa. Vastaavan ryhmän osuus perusjoukosta oli noin 30 % (kuvio 7). Sivutoimisten pienten (liikevaihto alle 10 000 euroa) elinkeinonharjoittajien osuus perusjoukosta oli noin 40 prosenttia, mutta riskiyrityksistä heidän osuutensa oli alle 20 prosenttia.

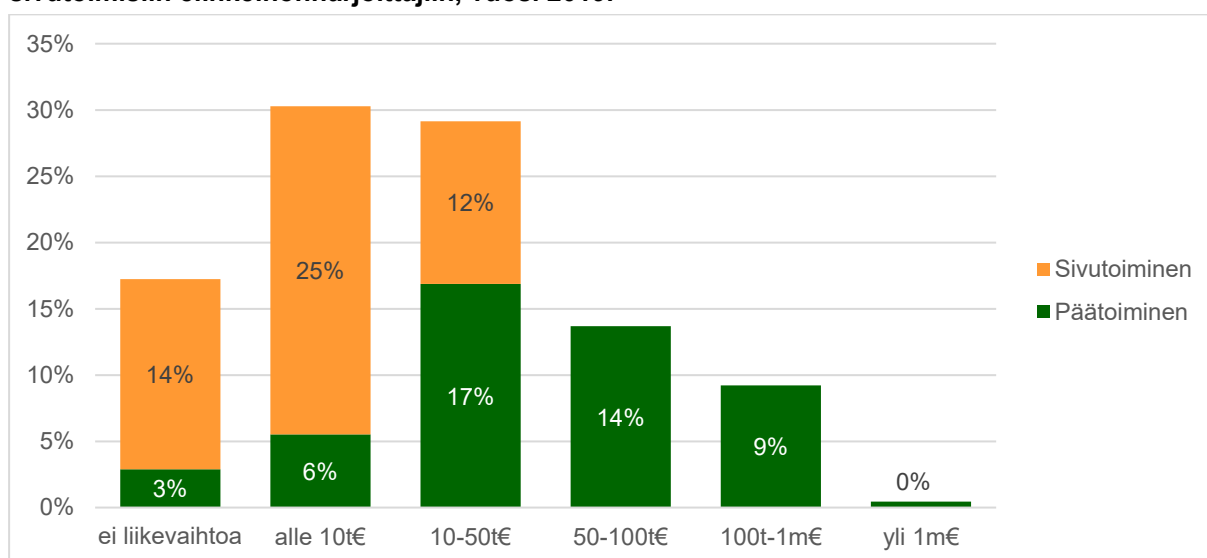
Näin ollen analyysin perusteella voimme todeta, että pimeän työn riski on kohonnut erityisesti päätoimisilla elinkeinonharjoittajilla, joiden liikevaihto on 10 000–100 000 euroa. Lisäksi vaikka sivutoimisia alle 10 000 euron liikevaihtoa esittäviä elinkeinonharjoittajia on runsaslukaisesti, heillä ei näytä olevan erityisen suurta riskiä pimeään työhön.

Kuvio 6. Riskiyritykset liikevaihtoluokittain sekä jaettuna päätoimisiin ja sivutoimisiin elinkeinonharjoittajiin, vuosi 2019.



Lähde: Verohallinto 2–6/2022.

Kuvio 7. Perusjoukon elinkeinonharjoittajat liikevaihtoluokittain sekä jaettuna päätoimisiin ja sivutoimisiin elinkeinonharjoittajiin, vuosi 2019.



Lähde: Verohallinto 2–6/2022.

7 Tulosten yhteenveto sekä verojen ja muiden maksujen menetykset

Tehtyjen analyysien perusteella elinkeinonharjoittajien tekemän ja teettämän pimeän työn arvo oli noin 250–450 miljoonaa euroa vuodessa. Tuloksen suurehko vaihteluväli kuvastaa estimointien epävarmuutta, jonka taustalla on ennen kaikkea tilastollisten menetelmien näkökulmasta vinoutunut ja pieni aineisto. Pimeästä työstä aiheutui arviolta noin 75–140 miljoonan euron vuosittaiset verojen ja palkkojen sosiaalivakuutusmaksujen menetykset ajalla 2018–2020.

7.1 Pimeän työn määrä sekä verojen ja muiden maksujen menetykset

Pimeän työn estimaatin avulla on arvioitu menetettyjen verojen ja muiden maksujen määrät. Veronmenetyksiä aiheutuu elinkeinonharjoittajan tuloverosta sekä palkansaajien ennakonpidätyksistä ja sosiaalivakuutusmaksuista.

7.1.1 Pimeän työn määrä

Selvityksessä on arvioitu elinkeinonharjoittajien pimeän työn määrää usealla eri menetelmällä. Taulukossa 15 on yhteenveto eri menetelmillä arvioidusta pimeän työn määrästä vuosina 2018–2020, joka oli noin 250–450 miljoonaa euroa vuodessa. Arvioitu pimeän työn määrä vastasi noin 8–14 prosenttia elinkeinonharjoittajien verotuksessa vahvistettujen elinkeinotoiminnan tulosten yhteismäärästä.

Taulukko 15. Elinkeinonharjoittajien pimeän työn estimaatit eri menetelmillä laskettuina (miljoonaa euroa), vuodet 2018–2020.

Menetelmä	2018	2019	2020
CEM	242	351	384
XGBoost	451	440	407
Profilointi (*)	362	362	362

(*) Tulos on kolmen vuoden keskiarvo

Lähde: Verohallinto 2–6/2022.

Työntekijöille maksettujen pimeiden palkkojen osuus oli noin 10 prosenttia ja yrittäjien osuus noin 90 prosenttia pimeän työn kokonaisestimaatista.

7.1.2 Verojen ja sosiaalivakuutusmaksujen menetykset

Elinkeinonharjoittajien pimeä työ aiheutti vuosittain arviolta noin 100 miljoonan euron verojen ja palkkojen sosiaalivakuutusmaksujen menetykset.⁷⁵ Veronmenetys on arvioitu käyttäen profiloinnin estimaattia, joka osui suurimpien ja pienimpien estimaattien väliin. Jos verojen menetys suhteutetaan muiden menetelmien, tuloksiin saadaan vaihteluväliksi noin 75–140

⁷⁵ Veronmenetyksen arvio on laskettu olettaen, että vuosittainen pimeiden palkkojen määrä on 362 miljoonaa euroa. Siitä elinkeinonharjoittajien ansiotulon osuus on 288 miljoonaa euroa, pääomatulojen osuus 37 miljoonaa euroa ja palkansaajien pimeän työn osuus niin ikään 37 miljoonaa euroa.

miljoonaa euroa vuodessa.⁷⁶ Summa käsittää arviot elinkeinonharjoittajien ansio- ja pääomatulojen menetyksistä sekä palkansaajien ennakonpidätyksistä ja sosiaalivakuutusmaksuista.

Veronmenetyksen laskemista on kuvattu edellä luvussa 1.3 ja liitteessä 4.

7.2 Tyypillinen pimeä elinkeinonharjoittaja

Tässä luvussa esitellään, millainen oli tyypillinen pimeää työtä tehnyt elinkeinonharjoittaja. Koneoppimis- ja profilointimenetelmissä jokainen y-tunnus pisteytetään arvioidun riskisyyden mukaan. Y-tunnuskohtainen riskiarvio⁷⁷ ja/tai pimeän työn estimaatti⁷⁸ mahdollistavat tulosten yksityiskohtaisen vertailun näiden kahden menetelmän välillä ja tekevät mahdolliseksi muodostaa kuvan tyypillisestä elinkeinonharjoittajasta. Vaikka eri menetelmien tuottamat kokonaisestimaatit ovat melko yhdenmukaisia, niin tarkemmassa tarkastelussa havaitaan eroja tuloksissa.

Riskisyyden arviointi perustuu kummassakin menetelmässä erehtyväiseen malliin, jonka taustalla on lukumäärältään vähäinen ja vinoutunut aineisto. Näin ollen kummankaan menetelmät tulokset eivät vastaa täysin todellisuutta. Tuloksia vertaamalla ja kriittisesti arvioimalla saamme kuitenkin havaintoja pimeää työtä tekevän elinkeinonharjoittajan keskeisistä ominaispiirteistä. Tulokset koskevat vuotta 2019, mutta ne ovat yleistettävissä koskemaan myös muita selvityksen vuosia.

On selvää, että kummassakin menetelmässä riskisiksi arvioitujen tai korkean pimeän työn estimaatin saaneiden elinkeinonharjoittajien joukkoon mahtuu paljon yrittäjiä, jotka ovat todellisuudessa täysin lainkuuliaisia. Virheitä tapahtuu myös toisin päin, eli mallit arvioivat monet tosiasiallisesti tulojaan pimittävät elinkeinonharjoittajat moitteettomiksi. Voidaan kuitenkin olettaa, että kokonaisestimaateissa virheet kumoavat toisiaan ja esimerkiksi riskisemmiksi arvioidut toimialat ovat myös aidosti riskisempiä.

7.2.1 Pimeää työtä tekevät elinkeinonharjoittajat liikevaihtoluokittain

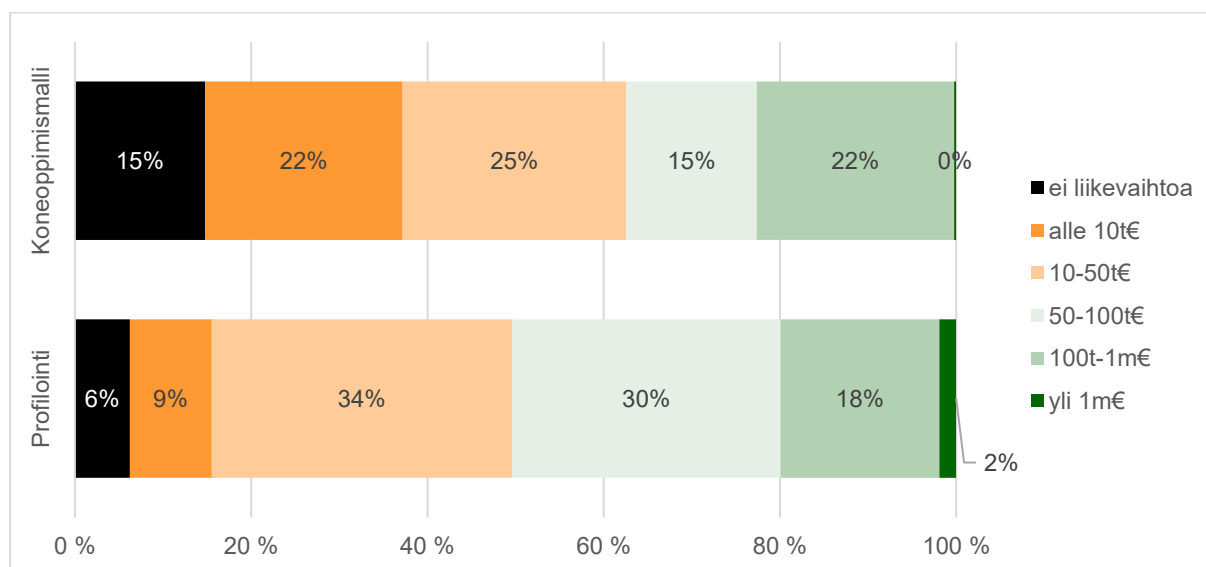
Kuviossa 8 on esitetty koneoppimismallin ja profilointimenetelmän tuottamien estimaattien jakautuminen elinkeinonharjoittajien liikevaihtoluokkiin. Kummassakin estimaatissa suurin luokka on 10 000–50 000 euroa. Lisäksi kumpikin menetelmä sijoittaa estimaatista noin 80 prosenttia alle 100 000 euron liikevaihtoluokkiin. Eroa havaitaan ennen kaikkea pienimpien elinkeinonharjoittajien osuuksissa. Koneoppimisen estimaatista 35 prosenttia sijoittuu alle 10 000 euron luokkiin, kun profiloinnissa tämä osuus on vain 15 prosenttia.

⁷⁶ Suhteutus on tehty pelkästään estimaattien suuruuksien perustella. Tässä ei ole huomioitu esimerkiksi estimaatin koon vaikutusta progressiiviseen veroprosenttiin.

⁷⁷ Profiloinnin tulos.

⁷⁸ Asiakaskohtaisen odottaman tulos.

Kuvio 8. Koneoppimismallin ja profiloinnin pimeän työn estimaatit liikevaihtoluokittain, vuosi 2019.



Lähde: Verohallinto 2–6/2022.

Pimeän työn estimaatteja on mielenkiintoista verrata aiemmin esitettyyn kuvioon 1 (luku 3.3), jossa on kokoluokittain kaikkien elinkeinonharjoittajien yhteenlaskettu lukumäärä, liikevaihto ja elinkeinotoiminnan tulos. Suurin pimeän työn estimaatti (profilointi 64 % ja koneoppimismalli 40 %) arvoitiin luokkiin, joissa liikevaihdot ovat välillä 10 000–100 000 euroa. Elinkeinoharjoittajien lukumäärästä alle kolmasosa, liikevaihdosta alle neljäsosa ja elinkeinotoiminnan tuloksesta noin kaksi kolmasosaa tuli tästä kokoluokasta (kuvio 1). Vaikka 10 000–100 000 euron liikevaihtoiset elinkeinonharjoittajat ovat lukumääräisesti ja liikevaihtoja tarkasteltaessa melko pieni joukko, niin kuitenkin heidän toiminnastaan muodostuu suuri määrä veronalaista tuloa. Kyseisen kokoluokan pimeä työ siis aiheuttaa huomattavaa verovajetta.

7.2.2 Pimeää työtä tekevät elinkeinonharjoittajat toimialan mukaan

Toimialavertailussa on huomattava, että koneoppimismallin toimialajakauma on tehty euro-määräisen estimaatin ja profiloinnin toimialajakauma riskiyritysten lukumäärän perusteella. Tästä johtuen taulukossa 16 vertaillaan ainoastaan kummankin menetelmän TOP10-toimialalistoja.

Talonrakentaminen, maaliikenne, majoitus, ravitsemistoiminta sekä vuokraus- ja leasingtoiminta nousivat kummallakin menetelmällä merkittävimpien pimeän työn toimialojen joukkoon. Tämän selvityksen perusteella nämä toimialat ovat siten ongelmallisimpia elinkeinonharjoittajien pimeän työn⁷⁹ näkökulmasta. Maaliikenteessä, ravitsemistoiminnassa ja talonrakentamisessa selvän enemmistön arvioitiin olevan päätoimisia elinkeinonharjoittajia (ks. kuvio 5).

Edellä mainittujen toimialojen lisäksi profilointi nosti riskitoimialoiksi muun muassa maatalouden ja kalatalouden, koneoppimismalli puolestaan kiinteistöalan ja varastoinnin. On hyvä muistaa, että toisilla toimialoilla käytetään ja tarvitaan ulkopuolista työvoimaa ja toisille

⁷⁹ Kattaa tässä selvityksessä elinkeinonharjoittajan salaamat tulot, perusteettomat kulut sekä veroilmoituksen kuluvähennykset ja tuloutuspuutteet ja vähäisissä määrin työntekijöille maksetut pimeät palkat.

toimialoille on tunnusomaista, että elinkeinonharjoittaja työllistää itsensä. Riskitoimialojen joukkoon nousi molempia.

**Taulukko 16. Merkittävimmät elinkeinonharjoittajien pimeän työn toimialat vuonna 2019. Profi-
loinnin ja koneoppimismallin tulosten vertailu.**

Toimiala	Profilointi #	Koneoppimismalli #
1 Kasvinviljely ja kotieläintalous	4	-
3 Kalastus ja vesiviljely	7	-
16 Sahatavaran ym. valmistus	9	-
41 Talonrakentaminen	6	5
45 Moottorijoneuvojen kauppa ja korjaus	8	-
47 Vähittäiskauppa	-	10
49 Maaliikenne ja putkijohtokuljetus	2	1
52 Varastointi ja liikennettä palveleva toiminta	-	6
55 Majoitus	3	4
56 Ravitsemistoiminta	1	2
62 Ohjelmistot, konsultointi ja siihen liittyvä toiminta	10	-
68 Kiinteistöalan toiminta	-	3
69 Lakiasiain ja laskentatoimien palvelut	-	9
75 Eläinlääkäripalvelut	-	8
77 Vuokraus- ja leasingtoiminta	5	7

Lähde: Verohallinto 2–6/2022.

7.2.3 Päätoimisten ja sivutoimisten elinkeinonharjoittajien pimeä työ

Tässä selvityksessä elinkeinonharjoittajien joukko jaettiin päätoimisiin ja sivutoimisiin elinkeinonharjoittajiin. Päätoimisia olivat ne, joiden tuloista yli puolet muodostui elinkeinotoiminnasta, ja sivutoimisia olivat ne, joiden tuloista alle puolet muodostui elinkeinotoiminnasta. Lisäksi päätoimisiin luettiin kaikki elinkeinonharjoittajat, joiden liikevaihto oli yli 50 000 euroa vuodessa.

Perusjoukossa oli lähes yhtä paljon päätoimisia ja sivutoimisia elinkeinonharjoittajia. Profi-
lointimenetelmällä tehdyn analyysin perusteella päätoimiset muodostivat kuitenkin noin kolme neljäsosaa ja sivutoimiset noin yhden neljäsosan pimeän työn estimaatista.

Pimeän työn riski oli todennäköisempi päätoimisten (10,2 %) kuin sivutoimisten (6,4 %) elinkeinonharjoittajien joukossa. Tämä on ymmärrettävää, sillä sivutoimisilla elinkeinotoiminta on usein pienimuotoista. Verotarkastuksissa havaitut virheet ovat keskimäärin melko suuria, ja tällaisten virheiden esiintyvyyden ei pitäisi olla kovin suuri siellä, missä toimintaa on vain vähän.⁸⁰

Vaikka virheiden esiintyvyys oli sivutoimisilla matalampi ja virheet olivat pienempiä kuin päätoimisilla, niin näistä seikoista huolimatta sivutoimisilla elinkeinonharjoittajilla pimeän työn osuus elinkeinotoiminnan tuloksista oli huomattavan korkea (27 %) verrattuna kaikkiin tai pelkästään päätoimisiin elinkeinonharjoittajiin (ks. luku 6.3.1). Tämä seuraa suoraan siitä, että sivutoimisten toiminta oli suppeampaa ja elinkeinotoiminnan tulokset, joihin sivutoimisten estimaattia verrattiin, olivat euromääräisesti pieniä.

⁸⁰ Toki ei sovi unohtaa, että osalla elinkeinonharjoittajista toiminta näyttää vähäiseltä, koska tuloja on pimitetty. Laadittuja tilastoja hallitsevat kuitenkin ne, joiden toiminta on aidosti vähäistä.

7.3 Estimaattien laskentaan liittyvistä epävarmuuksista

Monet seikat aiheuttivat epävarmuutta elinkeinonharjoittajien pimeän työn estimointiin. Erityisesti epävarmuutta aiheutui aineistoista, joita kuvattiin edellä luvussa 5.

Ensimmäinen verotarkastuksiin liittyvä ongelma on aineiston pieni määrä. Vähäinen havaintojen määrä lisää sattumanvaraisten virheiden riskiä. Todennäköisesti kuitenkin suurempi ongelma data-analyysin kannalta on tarkastuskohteiden riskiperusteinen valinta, mistä johtuen käytettävissä oleva aineisto on vinoutunut: virheitä löytyy suurella todennäköisyydellä ja ne ovat isoja.⁸¹

Tarkastusaineisto ei ylipäätään anna tasapuolista kuvaa elinkeinonharjoittajien joukosta. Tarkastukset keskittyvät pääosin liikevaihdoltaan suurempiin elinkeinonharjoittajiin, vaikka lukumääräisesti pieniä elinkeinonharjoittajia on paljon enemmän. Tarkastusten keskittyminen harvoille toimialoille on myös ongelmallista, koska kaikkien toimialojen erityispiirteet eivät välttämättä ole edustettuina tarkastusaineistoissa. Toimialojen välillä on huomattavia eroja muun muassa toiminnan laajuudessa sekä tulo- ja kulurakenteissa. Laajamuotoisessa yritystoiminnassa tarvitaan usein ulkopuolista työvoimaa, kun taas pienimuotoisemmasta toiminnasta elinkeinonharjoittaja vastaa yksin. Eroja toimialojen välillä tuo esimerkiksi myös käteisen rahan käyttö: joillakin toimialoilla käteiskauppa on vielä yleistä, kun toisilla se on hyvinkin vähäistä.

Luvussa 6.1 kuvattiin, kuinka **CEM-menetelmä** tasoittaa verotarkastusten valikointiharhaa etsimällä jokaiselle tarkastetulle elinkeinonharjoittajalle tarkastamattoman vastinparin tai pareja. Vastinparit etsitään luokittelemalla erilaisia ominaisuuksia ja vaatimalla, että vastinparit sijoittuvat samoihin luokkiin kaikkien ominaisuuksien suhteen. Näin vain pienelle ja tarkastetuja tarkkaan muistuttavalle tarkastamattomien osajoukolle päätellään pimeää työtä. Menetelmällä päättely saadaan rajattua aineistoon, jossa riskiperusteisen kohdevalikoinnin aiheuttama vinouma on poistettu.

CEM-menetelmä ei onnistu päättämään lainkaan pimeää työtä sellaisille toimijoille, jotka eivät muistuta tarkastettuja yrityksiä, ja tämä aiheuttaa estimaattiin harhaa alaspäin. Se ei myöskään pysty käsittelemään tarkastuksissa havaittujen virheiden kokoharhaa⁸², mistä aiheutuu tulokseen harhaa ylöspäin.

Profiloinnissa (luku 6.3) riskiperusteisen verotarkastusten valintaharhaa vähennetään jakamalla perusjoukko osiin erilaisten luokittelevien tietojen perusteella, kuten päätoimisuus ja sivutoimisuus sekä toiminnan laajuus (liikevaihto). Luokittelu on harvempaa kuin CEM-menetelmässä, mutta osaltaan se pienentää valikoinnin harhaa. Lisäksi monet selittävät muuttujat lasketaan suhdelukuina (esimerkiksi keskipalkka), mikä myös torjuu yksipuolisen

⁸¹ *Pimeän työn määrä – Mittaamisselvitys osakeyhtiöiden verotarkastustietojen perusteella* -selvityksessä havaittiin, että virheiden esiintyvyys ja keskimääräinen virheen koko ovat selvästi suuremmat riskiperusteisissa verotarkastuksissa kuin satunnaistarkastuksissa.

⁸² Tarkastuksissa pyritään löytämään suurimmat virheet.

kohdevalinnan vinoumaa. Tarkastusten keskittymistä harvoille toimialoille torjutaan sillä, että useiden selittävien muuttuja arvoja verrataan kunkin yrityksen toimialan mediaaniin.⁸³

Profiloinnin tuloksena syntyy arvio virheiden esiintyvyydestä perusjoukossa (virheiden lukumäärä jaettuna perusjoukon lukumäärällä). Tämä arvio muodostuu osittain menetelmän tuotamien tulosten perusteella, mutta siihen voidaan vaikuttaa myös asiantuntijavalintojen kautta. Tätä subjektiivista elementtiä voidaan pitää yhtä lailla menetelmän heikkoutena tai vahvuutena. Sillä voidaan torjua täysin epärealistisia tuloksia, mutta tulos voi ohjautua myös väärään suuntaan.

Virheiden kokoharhan takia euromääräiset estimaatit laskettiin virheiden mediaaneja käyttäen keskiarvojen sijaan. Mediaanit olivat pienempiä kuin keskiarvot ja antoivat uskottavampia estimaatteja.⁸⁴ Lopputulos on kuitenkin epävarma sikäli, että emme voi tietää onko kokoharhaa torjuttu riittävän tehokkaasti.

Asiakaskohtaisen odottaman koneoppimismallit oppivat tehokkaasti harjoitusaineistossa esiintyvät yhteydet pimeän työn ja elinkeinonharjoittajan ominaisuuksien välille. Mutta samalla ne oppivat myös riskiperusteisen valinnan aiheuttamaa harhaa ja oppivat etsimään samanlaisia kohteita, joita on verotarkastettu. Tärkein tapa torjua tätä harhaa on käyttää suurta määrää eri muuttujia, jotka kuvaavat elinkeinonharjoittajien riskisyyttä monipuolisesti myös tarkastamattomien elinkeinonharjoittajien osalta.

Hyödyntämällä säännönmukaisessa valvonnassa moitteettomaksi todettuja tapauksia saadaan myös korjattua ongelmia, jotka johtuvat verotarkastusten vähäisestä määrästä ja suuresta osumatarkkuudesta. Tarkastusten osumaprosentti tippuu lähemmäksi todellista pimeän työn esiintyvyyttä kaikkien elinkeinonharjoittajien keskuudessa. Isompi aineisto tarjoaa myös koneoppimismalleille enemmän havaintoja, joiden perusteella ne erottavat lainkuuliaiset ja veroja kiertävät elinkeinonharjoittajat toisistaan.

Käyttämällä algoritmeja ja niiden yhdistelmiä, jotka torjuvat tehokkaasti ns. ylisovittamista⁸⁵ voidaan myös torjua aineiston vinoumaa. Virheen kokoa estimoitaessa hyödynnetään päätöspuihin perustuvaa XGBoostia, joka on varsin robusti eli vakaa menetelmä. Toisin sanoen suurimmat tarkastushavainnot eivät vaikuta voimakkaasti ennusteisiin.

⁸³ Esimerkiksi selittävä muuttuja ei ole yrityksen keskipalkka vaan sen keskipalkan suhteellinen poikkeama yrityksen edustaman toimialan mediaanista. Jos yrityksen maksama keskipalkka on 10 000 € ja sen toimialalla keskipalkka on 11 000 €, niin yritys saa muuttujan keskipalkka arvoksi -0.091.

⁸⁴ Käytetyt virheiden mediaanit olivat jopa pienempiä kuin trimmatuista jakaumista lasketut virheiden keskiarvot. Tämä seikka havaittiin erityisesti sivutoimisten elinkeinonharjoittajien laskennassa.

⁸⁵ Ylisovittamisella (eng. overfitting) tarkoitetaan tilannetta, jossa malli oppii liian tarkkaan harjoitusaineistossa esiintyvät datan omituisuudet. Tällöin mallin kyky yleistää ennusteet harjoitusaineiston ulkopuolelle ovat heikot.

8 Johtopäätökset

Tässä raportissa tarkasteltiin elinkeinonharjoittajien pimeää työtä muodostamalla estimaatteja pimeän työn arvosta ja sen tuottamasta rahallisesta hyödystä vältettyjen verojen muodossa. Selvityksessä on arvioitu ensisijaisesti pimeän työn vaikutusta elinkeinotoiminnan verotettavaan tuloon sekä toissijaisesti arvioitu työntekijöille maksettujen pimeiden palkkojen määrää. Raportti täydentää Harmaan talouden selvitysyksikön tarkastelemaa pimeän työn kokonaisuutta Suomessa. On hyvä muistaa, että pimeä työ on vain yksi osa harmaan talouden laajasta kokonaisuudesta. Pimeää työtä tarkastelemalla ei arvioida harmaan talouden kokonaisuutta, vaan vain yhtä osaa siitä.

Suomessa elinkeinonharjoittajia on paljon ja valtaosalla toiminta on melko pientä. Noin puolet elinkeinonharjoittajista harjoittaa elinkeinotoimintaa pääsääntöisenä tulonlähteenään ja puolet sivutoimisesti esimerkiksi palkkatyönsä ohessa. Päätoimisetkin elinkeinonharjoittajat työllistävät usein vain itsensä ja harvalla on palkattua työvoimaa.

Elinkeinonharjoittajien pimeän työn määrä Suomessa on tehdyn arvion mukaan vuosittain noin 250–450 miljoonaa euroa. Tuloksen vaihteluväliä voidaan pitää siedettävänä, kun huomioidaan käytetyn aineiston epävarmuudet. Pimeä työ aiheuttaa veronmenetyksiä noin 75–140 miljoonaa euroa vuodessa, pääosin elinkeinonharjoittajien ansiotuloveroina.

Valtaosa elinkeinonharjoittajien pimeästä työstä aiheutuu siitä, että elinkeinotoiminnan tulos esitetään liian pienenä ja siksi veroa jää kertymättä. Vain kymmenesosa on perinteistä työntekijöille maksettua pimeää palkkaa. Suurin osa (3/4) veronvälttelystä näyttää aiheutuvan päätoimisten elinkeinonharjoittajien vilpillisestä toiminnasta. Erityisesti ravitsemusala ja kuljetusala ovat riskisiä veronvälttelylle. Näillä aloilla toimintaa harjoitetaan enimmäkseen päätoimisesti. 10 000–100 000 euroa liikevaihtoa ilmoittavat elinkeinonharjoittajat muodostavat noin kolmasosan elinkeinonharjoittajien lukumäärästä. Kuitenkin jopa puolet veronalaisesta elinkeinotoiminnan tuloksesta muodostuu heidän toiminnastaan, ja heidän arvioidaan tekevän myös huomattavan osan elinkeinonharjoittajien pimeästä työstä.

Verrattuna Euroopan komission vuonna 2017 esittämään arvioon elinkeinonharjoittajien pimeän työn määrä vaikuttaa melko maltilliselta. Harmaan talouden selvitysyksikön arvion mukaan elinkeinonharjoittajien vuosittainen pimeän työn määrä on noin 8–14 prosenttia verotuksessa vahvistetuista elinkeinotoiminnan tuloksista. Komission arvion mukaan Suomessa itsensä työllistäjien pimeän työn määrä oli jopa lähes puolet työpanoksesta. Harmaan talouden selvitysyksikkö ei arvioinut pimeän työn osuutta työpanoksesta, vaan pimeän työn euromäärä suhteutettiin elinkeinotoiminnan tuloksiin, toisin sanoen siihen tulon määrään, josta elinkeinonharjoittaja maksaa tuloveroja.

Ponnisteluja harmaan talouden ja pimeän työn vähentämiseksi on tarpeen jatkaa asian yhteiskunnallisen merkityksen takia. Harmaan talouden torjunnan toimenpideohjelmilla turvataan torjuntatyössä tarvittavien kehitystehtävien ja -hankkeiden systemaattisuus, jatkuvuus ja rahoitus.

9 Lähteet

Apteekkariliitto. (22. 3 2021). Apteekin verotus. Haettu 20. 7 2022 osoitteesta <https://www.apteekkariliitto.fi/apteekkitieto/verotus.html>

Sosiaali- ja (ei pvm). Sosiaalivakuutusmaksut. Helsinki, Suomi. Haettu 14. 10 2021 osoitteesta <https://stm.fi/sosiaalivakuutusmaksut>

Verohallinto. (23. 11 2021). Veroilmoituksen täyttöohje - liikkeen ja ammatinharjoittajan veroilmoitus. Helsinki, Suomi. Haettu 2. 6 2022 osoitteesta <https://www.vero.fi/yritykset-ja-yhteisot/verot-ja-maksut/liikkeen-tai-ammattinharjoittaja/ilmoittamisen-ohje/>

Williams, B. M. (2017). *An evaluation of the scale of undeclared work in the European Union and its structural determinants*. Luxemburg: Publications Office of the European Union. Retrieved toukokuu 30, 2022, from <https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&ved=2ahUKEwjsqaGB2ob4AhWOCBAIHVTuBdMQFnoECAcQAQ&url=https%3A%2F%2Fec.europa.eu%2Fsocial%2FblobServlet%3FdocId%3D19002%26langId%3Den&usg=AOvVaw2mqTIG-V5HuVfqWb2OGBID>

10 Luettelo kuvioista ja taulukoista

Kuvio 1. Elinkeinoharjoittajien lukumäärän, liikevaihdon ja tuloksen suhteellinen jakautuminen liikevaihtoluokkiin, vuosi 2019.....	12
Kuvio 2. Päätoimisten ja sivutoimisten elinkeinoharjoittajien lukumäärät liikevaihtoluokittain vuonna 2019.	15
Kuvio 3. Pimeän työn estimaatti liikevaihtoluokittain vuonna 2019.....	29
Kuvio 4. Elinkeinoharjoittajien pimeän työn odottaman keskiarvo toimialoittain vuonna 2019 koneoppimismallin mukaan.	30
Kuvio 5. Riskiyritysten osuus kaikista toimialan elinkeinoharjoittajista vuonna 2019. Jako päätoimisiin ja sivutoimisiin elinkeinoharjoittajiin.....	32
Kuvio 6. Riskiyritykset liikevaihtoluokittain sekä jaettuna päätoimisiin ja sivutoimisiin elinkeinoharjoittajiin, vuosi 2019.	33
Kuvio 7. Perusjoukon elinkeinoharjoittajat liikevaihtoluokittain sekä jaettuna päätoimisiin ja sivutoimisiin elinkeinoharjoittajiin, vuosi 2019.	33
Kuvio 8. Koneoppimismallin ja profiloinnin pimeän työn estimaatit liikevaihtoluokittain, vuosi 2019.	36
Taulukko 1. Palkan sivukulut ja niiden suuruus vuonna 2021.	6
Taulukko 2. Elinkeinoharjoittajat Verohallinnon rekistereissä vuosina 2018–2020.....	10
Taulukko 3. Elinkeinoharjoittajien lukumäärät liikevaihtoluokissa, vuodet 2018–2020.....	11
Taulukko 4. Elinkeinoharjoittajien tulot, vuosi 2019.....	12
Taulukko 5. Elinkeinoharjoittajien yleisimmät toimialat vuonna 2019.	14
Taulukko 6. Päätoimisten ja sivutoimisten elinkeinoharjoittajien tulot, vuosi 2019.....	15
Taulukko 7. Päätoimisten ja sivutoimisten elinkeinoharjoittajien suurimmat toimialat vuonna 2019.	16
Taulukko 8. Työnantajarekisterissä olevat ja satunnaisesti palkkoja maksavat elinkeinoharjoittajat: elinkeinoharjoittajien lukumäärät sekä palkkojen mediaanit ja keskiarvot vuosina 2018–2020.	17
Taulukko 9. Toimialat, joissa elinkeinoharjoittajia on verotarkastettu eniten. Tarkastettujen osuus kaikista toimialan elinkeinoharjoittajista vuosina 2018–2020.	24
Taulukko 10. Verotarkastuksissa havaittujen virheiden lukumäärät vuosina 2018–2020.	24
Taulukko 11. Verotarkastuksilla havaittujen virheiden suuruudet vuosina 2018–2020.	24
Taulukko 12. Arvio elinkeinoharjoittajien pimeästä työstä CEM-menetelmällä 2018–2020 (miljoonaa euroa).	27
Taulukko 13. Arvio elinkeinoharjoittajien pimeästä työstä koneoppimismenetelmällä 2018–2020 (miljoonaa euroa).....	28
Taulukko 14. Arvio elinkeinoharjoittajien pimeästä työstä profilointimenetelmällä, keskimäärin vuodessa ajalla 2018–2020 (miljoonaa euroa).....	31
Taulukko 15. Elinkeinoharjoittajien pimeän työn estimaatit eri menetelmillä laskettuina (miljoonaa euroa), vuodet 2018–2020.	34
Taulukko 16. Merkittävimmät elinkeinoharjoittajien pimeän työn toimialat vuonna 2019. Profiloinnin ja koneoppimismallin tulosten vertailu.	37

11 Liitteet

Liite 1. CEM-kaltaistaminen

Liite 2. Yrityskohtaisen odottaman laskenta

Liite 3. Profilointi

Liite 4. Ansiotulon verovajeen laskeminen

Liite 1. CEM-kaltaistaminen

Coarsened Exact Matching eli CEM on yksinkertainen ja tehokas kaltaistamisen menetelmä, jolla voidaan tasoittaa valikointiharhaa hakemalla kontrolliryhmästä (=tarkastamattomat elinkeinonharjoittajat) vastaavia havaintoyksilöitä kuin interventioryhmässä (=tarkastetut elinkeinonharjoittajat).⁸⁶ CEM-algoritmi toimii kahdessa vaiheessa.

Numeeristen muuttujien kategorisointi

Ensiksi CEM-algoritmissa kategorisoidaan kaikki valitut numeeriset muuttujat esimerkiksi kymmeneen eri tasaluokkaan ja annetaan kategoristen muuttujien olla sellaisenaan. Näin jokainen muuttuja muuttuu kategoriseksi, jossa kukin havainto kuuluu yhteen luokkaan kussakin muuttujassa. Tästä tulee algoritmin nimeen sana "coarsened" eli "karkeistettu". Numeeristen muuttujien kategorisointi mahdollistaa algoritmin seuraavan askeleen, jota kuvaa sana "exact" eli eksakti, tarkka.

Vastinparien etsiminen

Toisessa vaiheessa kullekin tarkastetulle elinkeinonharjoittajalle pyritään löytämään yksi tai useampi vastinpari, joka vastaa sitä täydellisesti kaikkien valittujen muuttujien ylitse. Esimerkiksi jos kaltaistus tehdään toimialan, palkansaajien lukumäärän sekä liikevaihdon ylitse, etsitään toimialaan D, liikevaihto-desiiliin 6 ja palkansaajien lukumäärä-desiiliin 4 kuuluvalla tarkastetulle elinkeinonharjoittajalle tarkastamattomien toiminimien joukosta yrittäjä, jonka vastaavat muuttujien arvot ovat vektorimuodossa esitettynä (D,6,4). Jos täysin eksaktia vastinparia ei löydy, paritusta ei tehdä. Jos useammalla tarkastetulla elinkeinonharjoittajalla on täysin toisiaan vastaavat arvot, niin ne etsivät yhteisiä vastinpareja tarkastamattomien joukosta. Kaltaisten elinkeinonharjoittajien etsintä tehdään kaikkien verovuosien ylitse eli vuoden 2018 verotarkastetun yrityksen tiedot voivat mennä niin vuosien 2019 kuin 2020 sitä muistuttaville tarkastamattomille yrittäjille.

Paritus tehdään 13 muuttujan ylitse, jotka on luotu pääkomponenttianalyysillä. Muuttujien määräksi valittiin 13 siksi, koska 13 ensimmäistä pääkomponenttia selittää yli 50 % alkupe-
räisessä 62 muuttujan aineistossa esiintyvistä varianssista. Jokainen näistä 13 muuttujasta kategorisoidaan 8 eri luokkaan, jotka perustuvat muuttujien tyypilliseen tilastolliseen jakaumaan.

Tietojen kopiointi kontrolliryhmään

Kun tarkastetuille elinkeinonharjoittajille on löydetty yksi tai useampi vastinpari, kopioidaan näille tarkastetun elinkeinonharjoittajan tarkastuksen lopputulos. Jos ryhmässä on kaksi tai useampia tarkastettuja elinkeinonharjoittajia, kopioidaan näiden vastinpar(e)ille tarkastustulosten lopputulosten keskiarvo.

⁸⁶ Kun interventio- ja kontrolliryhmät eivät ole valikoituneet satunnaisesti, koitetaan kaltaistamismenetelmällä löytää sellaisia vastinpareja, jotka vastaavat toisiaan ja ovat siten vertailukelpoisia. Toisin sanoen, jos on esimerkiksi verotarkastettu yritys, jonka liikevaihto on 5 miljoonaa ja toimialakirjain on D, menetelmällä pyritään löytämään vastaava yritys tarkastamattomien joukosta.

Menetelmän hyviä ja huonoja puolia

CEM-menetelmän hyvä puoli on sen yksinkertaisuus ja läpinäkyvyys. Pimeän työn määrä päätellään ei-tarkastettujen yritysten joukossa selkeästi; kopioidaan tiedot vastinparilta sellaisenaan. CEM vie myös kaltaistamisen äärimmäisen tarkalle tasolle, sillä vastinparin pitää algoritmin nimen mukaisesti olla eksakti kopio tarkastetusta yrityksestä.

Toisaalta menetelmä tekee naiiveja oletuksia kopioidessaan suoraviivaisesti pimeän työn määrän sellaisenaan tarkastetulta elinkeinonharjoittajalta. Tällä tavoin ei saada lainkaan mallinnettua pimeän työn esiintyvyyteen tai sen määrään liittyviä epävarmuuksia. Koska kussakin vastinparissa on parhaimmillaankin vain muutamia tarkastettuja yrityksiä, ei tilastollisia päätelmiä kyetä tekemään. Myöskään satunnaistarkastuksia ja niiden tarjoamaa informaatiota ei saada mitenkään hyödynnettyä.

Lisäksi menetelmä on altis isoille muutoksille riippuen montako ja mitä muuttujia kaltaistukseen käytetään, eikä muuttujien määrän tahi koostumuksen valintaan ole kovinkaan objektivisia laadun mittareita. Tässä selvityksessä päädyttiin pyrkiä hyödyntämään maksimaalinen määrä informaatiota minimaalisella määrällä muuttujia ja siksi valittiin pääkomponentti-analyysi datan esikäsittelyvaiheeksi.

Liite 2. Yrityskohtaisen odottaman laskenta

Tässä menetelmässä jokaiselle perusjoukon elinkeinonharjoittajalle päätellään erikseen kaksi suuretta, pimeän työn esiintymisen todennäköisyys sekä sen suuruus riskin toteutuessa. Kertomalla nämä keskenään saadaan jokaiselle elinkeinonharjoittajalle pääteltyä pimeän työn odottama. Arvio pimeän työn kokonaisarvosta perusjoukossa on elinkeinonharjoittajakohtaisten odottamien summa. Summaus voidaan tehdä jokaiselle vuodelle erikseen. Matemaattisesti menetelmä voidaan kuvata seuraavasti:

Asiakkaat: $1, \dots, N$

Pimeän työn esiintymisen todennäköisyys: $T_i \sim \text{Bernoulli}(p_i)$

Pimeän työn odotettu määrä: $M_i \sim \text{Gamma}(k_i, \theta_i)$

Pimeän työn odotusarvo: $E_i = T_i M_i$

Pimeän työn kokonaismäärä: $L^N = \sum_{i=1}^N E_i$

On hyvä huomioida, että saatua odottamaa ei voi suoraan käyttää arviona sille, kuinka paljon pimeänä teetettyä työtä löytyisi, jos elinkeinonharjoittaja tarkastettaisiin. Enemminkin se kertoo, että jos 1 000 täysin vastaavaa elinkeinonharjoittajaa tarkastettaisiin, kuinka paljon voitaisiin keskiarvoisesti olettaa pimeää työtä löytyvän per havainto.

Käytännössä asiakkaan riskisyyden arvio tehdään hyödyntämällä verotarkastuksien sekä säännönmukaisen valvonnan tuloksia, joista tiedämme, löytyikö asiakkaalta tarkastuksessa pimeänä teetettyä työtä vai ei ja jos kyllä, kuinka paljon tuloja on jäänyt ilmoittamatta. Verotarkastukset hyödynnetään sellaisenaan, mutta säännönmukaisen valvonnan aineistosta otetaan mukaan vain ne tapaukset, joilta virkailija on pyytänyt lisätietoja, mutta joille ei ole tehty muutoksia. Säännönmukaisen valvonnan tiedot täydentävät verotarkastusten pientä otoskokoja ja myös huomattavasti maltillistavat osumatarkkuutta. Malli, jonka näkemässä aineistossa reilusti yli puolet ovat olleet harmaan talouden toimijoita, ei voi tehdä realistisia arvioita pimeän työn esiintyvyydestä koko elinkeinonharjoittajien populaatiossa.

Menetelmässä hyödynnetään niin pimeän työn esiintyvyyttä (binäärinen 0/1 luokittelu) kuin niiden euromäärää vastemuuttujana (regressiomalli), ja rakennamme siis kaksi erillistä mallia pimeän työn odotusarvon päättelemiseksi.

Asiakaskohtainen pimeän työn todennäköisyys

Elinkeinonharjoittajan todennäköisyys tehdä pimeää työtä päätellään koko populaatiolle rakentamalla binäärisesti luokitteleva malli tarkastettujen elinkeinonharjoittajien populaation sisällä, joka pyrkii ennustavien muuttujien avulla erottelemaan pimeää työtä tekevät elinkeinonharjoittajat muista. Vasteena toimii binäärinen 0/1-muuttuja, jossa yhden euronkin kokoinen tarkastuslöydös lasketaan pimeän työn tekemiseksi. Valitsemalla sellainen joukko pimeän työn esiintyvyyttä ennustavia muuttujia, jotka voidaan yhtäläisesti päätellä sekä tarkastetuille että tarkastamattomille elinkeinonharjoittajille (kuten esimerkiksi Verohallinnolle ilmoitettujen palkkailmoitusten mukaan päätelty yrityksen työntekijöiden keskipalkka), voidaan myös tarkastamattomien elinkeinonharjoittajien riskisyyttä arvioida opitun

matemaattisen mallin avulla. Rakennetuilla malleilla voidaan siis arvioida koko laajemman populaation riskisyyttä, mutta tietyillä aineiston ja menetelmän asettamilla reunaehdoilla.

Pimeän työn esiintyvyyttä ennustamaan rakennettiin päätöspuihin perustuva XGBoost-malli sekä yhden kerroksen neuroverkon ns. ensemble-malli, joka yhdistelee useiden näiden kahden algoritmin tekemän erilaisen mallin ennusteita. XGBoost on hyvin tehokas ohjatun oppimisen algoritmi, joka perustuu siihen, että rakennetaan sarja pieniä päätöspuita, joista kukin päätöspuu koittaa korjata edellisen tekemiä virheitä. Se on epäparametrinen "musta laatikko"-algoritmi, jonka ennusteiden selittämiseksi tarvitaan erillisiä selkoälyn (=Explainable AI eli XAI) menetelmiä. Neuroverkot taas ovat malleja, joiden perusajatus perustuu ihmisaivojen rakenteeseen. Monet nykyiset tekoälysovellukset varsinkin tekstinkäsittelyn ja kuvantunnistuksen alalla pohjautuvat niin sanottuihin syviin neuroverkkoihin, joissa niin sanottuja datan käsittelyn kerroksia on useampia. Tässä tapauksessa tyydytään rakentamaan yksinkertainen yhden kerroksen neuroverkkomalli.

Malli toimii tarkastettujen elinkeinonharjoittajien populaation sisällä lähes hämmästyttävän hyvin, ja sen AUC on noin 0,95. Malli pystyy tekemään selvän eron tarkastettujen valkoisten ja harmaiden elinkeinonharjoittajien välille. Hyvin luultavasti mallin ennustevoimaa inflatoi se tosiasia, että aineisto koostuu niin tarkastetuista kuin säännönmukaisessa valvonnassa valvotuista yrityksistä, ja ainoastaan ensin mainittujen joukossa voi esiintyä pimeää työtä. Tosin selvityksen aikana tehdyt kokeilut pelkkään säännönmukaisen valvonnan aineistoon tuottivat todella hyviä tuloksia, joten valvottujen ja tarkastettujen alajoukoista malli selkeästi kykenee erottelemaan harmaan toimijan valkoisesta. Yleistys laajempaan joukkoon on kuitenkin aivan oma asiansa, jonka validointiin voidaan käyttää vain asiantuntijaosaamista. Testiaineistoon tehdyistä ennusteista ei siis voida vetää johtopäätöksiä siitä, kuinka hyvin malli toimii aidosti, kun kohdejoukkona on koko elinkeinonharjoittajien populaatio.

Asiakaskohtainen pimeiden palkkojen suuruus

Regressiomalli rakennetaan tarkastettujen elinkeinonharjoittajien kohdejoukon siihen osajoukkoon, joilta on tarkastuksessa löydetty pimeää työtä. Virheen suuruutta arvioidessa ei ole kannattavaa huomioida niitä valvottuja elinkeinonharjoittajia, joilla pimeää työtä ei ole havaittu. Jos nämä havainnot otettaisiin mukaan harjoitusaineistoon, tulisi virheen määrän arvio auttamatta epärealistisen pieneksi. Koska lopullinen pimeän työn odottama koostuu virheen koosta sekä sen todennäköisyydestä, tulee elinkeinonharjoittajan vähäriskisyys huomioitua jo riskin todennäköisyyden arvioissa.

Regressiomallin algoritmiksi valikoitui XGBoost. Selvityksessä kokeiltiin myös erilaisia ensemble-kokoonpanoja, mutta puhdas XGBoost näytti antavan realistisimmat arviot pimeän työn määrästä. Valvottujen harmaiden elinkeinonharjoittajien joukosta otetulle testidatan siivulle tehdyt ennusteet paljastavat, että regressiomalli toimii jopa yllättävän hyvin. Mallin keskimääräinen virhe MAE (Mean Absolute Error) on noin 13 700 euroa eli keskimääräinen ennuste osuu vajaan 14 000 euron päähän kohteestaan. Keskimääräinen tarkastustulos on vajaan 50 000 euroa, joten ennustetta voidaan pitää kelvollisen hyvänä.

Menetelmän vahvuudet ja heikkoudet

Asiakaskohtaisen riskisyyden päättelyn paras puoli on, että jokaiselle elinkeinonharjoittajalle saadaan yksilöllinen arvio sen pimeän työn teettämisen riskisyydestä sekä arvio sen koosta. Yksilötasolle viety arvio ottaa jokaisen elinkeinonharjoittajan ominaisuudet tarkalla tasolla huomioon ja mahdollistaa myös monipuolisen aineiston jatkojalostamisen. Menetelmän etuna on lisäksi se, että se on käytetyistä menetelmistä ainoa, josta voidaan saada objektivista palautetta sen toimivuudesta; tosin ennustetehoa voidaan testata ainoastaan tarkastettujen elinkeinonharjoittajien osajoukossa, jolloin malli operoi omassa, pienessä, vinoutuneessa todellisuudessaan, eikä ennustetehon metriikoita voida yleistää pätemään koko populaatiossa.

Menetelmän suurin heikkous on siinä, että se on altis harjoitusaineiston vinoumille, jotka kerätyvät koko elinkeinonharjoittajien populaatioon ennustettaessa. Siinä missä aiemmin esiteltyt kaltaistamisen menetelmät ovat nimenomaisesti suunniteltu poistamaan valikoitumisharhaa, joudutaan asiakaskohtaista odottamaa laskiessa pyrkiä vähentämään sitä erilaisin keinoin, esimerkiksi pyrkien hyödyntämään sellaisia muuttujia, joita ei käytetä suoraan tarkastuksen kohdevalinnassa.

Liite 3. Profilointi

Profilointimenetelmä on kehitetty ja ollut käytössä Verohallinnon Harmaan talouden selvitysyksikössä. Menetelmästä on käytetty myös nimeä riskiominaisuusvertailu. Sen avulla voidaan mm. laskea y-tunnuksille riskipisteitä erilaisten veronvälttämistapojen suhteen.

Riskipisteytys ja riskiyritysten valinta

Riskipisteytyksessä virheitä tehneiden elinkeinonharjoittajien (vertailujoukko) ominaisuuksia verrataan muihin elinkeinonharjoittajiin (perusjoukko). Ne perusjoukon elinkeinonharjoittajat, joilla on samankaltainen ominaisuusprofiili kuin vertailujoukon elinkeinonharjoittajilla, saavat korkeita riskipisteitä.

Jokainen ominaisuus pisteytetään laskentakaavalla (1). Jokaiselle elinkeinonharjoittajalle lasketaan henkilökohtaiset riskipisteet summaamalla toteutuneiden ominaisuuksien pisteet. Ominaisuus toteutuu elinkeinonharjoittajalla, jos sen arvo täyttää vaaditun raja-arvoehdon. Esimerkiksi raja-arvoehto voi olla: maksujen määrä on alle 1 000 €. Jos elinkeinonharjoittajan maksut alittavat 1 000 €, niin ominaisuus toteutuu ja elinkeinonharjoittaja saa tätä ominaisuutta vastaavat kaavalla (1) lasketut riskipisteet.

$$(1) \text{ Ominaisuuden pisteluku} = (P_{\text{ref}}) \times (1 - P_{\text{pop}}) \times (H_{\text{ref}} / H_{\text{pop}})$$

Missä

- Esiintyvyys (P) yleisesti = H / N
- H = osumien lkm (ne elinkeinonharjoittajat, joilla ominaisuus toteutuu)
- N = ko. joukon kaikkien yritysten lkm (vertailujoukko tai perusjoukko)
- P_{ref} = ominaisuuden esiintyvyys vertailujoukossa
- P_{pop} = ominaisuuden esiintyvyys populaatiossa
- H_{ref} = osumien lukumäärä vertailujoukossa
- H_{pop} = osumien lukumäärä populaatiossa

Ohjelmaan on syötetty suuri joukko erilaisia ominaisuuksia, mutta niitä kaikkia ei huomioida pistelaskussa. Jotta ominaisuus huomioitaisiin, sen tulee olla "yleisempi" vertailujoukossa kuin perusjoukossa. Pistelaskussa käytetty ehto on, että esiintyvyys vertailujoukossa (P_{ref}) tulee olla vähintään kaksinkertainen verrattuna perusjoukkoon (P_{pop}).

$$(2) \text{ Esiintyvyysehto} = P_{\text{ref}} / P_{\text{pop}} \geq 2$$

Itse raja-arvo valitaan siten, että se maksimoi esiintyvyyksien tulon ($P_{\text{ref}}) \times (1 - P_{\text{pop}})$ esiintyvyysehdon ollessa voimassa. Kaavan muodossa:

$$(3) \text{ Raja-arvo} = \max (P_{\text{ref}}) \times (1 - P_{\text{pop}}) \Leftrightarrow \text{ ehdolla } P_{\text{ref}} / P_{\text{pop}} \geq 2$$

Käytännössä pisteytys toimii siten, että ne perusjoukon elinkeinonharjoittajat, joilla on samankaltainen ominaisuusprofiili kuin vertailujoukolla, saavat korkeita riskipisteitä. Korkeita riskipisteitä saaneet perusjoukon elinkeinonharjoittajat muodostavat niin sanotun riskiyritysten joukon. Ominaisuudet, jotka esiintyvät yleisesti sekä perusjoukossa että vertailujoukossa, eivät kasvata riskipisteitä.

Lukumääräarvion muodostaminen

Lukumääräarvio muodostuu kahdesta osasta, joista toisen muodostaa edellä mainitut (a) havaitut riskiyrietykset. Toinen osa muodostuu niistä elinkeinonharjoittajista, jotka (b) jäävät pisteityksessä "piiloon".

(a) Edellä jo todettiin, että riskiyrietykset ovat korkeita riskipisteitä saaneita elinkeinonharjoittajia. Lukumääräarvion kannalta keskeinen ongelma on, mihin valintaraja eli niin sanottu katkaisukohta asetetaan. Mistä tiedämme, kuinka paljon perusjoukon elinkeinonharjoittajia tulee valita lukumääräarvioon?

Katkaisukohtat valitaan erikseen jokaiselle laskennan osajoukolle.⁸⁷ Valinnassa käytetään apuna lineaarista regressiota, jossa selittävänä muuttujana on tunnusluku ("erotteluvoima"), joka kuvaa vertailujoukon ja perusjoukon riskipistejakaumien päällekkäisyyttä. Vastemuuttuja on suhdeluku, joka kuvaa katkaisuun kuuluvien osuutta vertailujoukossa. Mitä paremmin menetelmä onnistuu pisteityksen avulla erottelemaan perusjoukon ja vertailujoukon toisistaan, sitä suurempi osa vertailujoukon yrityksistä ylittää katkaisurajan. Kääntäen, jos erottelu on heikkoa, niin iso osa vertailujoukosta saa matalia pisteitä perusjoukon tavoin.

Malli tarjoaa luontevia katkaisukohtia⁸⁸, mutta asiantuntijan valinnoilla voidaan vaikuttaa regressiosuoraan. Esimerkiksi asiantuntija voi arvioida eri osajoukkojen välistä riskisyyttä tai sitä kuinka paljon osajoukkojen riskisyyden tulisi vaihdella. Vertailujoukolle valittua katkaisurajaa sovelletaan myös perusjoukkoon, ja näin muodostuu riskiyrietysten joukko.

(b) Useimmiten tilanne on pisteityksessä sellainen, että osa kohdejoukon elinkeinonharjoittajista jää katkaisurajan alapuolelle. Toisin sanoen ne eivät ylitä kynnystä, jolla riskiyrietykset valitaan. Tällöin on syytä olettaa, että myös perusjoukossa on elinkeinonharjoittajia, jotka ovat "harmaita" mutta eivät tule valituksi riskiyrietysten joukkoon. Lukumääräarvio sisältää arvion myös näistä piiloon jääneistä perusjoukon elinkeinonharjoittajista. Tällöin oletetaan, että perusjoukossa ja vertailujoukossa piiloon jääneiden osuus on sama. Vertailujoukosta tämä suhdeluku tunnetaan, joten sitä voidaan soveltaa myös perusjoukkoon.

Menetelmän hyviä ja huonoja puolia

Estimoinnissa riskiperusteisen verotarkastusten valintaharhaa vähennetään jakamalla perusjoukko osiin luokittelevien tietojen perusteella, kuten erilaiset virheet sekä toiminnan laajuus (liikevaihto). Lisäksi monet selittävät muuttujat lasketaan suhdelukuina (esimerkiksi keskipalkka), mikä osaltaan torjuu kohdevalinnan kokoharhaa. Tarkastusten keskittymistä harvoille toimialoille torjutaan sillä, että useiden selittävien muuttujien arvoja verrataan kunkin yrityksen toimialan mediaaniin.

Lukumääräarvio muodostuu osittain menetelmän tuottamien tulosten perusteella, mutta siihen voidaan vaikuttaa myös asiantuntijavalintojen kautta. Tätä subjektiivista elementtiä voidaan

⁸⁷ Osajoukkoja muodostuu käytettyjen taustaluokittelujen avulla: koko, tyyppi eli tekoapa ja näiden yhdistelmä koko-tyyppi.

⁸⁸ Usein katkaisu tehdään X- eli riskipisteakselilla kohdasta, jossa perusjoukon lukumäärää kuvaava jakauma kääntyy jyrkkään kasvuun.

pitää yhtä lailla menetelmän heikkoutena tai vahvuutena. Sillä voidaan torjua täysin epärealistisia tuloksia, mutta tulos voi ohjautua myös väärään suuntaan.

Liite 4. Ansiotulon verovajeen laskeminen

Profiloinnin kokonaisestimaatti jaettiin veronmenetyksen laskemista varten kolmeen osaan: elinkeinonharjoittajien ansiotulo ja pääomatulo sekä palkansaajien pimeä työ. Veronmenetykset elinkeinonharjoittajien pimeistä pääomatuloista ja palkansaajien pimeästä työstä lasketaan suoraviivaisesti kertoimilla vastaavista pimeiden palkkojen estimaateista.⁸⁹ Sen sijaan elinkeinonharjoittajien ansiotulojen verovaje vaatii monimutkaisemman laskutoimituksen.

Ansiotulot verotetaan valtionverotuksessa progressiivisen asteikon mukaan, eli tulojen noustessa myös veroprosentti nousee. Progressiivinen veroprosentti huomioitiin verovajeen laskennassa seuraavasti: ansiotulon verovaje laskettiin jakamalla ensin estimaatti (taulukko 17, sarake 2) luokkiin, jossa huomioitiin liikevaihdon mukainen koko ja elinkeinonharjoittajien päätoimisuus tai sivutoimisuus (sarake 1). Eri luokkiin kuuluville elinkeinonharjoittajille laskettiin keskimääräiset ansiotulot (sarake 3),⁹⁰ johon lisättiin arvioitu pimeän työn mediaani (sarake 4). Näin muodostettiin se ansiotulojen taso (sarake 5), jonka mukaan veroprosentti määräytyi. Veroprosentilla (sarake 6) kerrottiin pimeän työn estimaatit luokittain ja siten saatiin arvio menetetyistä ansiotuloveroista (sarake 7).

Taulukko 1. Veronmenetyksen laskeminen ansiotulojen pimeän työn estimaatista.

1. Luokittelu: päätoimisuus + liikevaihto	2. Ansiotulojen osuus pimeän työn estimaatista	3. Veronalaiset Ansiotulot ka.	4. Virheen mediaani	5. Arvioitu ansiotulo [3 + 4]	6. Vero%	7. Menetty vero [2 x 6]
päätoim. 0€	1 993 000	4 300	19 300	23 600	0,171	341 000
päätoim. alle 10t€	23 860 000	5 400	19 300	24 700	0,180	4 295 000
päätoim. 10-50t€	78 934 000	14 600	19 300	33 900	0,238	18 786 000
päätoim. 50-100t€	82 200 000	31 100	23 400	54 500	0,333	27 373 000
päätoim. 100t-1m€	45 732 000	46 100	28 000	74 100	0,385	17 607 000
päätoim. yli 1m€	3 941 000	161 300	64 000	225 300	0,512	2 018 000
sivutoim. 0€	15 010 000	65 098	20 000	85 100	0,400	6 004 000
sivutoim. alle 10t€	12 414 000	42 110	6 500	48 600	0,314	3 898 000
sivutoim. 10-50t€	24 072 000	42 169	14 000	56 200	0,340	8 185 000
YHTEENSÄ	288 156 000					88 505 000

Lähde: Verohallinto 2–6/2022.

⁸⁹ Pääomatulojen estimaatti kerrotaan kertoimella 0,3 ja palkansaajien pimeän työn estimaatti kertoimella 0,28.

⁹⁰ Keskimääräiset ansiotulot laskettiin profilointimenetelmän valitseville riskitapauksille. Vertailun vuoksi keskimääräiset ansiotulot laskettiin myös koko perusjoukon elinkeinonharjoittajille. Riskitapausten ansiotulot olivat keskimäärin matalammat kuin koko elinkeinonharjoittajien perusjoukossa.