Statistisk sentralbyrå
Statistics Norway

Public disclosure of tax information: Compliance tool or social network?

Daniel Reck, Joel Slemrod, Trine E. Vattø

TALL SOM FORTELLER

DISCUSSION PAPERS

975

Discussion Papers No. 975, March 2022
Statistics Norway, Research Department

Daniel Reck, Joel Slemrod, Trine E. Vattø

Public disclosure of tax information: Compliance tool or social network?

Abstract:
We conduct the first-ever study of actual searches done in a public tax disclosure system, analyzing about one million searches done in 2014 and 2015 in Norway. We characterize the social network these searches comprise, including its degree of homophily and reciprocation, and the demographics of targets and searchers. About one-fourth of searches occur within identifiable household and employment networks. Most searchers target people similar to themselves—homophily in network parlance—but young, low-income searchers also target older, successful people and celebrities. A causal research design based on the timing of searches relative to tax filing uncovers no evidence that, upon discovering they were targeted, targets subsequently increase their reported income. The evidence suggests that social comparisons motivate the bulk of searches rather than tax compliance. However, public disclosure may deter evasion even when compliance-motivated searches are rare in equilibrium.

Keywords: Public disclosure, social network, tax compliance

JEL classification: H26, D83, D85

Acknowledgements: We are grateful for collaboration at an early stage of this project from Camilla Karto Abrahamsen. We are thankful for to Oslo Fiscal Studies for financial support and to the Norwegian Tax Administration for access to the data. We also thank for helpful comments Annette Alstadsæter, Guy Michaels, Thor O. Thoresen, and seminar participants at the Bank of Israel, Bocconi University, and the London School of Economics.

Address: Trine E. Vattø, Statistics Norway, Research Department. E-mail: trine.vatto@ssb.no

Daniel Reck, London School of Economics, d.h.reck@lse.ac.uk

Joel Slemrod, University of Michigan, jslemrod@umich.edu

Discussion Papers comprise research papers intended for international journals or books. A preprint of a Discussion Paper may be longer and more elaborate than a standard journal article, as it may include intermediate calculations and background material etc.

The Discussion Papers series presents results from ongoing research projects and other research and analysis by SSB staff. The views and conclusions in this document are those of the authors

Abstracts with downloadable Discussion Papers
in PDF are available on the Internet:
http://www.ssb.no/en/forskning/discussion-papers
http://ideas.repec.org/s/ssb/dispap.html

ISSN 1892-753X (electronic)

Sammendrag

Norge har lange tradisjoner for offentliggjøring av skattelister, med begrunnelse i at åpenhet og transparens er viktig for tilliten til skattesystemet. Samtidig har man ønsket å redusere personvernulempene ved å innskrenke spredningen av informasjon fra skattelistene på internett. Et ledd i dette var at det, fra og med inntektsåret 2013, ikke lenger var mulig å søke anonymt på skatteetatens sider, men at den som ble søkt på kunne få informasjon om hvem som søkte ved å logge seg inn på skatteetatens sider.

I denne studien benytter vi unike data fra Skatteetaten for det første året med ikke-anonyme personsøk i skattelistene; Til sammen har vi informasjon om rundt 1 million søk for 2013. Ved å koble søkene med registerdata, kan vi studere hva som karakteriserer søkere og de som blir søkt på, samt kartlegge relasjoner mellom de to partene.

Studien bidrar til den internasjonale litteraturen på to felt. For det første knytter vi våre funn opp mot litteraturen som omhandler sosiale nettverk. Vi finner blant annet at omtrent en fjerdedel av søkene skjer innenfor identifiserbare husholdnings- og arbeidsnettverk. De fleste søkere retter seg mot personer som ligner dem selv – såkalt homofili i sosiale nettverk – men unge søkere med lav inntekt er også opptatt av å finne opplysninger om eldre, «vellykkede» personer og kjendiser.

For det andre bidrar vi til litteartruen om hvorvidt offentlige skattelister påvirker skatteunndragelse. I likhet med flere andre analyser av skatteunndragelse, antar vi at siden selvstendig næringsdrivende selvrapporterer inntekt, er det større muligheter for å unndra skatt i denne gruppen. Vi bruker paneldata og forsøker å avdekke en årsakssammenheng ved å sammenlikne inntektsrapportering i neste års skattemelding (daværende «selvangivelse») for de som blir søkt på rett før og rett etter fristen for innlevering av skattemeldingen. Vi finner ingen signifikante effekter på rapportert inntekt av å bli søkt på, verken for selvstendig næringsdrivende eller for lønnstakere.

Vi konkluderer, ikke overraskende, med at mesteparten av søk i skattelistene ser ut til å være motivert av sosiale sammenligninger av inntekt og formue. Selv om vi ikke finner noen direkte mekanisme mellom søk på personer og deres rapportering av inntekt, kan vi likevel ikke se bort ifra at muligheten for å bli søkt på, kan virke disiplinerende i seg selv og dermed bidra til å forebygge skatteunndragelse.

1. Introduction

Several countries offer their citizens the opportunity to learn about the reported income (and, in some cases, wealth) and tax liability of their fellow citizens. Public disclosure of tax return information is usually justified as helping to ensure better tax compliance and to make the tax system more transparent. Even if such tax policy objectives underlie public disclosure, citizens likely make use of the information to satisfy their curiosity about others as well as to benchmark their own financial and tax situations against others.

In this paper, we study the micro-structure of how people use public tax information. There is a recent literature, reviewed in detail below, studying the impact of tax-return disclosure on a number of outcomes, but no research has examined who searches and who is targeted. This paper presents the first analysis of citizen-to-citizen search patterns, making use of newly available data from Norway. In Norway, with some limitations, people may search via the Internet to see what other Norwegians declare as taxable income, taxable wealth and tax liability. We obtained data on every search done in 2014 and 2015 querying tax year 2013 information—who searched for whom, and when. We merge this information with administrative records of individual demographic characteristics to construct a cross-sectional dataset of the network of all searches, allowing us a unique opportunity to study the consequences of public tax disclosure. An important feature of searches done in this period is that the targets of searches could observe, by logging into their account on the tax authority's website, who had searched for them. Thus, they could reciprocate the search and/or increase their reported income if they were concerned about whistleblowing that would reveal tax evasion to the authorities. Naturally, this fact also implies that our findings characterize searches under a non-anonymous search regime, and we do not know how the results would differ under anonymity.

We use descriptive and causal analysis to understand the main motivations for searches. Based on existing literature, we focus on two broad possibilities. First, individuals may use searches to facilitate social comparisons, which could explain the estimated effects of Norwegian public disclosure on subjective well-being (Perez-Truglia, 2020) and job quitting (Rege and Solli, 2013). Second, individuals may use searches to check whether others are truthfully reporting to the tax authority, which could explain the effects of public disclosure on tax reporting behavior (Bø, Slemrod, and Thoresen, 2015). In the end, we conclude that the bulk of searches are motivated by some form of social comparison. As such, we also interpret demographic patterns in searches as informative about to whom individuals choose to compare themselves. Despite a large literature on social comparisons, the evidence on which actual individuals or groups form the basis for such comparisons is scant, consisting of only survey evidence (Clark and Senik, 2010, Perez-Truglia, 2020). Our data provide a unique opportunity to shed light on this question with observational data.

We begin by characterizing in detail the social network these searches comprise and the characteristics of the individuals initiating a search (*searchers)* and the individuals whose information was queried by a searcher (*targets*). Our main findings regarding social comparisons involve the joint distribution of searcher-target characteristics. We find that approximately one-fourth of searches occur within identifiable household or employment networks. Searcher-target pairs are far more likely to have similar characteristics than two random individuals from the population, for virtually any characteristic we observe, a property known as *homophily* in network analysis. Our findings of substantial searching within household and employment networks, and strong homophily are both consistent with the survey evidence on social comparisons in Clark and Senik (2010) and Perez-Truglia (2020). We also document that 6.3 percent of searches are reciprocated, meaning that the targets search for those who searched for them. Frequent reciprocity is reminiscent of the "reciprocity norms" studied by Cullen and Perez-Truglia (2018b). In addition, we document one more data pattern that is not covered by any prior literature: young, low-income searchers frequently target older, highly successful people and celebrities. This is suggestive of an "aspirational" motive for social comparisons, perhaps along the lines of Hirschman and Rothschild (1973).

We then turn to assessing the role of the disclosure system in ensuring tax compliance. We find that the tax information of self-employed people is, ceteris paribus, substantially more likely to be targeted. Given that in Norway third-party information reporting severely limits evasion possibilities for most employees, this suggests that potential whistleblowing is a non-trivial motivation for searching. Nevertheless, about 90 percent of searches target wage earners, who have little capacity for evasion. Finally, we conduct a well-identified analysis of how being searched for affects subsequent income reporting. The research design leverages the fact that searches before tax filing in a given year can influence reporting behavior in that year, while searches after tax filing cannot. We find small and insignificant effects of being targeted on tax reporting behavior, even for self-employed individuals. From this we infer that being targeted does not promote tax compliance for the vast majority of individuals.

Overall, the evidence suggests that much of the utility of public tax information to individuals derives from the opportunity to learn about the incomes of others for non-tax reasons, most plausibly to engage in social comparisons. Nevertheless, much of the effect of public disclosure on tax compliance may be coming from the *availability* of information rather than whether information is actually accessed in equilibrium.

2. Related Literature

This research touches on many different active literatures in economics. In what follows, we briefly discuss these connections.

2.1 Networks in theory

The pattern of connections among people through Internet searches is, of course, a network, and networks are a lively field of study among theorists and empirical researchers, as recently surveyed in Jackson (2008, 2014) and Jackson, Rogers, and Zenou (2017). They stress that the full network of relationships affects how information spreads and how people behave. This literature has emphasized the importance of, inter alia, network *density* (the proportion of potential connections that are actual connections); *homophily* (the tendency of nodes in a network to link with similar others); *clustering* (the tendency of nodes to cluster together); *centrality* (which nodes are most connected to others); and the endogeneity of network formation. It provides several related measures of centrality in particular. How many inbound links one has (whether a person is a target, the number of times a person is targeted by search) is known in the network literature as *in-degree centrality*, and how many outbound links one has (whether a person searches, the number of searches a person does) is a measure of *out-degree-centrality*. In-degree centrality is associated with prestige, popularity, and prominence, while out-degree centrality is associated with influence and gregariousness or, in our context, curiosity or meddlesomeness.[1] Our analysis characterizes all of these features of the network of searches for tax information.

Searching for tax information with the possibility of whistleblowing upon encountering suspected tax evasion is an example of what Fehr and Gächter (2000) call negative reciprocity – when people take "revenge even in interactions with complete strangers and even if it is costly for them and yields neither present nor future material rewards." They discuss the extensive experimental studies that document the presence of negative reciprocity in many settings.

[1] We expect that in the tax disclosure network the correlation between one's in-degree centrality and out-degree centrality is relatively low—the curious are different folks from the prominent. As such we do not make use of more sophisticated centrality measures, which would use, for instance, which individuals are more commonly searched for by oft-searched-for individuals.

2.2 Social comparisons

Making tax information public creates a new avenue for social comparisons. A large literature has investigated the impact on one's own well-being and behavior of learning more about others' income by virtue of social comparisons. The title of a seminal work on the subject (Frank, 1985), "Choosing the Right Pond," alludes to the fact that individuals may choose with whom to compare. However, the existing empirical literature is primarily concerned with the causal effect of (learning about) the outcomes of a specified reference group on well-being and behavior, rather than the choice by individuals of a comparison group.

Prior literature identifies two primary reasons why learning others' income might affect one's own well-being: the "relative income hypothesis," and "tunnel effects." The relative income hypothesis channel posits that believing that one is better (or worse) than some other(s) directly makes the individual feel better (or worse) about their own circumstances. Following classic work by Duesenberry (1949), a large empirical literature examines whether such causal effects exist. For example, Luttmer (2005) finds that, holding own income fixed, living near higher earners predicts lower levels of subjective well-being. Perez-Truglia (2020) shows that the increased transparency of the Norwegian income-tax-return data beginning in 2001 was associated with a 21-29 percent increase in the gap in subjective well-being measures between richer and poorer individuals. He also finds evidence that, holding one's income constant, subjective well-being changes with the mean income in a reference group.

An alternative channel is proposed by Hirschman and Rothschild (1973): learning others' incomes might be informative for one's own future prospects. Hirschman and Rothschild used a "tunnel" metaphor to illustrate the idea. A driver stuck in traffic in a two-lane tunnel observes that traffic in the other lane starts to move faster. The driver could feel worse about his or her own situation by comparison to drivers in the other lane – the type of comparison studied in the relative income hypothesis literature. But seeing the other lane move could alternatively signal that our driver's lane will start moving soon, elevating the driver's mood. This channel has received comparatively less attention in the literature than the relative income hypothesis, but suggestive evidence for such tunnel effects is documented by Senik (2004, 2008) and, in a laboratory setting, by Konrad, Lang, and Morath (2015).

A closely related literature examines whether learning others' income affects labor market behaviors (Card et al., 2012; Cullen and Perez-Truglia, 2018a). The findings of these studies suggest that learning that coworkers' salaries are relatively higher decreases work effort and/or job satisfaction. Rege and Solli (2013) find that the disclosure of tax records in Norway increased the probability of quitting for workers with lower salaries within their occupation. We note that either of the above mechanisms for how social comparisons affect subjective well-being could explain these effects.

What do these patterns imply for the motivation underlying comparisons? Formalizing a prediction for the composition of comparison groups from prior work on the relative income hypothesis is difficult. In simple models of social comparisons, and more general theories of reference dependence, utility is lower when the outcomes of the reference group improve (e.g., Leibenstein, 1962; Kahneman and Tversky, 1979). This suggests that individuals engaging in a motivated decision of whom to compare to would target those with whom the searchers expect to compare favorably (Reck and Seibold, 2021). A naive theory of chosen comparison groups along these lines would make an extreme prediction: individuals will compare themselves to the poorest, most miserable people they can find. This seems unrealistic, so there must be some constraint, as yet not formalized, to the composition of individuals' comparison groups for a theory along these lines to make sense. In contrast, the tunnel effects idea provides an imprecise, but possibly useful, prediction: individuals should compare themselves to those whose incomes are the most informative about their future prospects. It is, however, difficult to directly test this proposition given that we cannot observe people's beliefs about which other individuals' incomes would be considered to be most informative.

In summary, neither of the two channels can be directly tested with observational data, but the micro-demographics of search provide some clues. Presumably, both hypotheses can explain why individuals would search for others similar to themselves, especially their coworkers. Under the relative income hypothesis, homophily in the network of searches must derive from the constraints on the composition of the reference group: perhaps people only derive relative income payoffs from comparing to those with whom they frequently interact. Under the tunnel effects hypothesis, the incomes of peers are plausibly informative about individuals' potential future prospects.

The fact that young, low-income people search for higher-income people and stars in the network, however, is difficult to rationalize under the relative income hypothesis as it is classically posed. Doing so would be heavily penalized by the disutility of comparing oneself to highly well-off individuals. The tunnel effects hypothesis provides a plausible explanation: these may be aspirational searches from young people who hope to attain the level of success of their targets someday. However, frequent searches for stars in particular only fit this narrative if these young people believe they might become similarly successful someday (perhaps by emulating them). A natural alternative is that learning about celebrities – even their taxable income – provides some kind of entertainment value.[2]

2.3 The role of public disclosure in tax compliance

The canonical model of tax evasion presumes that taxpayers behave as free riders such that no one contributes willingly to the public good absent enforcement. What limits evasion and ensures revenue collection is the threat of punishment of evasion—a deterrence motivation. The deterrence model takes as exogenous taxpayers' perceptions of the chance of getting caught, perhaps conditional on one's own and others' behavior. How taxpayers form these perceptions is, however, not well understood. Hoopes, Reck, and Slemrod (2014) examine one information diffusion mechanism, Internet-based searches for tax information. In part, the word gets out through communication among taxpayers – social learning.

Several recent studies have investigated the impact on tax compliance of public tax disclosure regimes—in Australia, Japan, Norway, and Pakistan—without investigating the micro interactions, network and other, that connect the regime to the perceptions of taxpayers about the enforcement environment. The studies of Japan and Australia uncover no evidence of disclosure reducing noncompliance, but do find that some taxpayers take actions to avoid the disclosure, which in these countries applies only to taxpayers above an income threshold. However, the studies of Norway and Pakistan did find that disclosure increased reported income for self-employed people.[3] Of most relevance for this paper is the study by Bø, Slemrod, and Thoresen (2015) of the Norwegian public tax disclosure system. They examine the effect on income reporting of the disclosure data being made accessible on the Internet, making use of the fact that prior to 2001, in some municipalities, tax information was already distributed widely through locally produced paper catalogs. The analysis reveals that there was an approximately 3 percent higher average increase in reported income among business owners living in areas where the switch to Internet disclosure represented a large change in access. Of note is that the Bø, Slemrod, and Thoresen study covered a period when searches could be made anonymously, while this paper concerns a period after this anonymity had been removed.

[2] A small literature in psychology studies the relationship between celebrity admiration and subjective well-being, with a number of studies concluding that such admiration decreases well-being. For example, Aruguete et al. (2019) find that celebrity admiration is negatively correlated with some predictors of life satisfaction, but they find the opposite for a measure of curiosity, which predicts life satisfaction is positively correlated with admiration for celebrities.

[3] The studies are Hoopes, Robinson, and Slemrod (2018) for Australia; Hasegawa et al. (2011) for Japan; and Slemrod, Ur Rehman, and Waseem (2020) for Pakistan.

There is some research about the effect of one link in the relationship of public disclosure to compliance, tax whistleblowers. Amir, Lazar, and Levi (2018) find that Israel's tax collections significantly increased after the introduction of a whistleblowing mechanism, in spite of a small direct tax take from evasion uncovered via the whistleblowing itself.[4] In support of the hypothesis that more salient deterrence led to the increase in tax collections, the authors note that collections increased in industries with high tax-evasion risk, but not in industries with low tax-evasion risk. Furthermore, the increase in tax collections occurred in corporations, where the timing and magnitude of tax payments are more discretionary, but not from employees.

3. The Income Tax and Public Tax-Return Disclosure in Norway

3.1 The Norwegian income tax system

Norway has a dual income tax system, with a graduated rate structure for labor income and pensions and a flat rate tax on capital income. In the period we study, the general income tax was 27 percent, with two additional brackets of 36 percent and, at the top, 39 percent for labor income. If one includes the national insurance contribution of 8.2 percent, the combined top marginal tax rate on salary income was 47.2 percent. There was also an annual wealth tax, whose top marginal tax rate was 0.85 percent on net assets exceeding NOK 1.48 million (about $191,000).[5]

Employers must withhold and remit tax for employees. In March/April following the tax calendar year, the tax authority provides people with a "pre-filled" tax return that lists what the tax authority knows from third-party information reports regarding income, deductions, assets, and debts. Because the employer is responsible for reporting to the tax authority the salary paid to every employee, the opportunities for a normal wage earner to underreport income are limited relative to the self-employed.

3.2 Public tax disclosure in Norway

Norway has a long history of public disclosure of information from income tax returns, going back at least to the middle of the nineteenth century. In earlier times, citizens could visit the local tax office or the city hall and look through a book that contained information about each taxpayer in the local area.

[4] A related literature also shows that individuals sometimes behave more pro-socially when they feel that they are being observed (e.g., Bursztyn and Jensen, 2015; Perez-Truglia and Troiano, 2018).

[5] The tax value of real estate assets is on average approximately half of the market value, and one-quarter if the house is the taxpayer's primary residence.

Access was limited to regular working hours for three weeks after tax assessment was finished, usually in mid-October. Persons were listed by name and address, along with key measures from the income tax return.

In the fall of 2001, a national newspaper offered online access to tax information for the whole population through the web version of the newspaper, and soon all of the major national newspapers followed. Not long afterward, the Norwegian government regulated these searches. As of 2004, only the tax agency was permitted to publish the raw data. From 2004 to 2006, the searches were confined to three weeks following the release of the data, but the number of searches was not restricted. Beginning in 2011, individuals were required to log in to the tax agency's website to conduct searches through a personalized log-in system for accessing online public services, which involved a pin-code and a password. Consequently, only Norwegian taxpayers could conduct searches after 2011. Searches were limited to 500 per month.

Beginning in October 2014, when tax records for the tax year 2013 were made available for searches, taxpayers could learn whether someone else searched for them, while previously searches were anonymous. On the website of the Norwegian tax authority one could access a list of who had searched for oneself. The end of anonymity corresponded with a drop of 85 percent in the aggregate number of searches.[6] The Tax Director of the Norwegian Tax Administration, Hans Christian Holte, characterized this change as "tak[ing] out the Peeping Tom mentality" and discouraging criminals from searching for wealthy people to target. He also stated, "We like people to do searches which could help us in investigating tax evasion and the amount of tips that we get has not gone down,"[7] implying that the authorities saw compliance-motivated searches as important, and that the deterrence function of disclosure was not diminished by non-anonymity. Our dataset of searches comes from the period after the anonymity of searches. This fact naturally introduces some caveats into our interpretation of search patterns; as we have no data on searches before anonymity, we do not know how the pattern of searches might be different in the prior, non-anonymous regime.

The Appendix display screenshots of the query process on the tax authority's webpage. Upon logging in, individuals can search for others' information by first and last name, or last name and year of birth. They then select the individual they wish to learn about, and the website provides that person's taxable

---

[6] Ministry of Finance (2014), cited in Perez-Truglia (2020). As noted in Perez-Truglia (2020, fn. 20), some individuals started selling a search service to allow users to search under their names and thus effectively preserve their anonymity. One company offered an anonymous search for NOK950 (about $120) per search.

[7] Bevanger (2017).

income, net wealth and assessed taxes.[8] Also available are birth year, and the postal code and city of the individual's registered residence.

Because they are more likely to go online to alter their pre-filled return, a self-employed person is arguably more likely to notice that someone has searched for their information. Self-employed taxpayers need to log in to the tax authority's website to report their own tax information. Wage earners without self-employment income, in contrast, typically have all their income sources pre-filled on their tax return. They have fewer reasons to log in in order to modify their tax return, but they may log in to learn how much tax is due to be remitted or rebated, or to provide information that documents deductions. Overall, 93 percent of self-employed, but only 15 percent of employees, modify their returns to some degree.

4. Who Searches, and For Whom Do They Search?

We begin by characterizing the volume and nature of searches. We do so using data on all searches for tax year 2013, which was available for searches from October, 2014 to August, 2015. Recall that we refer to individuals initiating a search as *searchers*, and individuals whose information was queried by a searcher as *targets*. The total number of observations amounts to 1,316,091 searches. In calculating all the figures we report below, we make three sample selection restrictions. First, we drop any search observations where either the searching or target individual cannot be identified in the income registers, which accounts for 1.9 percent of searches. Second, we drop all but one of multiple searches by the same searcher for the same target; this results in an additional 10.6 percent of searches being dropped. Most of the multiple searches appear on the same date, often within seconds of each other (possibly instances of inadvertent re-clicking). Third, we exclude all searches by under-aged individuals in 2013 (<18), which amounts to 7.9 percent of searches.[9] When below we make additional sample restrictions, we mention them.

4.1 Frequency of search: Network density

The three sample restrictions described above leave us with 969,804 searches between October 2014 and August 2015. To put this number into perspective, in 2013 the adult (18 and over) population in

[8] General income is a net income concept (taxed at 27 percent) including all types of taxable income, after the deduction of all deductible expenses. Net wealth is the basis for the wealth tax. Assessed tax is the sum of general income tax, surtax on personal income and wealth tax.

[9] Individuals above the age of 16 years old can search, but only for individuals at least 18 years old in 2013. This asymmetry is the reason for restricting attention to the adult population (≥18 years old) in the main analysis.

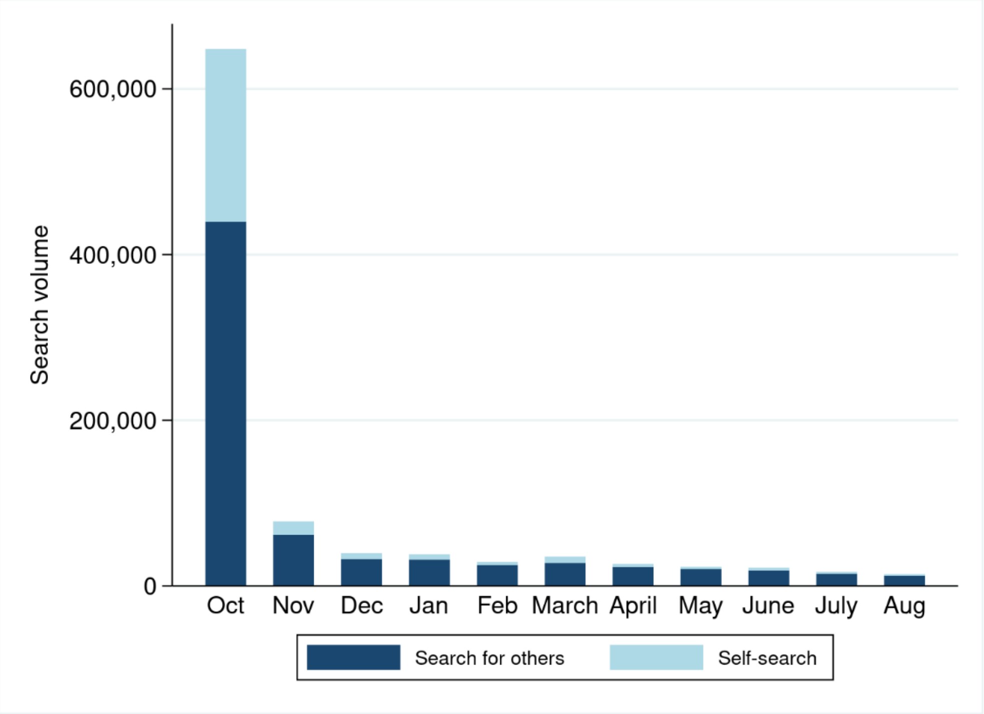
Norway in 2013 was 3,983,896. Of the total number of searches, 262,078, or 27 percent, were the results of people searching for themselves.

Figure 1 shows the distribution of searches by month in this period. A majority of the searches, 66.9 percent, occurred in the first month that information from the new tax year was available, October of 2014.[10] Self-searches are especially concentrated in October, suggesting that when search data are made available and publicized, many people search for themselves to see what others can see about them, a phenomenon akin to Googling oneself. Figure A-1 in the Appendix reports more details on the exact date of searches, which reveals that, even within October, the number of searches is skewed to the opening of the search process. More than half--54 percent--of the October searches occurred on the first day that search was available, October 17. There is also a slight blip up in March of 2015, when (on the 19th) the pre-filled returns for the 2014 tax year became available; some taxpayers were likely checking the 2013 information while logging in to view their own pre-filled tax information for 2014.[11]

[10] See Table A-1 in the Appendix.

[11] Notably, 43 percent of searches on March 19th were the results of people searching for themselves (against 27 percent overall).

Figure 1. Searches by Month in 2014 and 2015



Notes: This figure plots the number of searches conducted in each month for which tax year 2013 information was available, from October 2014 to August 2015, using the full sample of searches. Searches are divided into searches for one's own information and search for others' information.

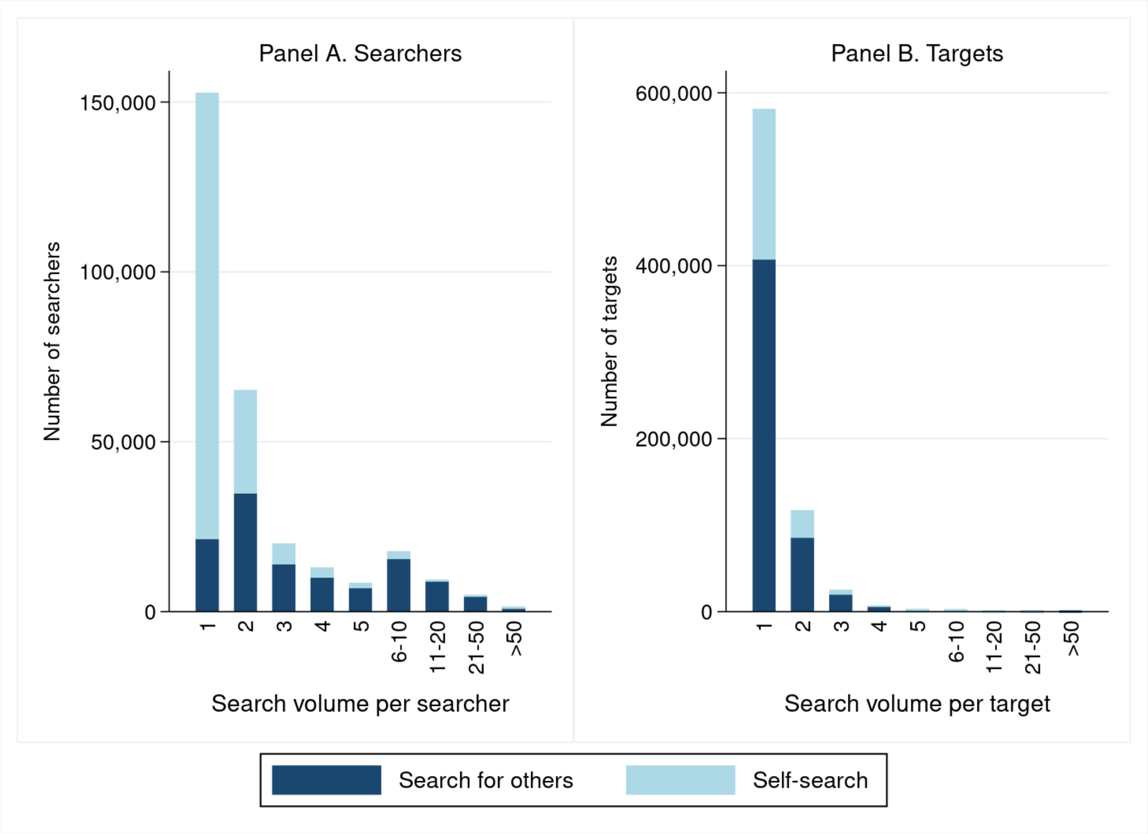
4.2 Concentration of search: Network clustering

Because people can search for multiple targets, the number of searches overstate the number of distinct searchers and distinct targets. The number of searchers was less than one-third of the number of searches, 292,417, so that on average each searcher made approximately three searches. About 7.3 percent of the adult population did at least one search. Notably, almost all of the searchers also searched for themselves (262,078 individuals). The number of distinct targets was much higher, 735,071, or 18.5 percent of the adult population. Of those, 561,116 individuals were targeted at least once by someone other than themselves.

Figure 2 provides information about the number and concentration of distinct searchers and targets. Panel A shows that 152,737 searchers (about half of all searchers) made only one search, and that 1,147 searchers (0.4 percent), searched for more than fifty distinct Norwegian taxpayers.[12]

[12] We can not rule out that some individuals may offer to allow others to search under their names and thus effectively preserve the others anonymity. But it does not seem like this practise is widespread.

Figure 2. Number of Searches per Searcher or Target



Notes: This figure plots the number of searchers and targets in the full sample by the month the search was conducted. Searches are divided into searches for one's own information and search for others' information.

Panel B shows that the search targets are less concentrated than the searchers. 581,254 targets, just under 80 percent of all targets, are targeted by only one searcher. And less than 1 percent of all targets are targeted by more than three searchers. Nevertheless, there are some "star" targets. 88 individuals were targeted by more than 50 Norwegians, and ten individuals were targeted by more than 500 individuals, with 1,087 Norwegians targeting the most "popular" Norwegian.[13]

4.3 Who searches? Out-Degree centrality

We next link search data to demographic information in the population register in order to characterize who searches, and who is targeted. As discussed in Section 2.1, the characteristics of frequent searchers is informative about what types of people have high out-degree centrality in the search network, and the characteristics of frequent targets is informative about in-degree centrality. In

[13] Of the top 10, three are billionaire business people, three are politicians, two are bloggers, one is a singer, and one is an athlete.

addition to giving us a broad overview of the network of searches, the characteristics of searchers and targets also provide some insight into searchers' motivations for searching.[14] If searches are related to interest in tax evasion, we should expect searches to be concentrated in populations where tax evasion is more common. As we elaborate on in Section 6, this logic suggests that self-employed individuals should be more likely to be targets than employees.

Table 1 shows that, compared to the adult population, on average those who target for other people (labeled "searchers for others") are much younger, much more likely to be male and a wage earner, much less likely to be married or have immigrant status; they have slightly higher than average income but are slightly less wealthy. Nearly all of these statements also apply to self-searchers who, however, have notably higher income. If we limit our attention to those who do at least ten or more searches (not shown), we find that this group is even younger (30 years), more male (0.72), but are more likely to be an immigrant (0.13) and have lower education (11.9) and low income (44). There is also an even larger share of students (0.26), especially upper-secondary and undergraduate students.

[14] Note that, during this period, Internet penetration in Norway was 96 percent, so differential access to the Internet is unlikely to explain a significant amount of the demographic variation in search behavior. See https://www.statista.com/statistics/631917/norway-access-to-the-internet/.

Table 1: Mean Characteristics of Searchers and Targets, Compared to the Overall Adult Population

	Adult Population	Searchers for others	Self-searchers	Targets by others
Age	47.50	36.73	38.31	43.17
Male	0.50	0.63	0.66	0.61
Single households	0.26	0.35	0.37	0.29
Married couples	0.45	0.37	0.35	0.43
Years of education	12.06	12.55	12.75	12.39
Immigrant	0.14	0.09	0.07	0.10
Residence in densely populated area	0.30	0.32	0.34	0.31
Student	0.07	0.15	0.12	0.09
Wage earner	0.71	0.88	0.89	0.82
Self-employed	0.086	0.099	0.105	0.121
Unemployed/disabled/soc. welfare	0.14	0.10	0.09	0.10
Old-age pensioner	0.19	0.06	0.06	0.12
Income percentile	50.50	57.09	61.84	58.85
Wealth percentile	50.50	44.21	43.60	48.00
Number of observations	3,983,896	161,045	262,078	561,116

Notes: The table presents mean characteristics for the overall adult population and for three sub-populations: "searchers for others," "self-searches," and "targets by others." "Searchers for others" are individuals searching at least once for someone other than themself. "Targets by others" are individuals who have been targeted at least once by someone other than themself. "Self-searchers" are all individuals who searched for themself. An individual may be present in more than one of these sub-populations. All reported characteristics are based on registered information from 2013. Male, single households and married couples are all indicator variables (0/1). Immigrants are defined in Norwegian registers as persons born abroad of two foreign-born parents and four foreign-born grandparents. Residence in a densely populated area is an indicator variable, which is 1 if the individual is registered as living in one of the five largest cities in Norway (or the area around these cities). Students are defined as those receiving grants from the State Educational Loan Fund ("Lånekassen"). Wage earners are defined as having positive wage income. Self-employed are defined as having non-zero income from self-employment. Unemployed/disabled/soc. welfare have positive income from at least one of these sources. Old-age pensioners have positive pension income from the National Insurance Scheme (age 67 and above). Gross income ("samlet inntekt") and net wealth ("netto formue") are used to categorize individuals into income percentiles / wealth percentiles based on the income/wealth distribution of the overall adult population.

Because several of these characteristics are correlated, we next look at multiple regression analyses of the association of search behavior with demographic characteristics, the results of which are shown in Table 2. Column (1) presents the estimates from a linear probability model of whether someone does at least one search on someone other than themselves, as a function of their demographics.[15] Most of the demographic patterns apparent in the summary statistics shown in Table 1 are also visible in the multivariate regression analyses. These effects should be compared with a mean search probability of about 4 percent. The probability of search declines with age, with a decreasing absolute slope. Men are 1.4 percentage points more likely to search than women, other things equal, and being married is associated with a 0.5 percent higher probability of search. A higher search probability has a positive

[15] A probit specification yields qualitatively very similar results.

partial association with both income and wealth. Being self-employed is associated with a 0.4 percent higher probability of doing at least one search. Column (3) of Table 2 shows that all of these patterns also appear when the dependent variable is the number of searches. Column (2) concern self-searchers, where most but not all of the same patterns emerge. In contrast to search for others, married people and wage earners are less likely to self-search, while more educated people are more likely to do so.

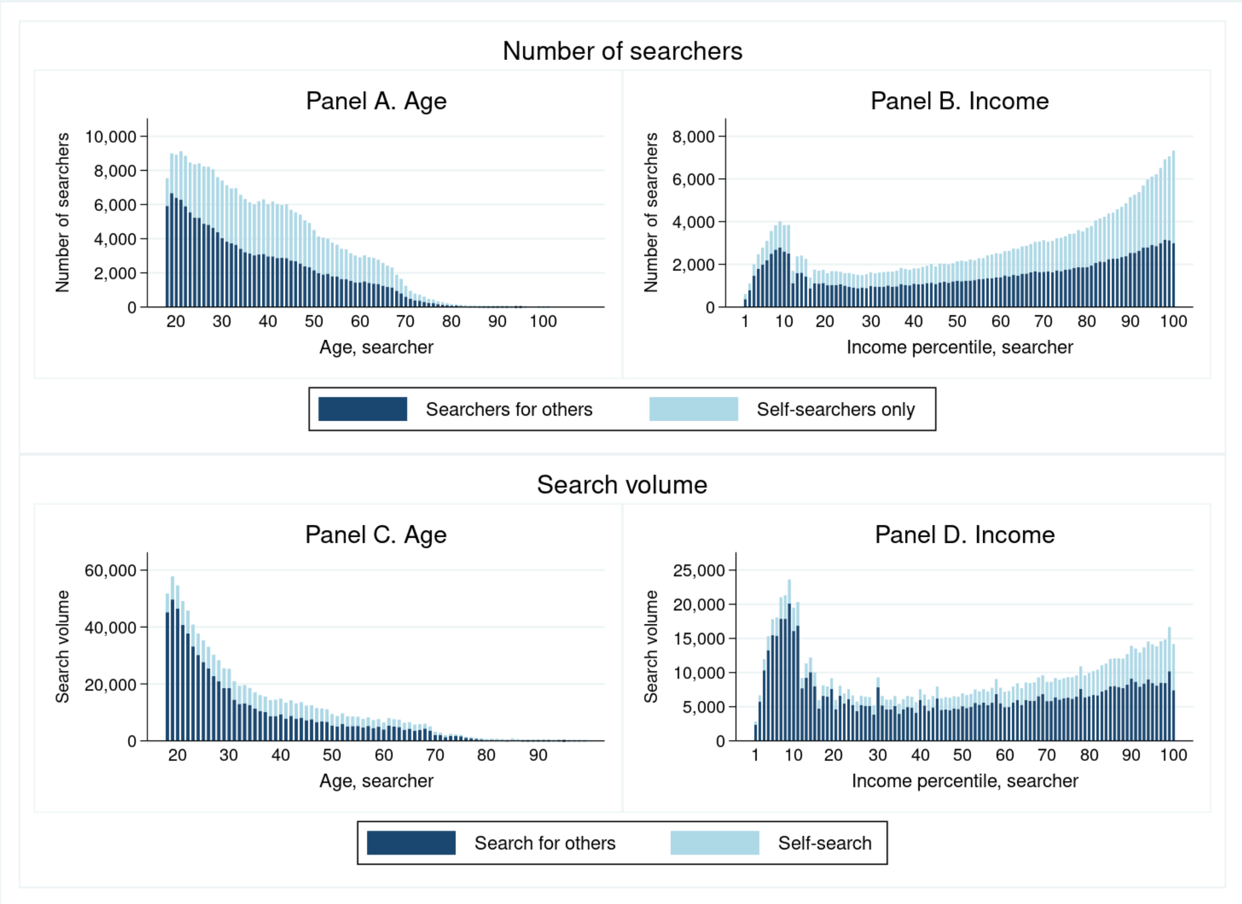
Moving beyond the mean characteristics, Figure 3 illustrates the distribution of searchers and search volume by two important searcher characteristics, income and age. Panel A shows that the younger people are the more likely they are to be a searcher. Panel B shows that both very low income and individuals with higher income are more likely to be a searcher. High-income searchers search much more often only for themselves. Panel C shows that the search volume is much higher for younger individuals, suggesting that the volume of searches for each searcher is significantly higher for the youngest searchers. Panel D shows that the search volume is much higher for low-income individuals, in line with the age pattern of searches as there is a strong positive correlation between age and income shown in Figure A.2 in the Appendix. However, comparing Panels C and D reveals that search volume rises with income starting at about the 35^{th} income percentile, although no such non-monotonicity appears with respect to age. All in all, there seems to be two main types of searchers: the very young (18-25 years old) with low income (often lower-degree students) searching several targets; and middle-age people (26-62 years old) with high income and more selected search behavior (searching for themselves and one or only a few others).

Table 2: Regression Analysis of Searcher Characteristics

Dependent variable	(1) Searcher for others (0/1)	(2) Self-searcher (0/1)	(3) Number of searches
Age/10	-0.0532***	-0.0622***	-0.3552***
	(0.0003)	(0.0004)	(0.0052)
Age, squared /1000	0.0378***	0.0416***	0.2757***
	(0.0003)	(0.0003)	(0.0045)
Male	0.0140***	0.0271***	0.1254***
	(0.0002)	(0.0002)	(0.0038)
Married	0.0053***	-0.0068***	-0.0058**
	(0.0002)	(0.0003)	(0.0026)
Years of education / 10	-0.0001	0.0069***	-0.0163**
	(0.0004)	(0.0005)	(0.0064)
Immigrant	-0.0233***	-0.0411***	-0.0870***
	(0.0003)	(0.0003)	(0.0042)
Residence in densely populated area	-0.0008***	0.0035***	0.0061
	(0.0002)	(0.0003)	(0.0041)
Pensioner/Disability/Unemployed	0.0025***	0.0013***	0.0134**
	(0.0003)	(0.0003)	(0.0058)
Wage earner	0.0021***	-0.0030***	0.0540***
	(0.0003)	(0.0003)	(0.0054)
Self-employed	0.0037***	0.0043***	0.0221***
	(0.0004)	(0.0005)	(0.0048)
Income percentile / 10	0.0051***	0.0109***	0.0064***
	(0.0001)	(0.0001)	(0.0010)
Wealth percentile / 10	0.0007***	0.0002***	0.0048***
	(0.0000)	(0.0001)	(0.0006)
Constant	0.1572***	0.1842***	1.01727***
	(0.0010)	(0.0012)	(0.0151)
Number of observations	3,983,896	3,983,896	3,983,896
R-squared	0.0251	0.0420	0.0030

Notes: This table reports the results of OLS regression analyses of individuals' searching behavior on observed characteristics of the searcher. The regression sample is the overall adult population (at least 18 years old). In column (1), the dependent variable is 1 if searched at least once for someone others than themself, and 0 otherwise. In column (2), the dependent variable is 1 if searching for themself, and 0 otherwise. In column (3), the dependent variable is the number of searches (excluding self-searches) an individual conducted. All individual characteristics are based on registered information from 2013. See the note to Table 1 for details on the construction of right-hand-side variables. Standard errors are provided in parentheses below point estimates. * p<0.10, **p<0.05, ***p<0.01.

Figure 3: Number of Searchers and Search Volume by Searcher's Age and Income



Notes: Panel A and B depict the number of distinct individual searchers by the age and the income percentile of the searcher, respectively. Panel C and D depicts the number of searches by the age and the income percentile of the searcher, respectively.

4.4 Who is searched for? In-Degree centrality

The last column of Table 1 presents the average demographic characteristics of those who have been targeted at least once by someone other than themself. It reveals that targets are younger than the average Norwegian, but older than the average searcher. Targets are about as equally male as the average searcher. They are more likely to be married: 43 percent of targets are married, compared to just 37 percent for those who search for others. Targets also have higher income compared to either searchers for others (but lower than self-searchers) or the overall population, and are more likely to be self-employed.

Table 3 presents the results of multivariate linear probability regression analyses of who is targeted at least once by someone other than themself, as well as how many times someone is targeted.

The results confirm that higher-income individuals are more likely to be a target, other things equal; the estimated coefficient on income percentile in Table 3 implies that the probability that someone in the 90th percentile is targeted is 12.9 percent higher (8 x 0.0161) than is someone in the 10th percentile. Note that income is more strongly associated with being targeted than with searching: the estimated effect of income on being a target is more than three times higher than the estimated effect of income on the probability of being a searcher, as shown in Table 2. Table 3 also reveals that the self-employed are much more likely to be targets, a topic we revisit in Section 6.

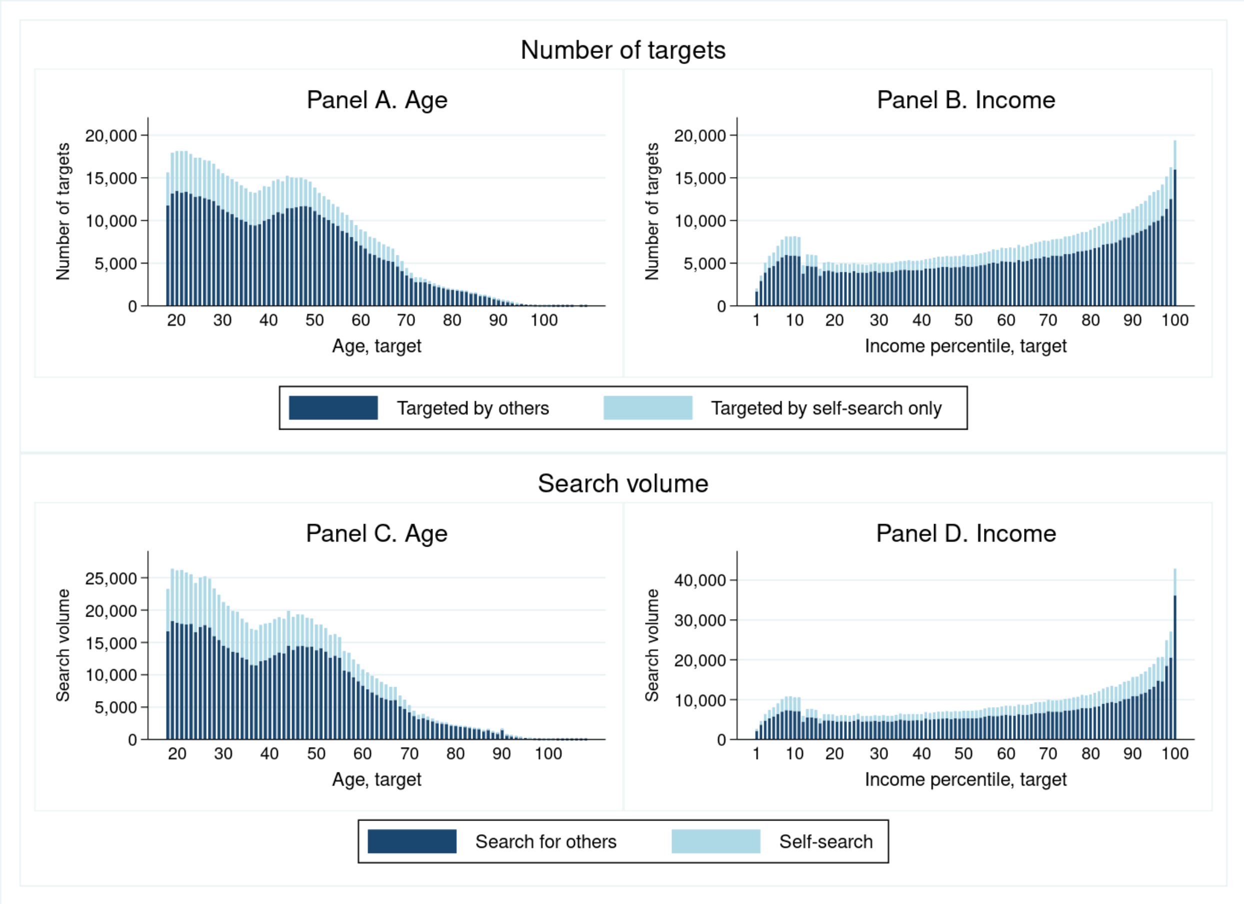
Figure 4 illustrates the distribution of targets and volume of search by targets' income and age. The patterns are broadly similar to those of searchers shown in Figure 3, but with some noteworthy differences. Middle-aged taxpayers are more likely to be targets compared to being searchers, as are very high-income people. The bump in targets for very low-income people is present, but is not nearly as substantial compared to searchers. Above this small bump, the distribution of target income exhibits a relatively steep positive gradient. This indicates that both young, low-income searchers and middle-aged, high-income searchers search for high-income targets.

Table 3: Regression Analysis of Target Characteristics

	(1)	(2)
Dependent variable	Targeted (0/1)	Number of searches
Age/10	-0.0854***	-0.1530***
	(0.0006)	(0.0052)
Age, squared /1000	0.0635***	0.1167***
	(0.0005)	(0.0047)
Male	0.0371***	0.0546***
	(0.0004)	(0.0020)
Married	0.0103***	0.0151***
	(0.0004)	(0.0016)
Years of education / 10	-0.0250***	-0.0598***
	(0.0007)	(0.0050)
Immigrant	-0.0478***	-0.0630***
	(0.0005)	(0.0011)
Residence in densely populated area	-0.0013***	0.0136***
	(0.0004)	(0.0024)
Pensioner/Disability/Unemployed	-0.0187***	-0.0288***
	(0.0005)	(0.0012)
Wage earner	0.0013***	-0.0095***
	(0.0005)	(0.0022)
Self-employed	0.0423***	0.0880***
	(0.0007)	(0.0063)
Income percentile / 10	0.0161***	0.0287***
	(0.0001)	(0.0008)
Wealth percentile / 10	0.0021***	0.0050***
	(0.0001)	(0.0005)
Constant	0.3069***	0.4841***
	(0.0016)	(0.0135)
Number of observations	3,983,896	3,983,896
R-squared	0.0361	0.0048

Notes: This table reports the results of OLS regression analyses of targeted individuals on observed characteristics of the target. The regression sample is the overall adult population (at least 18 years old). In column (1), the dependent variable is 1 if targeted at least once by someone other than themselves, and 0 otherwise. In column (2), the dependent variable is the number of times targeted (excluding targeted by self-search). All individual characteristics are based on registered information from 2013. See the note to Table 1 for details on right-hand-side variable construction. Standard errors are provided in parentheses below point estimates. * p<0.10, **p<0.05, ***p<0.01.

Figure 4: Number of Targets and Search Volume by Targets' Age and Income



Notes: Panel A and B depicts the number of distinct targets (counting individuals) by the age and the income percentile of the target, respectively. Panel C and D depicts the number of searches by the age and the income percentile of the target, respectively.

5 Who Searches for Whom?

Having looked at who searches and who is targeted, we next examine the patterns of who searches for whom, using characteristics of both searchers and targets. As discussed above, one way to interpret these empirical patterns is in terms of social comparisons: which searchers target which searchers could help us understand to whom individuals prefer to compare themselves. Here, we analyze identifiable social networks of households and employment, examine the degree of homophily in search, and investigate the joint distribution of particular characteristics between searchers and targets. The last of these, together with our analysis of who searches for stars, provides evidence for the more aspirational types of searches described above. Finally, we characterize reciprocal search behavior.

5.1. Search within identifiable social networks of households and employment

In this subsection, we look more closely into how many of the "searches for others" (excluding self-search) occur within identifiable household or employment networks.

Search within households

We find that, out of 707,726 searches (excluding self-search), 86,851 searches (12.3 percent) occur within households.[16] Of the number of distinct searchers (161,045 in total), 69,360 searched for a member of their own household, and 122,785 searched for someone outside of their household. Of the number of distinct targets (561,116 in total), 84,386 were targeted by a member of their own household, and 492,179 were targeted by someone outside of their household.[17] A regression analysis (not reported) of searches within households reveals that women are more likely to be targeted by spouses, but not generally. Notably, income matters a lot less for spousal searches.

Search within employment networks

We find that, out of 707,726 searches (excluding self-search), 112,369 searches (15.9 percent) occur within an employment network.[18] Overall, 25.9 percent of searches (excluding self-search) occur within identifiable networks of either household or employment.[19]

More details on how search within identifiable networks of household and employment is distributed over age and income are provided in Figure 5.

Figure 5, Panel A reveals that the very youngest (below age 25) and mid-age (peak 50-60 years old) search relatively more often within-household.[20] Whereas within employment network searches are mostly concentrated in the age group 20-50.[21] When sorted by the target's age in Panel C, we see a similar pattern, but note that there are more within-employment-network targets in the age group 50-

[16] Within-household searches consist of persons resident in the same dwelling and related to each other as spouse, registered partner, cohabitant, and/or parent and child (regardless of the child's age).

[17] The sum of the two categories exceeds the total because someone could have been targeted by a household member and by a non-household member.

[18] Employment networks are established based on information about all employers for each individual in 2013 (the tax year) and 2014 (the search year). If any of the employers are the same for the searcher and the target, it is regarded as a search within an employment network.

[19] When excluding searches within households, 15.5 percent of searches occur within employment networks.

[20] This pattern could occur if parents and children are more likely to search for each other, but they are registered as belonging to the same household only when the child is around 20 years old and the parent is about 50 years old.

[21] We note that one motive for search could be to learn about the income of e.g. potential employees or home renters. Some searches within employment networks could be such searches. Unfortuantely, we are limited in our capacity to infer from the data whether a search occurs for this particular reason.

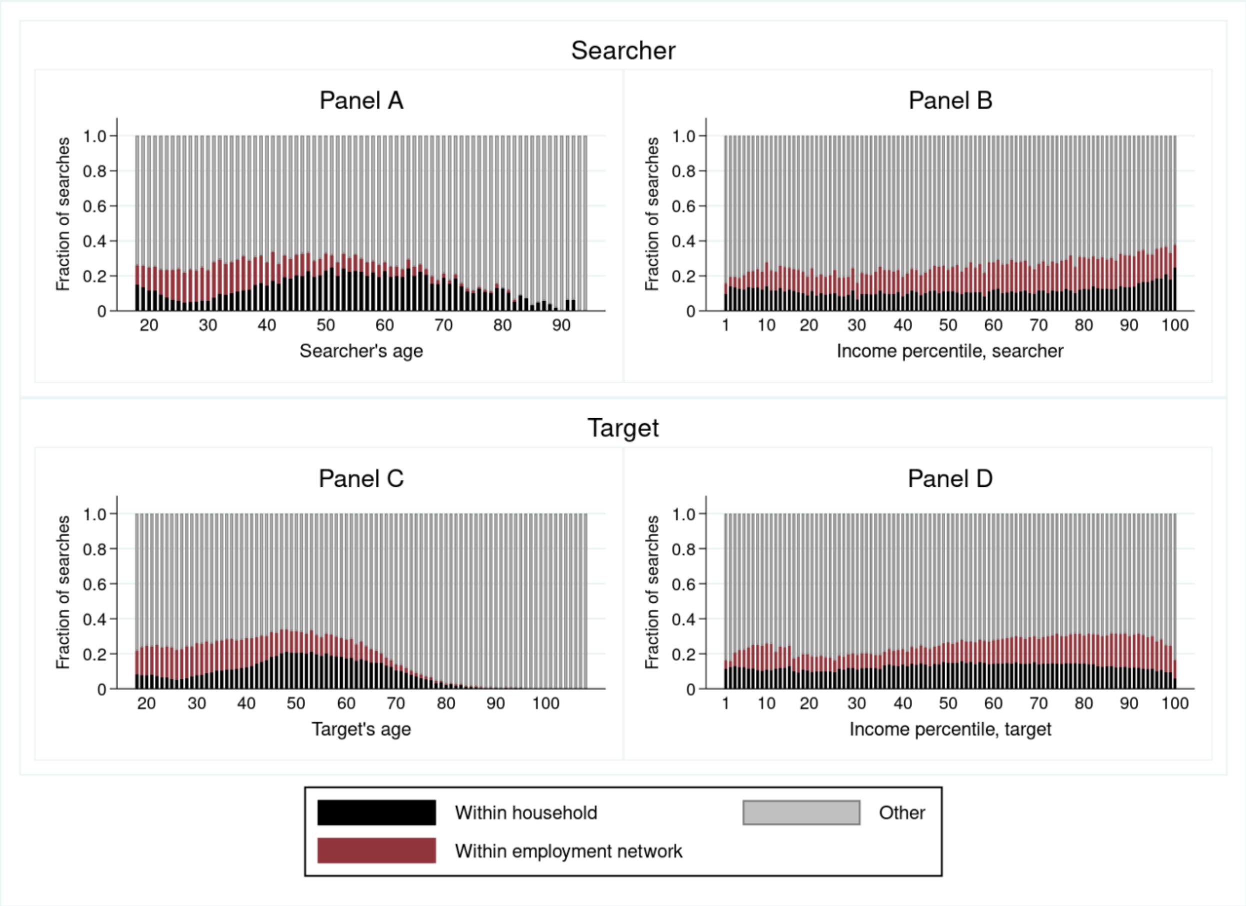
70 than there are searchers. The fraction of searches within-household and within-employment network are rather evenly distributed across the income distribution of the searcher in Panel B. Similar patterns are found along the income distribution of targets in Panel D.

5.2. Homophily in search

A ubiquitous finding about other social networks, such as networks of friends, is that people tend to be more frequently linked to others that are similar to themselves, a network characteristic denoted homophily.[22] If there is substantial homophily in the tax search network, then different groups are quite isolated from each other and they may acquire only "local" information—information on people like them—about reported income and wealth. We test for homophily along several dimensions, and begin with binary homophily, where each searcher and target is defined as being in one of two categories. In these analyses, we exclude self-searches and household searches. Table 4 summarizes these results, by comparing the actual observed probability of a given searcher-target identity compared to the probability of a random pair of a searcher and target in our sample. The final column shows the ratio of the observed probability to the random probability. The higher is that ratio, the greater is homophily. Note, though, that these ratios cannot always be meaningfully compared across the rows that represent characteristic categories, because for well-represented groups this statistic cannot greatly exceed one.

[22] McPherson, Smith-Lovin, and Cook (2001).

Figure 5: Composition of Searches by Social Networks of Households and Employment. Fraction of searches



Notes: The panels present the fraction of searches that occur within household, and within the same employment network. The network of households consists of persons resident in the same dwelling and related to each other as spouse, registered partner, cohabitant, and/or parent and child (regardless of the child's age). Information about employment network is established based on information about all employers for each individual in 2013 and 2014. If any of the employers are the same for the searcher and the target, it is regarded as a search within an employment network. Panel A and B depicts the fraction of searches by the age and the income percentile of the searcher, respectively. Panel C and D depicts the fraction of searches by the age and the income percentile of the target, respectively.

Table 4: Homophily in Search

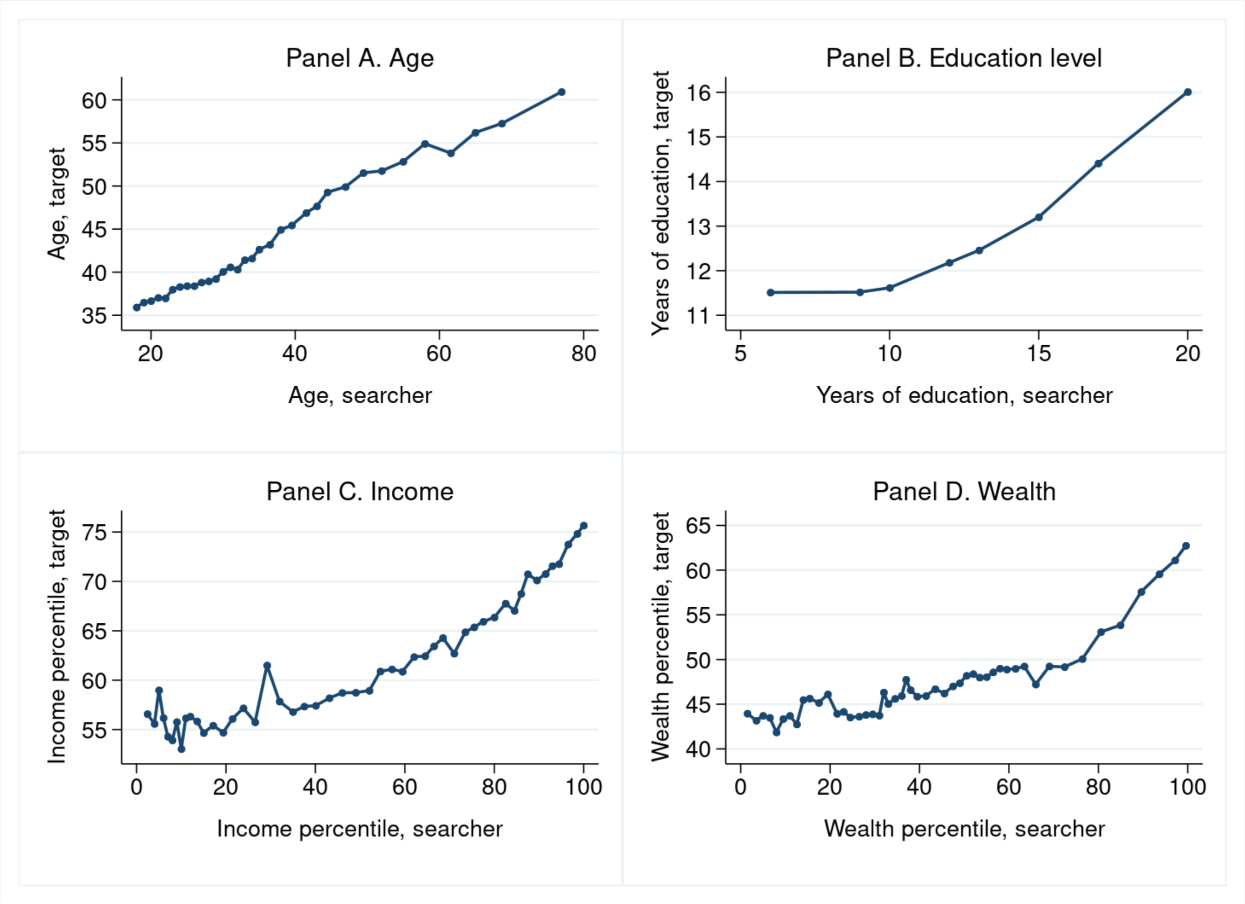
	Observed probability	Random probability	Ratio
Both males	49.4	46.7	1.1
Both females	12.6	10.0	1.3
Same municipality	45.8	3.3	14.0
Same age	19.7	6.5	3.1
Both immigrants	6.0	1.1	5.6
Neither immigrants	85.4	80.5	1.1
Both students	4.3	2.0	2.2
Neither students	73.3	70.9	1.0
Both self-employed	2.1	1.1	1.8
Neither self-employed	79.8	78.8	1.0
Same main employer	9.2	0.1	152.5
Same employment network	15.5	0.2	79.2
Same education level	33.1	22.5	1.5
Same education field	14.8	8.4	1.8
Same education level & field	8.1	2.8	2.9
Same income level	15.9	9.9	1.6
Number of observations= 620,875			

Notes: This figure reports the probability that searcher and target share a given characteristic for several different characteristics. The first column reports these probabilities in the population, excluding self-searches and searches within a household. We contrast this with the probability that a random pair of searchers and targets share this characteristic in the second column. This means that, for example, because there are many more male searchers and targets in the data, the probability of a random female-female pair is less than 0.25, and the probability of a random male-male pair is greater than 0.25. Same age is defined as the target and searcher age being 1 year apart or less (in either direction). Education level has 9 categories. Education field has 7 categories (excluding "general" and "unknown"). Same income level is defined as less than a 5-percentile difference between searcher and target income rank. Same employment network is defined as in Figure 5 above.

According to the results in Table 4, Norwegians are much more likely to search for tax information about people who live within their own municipality. As many as 45.8 percent of searches (excluding self-search and search within the same household) are for people who live within the same municipality; this compares to just 3.3 percent of random searcher-target pairs who reside in the same municipality. A highly intensive dimension of binary search homophily is by employer. While random searcher-target pairs have the same main employer just 0.1 percent of the time, 9.2 percent of such pairs are between pairs of people with the same main employer, over hundred times more likely than random. Immigration status also exhibits substantial search homophily, as do self-employed taxpayers.

We can say more about several dimensions of search homophily because we can divide searchers and targets into finer categories. Panel A of Figure 6 offers an alternative look at age homophily, where the average age of the target is shown for each bin of searcher age; a clear positive relationship emerges.

Figure 6: Homophily in Search by Age, Education Level, Income and Wealth



Notes: This figure presents binned scatterplots of the mean characteristics of a target characteristic, by the same characteristic of the searcher. We exclude self-searches and searches within household. Panel A through Panel D present results for age, education level, income and wealth, respectively. Years of education is computed by the highest registered education level of the individual. Income percentile refers to the income percentile among the overall adult population. Wealth percentile refers to the net wealth percentile among the overall adult population.

Panel C in Figure 6 bins searchers by their income percentile and then shows the average income percentile of the people searched for by searcher income bin.[23] It shows that, on average, people search for people with relatively high income; the average percentile of the targets is always above 50. Also

[23] Of course, age and income homophily are not independent phenomena, as age and income are highly correlated. In Appendix Figure A-2, we illustrate this issue by plotting the fraction of searchers by age group against the income percentile of the searcher. Sure enough, low-income searchers are predominantly young people, while high-income searchers are not.

for education level, in panel B, and for wealth, in panel D, there is a clear positive relation between searchers' and targets' characteristics.

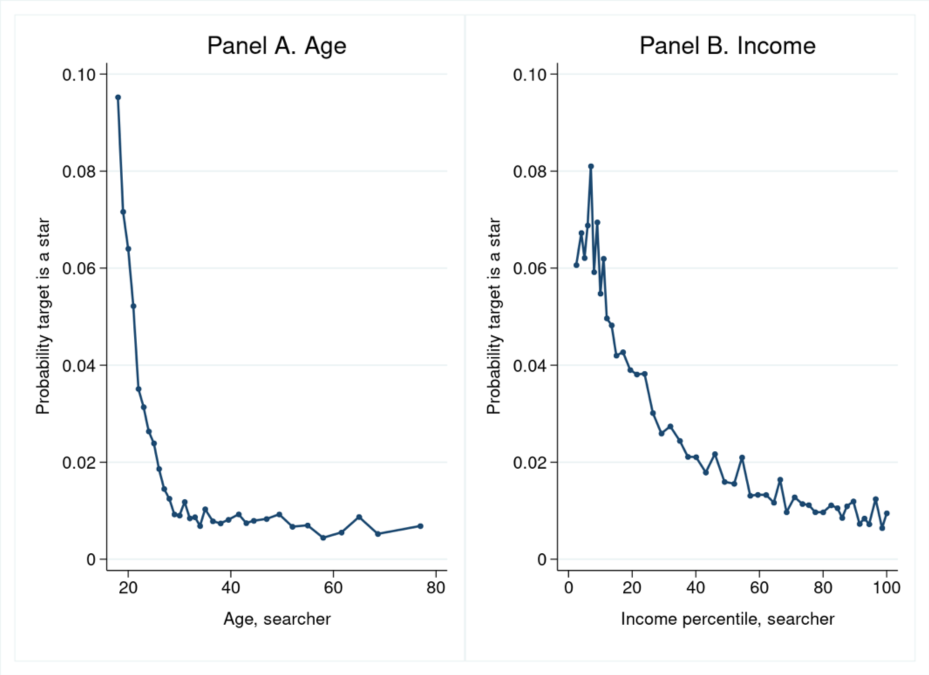
5.3. Beyond homophily: Aspirational searches

We next provide descriptive evidence that young, low-income people also frequently conduct what we call "aspirational searches," reflecting our discussion in Section 2.2. Figure 7 demonstrates that searches for stars are more common among young people at the bottom of the income distribution, where a star is defined as someone targeted by at least 50 searchers.

Figure 8 depicts the distribution of target income conditional on searcher income with a heat map, allowing one to go beyond just the mean of target income conditional on searcher income. It shows that low-income people mostly search for other low-income people and very high-income people. Middle-to-high income people tend to search for people with income closer to their own income, consistent with Figure 6, Panel C. There is a very low probability of search in the bottom-right quantile of the graph: high-income people rarely search for low-income people.

The bimodal distribution of target incomes conditional on low searcher income is the most significant instance we see of non-homophily. Overall, searcher and target characteristics are correlated, but at the bottom of the income distribution we also observe the opposite. In the next several results, we try to understand what drives this pattern. Doing so helps shed light on the different types of relative income comparisons illustrated by Hirschman and Rothschild's tunnel metaphor discussed in Section 2.2. Rather than study the effect of learning another's income on well-being, like most of the literature, we try to understand to whom individuals choose to compare themselves.

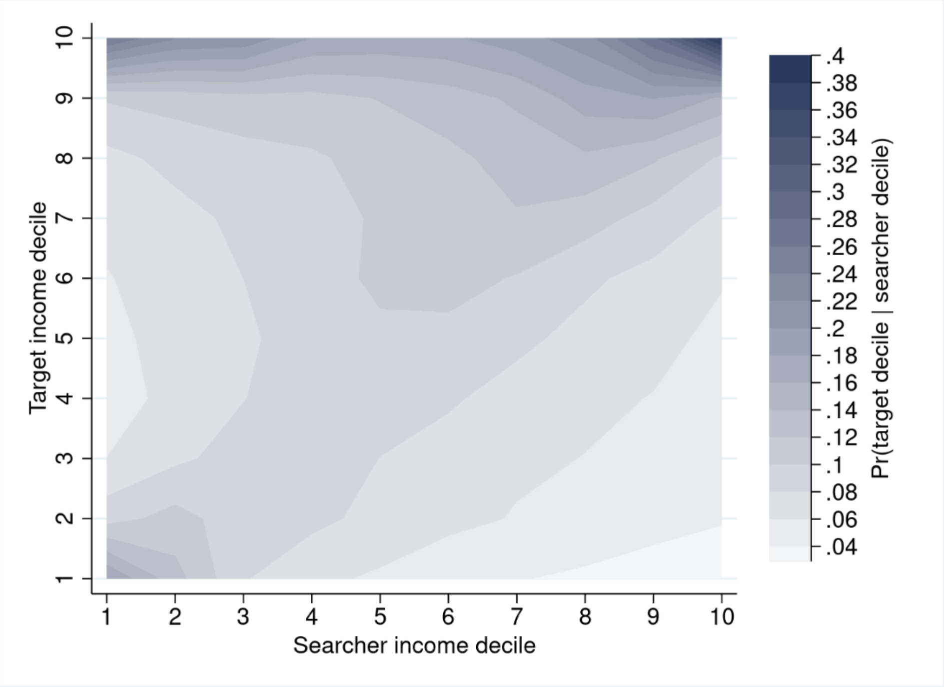
Figure 7: Probability of Searching for a "Star," by Searcher's Age and by Searcher's Income



Notes: Panel A and Panel B present a binned scatterplot of the probability that the target is a star by the age and the income percentile of the searcher, respectively. We define a star as an individual targeted by at least 50 searchers (distinct individuals). The income percentile of the searcher in Panel B refers to the income percentile in the overall adult population.

The heat map of Figure 8 summarizes a key finding of our paper: everyone looks up people similar to themselves, perhaps in their social network, but young, low-income people also disproportionately look up highly successful people, as is also evident in Figure 7. This fact also appears to be the main driver of apparent non-linearity in the mean searcher-target income gradient from Figure 6, Panel C. What makes very low-income people especially likely to search for high-income people? In Panels A and B of Figure 9 we pursue this question further by examining the composition of different types of searches within bins of searcher income, and we decompose the overall relationship in Panel C by different types of searches. We use what we already know from our earlier analysis that certain relationships between targets and searches are especially common, including searches for coworkers, geographic neighbors, and people with a similar education level.

Figure 8. Heat Map of Searcher and Target Income

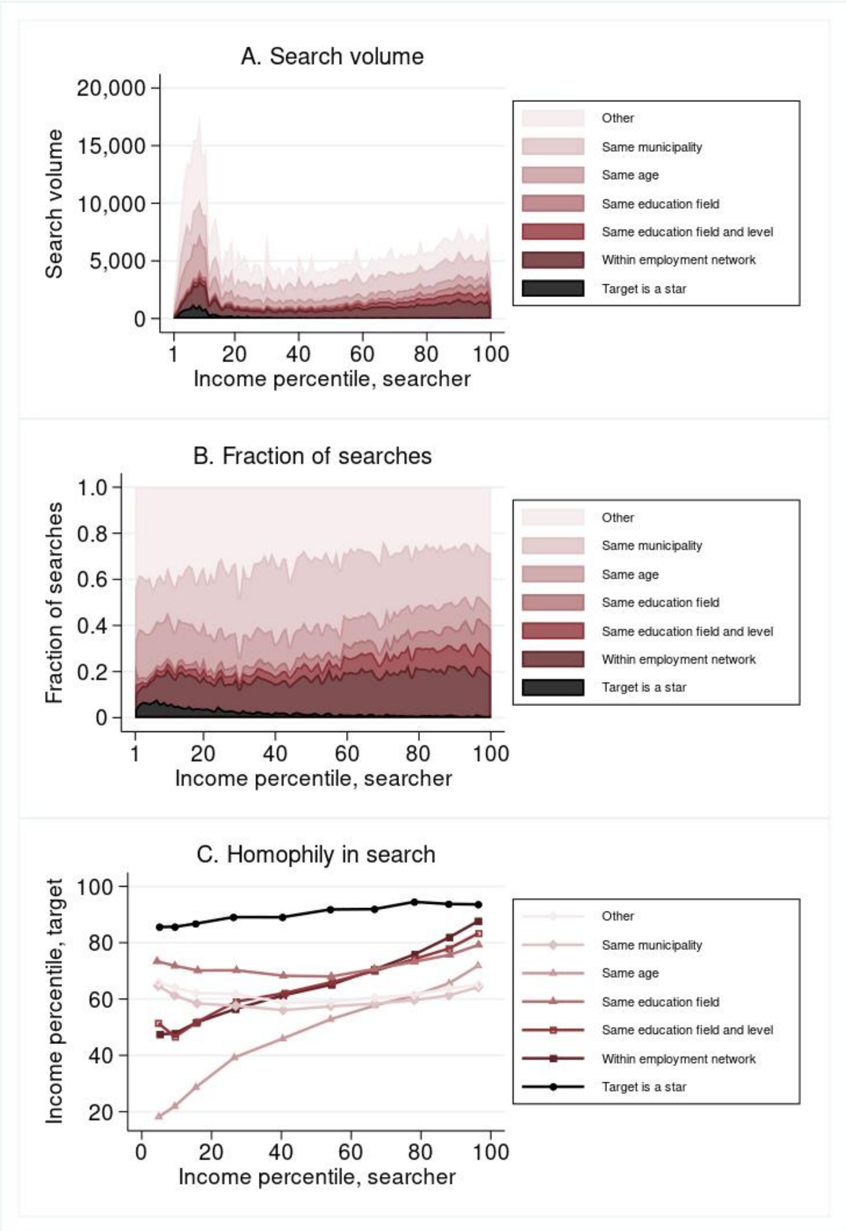


Notes: The figure presents a heatmap of the distribution of target incomes conditional on searcher income. Darker colors represent higher probabilities for searching for targets of the given income decile, within the given searchers' decile. The income of searchers and targets are categorized into income deciles (1-10) defined by the overall adult population.

First, note from Panel A that low-income people are engaged in more searching overall, and for almost every type of search we examine. To better compare the composition of searches throughout the searcher income distribution, in Panel B we normalize by the number of searches within an income bin, and note several patterns that are consistent with our story. First, searchers at the bottom of the income distribution disproportionately search for stars. As stars have high incomes, the fact that searches for stars are so common at the bottom of the distribution can help explain the pattern of Figure 6. In contrast, searches within employment networks are much less common for very low income searchers.

Panel C of Figure 9 plots the searcher-target income gradient for different types of search. We observe a steeper positive gradient when searchers and targets are in the same employment network, when they have similar ages, and when they have the same educational credentials and field of education. The pattern is somewhat weaker for cases where the searchers and targets live in the same municipality or when they share the same field of education with another education level, though. Overall, these findings are nevertheless consistent with the notion that most of these types of searches are being done by people searching for demographically similar people in their social network, so that searcher and target incomes are more strongly positively related. Naturally, targeted stars' incomes are high everywhere, but recall that searches for stars were more common at the bottom. Overall, the results in Figure 9 are consistent with our proposed explanation for the bimodality of target incomes at low searcher incomes we observe in Figure 8.

Figure 9. Decomposing Searches by Searcher's Income



Notes: The panels present the search volume (A), the fraction of searches (B), and the income percentile of the target (C), respectively, by the income percentile of the searcher. Searches (excluding self-search and within household search) are divided into 7 categories by the following order: 1) Target is a star; 2) Within employment network; 3) Same education field and level; 4) Same education field; 5) Same age; 6) Same municipality; 7) Other. See the notes to Table 4 and Figure 7 for detailed definitions of these categories. "Other" refers to searches that do not fit into any of the six categories.

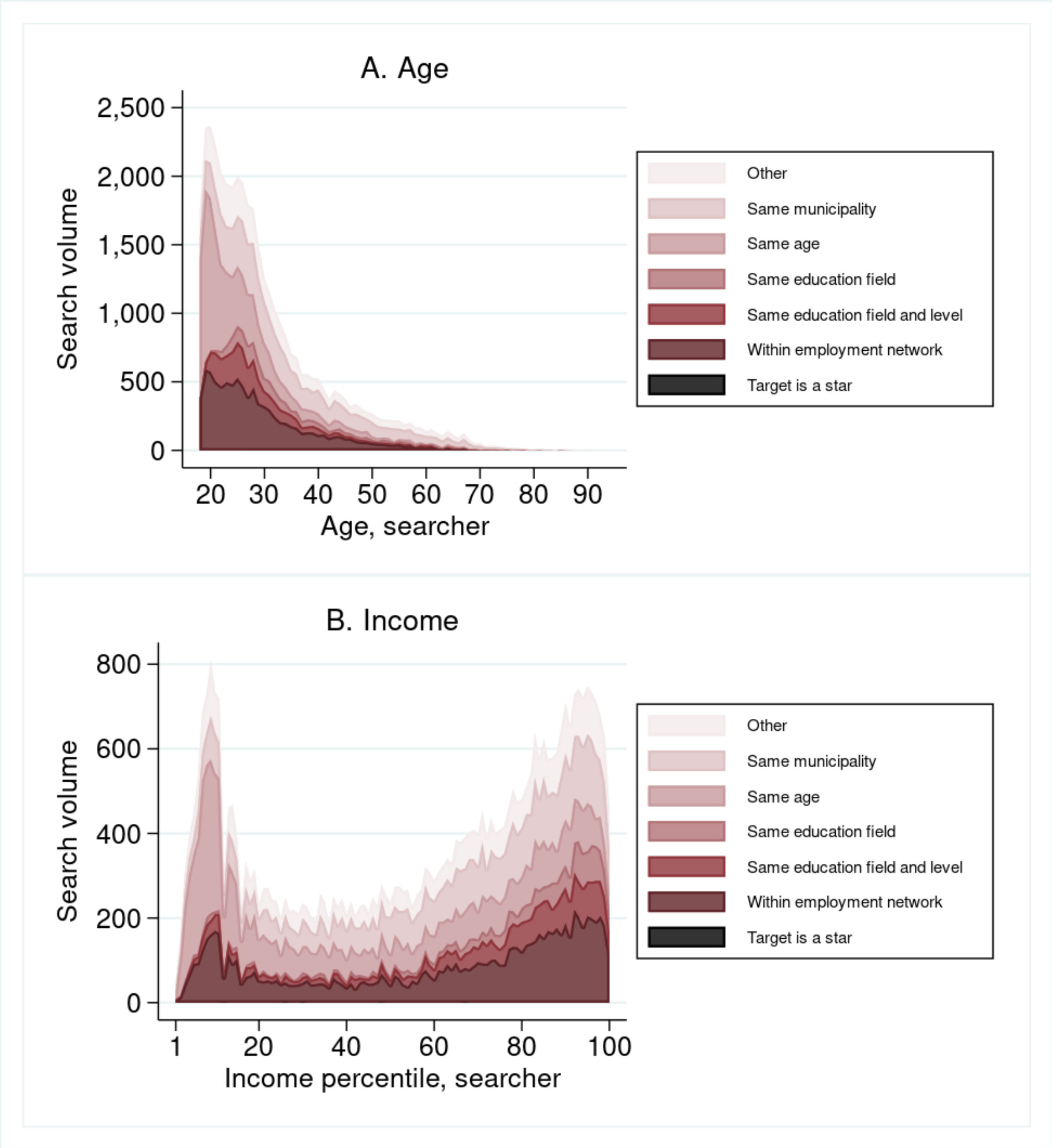
5.4 Reciprocal searching

The reciprocal nature of searches is of interest in part because it has been shown to affect the ability of a group to monitor and enforce behaviors. For example, if there is substantial clustering a cheated individual might inform other people who are also involved in relationships with the cheating agent, who can aid in retribution and punishment. Additionally, Cullen and Perez-Truglia (2018b) show that one coworker revealing their income to another creates an expectation that the other coworker will do the same; the extent of reciprocal searching helps us understand the extent to which this "reciprocal norm" affects search behavior.

Excluding search within households, in 6.3 percent of searches the target also searched for the searcher. This compares to a random searcher-target pair occurring in 0.11 percent of cases, so that reciprocal searching was 60 times more likely than random. It is interesting to note that 44.5 percent of reciprocal searches occur on the same date (as opposed to 11.4 percent for a random searcher-target pair), and 70 percent of reciprocal searches happened within 4 days (compared to 27.4 percent for a random pair).

The nature of reciprocal searches by searcher's age and searcher's income is described in Figure 10. Reciprocity is skewed toward young taxpayers. Regarding income, the pattern reflects some of the patterns shown earlier. Compared to Panel A of Figure 9, Panel B of Figure 10 reveals that there are relatively more reciprocal searches among high-income taxpayers. In both cases, the fraction of searchers within employment networks and the fraction of searches where the searcher and target are of the same age are notably higher. Unsurprisingly, the fraction of reciprocal searches involving stars is negligible, as Norwegian billionaires, politicians, celebrities, and star athletes apparently have no interest in the tax information of the average Anne or Jan who searches for their information.

Figure 10. Decomposing Reciprocal Searches



Notes: Panel A and Panel B decompose reciprocal searches by the age and the income percentile of the searcher, respectively. We define reciprocal searches as both searches conducted by a searcher-target pair, where both searched for one another. As in Figure 9, reciprocal searches (excluding within household searches) are divided into 6 categories by the following order: 1) Within employment network; 2) Same education field and level; 3) Same education field; 4) Same age; 5) Same municipality; 6) Other. Unlike the previous figure, we do not separate out searches for stars; virtually no reciprocal searches involve a star.

6. Tax Disclosure and Tax Compliance

One important justification of public disclosure of tax returns is that it constrains tax evasion, because potential evaders are concerned that others will encounter suspiciously low reported income (or wealth) and report their concerns, and perhaps evidence, to the tax authorities—i.e., they will become whistleblowers. The fact that we find strong evidence of homophily is not inconsistent with this kind of search happening, as potential whistleblowers may have more information about, and more interest in monitoring, people like themselves. A small-town hair salon owner may be especially interested in discovering whether her principal local competitor is gaining an unfair competitive advantage by evading income taxes.[24] Clearly, however, the evidence we have described in this paper suggests that enforcing tax compliance is not the only motivation to search the publicly available tax information. For example, the fact that many Norwegians, and especially young, low-income people, often search for information about celebrities is unlikely to reflect tax compliance concerns.

As discussed in Section 2.3, there is, though, credible evidence from other research—including about Norway—that public disclosure of tax information increases tax compliance of those with significant latitude for tax evasion, in particular self-employed people. The effect on tax compliance likely comes from the perceived *threat* of whistleblowing, which is not well measured by the extent of such information provided in equilibrium or the tax collections directly tied to information from whistleblowers.[25] Indeed, a well-known feature of many game-theoretic models is that in equilibrium, agents may never follow through on a threat if other agents believe that the threat is credible and respond accordingly. More concretely, Kleven, Kreiner, and Saez (2016) describe a model where a whistleblower threat deters tax evasion but, in equilibrium, whistleblowing seldom occurs. Nevertheless, it is of interest to assess the extent to which searches are motivated by concerns about the tax compliance of others, and to what degree being targeted for a search increases tax compliance. In what follows we examine each of these two issues.

6.1 How much search is tax-motivated?

We can calculate an upper bound of compliance-related searches by noting that in Norway, as in most developed countries, third-party information reporting severely limits evasion possibilities for most

[24] It is conceivable that the public disclosure scheme erodes tax compliance by facilitating a race to the bottom—learning about the surprisingly low reported taxable income tax someone else is apparently "getting away with" could induce more aggressive tax reports by the searcher. This avenue of influence has not been pursued in the literature, and we are grateful to a referee for suggesting it.

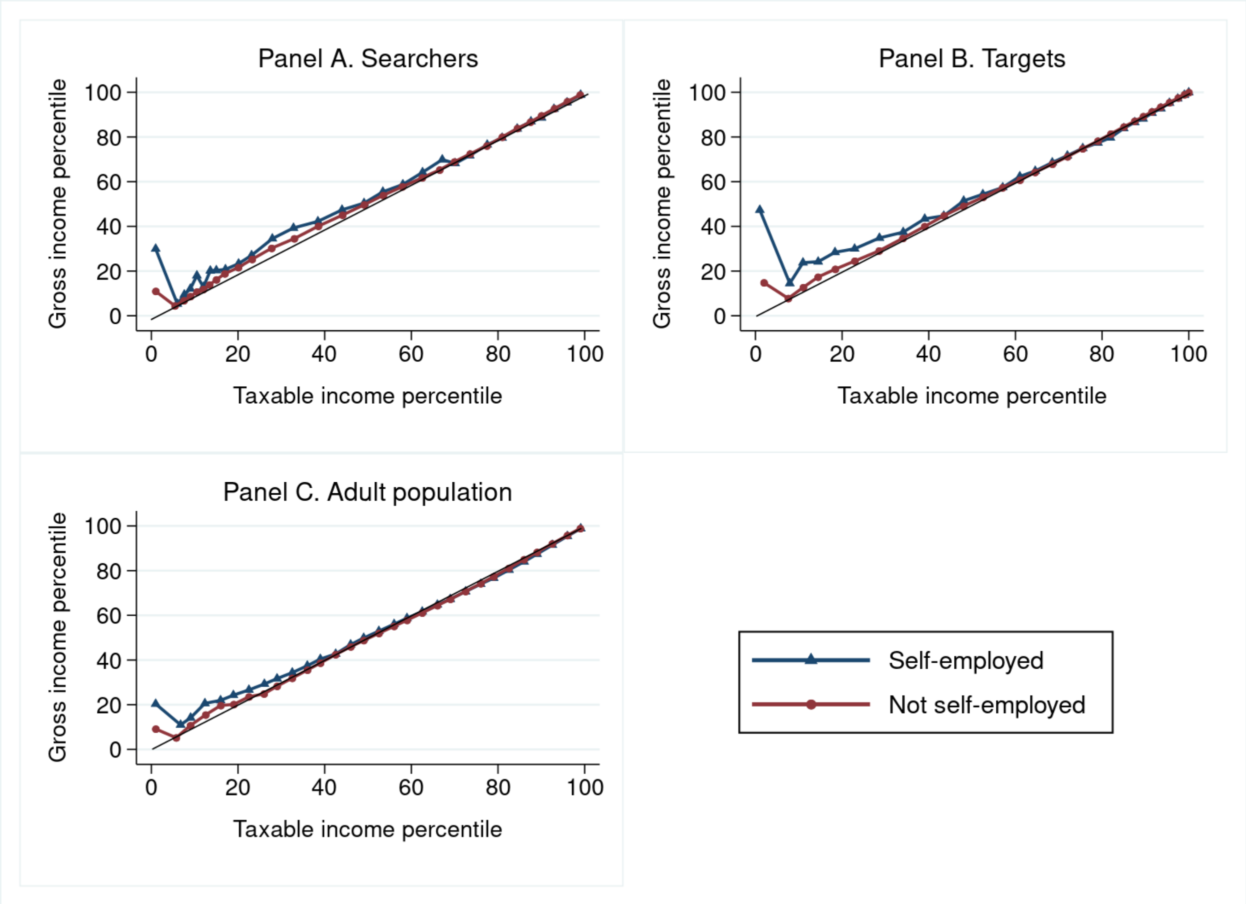
[25] Information on the extent and nature of whistleblowing would, nevertheless, be of substantial interest. We requested, alas unsuccessfully, data about tax evasion tips in Norway.

employees, but not as much for the self-employed, as the results of Bø, Slemrod, and Thoresen (2015) suggest. Recall from Table 1 that self-employed people comprise 12.1 percent of targets, although they make up just 8.6 percent of the adult population. If compliance-related searches target only self-employed people, and all searches for self-employed people are tax-compliance-motivated, then the fraction of such searches is at most 12.1 percent. But some self-employed people are certainly targeted even in the absence of any suspicion regarding tax evasion or thought to informing the tax authorities about any suspicious tax reporting behavior. To get a sense of this, Table 3 shows that a self-employed person has a statistically significant 4.2 percentage point higher probability of being targeted. We thus estimate that the counterfactual targeting of self-employed people would be 7.9 percent (12.1 minus 4.2) in the absence of tax-compliance-related search, and about 4.2 percent of all searches in Norway were tax-compliance-related. This estimate depends on an assumption that, ceteris paribus, the only reason that self-employed people are targeted disproportionately is to learn about their likely tax compliance. If self-employed people are more (or less) likely to be search targets for non-compliance-related reasons that are not controlled for in the regression of Table 3, then the 4.2 percent figure is subject to error.

In sum, that the tax information of self-employed people is substantially more likely to be targeted is consistent with the idea that potential whistle blowing is a non-trivial motivation for searching. At the same time, we conclude that the great majority of searches of the public disclosure system are not motivated by potential whistleblowing. This does not necessarily indicate that the deterrence effect of *potential* whistle-blowing is small. Much of the effect of public disclosure on tax compliance may be coming from the availability of information rather than whether information is actually accessed in equilibrium, so the volume of whistle-blowing-related searches is not directly informative about the deterrent effect of public disclosure.

We have one further piece of evidence of the role of compliance-related searches. Our focus so far in this paper has been on reported gross (i.e., before deductions) income, but self-employed people have much more latitude to deduct certain expenses to obtain a lower taxable income. Figure 11 demonstrates that self-employed taxpayers have more possibilities to deduct expenses to obtain a lower taxable income, compared to the case for non-self-employed taxpayers. As Panel B shows, this is especially the case for self-employed targets, which is consistent with people searching more often for people with relatively low reported *taxable* income, which may possibly indicate tax evasion or avoidance in the form of overstated deductions.

Figure 11: Gross Income vs Taxable Income for Searchers, Targets and the Full Population



Notes: The panel presents binned scatterplots of individuals' gross income percentile by the same individuals' taxable income percentile. Panel C presents figures for the overall adult population (at least 18 years old), whereas Panel A and Panel B restrict attention to searchers and targets, as defined above. All panels include separate binned scatterplots for self-employed and not self-employed individuals (everyone not self-employed), where self-employed are defined as individuals with non-zero (positive or negative) income from self-employment.

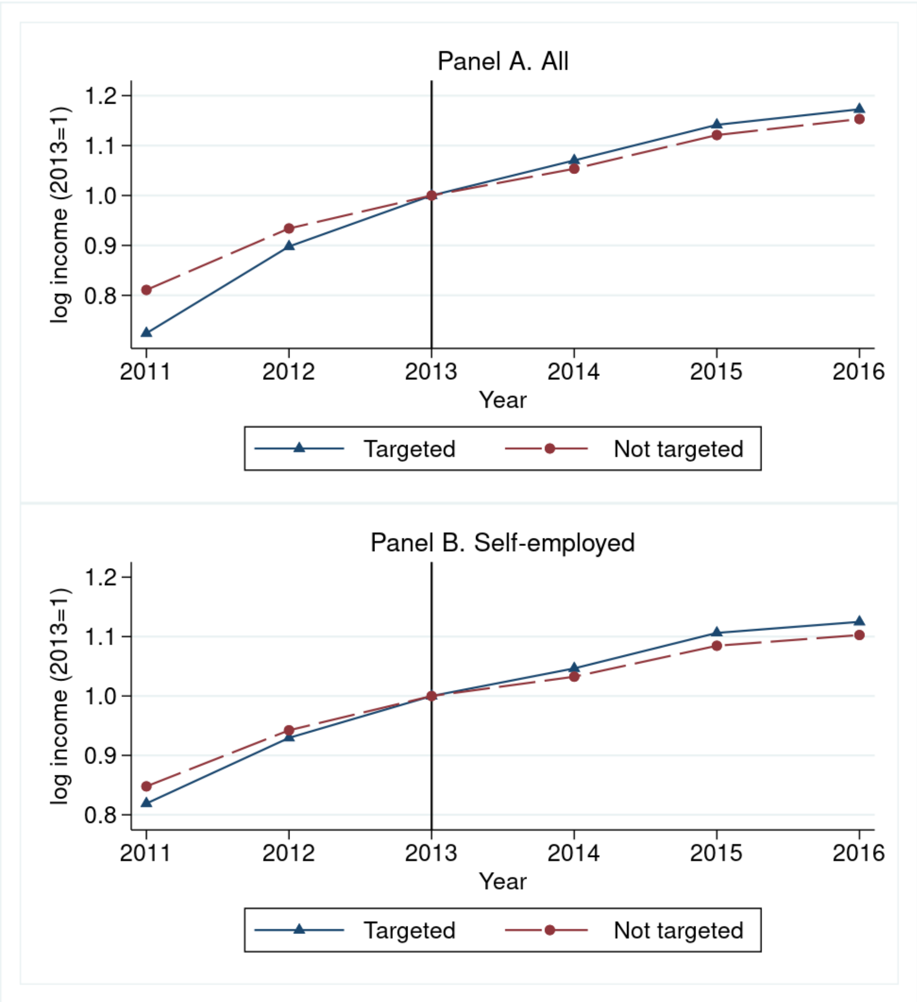
6.2. Does being targeted change tax reporting behavior?

Getting searched may send a message to some targets: "I'm watching you." It might suggest that whistleblowing will ensue, depending on what is learned from the search and what is already known by the searcher about the target. In the absence of micro data on whistleblowers, we cannot directly study the connection between information search and subsequent whistleblowing activity. We can, though, analyze whether targets' tax reporting behavior changes after being targeted, recognizing that a change in reported income reflects both any change in true income as well as any change in the extent of noncompliance.[26]

[26] For this analysis, we exclude self-searches and searches between individuals living in the same household.

We begin by studying aggregate time-series data on income reports for search targets versus non-targets. The upper panel of Figure 12 demonstrates that the trends of income reporting are slightly different between those targeted and not targeted throughout our period of observation. Apparently, searchers tend to target those whose income is growing faster than non-targets. We do not see any particular divergence of this trend in 2014, which is suggestive of little to no causal effect of being targeted on tax reporting behavior. In the lower panel of Figure 12, we condition on both targeted and not targeted individuals being self-employed in 2014. This selection isolates the group for whom a causal tax compliance effect of being targeted is more likely. Conditioning on being self-employed eliminates some (but not quite all) of the differential trend between the targeted and not targeted before 2014, but we still see no differential break from trend after 2014. In short, we are unable to entirely rule out selection bias in an analysis comparing those who are and are not targeted for search, but the results suggest that the causal effect of interest, if it exists, is likely small.

Figure 12: Time-Series Plots. Targeted versus Not Targeted



Notes: This figure displays the evolution over time of the mean of log income in the given year for targets and non-targets. To make the time series comparable, we subtract from each series the mean of log income in 2013 and add 1. Panel B does the same analysis restricting to individuals who were self-employed in 2013. We add a black vertical line at 2013. Tax reporting in 2013 or earlier happens before searches occur while reported income in 2014 and later can possibly be affected by being targeted.

We next implement a more sophisticated causal design to test for compliance effects. We address the potential selection bias from comparisons between those targeted and others by examining the change in reporting behavior only of those taxpayers who were targeted by a search, differentiated by the *timing* of the search. In particular, we examine whether the first search for that taxpayer occurred before or after when their tax return was filed.

The idea behind this design is that if a search occurs before filing, the taxpayer upon logging on to fill out her return would be able to see that the search has occurred and respond to it, potentially by increasing reported income; if, on the other hand, the search occurs after filing, it would be too late to increase reported income (for the current tax year). Our identification strategy presumes that the nature of searches and targets of searches occurring before or after filing are otherwise comparable, an assumption we can explore with placebo tests and pre-trend analysis.

Table 5 shows the results of a specification designed to reveal the effect of being targeted on the tax year 2014 income report. First, we restrict the sample to only individuals who were targeted by a search for the first time either two months before or two months after the filing date, to minimize the possibility of selection bias based on the timing of search. We estimate a regression with individual and year fixed effects, where the dependent variable is the log of reported taxable income, first for all individuals, and then separately restricting the sample to self-employed individuals. Observations with zero or negative values for reported income in any year are discarded, so the regression includes only individuals declaring positive income in each tax year, from 2011 to 2016. We include individual fixed effects to account for the fact that the timing of search may be correlated with unobservable determinants of income; no interactions referring to 2013 are included, so estimated coefficients refer to the effect relative to 2013.

Table 5: OLS Regression Analyses of the Effect of Being Targeted for on Reported Income

	(1)	(2)	(3)	(4)
Dependent variable		log taxable income		
Sample	All	Self-employed	All	Self-employed
Year 2011 * targeted before filing	0.0070	0.0219		
	(0.0103)	(0.0222)		
Year 2012 * targeted before filing	-0.0033	0.0213		
	(0.0083)	(0.0185)		
Year 2014 * targeted before filing	**0.0000**	**0.0110**		
	(0.0077)	**(0.0185)**		
Year 2015 * targeted before filing	-0.0050	0.0053		
	(0.0091)	(0.0198)		
Year 2016 * targeted before filing	-0.0030	0.0022		
	(0.0099)	(0.0214)		
Year 2011 * targeted before deadline			-0.0049	0.0012
			(0.0104)	(0.0234)
Year 2012 * targeted before deadline			-0.0052	0.0126
			(0.0084)	(0.0205)
Year 2014 * targeted before deadline			**-0.0027**	**0.0183**
			(0.0078)	**(0.0195)**
Year 2015 * targeted before deadline			-0.0061	0.0136
			(0.0092)	(0.0211)
Year 2016 * targeted before deadline			-0.0020	0.0315
			(0.0100)	(0.0232)
Year fixed effects	yes	yes	yes	yes
Individual fixed effects	yes	yes	yes	yes
Number of observations	339,316	37,237	338,305	36,566
R-squared	0.0091	0.0088	0.0090	0.0091

Notes: This figure reports the results of regression analysis on the effect of being targeted on reported income. The regression sample includes all individuals targeted for the first time within the time window of +/- 60 days around the filing date (columns 1 and 2) or the deadline for filing (in columns 3 and 4). Self-employed are defined as individuals with non-zero (positive or negative) income from self-employment in year 2013. The dependent variable is the log of taxable income ("*alminnelig inntekt*"), which is tax base for income taxes.

The effects of interest in Table 5 are the year dummy variables interacted with the variable *targeted before filing*, which is a dummy variable that takes a value of one if the first search for tax year 2013 information was conducted in the two months window before the filing date for tax year 2014, and zero if the search occurred in the two months window after the filing date. If being targeted during the filing window increases reported income in 2014, we would expect to see that the estimated coefficient on the interaction term to be higher in tax year 2014 than in other years. Note that, in years before tax year 2014, the taxpayer could not know whether she had been targeted in the later year, so the pre-filing versus post-filing behavior would not be affected by learning of a search. Observing whether these effects are statistically different from zero is a pre-period test of the validity of the research design: if there are confounding differences between those targeted before or after filing throughout the period, these are likely reflected in pre-period trends in income (as in Figure 12).

Table 5 reveals that the estimated effect of a search before filing (*year 2014 * targeted before filing*) on the log of reported income is not statistically significant, either for all taxpayers or for only self-employed individuals. The causal effect of being targeted on income reporting is small and statistically indistinguishable from zero.[27] The point estimate for the self-employed is positive, as a compliance effect would suggest, but the *t*-statistic attached to this estimate is barely above one, and the point estimate is small, suggesting that being targeted increases incomes by perhaps 1.1 percent for the self-employed.

A concern with this specification is that when one files is endogenous. Indeed, being targeted could directly affect when one files. Taxpayers can file a return and then revise it several times before the filing deadline. The filing date we observe in such instances would be the last date the individual revises their return. An individual could file before they were targeted, see that they were targeted, and then file again (perhaps because of being targeted and wishing to report more truthfully). Such an individual would be classified as “targeted after filing” in Table 5 when, in reality, they first filed before they were targeted and did respond to the policy, which could introduce selection bias into our estimates.

We address this concern with an instrumental variables design leveraging the fact that searches after the filing deadline could not affect the timing of filing. We therefore use whether the search occurred before the filing deadline as an instrument for whether the search occurred before the individual filed. For regular wage earners and pensioners, the filing deadline was April 30th, while the filing deadline for self-employed people was May 31st. Columns (3) and (4) of Table 5 implement the reduced form of this design, replacing “targeted before filing” with “targeted before deadline” as the main treatment variable. We find very similar results, suggestive of little to no effect of search on reported income.

For simplicity, we do the actual instrumental variables analysis cross-sectionally, with the change in log income from 2013 to 2014 as the outcome of interest. The OLS results, shown in columns (1) and

[27] The same null result obtains if we replace log income with an inverse hyperbolic sine transformation of the reported income, or use as an outcome variable the presence of any self-employment income.

(2) of Table 6, are similar to the year 2014 interactions in Table 5, suggesting a small and insignificant effect of being targeted on compliance. The first-stage results in columns (3) and (4) of Table 6 show that, unsurprisingly, being targeted before the filing deadline is highly predictive of being targeted before filing. Finally, in columns (5) and (6), we report the IV estimates. As in prior specifications, the point estimate of the effect of being targeted for the self-employed remains small and statistically insignificant.

All in all, our analysis of the aggregate time series and micro search data does not generate compelling evidence that being targeted increases reported incomes. Taken together with the results of Section 6.1 about the prominence (or more specifically, the lack of prominence) of searches for those with more easily evadable self-employment income, we conclude that we cannot confidently ascribe to individual searches a compliance-increasing intent or outcome. As discussed above, this conclusion is not necessarily inconsistent with the findings of other research that the public disclosure *system* increases compliance because the compliance effect of this system on tax compliance may derive from the *availability* of information rather than the extent to which this information is actually accessed in equilibrium.

Table 6: IV Analyses of the Effect of Being Targeted on Reported Income

	(1)	(2)	(3)	(4)	(5)	(6)
	OLS		First stage		IV	
Dependent variable	log taxable income 2014- log taxable income 2013		Searched before filing		log taxable income 2014- log taxable income 2013	
	All	Self-employed	All	Self-employed	All	Self-employed
Searched before filing	0.0019	0.0135			-0.0001	0.0168
	(0.0075)	(0.0172)			(0.0083)	(0.0312)
Searched before deadline			0.9257	0.6532		
			(0.0016)	(0.0106)		
Number of observations	57,111	6,135	56,929	5,305	54,695	5,014
R-squared	0.0000	0.0000	0.8485	0.4160	.	.
F-statistic			>99,999	3778.3		

Notes: Columns (1) and (2) report OLS estimates of the effect of being searched for before filing on reported taxable income in 2014, using the first difference in log taxable income between 2014 and 2013. Note that the individual fixed effects included in the specification from Table 5 are accounted for by using such a difference specification, in these columns and in columns (5) and (6). These results include all individuals targeted for the first time within the time window of +/- 60 days around the filing date. Columns (3) and (4) present the first stage of a 2SLS regression, where searched before filing is the endogenous variable and searched before deadline is the exogeneous variable. Columns (5) and (6) present the results of the second stage of a 2SLS regression, where the predicted value of searched before filing from columns (3) or (4) is used to estimate the causal effect of being targeted on reported income growth from 2013 to 2014. Searched before filing/deadline are indicator variables of 0 (targeted after filing) or 1 (targeted before filing). Even-numbered columns restrict the sample to self-employed individuals, defined as individuals with non-zero (positive or negative) income from self-employment in year 2013.

7. Conclusions

People query public information on the affluence of others for manifold reasons. Some people may feel envious, even jealous, of other people's level of affluence. Others may feel happy, even elated, from learning that their own affluence exceeds that of others. Still others may wish to learn about just how successful their role models are, or maybe they are just plain curious. Perhaps others are private rule enforcers, even at a cost to themselves, who resent that some people get away with the tax equivalent of murder—evasion. Public disclosure of tax returns, a policy motivated in part by its perceived effect on tax compliance, also facilitates the satisfaction of curiosity about those in one's social network and, especially for low-income people, about the rich and famous, as well as aiding income comparisons motivated by envy or collecting information about compensation in other jobs.

In this paper we shed light on the motivation for tax searches by undertaking the first-ever analysis of the actual searches done in a public tax disclosure system, comprising over 1 million searches done in 2014 and 2015 in Norway. We characterize the social network these searches comprise, including its degree of homophily, reciprocation, the extent to which searches occur in identifiable social networks, and the characteristics of searchers and targets. We find that about one-quarter of searches occur within identifiable social networks of households and employment. More broadly, searchers tend to target people similar to them, but young low-income people also non-trivially target older successful people and celebrities.

We also study the extent of compliance related search and the effects of searches on target tax compliance, primarily focusing on the self-employed because third-party information reporting in Norway severely limits evasion possibilities for most employees. The tax information of self-employed people is, ceteris paribus, significantly more likely to be targeted, and potential whistleblowing may be a motive for many of these searches. However, at least 90 percent of searches are highly unlikely to be driven by whistle-blowing and tax compliance motivations, and our causal analysis of the effect of search on targets' income reporting finds small, statistically insignificant effects, even for self-employed targets. This evidence suggests that in equilibrium, compliance-motivated searching does not occur frequently, and searches themselves have minimal effects on compliance. However, these findings are not directly informative about the impact of public disclosure on tax compliance overall, as much of the deterrent effect may be coming from the availability of information rather than whether information is actually accessed in equilibrium.

References

Amir, Eli, Adi Lazar, and Shai Levi. 2018. “The Deterrent Effect of Whistleblowing on Tax Collections.” *European Accounting Review* 27.5: 939-954.

Aruguete, Mara S., Ho Huynh, Lynn E. McCutcheon, Blaine L. Browne, Bethany Jurs, and Emilia Flint. 2019. "Are Measures of Life Satisfaction Linked to Admiration for Celebrities?" *Mind & Society* 18.1: 1-11.

Bevanger, Lars. 2017. “Norway; The Country Where No Salaries Are Secret.” *BBC News*, accessed at https://www.bbc.com/news/magazine-40669239.

Bø, Erlend E., Joel Slemrod, and Thor O. Thoresen. 2015. “Taxes on the Internet: Deterrence Effects of Public Disclosure.” *American Economic Journal: Economic Policy* 7.1: 36-62.

Bursztyn, Leonardo, and Robert Jensen. 2015. “How Does Peer Pressure Affect Educational Investments?” *Quarterly Journal of Economics* 130.3: 1329-1367.

Card, David, Alexandre Mas, Enrico Moretti, and Emmanuel Saez. 2012. “Inequality at Work: The Effect of Peer Salaries on Job Satisfaction.” *American Economic Review* 102.6: 2981-3003.

Clark, Andrew E., and Claudia Senik. 2010. “Who Compares to Whom? The Anatomy of Income Comparisons in Europe.” *Economic Journal* 120.544: 573-594.

Cullen, Zoë, and Ricardo Perez-Truglia. 2018a. “How Much Does Your Boss Make? The Effects of Salary Comparisons.” Working paper No. 24841, National Bureau of Economic Research.

Cullen, Zoë, and Ricardo Perez-Truglia. 2018b. “The Salary Taboo: Privacy Norms and the Diffusion of Information.” Working paper No. 25145, National Bureau of Economic Research.

Duesenberry, James S. 1949. *Income, Saving and the Theory of Consumer Behavior*. Cambridge, MA: Harvard University Press.

Fehr, Ernst, and Simon Gächter. 2000. “Fairness and Retaliation: The Economics of Reciprocity.” *Journal of Economic Perspectives* 14.3: 159-181.

Frank, Robert H. 1985. *Choosing the Right Pond: Human Behavior and the Quest for Status*. Oxford: Oxford University Press.

Hasegawa, Makoto, Jeffrey Hoopes, Ryo Ishida, and Joel Slemrod. 2012. “The Effect of Public Disclosure on Reported Taxable Income: Evidence from Individuals and Corporations in Japan.” *National Tax Journal* 66.3: 571-608.

Hirschman, Albert O. and Michael Rothschild. 1973. “The Changing Tolerance for Income Inequality in the Course of Economic Development.” *Quarterly Journal of Economics* 87.4: 544-566.

Hoopes, Jeffrey, Daniel Reck, and Joel Slemrod. 2014. “Taxpayer Search for Information: Implications for Rational Attention.” *American Economic Journal: Economic Policy* 7.3: 177-208.

Hoopes, Jeffrey, Leslie Robinson, and Joel Slemrod. 2018. “Public Tax-Return Disclosure.” *Journal of Accounting and Economics* 66.1: 142-162.

Jackson, Matthew O. 2008. *Social and Economic Networks*. Vol. 3. Princeton: Princeton University Press.

Jackson, Matthew O. 2014. “Networks in the Understanding of Economic Behaviors.” *Journal of Economic Perspectives* 28.4: 3-22.

Jackson, Matthew O., Brian W. Rogers, and Yves Zenou. 2017. “The Economic Consequences of Social Network Structure.” *Journal of Economic Literature* 55.1: 49-95.

Kahneman, Daniel and Amos Tversky. 1979. “Prospect Theory: An Analysis of Decision under Risk.” *Econometrica* 47.2: 263-292.

Kleven, Henrik Jacobsen, Claus Thustrup Kreiner, and Emmanuel Saez. 2016. “Why Can Modern Governments Tax So Much? An Agency Model of Firms as Fiscal Intermediaries.” *Economica* 83.330: 219-246.

Konrad, Kai A., Harald W. Lang, and Florian Morath. 2015. “A Glance into the Tunnel: Experimental Evidence of Expectations versus Comparison Considerations.” VfS Annual Conference 2015 (Muenster): Economic Development - Theory and Policy 113017, Verein für Socialpolitik, German Economic Association.

Leibenstein, Harvey. 1962. “Notes on Welfare Economics and the Theory of Democracy.” *The Economic Journal* 72.286: 299–319.

Luttmer, Erzo F. P. 2005. “Neighbors as Negatives: Relative Earning and Well-Being.” *Quarterly Journal of Economics* 120.3: 963-1002.

McPherson, Miller, Lynn Smith-Lovin, and James M. Cook. 2001. “Birds of a Feather: Homophily in Social Networks.” *Annual Review of Sociology* 27.1: 415-444.

Ministry of Finance (Norway). 2014. Høringsnotat (Consultation note), January 24.

Perez-Truglia, Ricardo. 2020. “The Effects of Income Transparency on Well-Being: Evidence from a Natural Experiment.” *American Economic Review* 110.4: 1019-1054.

Perez-Truglia, Ricardo, and Ugo Troiano. 2018. "Shaming Tax Delinquents." *Journal of Public Economics* 167.11: 120-137.

Reck, Daniel, and Arthur Seibold. 2021. “The Welfare Economics of Reference Dependence.” Working paper, London School of Economics.

Rege, Mari, and Ingeborg F. Solli. 2013. “Lagging Behind the Joneses: The Impact of Relative Earnings on Job Separation.” Working paper, University of Stavanger.

Senik, Claudia. 2004. “When Information Dominates Comparison: Learning from Russian Subjective Panel Data.” *Journal of Public Economics* 88.9-10: 2099-2123.

Senik, Claudia. 2008. “Ambition and Jealousy: Income Interactions in the ‘Old’ Europe versus the ‘New’ Europe and the United States.” *Economica* 75.299: 495-513.

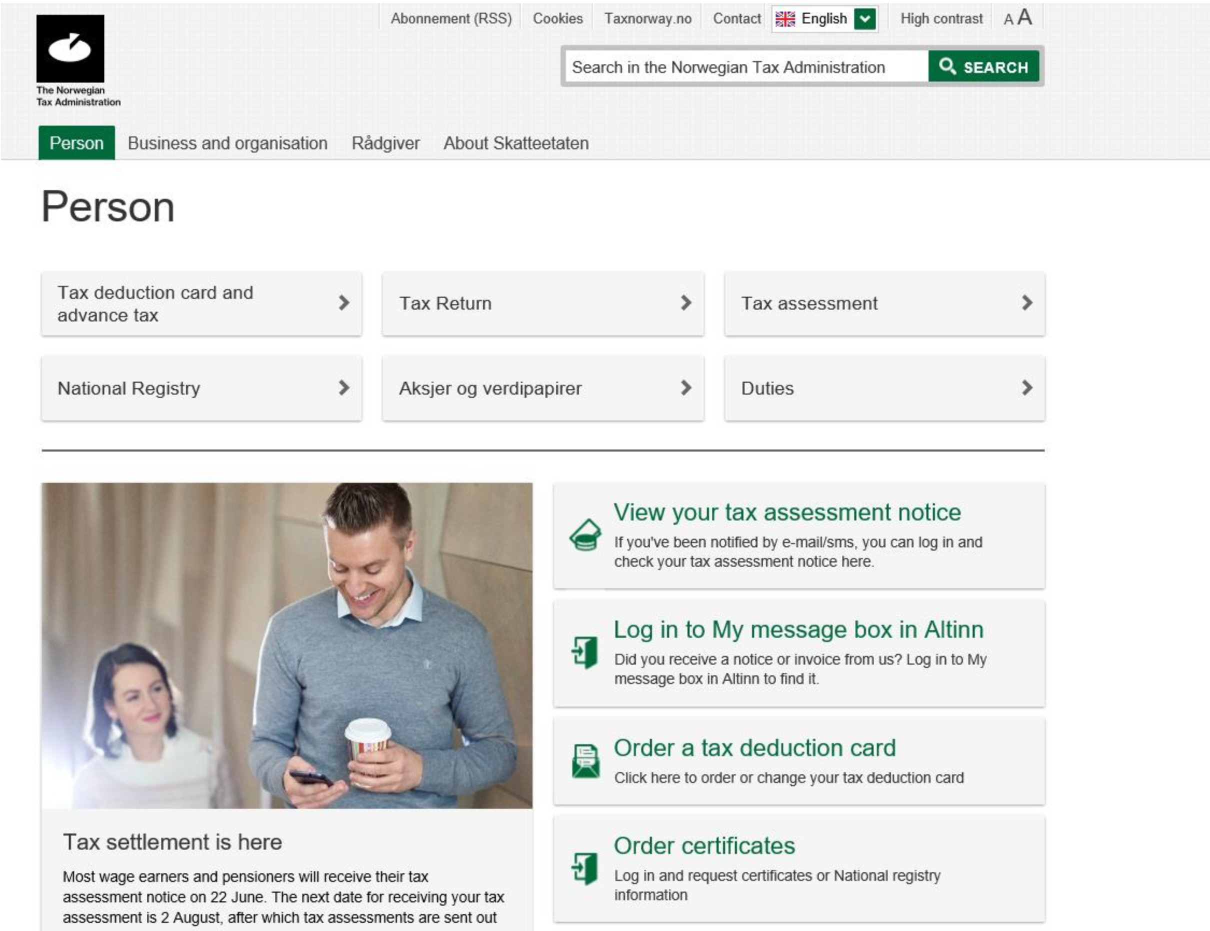
Slemrod, Joel, Obeid Ur Rehman, and Mazhar Waseem. 2020. "How Do Taxpayers Respond to Public Disclosure and Social Recognition Programs? Evidence from Pakistan." *The Review of Economics and Statistics*: 1-44.

Appendix A: Screen Shot of Search Website

Screenshots Public tax information:

Accessed at www.skattetaten.no.

The first few screenshots show the website as one navigates to the place where one can search for others.



Abonnement (RSS) | Cookies | Taxnorway.no | Contact | English | High contrast | A A

The Norwegian Tax Administration

Search in the Norwegian Tax Administration | SEARCH

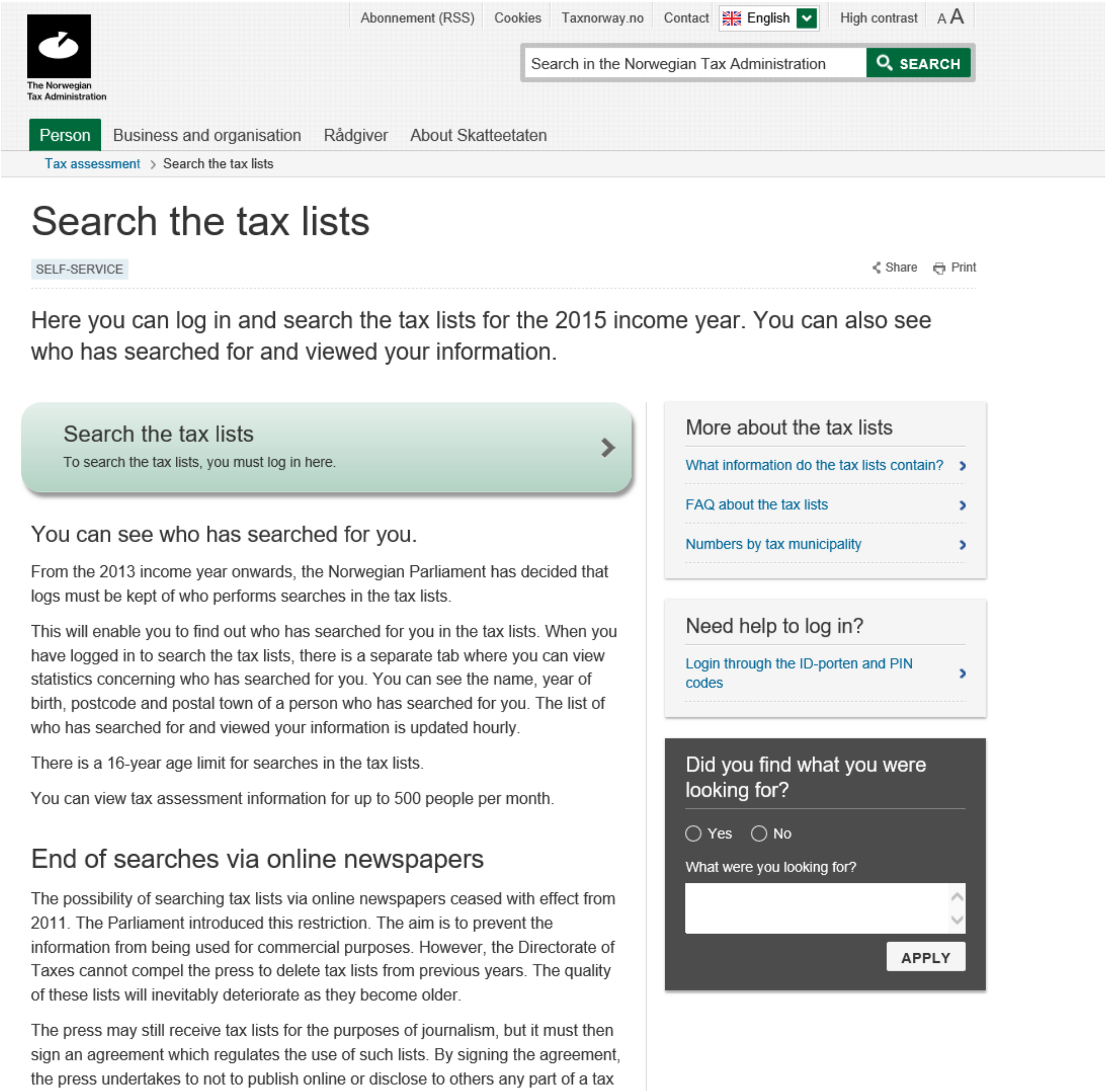
Person | Business and organisation | Rådgiver | About Skatteetaten

› Tax assessment

Tax assessment

Share | Print

- View your tax assessment notice ›
- Underpaid tax, interest and payment ›
- E-invoices for tax payment ›
- Appeal ›
- Calculate your tax ›
- Search the tax lists ›



Abonnement (RSS) | Cookies | Taxnorway.no | Contact | English | High contrast | A A

The Norwegian Tax Administration

Search in the Norwegian Tax Administration SEARCH

Person | Business and organisation | Rådgiver | About Skatteetaten

Tax assessment > Search the tax lists

Search the tax lists

SELF-SERVICE

Share Print

Here you can log in and search the tax lists for the 2015 income year. You can also see who has searched for and viewed your information.

Search the tax lists

To search the tax lists, you must log in here.

You can see who has searched for you.

From the 2013 income year onwards, the Norwegian Parliament has decided that logs must be kept of who performs searches in the tax lists.

This will enable you to find out who has searched for you in the tax lists. When you have logged in to search the tax lists, there is a separate tab where you can view statistics concerning who has searched for you. You can see the name, year of birth, postcode and postal town of a person who has searched for you. The list of who has searched for and viewed your information is updated hourly.

There is a 16-year age limit for searches in the tax lists.

You can view tax assessment information for up to 500 people per month.

End of searches via online newspapers

The possibility of searching tax lists via online newspapers ceased with effect from 2011. The Parliament introduced this restriction. The aim is to prevent the information from being used for commercial purposes. However, the Directorate of Taxes cannot compel the press to delete tax lists from previous years. The quality of these lists will inevitably deteriorate as they become older.

The press may still receive tax lists for the purposes of journalism, but it must then sign an agreement which regulates the use of such lists. By signing the agreement, the press undertakes to not to publish online or disclose to others any part of a tax

More about the tax lists

- What information do the tax lists contain?
- FAQ about the tax lists
- Numbers by tax municipality

Need help to log in?

- Login through the ID-porten and PIN codes

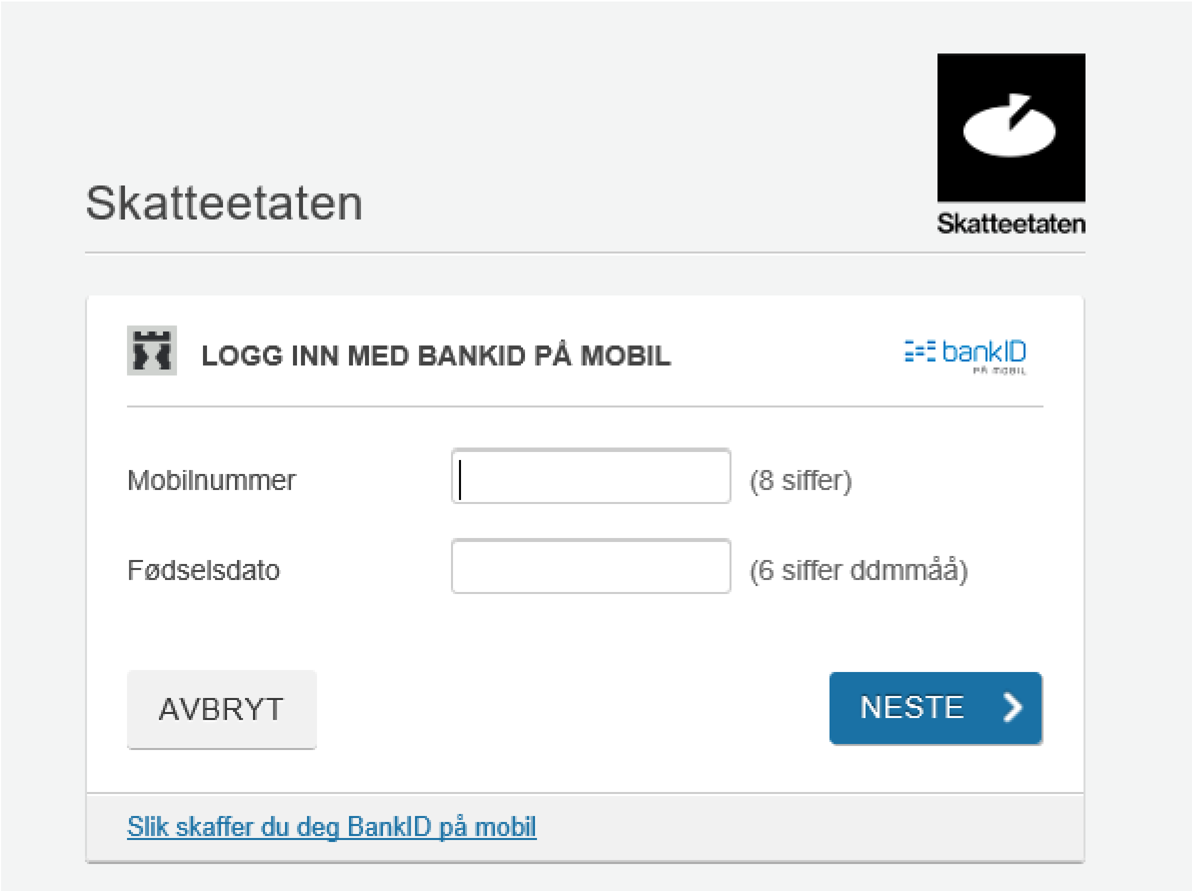
Did you find what you were looking for?

Yes No

What were you looking for?

APPLY

The above are English versions of web pages that are also available to taxpayers in Norwegian. After this page, there is no option of an English translation. From this point, one logs in by entering one’s phone number and date of birth, and verifying one’s identity via a SMS message sent to one’s phone, which is registered with the government. There are also other ways to log in, but this method is the simplest one for most people.



Here, after logging in, one can enter the name of the person he or she wishes to search for. One must enter at least first+last name or last name+year of birth (need to click on "advanced search") in order to get any search results.

When clicking on *søkestatistikk* (search statistics), one can see who searched for you:

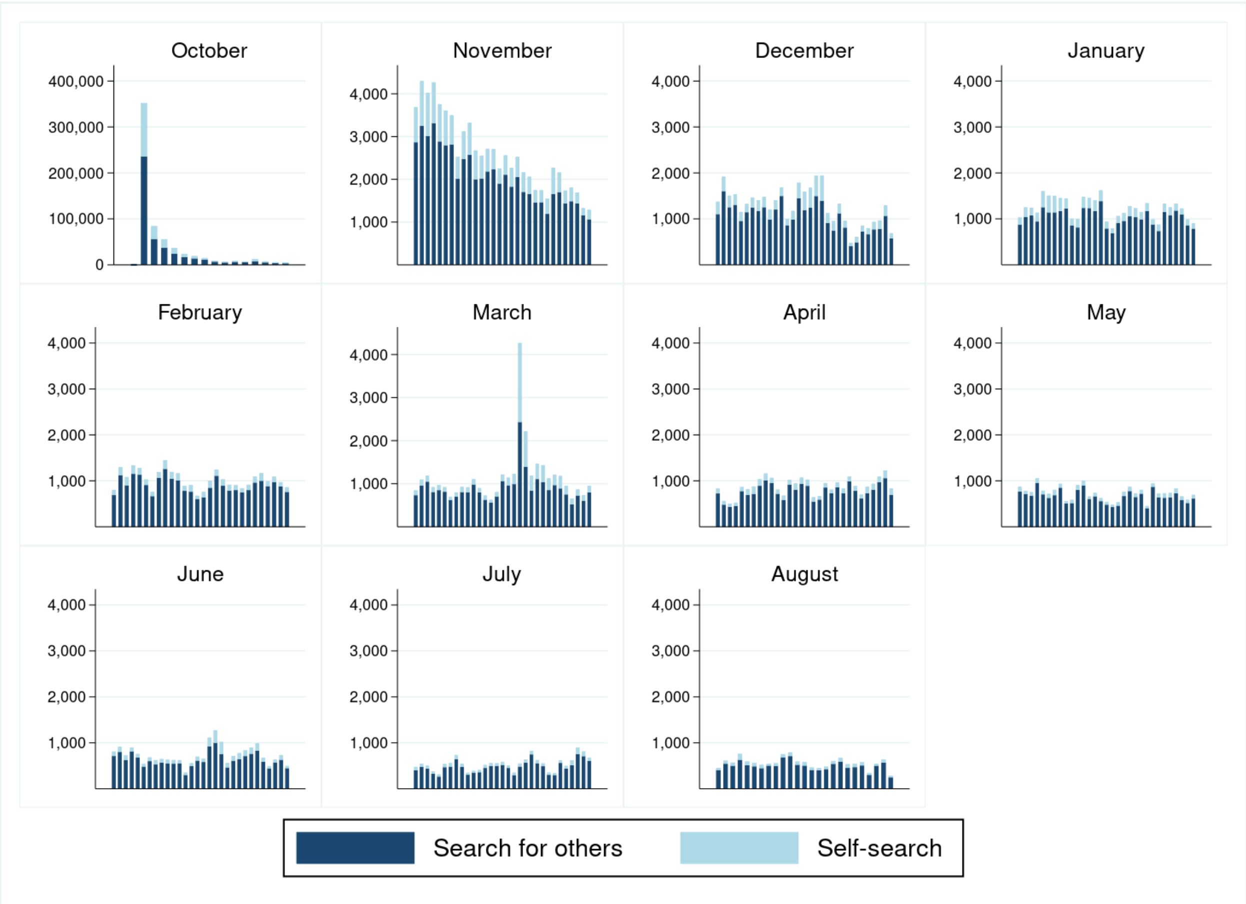
Appendix B: Tables

Table A.1: Searches by Month, 2014-2015

	Observations	Percent of Total
October	648,288	66.9
November	77,910	8.0
December	39,558	4.0
January	37,962	3.9
February	28,851	3.0
March	35,256	3.6
April	26,318	2.7
May	22,873	2.4
June	21,949	2.3
July	16,813	1.7
August	14,026	1.5

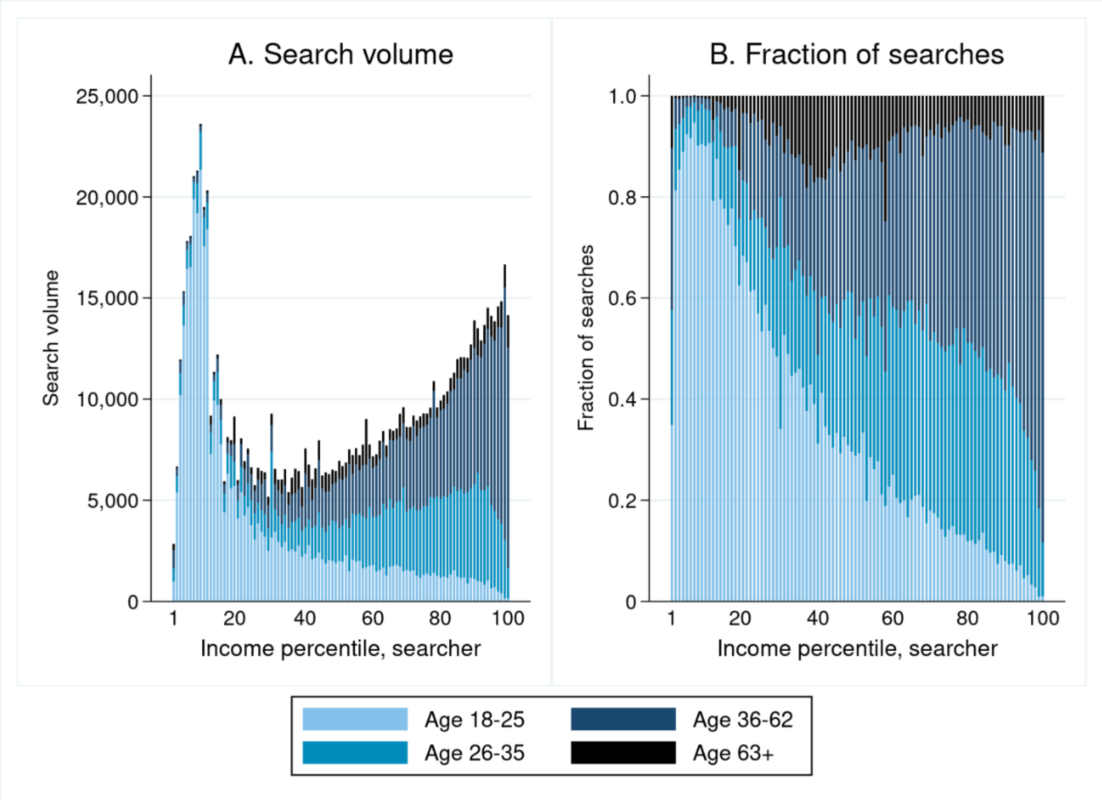
Notes: All searches in our sample, by the month the search was conducted.

Figure A-1. Daily Number of Searches by Month



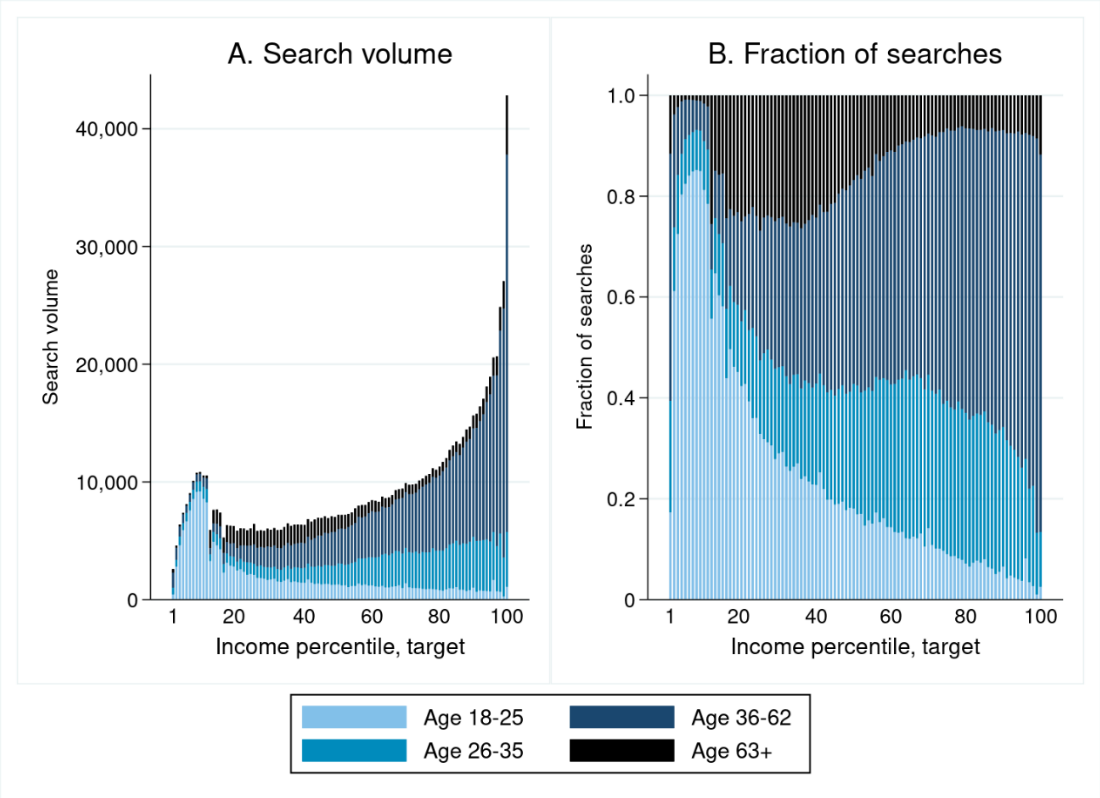
Notes: All searches in our sample, by the month and the day the search was conducted. Searches are categorized into self-search and search for others.

Figure A-2: Searcher's Age Distribution of Search Volume by Searcher's Income Percentile



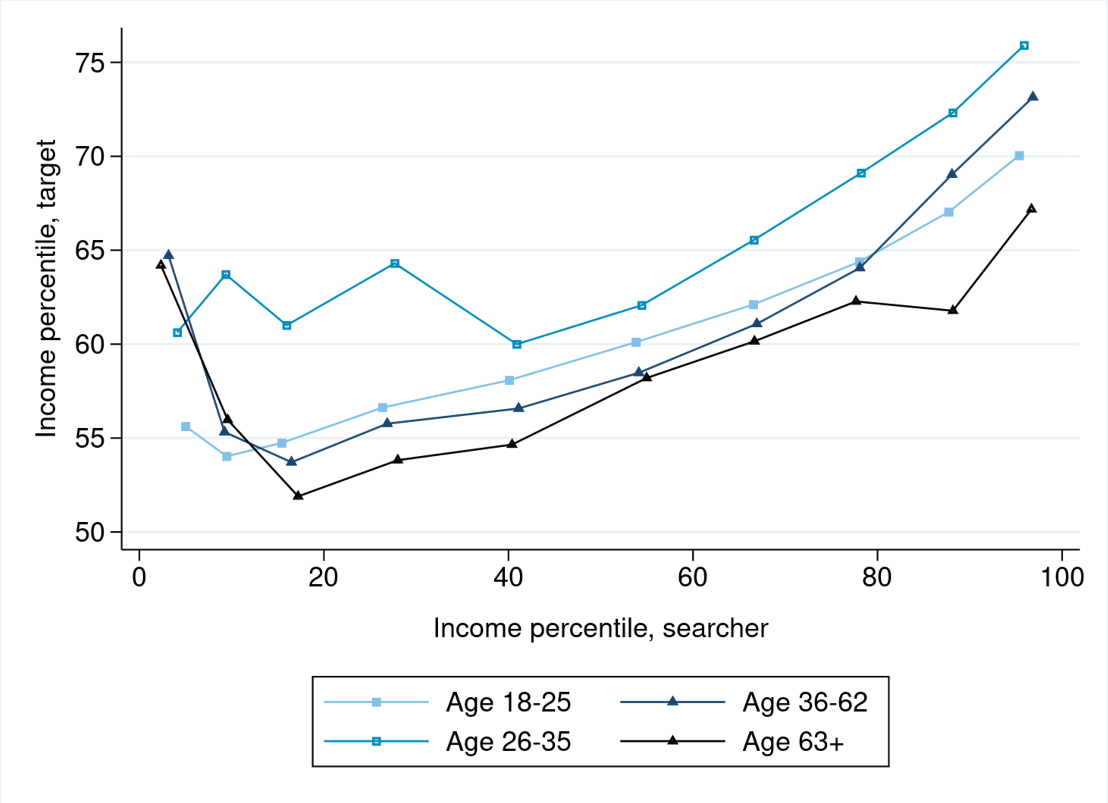
Notes: The panels present a decomposition of search volume (Panel A) and the fraction of search volume (Panel B), respectively, in four age categories of the searcher by the income percentile of the searcher.

Figure A-3: Target's Age Distribution of Search Volume by Target's Income Percentile



Notes: The panels present a decomposition of search volume (Panel A) and the fraction of search volume (Panel B), respectively, in four age categories of the target by the income percentile of the target.

Figure A-4: Decomposition of the Income Gradient by Searcher's Age



Notes: The figure depicts a binned scatterplot of the income percentile of the target by the income percentile of the searcher, where searches (excluding self-search and within household search) are divided into four age categories of the searcher.