Social networks and tax avoidance: evidence from a well-defined Norwegian tax shelter

Annette Alstadsæter[1] · **Wojciech Kopczuk**[2] · **Kjetil Telle**[3]

Abstract

In 2005, over 8% of Norwegian shareholders transferred their shares to new (legal) tax shelters intended to defer taxation of capital gains and dividends that would otherwise be taxable in the aftermath of a reform implemented in 2006. Using detailed administrative data, we identify family networks and describe how take-up of tax avoidance progresses within a network. A feature of the reform was that the eligibility to set up a tax shelter changed discontinuously with individual shareholding of a firm and we use this fact to estimate the causal effect of availability of tax avoidance for a taxpayer on tax avoidance by others in the network. We find that eligibility in a social network increases the likelihood that others will take-up. This suggests that taxpayers affect each other's decisions about tax avoidance, highlighting the importance of accounting for social interactions in understanding enforcement and tax avoidance behavior, and providing a concrete example of optimization frictions in the context of behavioral responses to taxation.

✉ Wojciech Kopczuk
  wojciech.kopczuk@columbia.edu

  Annette Alstadsæter
  annette.alstadsater@nmbu.no

  Kjetil Telle
  kjetilelias.telle@fhi.no

1 Norwegian University of Life Sciences, Ås, Norway

2 Columbia University, New York, USA

3 Statistics Norway and Norwegian Institute of Public Health, Oslo, Norway

 Springer

1 Introduction

The standard public finance approach to analyzing tax-influenced economic decisions presumes a well-informed taxpayer who makes rational decisions while understanding the important features of the economic environment. This paradigm has long been considered non-satisfactory in the context of tax evasion, where the standard (Allingham and Sandmo 1972) model overpredicts the extent of cheating (see Andreoni et al. 1998; Slemrod and Yitzhaki 2002, for surveys of the literature). Recent empirical work also recognizes that behavioral responses are sometimes puzzlingly small and inconsistent across different contexts.[1] One potential direction for reconciling theory and evidence on non-compliance is to provide a more realistic characterization of the economic environment. The objective of the current paper is to provide empirical evidence regarding a particular class of explanations for tax-motivated behavior: whether tax avoidance spreads within social networks. Our results show that tax avoidance runs in the family.

Beyond attempts to improve characterization of the incentives faced by individuals, the recent development is to postulate an existence of optimization frictions that may stop individuals from pursuing otherwise optimal tax adjustments (e.g., Chetty et al. 2011; Chetty 2012; Kleven and Waseem 2013). This is a useful abstraction that potentially allows for explaining inconsistencies in observed empirical patterns, but it encompasses many possibilities: optimization frictions may be due to behavioral biases, lack of information, monetary or time adjustment costs or non-standard preferences. These varying possibilities might have very different policy implications so that discriminating between them is very important. Furthermore, there are two related, but distinct reasons to consider frictions. On the one hand, one may be interested in developing a better understanding of individual behavior. On the other hand, frictions are a potential source of *heterogeneity* in behavior in the population. Our findings of social interactions in the tax avoidance context provides evidence for both of these lines of thinking: for networks to matter, individual optimization has to depend on their characteristics; at the same time, by their very nature, networks are heterogeneous and hence generate differences in behavior of otherwise similar individuals.

We focus on a particular and natural choice of an exogenous social network: family members. There are many channels through which the tax minimizing behavior may spread, such as information about costs and benefits, awareness of tax avoidance strategies, and perception of social acceptance. We do not have data to separate the

---

[1] For example, Saez (2010), Chetty et al. (2011) and Kleven and Waseem (2013) show that elasticities implied by the number of taxpayers who are bunching at the kinks of income tax schedule are very small, Chetty et al. (2009) and Finkelstein (2009) show evidence consistent with "salience" of tax incentives playing a role, Jones (2012) shows that taxpayers do not adjust withholding to reduce refunds, and Chetty et al. (2014) show that only a small number of taxpayers makes active saving decisions. Related, a large literature shows the importance of default options in retirement programs, see Madrian and Shea (2001) and the literature that followed. Duflo and Saez (2003) provide evidence for the role of social interactions in retirement planning, Dahl et al. (2014) provide evidence for social interactions in the take-up of welfare benefits, and Currie (2006) provides evidence for imperfect take-up of social benefits.

relative importance of such channels, and our analysis will thus provide estimates of their combined influence on behavior.[2]

Empirical work on tax avoidance and evasion faces a lot of challenges, for example due to difficulty in observing the outcomes: participation in and extent of tax avoidance/evasion. We can sidestep this problem because of the existence of a well-defined tax shelter that is observable in our data; we provide more details below. Approximately 8% of Norwegian firm owners adopted this particular tax shelter during the second half of 2005. We can also observe the precise timing of adoption and hence analyze its dynamics. We utilize very detailed administrative data covering the universe of Norwegian firms, individuals and shareholders, and we are thus able to link firms with their individual owners. And, importantly, our data allows for constructing extended family networks that we can then use to identify spillover effects.

We begin by showing that decisions to pursue legal tax avoidance are correlated within extended family networks: Early adoption in family network predicts own subsequent take-up. Furthermore, the precise timing of adoption is linked: our evidence reveals significant increase in adoption within the week after take-up of a network member. This is robust to controls and provides suggestive evidence that take-up in the network stimulates own adoption.

While timing evidence is suggestive, it does not nail causality. In order to establish causal evidence that avoidance spreads within networks, we exploit the presence of a discontinuity in taxpayers' eligibility for setting up a tax shelter at 10% ownership in a regression discontinuity design. We show that this discontinuity affects own tax avoidance, and then we establish that it also affects tax avoidance of taxpayers in the family network (who are not necessarily themselves on the 10% margin). In other words, similar taxpayers who have similar family networks, pursue different decisions as the result of a slight difference in characteristics of one of their family members that discontinuously changes availability of avoidance for that family member (rather than the individual itself). We interpret this evidence as providing a concrete example of an optimization friction (driven by characteristics of the network) that is responsible for generating heterogeneity in taxpayer behavior with real tax consequences.[3]

There is a variety of social networks to choose from when studying social interactions, such as family, colleagues, schools, sports, church attendance, shareholders, accountants, board members, neighborhoods, etc. The current paper studies the impact of predetermined family networks for participation in legal tax avoidance. Bohne and Nimczik ([2018](#)) study the dynamics of legal tax avoidance within networks of firm and employees. They document the take-up of legal individual tax deductions for per-

---

[2] There is a related literature on the association between tax morale and illegal tax evasion, where tax morale denominates a group of intrinsic factors in the individual's compliance decision. See Luttmer and Singhal ([2014](#)) for an overview.

[3] Recent work on tax compliance emphasizes the importance of third-party reporting (Kleven et al. [2011](#)), attachment to the financial sector (Gordon and Li [2009](#)) and arms-length transactions (Kopczuk and Slemrod [2006](#)) as factors limiting the extent of evasion. While this strand of work suggests that administrative environment plays a very important role in tax compliance, it does not fully account for the empirical patterns that suggest that taxpayers who face seemingly similar circumstances often make different tax decisions. Indeed, in the strongest piece of evidence so far on the importance of third party reporting, Kleven et al. ([2011](#)) find that while accounting for third party reporting is extremely important for understanding patterns of compliance, only about 40% of taxpayers who are able to cheat do so.

sonal expenses spreads as workers and accountants switch firms. There is also a small, but growing, literature studying the role of various networks for agents' decisions to participate in illegal tax evasion, where the conclusions correspond to the conclusions in the current paper: tax evasion spreads within networks. Pomeranz (2015) highlights that networks matter in the context of VAT-evasion and finds evidence of spillover effects in a firm's trading network. Boning et al. (2018) find that tax compliance by firms is affected by IRS-interventions in the firm's network, through a shared tax preparer, geography, or a parent-subsidiary relationship. Paetzold and Winner (2016) use variation in job changes to identify the spillover effects from the work environment on the individual compliance decisions. They find that job changers moving to companies with a higher fraction of cheaters increase their cheating, while movers to companies with a lower fraction of cheaters tend not to alter their reporting behavior. To our knowledge, the only other (and concurrent to our work) paper that uses family relations to study the effect of norms and social interactions on the participation in tax minimization is Frimmel et al. (2018). They use Austrian data on claimed commuter tax deductions, where they can actually check the commuting distance and determine whether the deduction was rightfully or wrongfully claimed, the latter constituting tax evasion. By studying father-son pairs they find that tax evasion runs within the very close family. However, where Frimmel et al. (2018) study the intergenerational transmission of illegal tax evasion behavior, we study how legal tax avoidance behavior spreads within broad family networks.

The plan of the paper is as follows. In the next section, we describe Norwegian tax policy and the reform that gives rise to the research design in this paper and in Sect. 3 we describe our data. In Sect. 4 we show descriptive evidence on timing and find that the take-up in the network accelerates overall take-up. Section 5 is devoted to the empirical strategy. Our main results are in Sect. 6, where we present regression discontinuity-based evidence of the effect of the 10% rule that is the source of the discontinuity on individual take-up, followed by demonstrating the spillover effect in the network as well as timing effects. Conclusions are in the final section.

2 The 2006 reform and tax sheltering opportunities

Under the Norwegian dual income tax in effect as of 1992, capital gains realized by both individuals and corporations were subject to the basic tax rate of 28% (that applied also to corporate, capital and labor income). Dividends were tax exempt on both individual and corporate levels.[4] A shareholder income tax implying 28% tax on dividends for personal shareholders was announced in 2003 and introduced effective as of January 1, 2006. As one would expect, this led to massive avoidance responses. Dividends to personal shareholders were extraordinarily high in 2005, but plummeted

---

[4] This structure provided incentives for income shifting toward capital income tax base and to prevent it, the split model (1992–2005) imputed a return to the owners' labor effort in the firm, which was taxed as wage income. The split model applied to sole proprietors and corporations with 2/3 or more of shares held by active owners or where active owners were entitled to 2/3 or more of dividends. The split model and the incentives for income shifting are analyzed by Lindhe et al. (2004), Alstadsæter (2007) and Thoresen and Alstadsæter (2010).

post-2005.[5] During the transition period, the tax on realized capital gains on shares for corporate shareholders was removed without warning on March 26, 2004.[6] These changes unambiguously strengthened the incentive to own shares in a firm through another entity rather than directly.

Indirect ownership in general allows for separating two decisions: extracting resources from a firm and the ultimate transfer to the individual. Such a separation can have non-tax-related benefits to the owners such as shielding personal assets from third parties (creditors, family members) in a holding company, as well as tax-related benefits such as tax-free consumption within a (holding) firm without bearing the economic risk associated with the activity of the original firm (see Alstadsæter et al. 2014, for evidence of this type of tax planning), and arbitrage between personal and corporate taxation.

Founding a holding company implies some costs. A modest registration fee is to be paid at foundation, and at the time of the reform the annual accounts were to be approved by an auditor, which generated added costs. There are also time costs for the shareholder in keeping and submitting the accounts, or alternatively, out of pocket costs if she chooses to employ outside help by an accountant for this. However, in the case of a holding company with little or no additional activity, these auditor and accountant fees are modest due to the lack of complexity of the accounts. In addition, there was a minimum equity requirement of NOK 100,000. Another potential cost of establishing a holding company is advisor/lawyer costs if seeking advice in how to pursue legal tax avoidance strategies. These costs would be reduced or even not occur if someone in the shareholder's network already has established a holding company and can share these insights for free.

Following the reform, tax exempting capital gains and dividends for corporate owners creates some additional and very important advantages due to deferral of taxation. First, the deferral enables tax-free growth of assets within the holding company. Second, it enables pooling of losses and gains from various enterprises on the holding company level. The shareholder income tax does not allow the rate-of-return allowance (see footnote 5) to be transferred across different types of shares, and at realization,

---

[5] The shareholder income tax was first proposed by an advisory committee on February 6, 2003. A revised version was presented by the government on March 26, 2004, and sanctioned by the Parliament on June 11, 2004, to be introduced on January 1, 2006. The shareholder income tax ensures "equal" tax treatment of all personal owners of corporations, independent of ownership composition. It levies a tax of 28% on all personal shareholders' income from shares, including both dividends and capital gains. Under the shareholder income tax, the risk-free return to the share, the so-called rate-of-return allowance (RRA), is tax exempt. If received dividends are less than the RRA, the remaining amount is added to the imputation basis of the share for the calculation of future RRAs. The unused RRA is carried forward and added to the imputed RRA in the following year. The share-specific RRA cannot be transferred between different types of shares and only owner at the end of the year benefits from the calculated RRA for that year. Dividends paid to corporations were tax exempt at the introduction of the model, as were corporations' capital gains from realization of shares. See Sørensen (2005), Alstadsæter and Fjærli (2009), Alstadsæter et al. (2014) and Alstadsæter et al. (2016) for more information on the shareholder income tax and responses to it.

[6] Anecdotal evidence that this was not expected by the business community is the fact that one of the nation's richer investors on March 25, 2004 sold shares in a corporation that he owned indirectly through his investment company. Christian Sveeas' investment company Kistefoss sold its 6.5% stake in the online price comparing service Kelkoo to Yahoo on March 25, 2004. This resulted in a taxable capital gain of 235 million NOK, and capital gains taxes of 63 million NOK or appr. 10 million USD. Had this sales contract been signed one day later, the capital gain would be tax exempt.

unused rate-of-return allowances are lost at shareholder level. At company level, there is no dividend tax and thus no unused allowance to be lost at realization. This will then increase the total allowance of the owner of the holding company, as the individual shareholder's allowance is based on her share of the external equity in the holding company, which is unaffected by this transaction. Third, the investor may make a policy bet on the dividend tax to be removed in the future. Alstadsæter et al. (2014) and Alstadsæter et al. (2016) show that the reform led to very large changes in tax reporting behavior of business owners. From our perspective, the key point is that individuals have a stronger incentive to own firms indirectly after the reform; the *prima facie* evidence of it being so is the massive number of conversions that took place.

For the existing firms, switching from direct (individual) to indirect (holding company) ownership should in principle require transferring/sale of existing shares and would trigger capital gains tax liability. In order to level the playing field between individual and corporate investors, the so-called Transition Rule E was introduced, which under certain conditions enabled an individual to transfer his/her shares in an existing firm to a holding company during 2005 without triggering capital gains tax.

The Transition Rule E was first proposed on November 19, 2004, and sanctioned on December 10, 2004. It removed capital gains tax liability when an individual shareholder transfers all her shares in a firm to a newly founded corporation*,* given that this new holding company in the end holds at least 90% of the shares in the transferred company and the compensation is in the form of shares in the new corporation. The new holding corporation had to be founded and a report sent to the company register by December 31, 2005. It turned out that this transition rule was restrictive and relatively few shareholders could utilize it. A more liberal version of the Transition Rule E was proposed on May 13, 2005, and later sanctioned on June 17. Under this new version, the 90% threshold was reduced to 10%. We will refer to a holding corporation that was founded during 2005 in response to the Transition Rule E as a tax shelter or an E-firm.

An individual who already owns 10% of shares in a firm was in a position to establish an E-firm alone with no additional adjustments, while an individual who owned just below 10% would have to either buy or coordinate with others, inducing increased coordination costs, both in time and loss of control over future payout policy from the tax shelter.

To summarize: prior to 2004, the incentive to own corporations directly was fairly strong because corporate capital gains were subject to taxation (thereby resulting in multiple layers of taxation in case of corporate ownership before reaching personal owners), while dividends were tax exempt in any case. As of March 2004, neither corporate capital gains nor dividends were subject to the tax. As a result, indirect ownership of a firm allowed for deferral of taxation of capital gains until the holding company is sold. The incentive for indirect ownership was significantly strengthened by the introduction of individual-level dividend taxation as of January 1, 2006. For the existing ownership stakes, taking advantage of these deferral opportunities should in principle require realizing capital gains and triggering tax liability, but the Transition Rule E provided an opportunity to convert to indirect ownership without the tax. The main purpose of holding companies set up under Transition Rule E appears to be to work as a tax shelter intended to defer taxation, and alternatives to achieve the same

outcome would be costly. During 2005, 16,483 holding corporations were set up and approximately 9% of existing non-listed firms at the end of 2004 had at least some of the owners electing to transfer their stake to a holding company. Figure 1 shows the timing of adoption of firms that we classify as being set up under the transition rule. Adoption was slow at the very beginning and increased rapidly toward the end of 2005, just before the opportunity to take advantage of it expired.

3 Dataset description

We use very detailed administrative data covering the *universe* of Norwegian firms, individuals and shareholders. Every resident in Norway is provided a unique personal identifier that is present in all databases, enabling us to follow every individual over time and across datasets. The same holds for firms. The shareholder register contains records of every shareholder (firms and individuals) of every Norwegian corporation for 2004–2008 at year end. For our sample, we include all individual share holders in 2004 who resided in Norway, owned shares of a Norwegian non-listed corporation with less than 100 individual owners and are not sole proprietors.[7]

We can also identify holding companies that were set up during 2005 through the sector code assigned to them by Statistics Norway, determine their ownership structure and holdings.[8] Because we observe this information for a number of subsequent years, we can also trace changes in the ownership structure such as transfers of an existing firm to a holding company. Importantly for our analysis, we know the exact date when each firm (holding companies included) was registered.

The shareholder register was established in 2004, and we do not have information on firms' dividend distributions, or on individuals' ownership shares, prior to 2004. In order to avoid selection into eligibility for adoption of E-firm during the second half of 2005 by increasing ownership in a firm beyond 10%, we define eligibility pre-reform as of December 31, 2004. This means that we might have individuals in our treatment group that are not eligible to adopt E-firm during the last half of 2005, while others that we have in our control group may become eligible during 2005, due

---

[7] We can also follow indirect ownership—via other firms—but we opted to not use it for our 2004 running variable (ownership share) because each individual owner of such a pre-existing holding company may not have full control over shares and thus may not be in a position to take advantage of the E-firm rule. Correspondingly, allocating shares owned by a firm to its owners is likely to be somewhat arbitrary and introduce noise in the running variable. Having individuals below 10% mark also owning shares through corporate channels is one possible explanation for significant take-up for individuals who were not classified as eligible in 2004.

[8] Statistics Norway identified new holding corporations set up under the Transition Rule E by an existing sector code that was rarely used: NACE-code 65.238 "Portfolio Investments". A shareholder is defined to set up a tax shelter (Transition Rule E) if in 2005 she is an owner of a corporation with NACE-code 65.238 that was founded during 2005, and that in 2005 owns shares in a non-listed corporation in which the physical shareholder held at least 1% shares in 2004. At the beginning of 2005, there were a total of 1886 existing corporations with NACE-code 65.238, and during 2005, 16,483 new firms with this code were set up. 8.2% of all our sample of shareholders in 2004 set up a tax shelter (E-holding) in 2005. As the NACE-code 65.238 is an existing code, some of these new firms might be founded for other reasons than tax sheltering (i.e., utilization of Transition Rule E). Due to the low number of firms in this group at the beginning of 2005 error should be small.
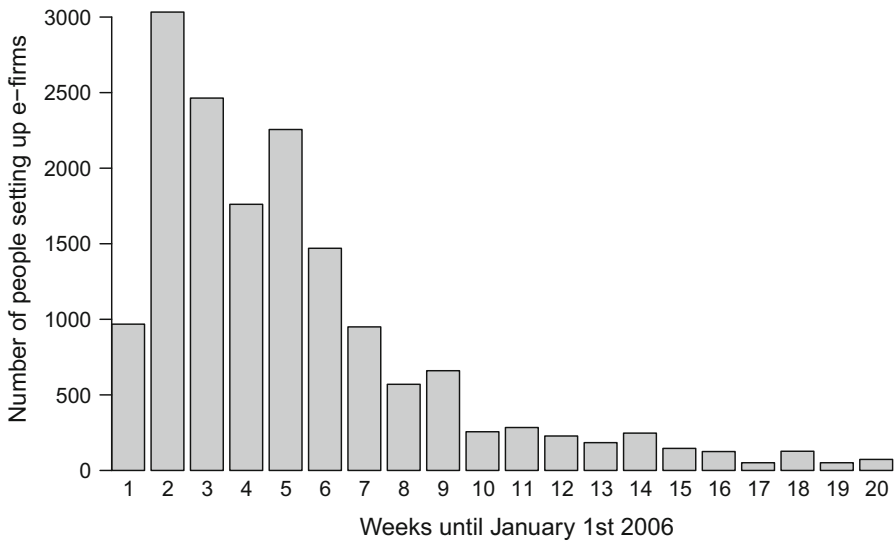
Fig. 1 Timing of setting up E-firms

to change in ownership. This might introduce some attenuation bias, making our job of identifying an effect more difficult. An additional potential complication might be that the accuracy of the shareholder register might be lower in the first year of 2004, due to start-up problems in reporting.

Using other register information we are able to link characteristics, both demographic (gender, age, marital status, immigrant status, education) and economic (including tax-related information such as gross and taxable income, dividend income, capital gains realizations).

To estimate the effect of a tax shelter being set up in shareholder *i*'s network on the likelihood that the shareholder himself adopts a tax shelter, we need to make operational a definition of the network. In this paper, we focus on a particular and natural choice of an exogenous network: family members. To do so, we identify the following family members of each shareholder in our 2004 sample: her direct (parents, children, siblings, spouse) relatives and direct relatives of the direct relatives.[9]

For the descriptive evidence of the timing of adoption of E-firm in the next section we will use the whole sample. Table 1 shows summary statistics for this sample, and we notice that three quarters of the shareholders are male, average age is 46, a vast majority are married and live in urban areas. Also, 8% of the shareholders established an E-firm, and 12% have a family member with an E-firm.

---

[9] More specifically, an individuals' family network consists of parents; grandparents; children (born in 2004 or earlier); children's spouse (married 2004 or cohabitant with common children); grandchildren; spouse (married as of 2004 or cohabitant with common children 2004); spouse's siblings; spouse's children (born in 2004 or earlier); spouse's parents; siblings; siblings' spouses (married as of 2004 or cohabitant with common children); siblings' children (born in 2004 or earlier); siblings' parents; aunts/uncles; aunts/uncles' spouses; cousins. We will usually exclude the spouse from the network because the relevant unit of observation may be a household rather than an individual. The results are robust to this restriction.

Table 1 Summary statistics

	#Obs	Mean	SD
Male	199581	0.747	0.435
Age	199581	46.473	13.285
Age < 30	199581	0.077	0.267
Age 30–40	199581	0.241	0.428
Age 40–50	199581	0.269	0.443
Age 50–70	199581	0.357	0.479
Age > 70+	199581	0.047	0.211
Married/cohabiting with children	199581	0.729	0.444
Number of children	199581	1.871	1.257
Immigrant	199581	0.037	0.189
Urban	199581	0.780	0.414
University	199581	0.341	0.474
High school	199581	0.550	0.498
Business education	199581	0.086	0.281
Self-employment	199581	0.234	0.424
Net worth, 2004 (NOK 1000s)	199581	1657.3	18800.0
Capital income, 2004 (NOK 1000s)	199581	366.1	6175.7
Total income, 2004 (NOK 1000s)	199581	757.2	6207.9
Total dividends in 2004 (NOK 1000s)	199581	297.2	6188.6
E-firm dummy	199581	0.082	0.274
E-firm before June 17	199581	0.001	0.026
E-firm before November 1	199581	0.010	0.098
E-firm before December 1	199581	0.032	0.175
E-firm in family network	199581	0.120	0.325
E-firm in family before June 17	199581	0.001	0.037
E-firm in family before November 1	199581	0.016	0.127
E-firm in family before December 1	199581	0.050	0.218

For the regression discontinuity analyses, we focus on observations around the 10% threshold; see Sects. 5.1 and 6 for details. A complication is that inspection of the ownership data reveals clustering of individuals at ownership shares that correspond to splitting shares of the firm as exact fractions. This is a potential threat to the continuity of characteristics of the underlying population and, hence, a possible threat to a practical implementation of regression discontinuity approach that requires that the outcome is smooth in the neighborhood of the threshold. It is indeed possible that observations that are bunched at these selected points are not similar to the neighboring ones—splitting shares equally is likely to be (and is in practice) correlated with many characteristics of individuals and firms. Therefore, for the regression discontinuity analysis in Sect. 6, we exclude exact fractional observations from the sample of analysis, as described in more detail in Sect. 6.1 and Appendix B. For the net-
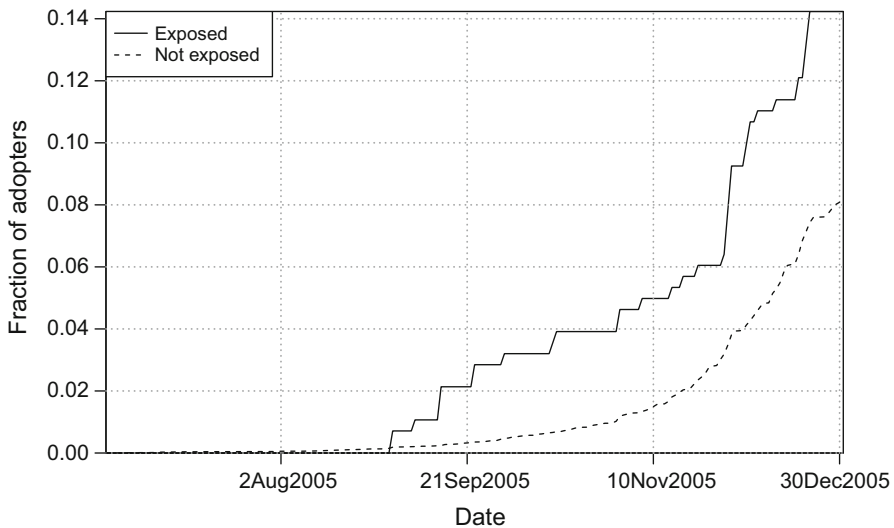
Fig. 2 Adoption by individuals with and without a family member adopting before June 17th 2005

work analyses, we also need to restrict the sample further to operationalize the family network variable, and to ensure that the assumptions of the regression discontinuity design are not violated by family members owning identical number of shares in the same firm; see Sects. 5.2 and 6.3 for details.

4 Timing of adoption

We start by illustrating the dynamics of setting up tax shelters in the data. Remember that prior to June 17, 2005, one needed 90% ownership to be eligible for an E-firm. From June 17 and onward, this eligibility criteria was reduced to 10%. There are few adoptions during the early period, and as also visible in Fig. 1, the timing of adoption is heavily concentrated toward the end of the period. This raises the possibility that these early adoptions may have influenced family members to also adopt a tax shelter.

Figure 2 shows the adoption of the tax shelter by individuals with ("exposed") and without ("not exposed") a family member setting up a tax shelter prior to June 17. Exposed individuals end up approximately 6 percentage points more likely to eventually set up a tax shelter. Furthermore, Fig. 3 shows that even conditional on ultimate adoption there are differences in timing—those who have exposed family members adopt earlier than others. These patterns do suggest that there is correlation between adoption by network members in the past and the individual's own adoption of the tax shelter. They also suggest that there may be an effect on timing: individuals in networks with early adopters are not just more likely to adopt in general, they also tend to adopt earlier than others.

We investigate these patterns more formally in a simple regression framework. Table 2 shows the results of regressing the E-firm adoption dummy on the indicator
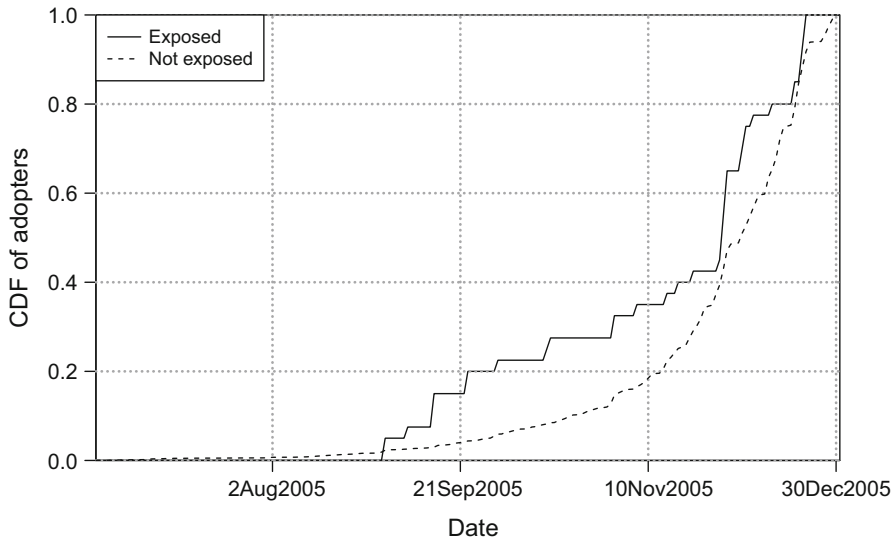
Fig. 3 Timing of adoption by individuals with and without a family member adopting before June 17th 2005 (conditional on ultimately adopting)

for having somebody in the family network adopting by a particular date, with various sets of controls. Only the coefficient on the network dummy is reported and each cell corresponds to a different regression. The first panel shows the results of regressing the dummy for ever setting up an E-firm on the dummies for having somebody in the network setting up by June 17, November 1, and December 1. Consistently with the descriptive graphs that we have just discussed, the results of baseline regressions with no controls show a strong effect in each case. In the second column, we control for a number of demographic characteristics: gender, immigrant dummy, urban dummy, self-employment status, education dummies, business/law education dummy, number of children and age dummies (decades). Including these controls does not have a strong effect on the estimated coefficient although many of them are individually very significant (not reported). The final column shows the effect of including economic controls: logarithms of total income, net worth, capital income and 2004 dividends. Inclusion of these variables reduces the estimated network coefficients but they do retain statistical significance. This indicates that early take-up in one's network correlates with individual economic characteristics that are relevant to take-up decisions, but that it works beyond them.

The effect of somebody else adopting may not be just on ultimate adoption but also on timing of adoption. To rudimentarily pursue it further, we note in the following two panels that adoption before December 1st is more robustly explained by family network adoption before November 1st and adoption before November 1st appears correlated with family network adoption pre-June 17. Especially in the latter case, the effect of economic controls on the estimated coefficient is weakened. This is consistent with the coefficient on early adoption picking up the effect of inducement over the short horizons, but at least partially reflecting the effect of correlation of early adoption in

Table 2 The effect of early adoption on take-up

Adoption	No controls	Demographic controls	Demographic and income controls	N
Outcome: Ultimate adoption				
Before June 17th	0.076***	0.079***	0.055***	199443
	(0.018)	(0.018)	(0.017)	
Before November 1st	0.059***	0.059***	0.038***	197642
	(0.005)	(0.005)	(0.005)	
Before December 1st	0.054***	0.054***	0.042***	193267
	(0.003)	(0.003)	(0.003)	
Outcome: Adoption before December 1st				
Before June 17th	0.037***	0.038***	0.028**	199443
	(0.011)	(0.011)	(0.011)	
Before November 1st	0.013***	0.013***	0.006**	197642
	(0.003)	(0.003)	(0.003)	
Outcome: Adoption before November 1st				
Before June 17th	0.021***	0.021***	0.018***	199443
	(0.006)	(0.006)	(0.006)	

Demographic controls include gender, immigrant dummy, urban dummy, self-employment status, education dummies, business/law education dummy, number of children and age dummies (decades). Economic controls include log of income, log of capital income, log of net worth and log of dividends, all as of 2004
**significance at 5% level; ***significance at 1% level

networks with economic characteristics that ultimately matter over a longer horizon. At the same time, it is interesting to note that demographic characteristics (while individually significant) do not seem to be correlated with early adoption.

Overall, these results suggest that while adoption of an E-firm is also correlated with many demographic characteristics, it does not seem that correlation of early adoption in the family networks is related to these factors. At the same time, it appears that the link between adoption and the network is less sensitive to the inclusion of controls as the horizon is reduced. This is intuitive: the impact of having someone in the network adopting should be on timing first of all, and while the effect may persist in the longer term, it is possible that it is hard to distinguish from the effect of other characteristics correlated with early adoption. This motivates a strategy that treats timing more carefully.

It's possible that taxpayers in the network are exposed to the same shocks (for example news) at the same time. But it is harder to make the case that individuals would happen to make similar decisions at similar time based purely on correlation in characteristics that are constant over time absent common shocks or interactions. We thus regress the dummy for taking up an E-firm in a particular week on having somebody in the network take-up a week before.

Figure 4 shows the results for family network based on simple OLS regressions. This is again a linear probability model and a hazard-like context. Week 1 corresponds to the last week before January 1, 2006 (and the right-hand side variable is adoption in
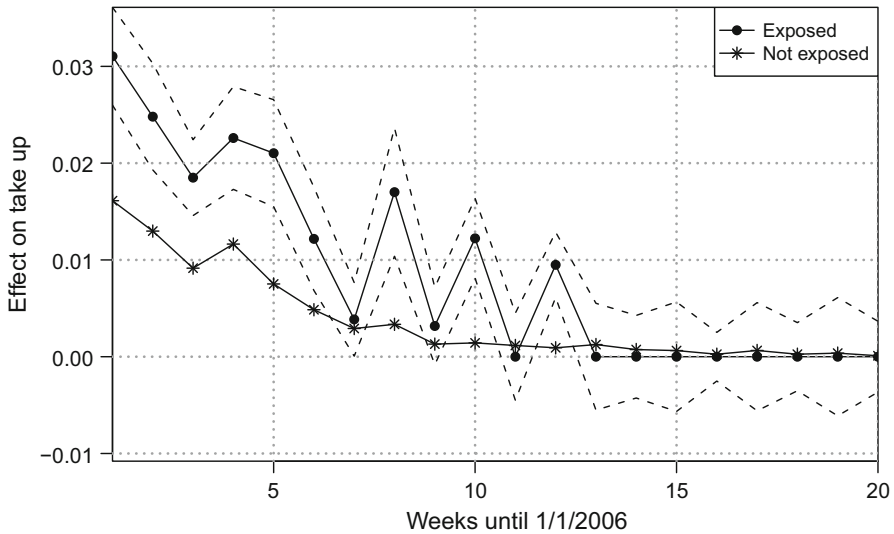
Fig. 4 The effect of a family member adopting a week earlier (OLS with no controls)

the family a week before that) and higher numbers correspond to earlier adoption. The figure shows the baseline effect (the constant from the OLS) that represents adoption of the tax shelter by individuals with no exposure in the family network in the preceding week, and the effect of those who were exposed last week (the sum of the constant and the coefficient on the exposure dummy), together with the 95% confidence interval for the latter. There is a significant effect for the last six weeks of the year and some weeks before that. At the longer horizon, the effect is gone. It is possible that a week is in the right ballpark of the timing of inducement effect late in the game, but is too short of a period earlier when there is no reason to rush.[10]

Overall, these results are suggestive of the relationship in timing and eventual take-up, but still may not be causal. We now turn to a regression discontinuity framework to investigate the network effects further using a research design that more readily lends itself to a causal interpretation. We will revisit timing effects again in Sect. 6.4.

5 Basic framework

Our core econometric framework can be described in two closely related equations—one for the individual i herself and one for the network member j that may affect ("treat") her

$$E_j = \alpha_j + \beta_j \cdot X_j + \gamma_j Z_j + \mu_j \tag{1}$$

$$E_i = \alpha_i + \beta_i \cdot X_j + \gamma_i Z_i + \mu_i \tag{2}$$

---

[10] Adding demographic and economic controls (as before) yields very similar results as those in Fig. 4. Using adoption in the last 3 weeks (rather than 1 week) yields smoother, but a bit smaller, estimates suggesting that recent take-up has stronger effect than take-up further out.

Equation 1 relates sheltering decision E to one's own incentives represented by X_j, and controlling for own characteristics Z_j. Equation 2 relates sheltering decision of an individual to her own characteristics Z_i and some characteristics X_j of the network member. We will refer to the individual i as "treated" individual and to individual j as "treating" individual.

In most cases we will use a dummy variable for setting up a tax shelter (i.e., E-firm) as the dependent variable and estimate specifications as linear probability models. Given that we will primarily focus on local effects in small (bounded) neighborhoods of the discontinuity point, this is not particularly restrictive. We will also occasionally investigate the timing of decisions by replacing E with adoption of the shelter in some period τ, E^τ, or using the timing of adoption t directly. Some of these specifications will be estimated using tobit and probit methods to address censoring (not everybody adopts before the deadline) or accommodate periods with very low adoption rates.

The establishment of an E-firm in the family network may directly affect individual i's likelihood of setting up an E-firm, but there are likely to be other channels at work too. Indeed, the fact that a network member is eligible to set up an E-firm may result in the collection of information about costs and benefits, and this information may affect individual i regardless of whether j ends up establishing an E-firm or not. The mere awareness of the tax shelter option in the network, or even perceptions of social acceptance in the network, can also affect the behavior of individual i. Thus, in Appendix A, we provide a simple theoretical framework for interpreting β_i and β_j. The non-zero value of β_i implies that the social interactions are present. Its value provides an indication of the magnitude of the effect that is not "structural"—it measures the responsiveness to the particular shock. Remark 3 in Appendix A (under assumptions leading up to it), provides a way to guide the interpretation of the ratio $\frac{\beta_i}{\beta_j}$. β_i being large relative to β_j indicates that either the interactions are very strong or that the awareness of sheltering opportunities of the family members that are influenced by the recipients of the shock is relatively low.

The most restrictive feature of our estimation equation may seem to be due to the fact that we include X_j for only a single other individual—we will discuss the interpretation below. As mentioned before, we implement a regression discontinuity design that relies on the feature of the reform that required a newly setup holding company to own at least 10% of shares of a firm. Hence, an individual who already owns 10% of shares in a firm was in a position to pursue this path alone with no additional adjustments, while an individual who owned just below 10% would have to either buy or coordinate with others. Consequently, we define $X_j = 1(S_j \geq 0.1)$ where S_j is individual shareholding in a firm.[11] Crucially, we have information about the exact number of shares that an individual owns in 2004 as well as the total number of

---

[11] Note that two individuals who own at least 10% combined, say 5% each, would need to coordinate and set up one common E-firm. As noted in Sect. 2 such an E-firm may be less attractive as it would require future cooperation in, e.g., payout decisions, but as is indicated by Figs. 7 and 8 and noted in Sect. 6 the setting up of such E-firms seems to have occurred in some cases. However, and unlike the discontinuity at 10%, two individuals cooperating could occur at ownership shares 5% and 5%, but also at 6% and 4%, or any other combination summing to at least 10%, implying that individuals just above and below 5% could easily be partly treated. In an attempt to restrict the complexity of the econometric approach, we have not tried to take advantage of such coordinated setups in this paper.

shares in a firm, so that we can (1) construct S_j exactly and (2) do so using information that precedes the reform and hence does not reflect the effect of the reform itself.

Our basic comparison is that of individuals just below and just above the 10% threshold; corresponding to Eq. 1. While, as we stated in Remark 2 of Appendix A, the response of the individual to this incentive is not a necessary condition for the presence of network effects, a combination of the lack of such evidence with the presence of network effects would certainly be surprising. Equation 1 is important for a number of other reasons. First, we will investigate subsamples with different propensities to set up an E-firm and expect that those where the direct effect is strongest, are also likely to exhibit stronger network effects. Second, our attempts to provide a structural interpretation of the estimates rely on comparison of direct and indirect effects.

Equation 2 specifies how the decision of a "treated" individual (i) is related to incentives (X_j) of her "treating" network member (j). Hence, the comparison is between individuals who happen to have in their networks somebody with just over 10% shares in a firm versus those that have in their networks somebody with just under 10% shares.[12]

There are many characteristics of individuals and the network that may matter as well in general. The regression discontinuity design allows to abstract from them as long as they do not change discretely at the 10% threshold. We will investigate this assumption for particular variables and will test sensitivity of results to including controls. Given that the assumptions for validity of the regression discontinuity design hold, controlling for such additional characteristics is not necessary for obtaining unbiased estimates of the effect of X_j on E_i.

We will investigate heterogeneity of the response by splitting the sample along some dimensions (such as history of dividends) and/or including interaction effects.

We also note that since any operationally available definition of a network is intrinsically arbitrary, our measure of the presence of a tax shelter within the network will not be fully correct if we do not properly classify individuals as members of a network. Thus, estimates of β may suffer from the attenuation bias if what one is interested in is the effect of *any* interactions. As long as assumptions for the validity of the regression discontinuity design hold, the estimates reflect though the average effect of exposure to eligibility for sheltering in *a family network*. While a concern in general, the downward bias due to mis-classification makes our task harder, but should not lead to spurious findings.

---

[12] Thus, we are not attempting to estimate the effect of a network member j setting up an E-firm on the likelihood that individual i herself sets up an E-firm, e.g., by treating X_j as an instrument for E_j in a first stage. Such an instrumental variable approach would require the additional assumption that there is no (conditional) effect of X_j on E_i, except for the one going through E_j. As noted above, however, the behavior in the family network may affect individual i's likelihood of setting up an E-firm in other ways than through network member j's actual establishment of an E-firm. Indeed, the fact that network member j is eligible to set up an E-firm may result in her collecting information about costs and benefits, and this information may benefit individual i regardless of whether j ends up establishing an E-firm or not. Moreover, and as we show in Appendix A, without adding more structure to our econometric model, it is not possible to separately identify the impact on the behavior of individual i of network member j setting up an E-firm, of network member j becoming more aware of costs and benefits thereof, or of interaction effects.

5.1 Unit of observation

Our running variable is defined on the level of shareholding. A shareholding in a particular firm k of a particular individual j may or may not be eligible for establishing an E-firm depending on whether it corresponds to less or at least 10% share. Any individual may have multiple shareholdings in multiple firms that may fall on either side of the threshold.

We want to avoid assumptions necessary to aggregate such information to the individual level. This is because aggregation disposes of potentially useful information and comes with practical concerns. For example, the largest share owned by a taxpayer in any firm is also a continuous variable to which the 10% discontinuity applies, but it ignores all smaller shareholdings that also correspond to discontinuous incentives (i.e., some taxpayers who own around 10% of a firm also turn out to own a higher share of some other firm). Various forms of averaging are incompatible with regression discontinuity design because they blur the running variable, so that there is no longer discontinuity in incentives of such a measure (e.g., at 10% of *average* shareholding).

Hence, instead, we usually represent our data on the shareholding level. That is, we are treating each (j, k) as a separate observation and use statistical correction (clustering) to correct for the dependence due to potential inclusion of multiple observations for the same person. As the size of the window around the threshold declines, the likelihood that more than one observation per individual is used declines and the distinction between individuals and shareholdings becomes irrelevant in the limit (and is of small consequence for standard errors in practice).

There is a corresponding issue that relates to the definition of the outcome variable. Setting up an E-firm can be defined on a shareholding level: an individual transfers shares of a particular firm to an E-firm and may choose to do so for some firms but not for others. We will show some evidence of the effect on the shareholding level, but will primarily focus on the outcomes defined on the *shareholder* (i.e., individual) level. That is, our outcome variable E_j represents whether an individual adopted any E-firm for any of her shareholdings. Hence, the unit of observation is (j, k), the corresponding running variable is $S_{j,k}$ but the outcome is E_j—constant for all k.

In the network context, we want to retain the same structure on the treatment level. The discontinuity is defined on the level of the shareholding of the treating network member, (j, k). The corresponding treatment affects all individuals i who are related to the network member j. Because networks overlap, there is no straightforward way of collapsing information to the whole network level. Instead, we treat each link $(i, (j, k))$ as a separate observation. As a result, a single shareholding k of person j gives rise to multiple observations for all individuals who are in the same network as j. We address the corresponding dependence by clustering standard errors at j level. There is also the possibility that person i gives rise to multiple observations corresponding to links with shareholdings of all her network members, but in practice this is of little concern because it is rare that the same person has multiple network members with shares falling into the same small interval around the threshold. Finally, as before, we define the outcome variable as setting up any E-firm so that it is the same for all observations corresponding to individual i.

5.2 Interpretation of the estimated coefficients

As we discussed, the unit of observation for our analysis is the (directed) network relationship and our baseline specifications Eqs. 1 and 2 include X_j only, rather than characteristics of all individuals in the family network (X_{-j}); see also Appendix A. In general, individuals may be influenced by many different network members

$$E_i = \alpha_i + g(X_j, X_{-j}) + \gamma_i Z_i + \mu_i \tag{3}$$

Suppose that $X_j \perp X_{-j}$ (i.e., that in our regression discontinuity context, the likelihood of being below/above the 10% threshold is uncorrelated in the network) and, counterfactually, that for each i we observe just one randomly selected individual j. In that case, our specification would estimate $\beta_i = \mathrm{E}\left[\frac{\partial g}{\partial X_j} \middle| Z_i\right]$—the local average treatment effect of exposing an additional network member to tax sheltering opportunities, with equal weights assigned to all individuals. In our application though, we include an observation for each network relationship (i, j) so that, instead, we weigh equally relationships rather than individuals.

This strategy makes it straightforward to pursue estimation using relationship data and, as long as the assumption $X_j \perp X_{-j}$ holds, it remains an unbiased estimator of treating an additional relationship (not an individual) in the network.

6 Regression discontinuity evidence

Our main identification strategy exploits differences in eligibility for setting up an E-firm. As discussed before, the newly created E-firm has to hold at least 10% of shares of the original firms. Hence, taxpayers who own at least that much can set up an E-firm without further complications while taxpayers who own less than 10% of shares have to either buy more or set up an E-firm in cooperation with others. However, an examination of the dataset shows bunching in many places, threatening the identification through regression discontinuity approach if not addressed.

6.1 Smoothness of the distribution

A closer inspection reveals that bunching is very systematic—it occurs at points that correspond to splitting shares of the firm as exact fractions.[13] Thus, for one, non-randomly distributed observations at bunching points differ from others because they correspond to firms that choose to split ownership in such a regular way and it is possible that observations that are bunched at these selected points are not similar to the neighboring ones—splitting shares equally is likely to be correlated with many characteristics of individuals and firms.[14]

---

[13] Online Appendix Fig. A1 shows the log of the number of observations in the full sample, by 0.1% bins.

[14] Beyond the number of observations, we found indication of non-smoothness for individual characteristics—this is expected, because these "fractional" observations are observationally different than others. For example, looking at individuals with ownership shares in the interval (0.05, 0.15), firms with

Hence, we proceed by eliminating exact fractions from the sample as explained in Appendix B. The outcome of this trimming procedure in terms of the number of observations is shown in Fig. 5.[15] While eliminating exact fractions removes a lot of bunching, we see from the figure that the density is still not completely smooth around the 0.25 share—our rules for eliminating fractions do not seem sufficient for dealing with that bunching. Tax rules pre-reform also provided an incentive to have *active* ownership below 2/3 in order not to be subject to the so-called split model that taxed part of profits at labor income tax rates—as a consequence, there are many examples of firms that assigned just over 1/3 stake to passive owners, in particular often dividing it further in half (e.g., among two children) and hence resulting in shareholdings of just over 1/6th—some of the irregularities are likely associated with that. Similarly, predetermined characteristics (measured in 2004) are also quite noisy, but this is so mostly away from the threshold and especially for shares above 1/6th (see Online Appendix Figs. A5, A6, A7 and A8). We draw two conclusions. First, the data around the 10% threshold appears reasonably smooth and we will limit the window around the threshold to at most of 0.05 on each side, where the case for smoothness of the distribution is strongest. Second, we will test robustness of the results by controlling for demographic characteristics. We thus proceed with this subsample in what follows.

In Sect. 6.3 we will show that the density and predetermined variables are similarly smooth around the threshold in the network analysis.

6.2 The effect of 10% rule on individual adoption

Figure 6 shows individual ownership share in 2004 (i.e., half a year before the 10% eligibility criterion was introduced) and the fraction of individuals setting up E-firms by 1% bins (starting at round percentage values, inclusive, e.g., [0.10, 0.11)). The unit of observation for this figure is a shareholding—an individual who owns shares in multiple firms corresponds to multiple shareholdings and hence multiple observations.[16] The adoption of the E-firm is defined at the individual level. Hence, the figure suggests that individuals who happen to have a shareholding that inches just above the 10% mark are more likely to set up an E-firm (overall, not just or solely for this particular shareholding). The figure illustrates a number of points that will be important

---

Footnote 14 continued

fractional shares are more female (0.69 vs. 0.75), younger (44.3 vs. 46.60) and have fewer owners (6.33 vs. 12.97). These differences in characteristics combined with discreteness of the distribution of "fractional" observations, generates non-smoothness of the overall distribution of characteristics in the full sample.

[15] While this procedure is necessary to apply the regression discontinuity approach, it introduces a natural limitation to the interpretation of our results: we are focusing on a subsample, so that the estimated effects are for the corresponding population only. We want to re-emphasize though that the procedure relies on a systematic selection rule based on pre-existing variable, so that it does not depend on the effect of any reform. The procedure does of course change the composition of the sample—that is precisely its objective—but we expect that the resulting subsample satisfies the necessary conditions for the regression discontinuity design.

[16] For completeness, Online Appendix Fig. A4 shows the analogous relationship using as outcome adoption of E-firm defined on shareholding (rather than individual) level. In the Appendix we also show the analogue of Fig. 6 in Figs. A2 and A3 using the full sample (i.e., before eliminating exact fractions). Figure A2 shows the likelihood of adopting E-firm for the particular shareholding, while Figure A2 shows the likelihood of adopting for any shareholding of the corresponding individual.
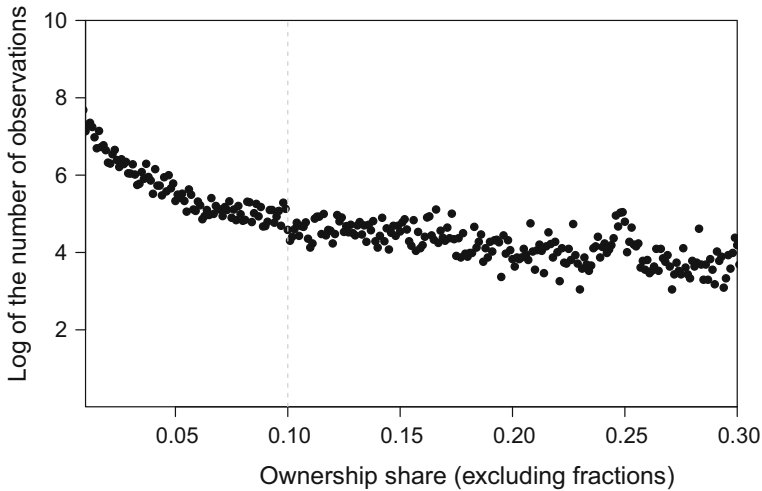
Fig. 5 Distribution of ownership shares in 2004 around the 10% threshold

below. First, there is an appearance of discontinuity at the 10% threshold but there is also enough variation in the data overall that careful testing is necessary to establish its presence.[17] Second, it is a "fuzzy" regression discontinuity design—E-firms are created by some individuals below the threshold (by coordinating with others, through additional purchases of shares during 2005 or because of imprecision in the running variable if there is corporate ownership) and take-up is far from universal above the threshold. Imperfect assignment implies that the estimated effects are very likely to be heterogeneous across different groups, since the take-up would depend on incentives, and we will investigate such heterogeneity. Third, the pattern of adoption is nonlinear over the whole support but reasonably linear in the neighborhood of 0.1; adoption increases significantly with shareholding until it reaches a plateau at around 0.2, above which around 20% of the population adopts (and the data is considerably noisier). Consequently, we will restrict analysis to a reasonably narrow neighborhood of the discontinuity point—in most cases, subsets of interval (0.05, 0.15)—where nonlinearity is not an important issue.

Figure 7 zooms in to the smaller region (0, 0.30), that more clearly displays the 10% threshold (with bins corresponding to 0.01 intervals). It also shows point-wise standard errors of the mean within a bin. The likelihood of taking up an E-firm jumps discontinuously at the 10% point. This is formally investigated in the top panel of Table 3. The baseline regression is a linear probability model of the dummy for taking up an E-firm on an indicator for being at or above the 10% mark in 2004 within a narrow band around the 10% point. The "flexible" controls specification additionally allows for linear (and possibly different) terms on the left- and right-hand side of the threshold. We show the effect in adoption on individual level and (of our main interest)

---

[17] The data also suggests a discontinuity at 5%, which might result from two individuals with 5% each adopting an E-firm together. As noted in footnote 11, and elaborated on in footnote 20, we are not using this variation in the analysis.
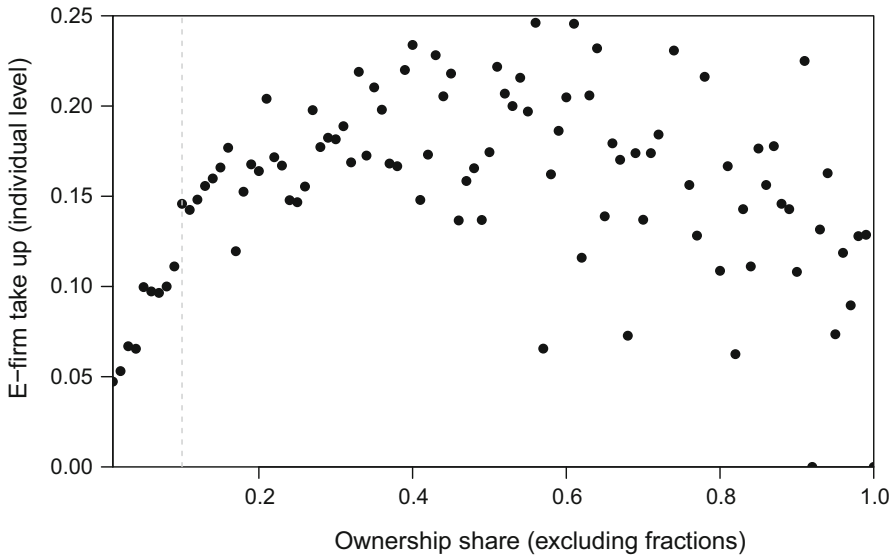
Fig. 6 Ownership share in 2004 and probability of ultimate adoption of the E-firm
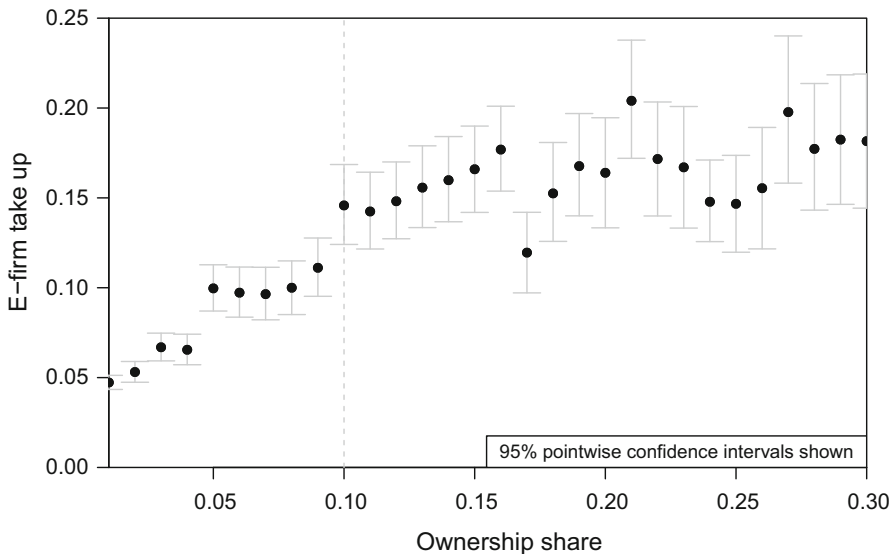
Fig. 7 Ownership share in 2004 and probability of ultimate adoption of the E-firm—individual level, excluding fractions

the effect on shareholder level. The results indicate that the discontinuity is present and statistically significant both if adoption is defined for shareholding and for an individual.[18] In particular, our preferred estimates (on shareholder level, using larger

---

[18] In what follows, the estimate of the magnitude of the discontinuity is sometimes sensitive to introducing flexible controls when very small window around the threshold is used, but it also corresponds to unrealistic

windows around the threshold) indicate that individuals just above the threshold are 4 percentage point more likely to adopt the E-firm, relative to the base of approximately 10.5 percentage points—nearly 40% increase.

Because all our regressions are estimated using a shareholding as the unit of observation, Table 3 also shows the number of unique individuals in the sample used for each specification (this is also the number of clusters for standard errors estimation). The number of individuals is generally very close (within 5%) of the number of shareholdings, because it is not very common that the same person owns shares in two different firms that happen to be close to 10%.

In the following panel we pursue basic robustness checks by including a set of individual controls—age, gender, number of individual owners and log capital. Inclusion of these additional controls has small impact on both estimates and standard errors, providing some comfort that composition differences are not driving the results.

While the evidence that the 10% ownership share matters for the decision to adopt the E-firm is robust, we are primarily interested in using this effect of own eligibility (cf. Eq. 1) to trace its implications in the network (cf. Eq. 2). We are more likely to be able to statistically trace such responses if the effect of own eligibility is strong. We further investigate subsamples in order to zoom in on a group, if any, that is particularly strongly affected.

Since the benefit of setting up an E-firm is due to reduction in taxation of capital gains or dividends, individuals and firms that generate capital income should be more likely to adopt. Hence, if we further restrict the sample to those shareholdings of individuals who received dividends in 2004 (i.e., pre-reform), results are noisier but arguably more pronounced (see Online Appendix Fig. A14), and there is no discernible effect for the remainder of subsample (Online Appendix Fig. A15). The formal results are shown in the third panel of Table 3 and the magnitude of the effect seems larger than for the full sample, so that despite this group including only about 1/3 of the original sample the t-statistics are of comparable magnitude (consistent with Online Appendix Fig. A15, there is no robust regression evidence of an effect for those with no dividends).

The final panel of Table 3 imposes an additional restriction on the sample by limiting it to those individuals who own firms that have over 1000 shares—a group for which the abstraction of "continuous" variation in ownership shares is more realistic. The estimated effects are large and robust despite much smaller sample size than before.[19]

Footnote 18 continued

estimates of the corresponding coefficients. Restricting the linear term to be the same on both sides of the discontinuity usually stabilizes the results in such cases. The alternative would be to use an automatic bandwidth-selection procedure. We opted to show estimates for a range of intervals in order to allow the reader to asses robustness of the results to variation in bandwidth size directly.

[19] Online Appendix Fig. A16 shows the likelihood of adoption on an individual level for that sample, and a clear jump at the threshold. Online Appendix Fig. A17 shows no jump at the threshold for the remaining individuals owning firms with less than 1000 shares. Restricting the sample to just those with over 1000 shares, with no dividends-in-the-past restriction, also strengthens results relative to the original sample.

Individuals can set up an E-firm either on their own or with others and the 10% rule makes it easier to pursue the latter. Online Appendix Figs. A18 and A19 show that the effect is very clearly driven by setting up E-firms on one's own, with little evidence that there is any decline in setting up E-firms with others. The E-firms stimulated by the 10% rule are single owner ones, with no evidence of crowdout of multiple-owner

Table 3 The effect of crossing 10% ownership on E-firm take-up

Range	Adoption for shareholding			Adoption by an individual		
	No controls	Flexible	N (Unique N)	No controls	Flexible	N (Unique N)
[0.095, 0.105]	0.034*	0.145***	1246	0.041*	0.155***	1246
	(0.019)	(0.044)	(1219)	(0.022)	(0.047)	(1219)
[0.090, 0.110]	0.027**	0.070**	2416	0.035**	0.073**	2416
	(0.013)	(0.029)	(2355)	(0.015)	(0.032)	(2355)
[0.080, 0.120]	0.047***	0.033*	5009	0.039***	0.041*	5009
	(0.009)	(0.019)	(4684)	(0.010)	(0.021)	(4684)
[0.070, 0.130]	0.052***	0.041***	7669	0.043***	0.039**	7669
	(0.007)	(0.015)	(7391)	(0.008)	(0.017)	(7391)
[0.050, 0.150]	0.055***	0.049***	13388	0.050***	0.040***	13388
	(0.005)	(0.011)	(12695)	(0.006)	(0.013)	(12695)
…with demographic controls						
[0.095, 0.105]	0.025	0.152***	1246	0.034	0.160***	1246
	(0.019)	(0.043)	(1218)	(0.021)	(0.046)	(1218)
[0.090, 0.110]	0.026**	0.062**	2416	0.034**	0.064**	2416
	(0.013)	(0.029)	(2355)	(0.014)	(0.032)	(2355)
[0.080, 0.120]	0.048***	0.029	5008	0.041***	0.036*	5008
	(0.009)	(0.019)	(4863)	(0.010)	(0.021)	(4863)
[0.070, 0.130]	0.057***	0.036**	7664	0.049***	0.034**	7664
	(0.007)	(0.015)	(7387)	(0.008)	(0.017)	(7387)
[0.050, 0.150]	0.063***	0.045***	13378	0.060***	0.036***	13378
	(0.005)	(0.011)	(12686)	(0.006)	(0.013)	(12686)
Individuals with dividends in 2004						
[0.095, 0.105]	0.109**	0.461***	390	0.118**	0.434***	390
	(0.046)	(0.095)	(384)	(0.047)	(0.098)	(384)
[0.090, 0.110]	0.086***	0.223***	767	0.098***	0.220***	767
	(0.031)	(0.066)	(764)	(0.032)	(0.068)	(764)
[0.080, 0.120]	0.070***	0.133***	1613	0.066***	0.135***	1613
	(0.020)	(0.044)	(1577)	(0.021)	(0.045)	(1577)
[0.070, 0.130]	0.088***	0.087**	2450	0.085***	0.082**	2450
	(0.016)	(0.035)	(2383)	(0.017)	(0.037)	(2383)
[0.050, 0.150]	0.084***	0.101***	4415	0.083***	0.096***	4415
	(0.012)	(0.026)	(4243)	(0.012)	(0.027)	(4243)
Dividends in 2004 and firms with at least 1000 shares						
[0.095, 0.105]	0.138***	0.453***	293	0.141**	0.425***	293
	(0.053)	(0.100)	(289)	(0.055)	(0.103)	(289)
[0.090, 0.110]	0.121***	0.240***	528	0.144***	0.227***	528
	(0.039)	(0.072)	(512)	(0.041)	(0.075)	(512)

Table 3 continued

Range	Adoption for shareholding			Adoption by an individual		
	No controls	Flexible	N (*Unique N*)	No controls	Flexible	N (*Unique N*)
[0.080, 0.120]	0.112***	0.156***	1072	0.112***	0.170***	1072
	(0.027)	(0.053)	(*1045*)	(0.028)	(0.055)	(*1045*)
[0.070, 0.130]	0.125***	0.132***	1614	0.128***	0.130***	1614
	(0.022)	(0.044)	(*1564*)	(0.023)	(0.046)	(*1564*)
[0.050, 0.150]	0.114***	0.139***	2909	0.121***	0.141***	2909
	(0.016)	(0.034)	(*2779*)	(0.017)	(0.035)	(*2779*)

The unit of observation is a shareholding (possibly multiple per person). The dependent variable is adoption of E-firm for a given shareholding (left side) or for a given individual for any shareholding (right side). Flexible specifications include linear terms on each side of the threshold. N is the number of shareholdings (possibly multiple per person), and *Unique N* is the number of individuals (cf. Sect. 5.1)
*significance at 10% level; **significance at 5% level; ***significance at 1% level

Overall, the results in this section clearly demonstrate that the 10% discontinuity played an important role in determining take-up of E-firms. Those with just over 10% share are much more likely to do so than those below, and the difference is both economically and statistically large. The effect is heterogeneous. It is there for those who are most likely to benefit from it—individuals who have the history of receiving dividends. While this is intuitive, it also indicates that either alternative means of setting up an E-firm (coordinating with others or purchasing additional shares) are costly enough or that the information about availability of the shelter is not there, so that those below 10% share who are otherwise similar do not adopt E-firms to the same extent. Hence, those that take-up E-firms as the result of the treatment would have either been uninformed about this option or found coordination too costly in the absence of the treatment.

6.3 Network effects

We now turn to the network level analysis by analyzing adoption of E-firms of an individual (i) as a function of ownership around the 10% share in a particular firm (k) of a *network member* (j). As discussed before, for this analysis, we focus on the data on the shareholding level so that each shareholding of a treating family network member (j, k) is a separate observation affecting the impacted individual. We focus on network members who fall into subsamples in which we showed evidence of a discontinuity in adoption: we exclude network members with fractional shares, and further zoom in on those receiving capital income and in firms with large number of shares. We do not impose any additional restrictions on individuals (i) themselves—the running variable (ownership share) is the property of the network member and she may affect family members regardless of their characteristics (though we will investigate heterogeneity).

Footnote 19 continued
E-firms, suggesting that setting up an E-firm with others was not the alternative entertained by the population complying with the treatment.
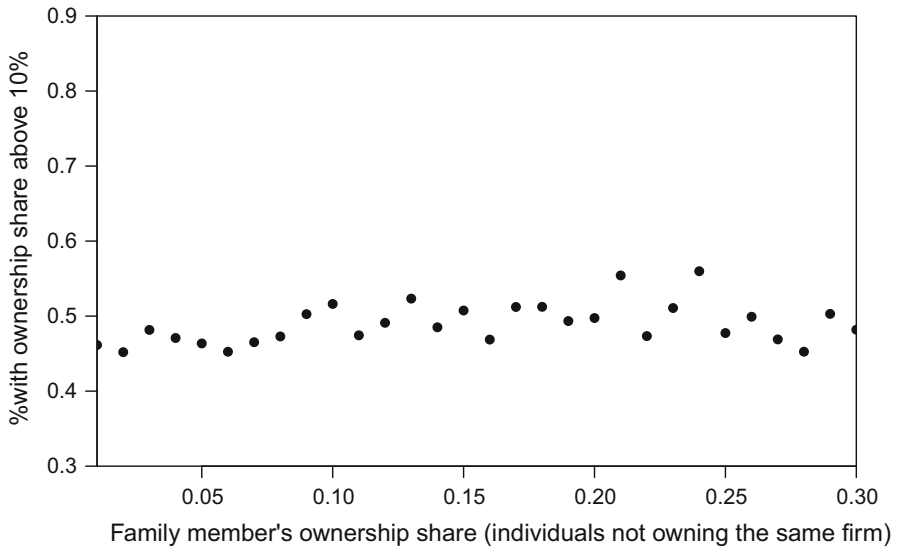
Fig. 8 Family member's ownership share in 2004 and number of individuals with at least 10% shares—individuals not owning the same firm
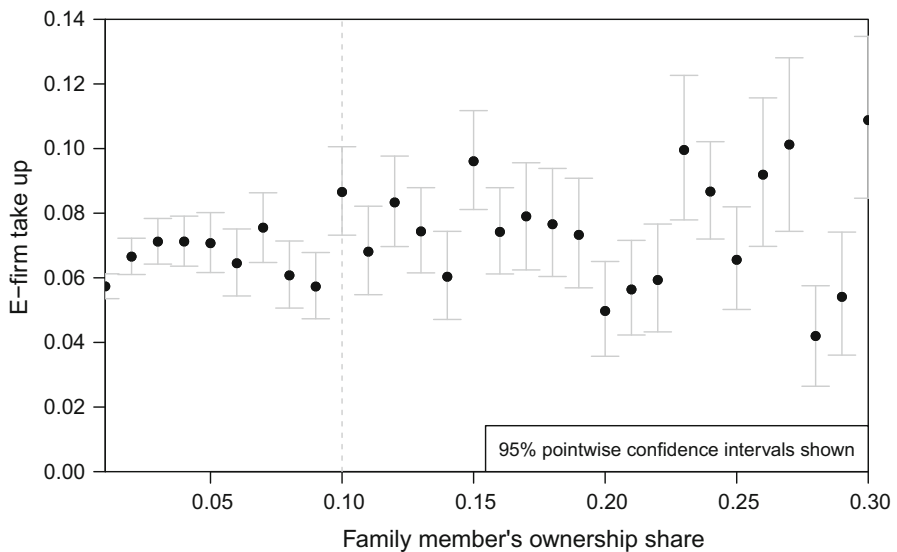
Fig. 9 Family member's ownership share in 2004 and ultimate adoption of the E-firm

Before proceeding further, we want to make sure that when we compare individuals with network members on either side of the 10% threshold, this is the only difference between those groups. Online Appendix Fig. A20 shows though that as the network member's share is crossing 10%, the share owned by the individual itself is more likely to be above 10% as well. It turns out that this is driven by family members owning identical number of shares in the same firm. Hence, in what follows, we restrict

attention to network links between individuals who do not own shares in the same firm (this is our $X_i \perp X_j$ orthogonality assumption). As Fig. 8 shows, in that subsample the likelihood of having a share above 10% sails smoothly through the threshold. We restrict attention to this subsample in what follows. Online Appendix Fig. A9 shows that the number of observations is smooth around the threshold, but, similar to what we see in Sect. 6.1, there is some noise, especially as the ownership share exceeds 1/6th. Similarly, predetermined characteristics (measured in 2004) are quite noisy, but this is so mostly away from the threshold and especially for shares above 1/6th (see Online Appendix Figs. A10, A11, A12, and A13). These patterns are similar to those discussed in the case of individual-level analysis in Sect. 6.1, and again we thus proceed by limiting the window around the threshold to at most of 0.05 on each side, and test robustness of the results by controlling for demographic characteristics. Beyond the necessity of exploiting discontinuity for identification purposes, restricting attention to network links between individuals who do not own shares in the same firm also has economic content: the interaction between "treating" and "treated" individuals is guaranteed not to take place in the context of the firm, but rather has to flow through other channels.

Figure 9 shows the discontinuity-based evidence of adoption elsewhere in the network on individual adoption,[20] and the top panel of Table 4 shows the corresponding estimates. The estimates of the discontinuity are generally significant and reasonably stable as the window around 10% is adjusted. The table also shows the number of unique treating network members j that underlie each specification—there are about half as many of them as all the observations. As we see in Table 3 that is to a small extent driven by the same person having multiple shareholdings close to 10%. The bulk of the difference is explained by the same network member treating multiple individuals in the network.

While the network effect may be present regardless of one's own ownership, individuals who already own at least 10% are already eligible for setting up an E-firm without any additional arrangements and hence may be more strongly affected. At the same time, by virtue of their eligibility, they are more likely to set up an E-firm regardless, so that the additional network incentive might be expected to be weaker for that reason. Zooming in on individuals who own at least 10% ownership share in any firm, strengthens the results (right columns in Table 4). The estimates for this group are larger, suggesting that the first effect dominates.

The second panel shows robustness of the results to inclusion of demographic controls—they are essentially unaffected. Overall, we observe that the estimated effect of a network member being eligible in Table 4 is roughly similar in magnitude to the effect of the individual herself being eligible (Table 3). Theory in Sect. 5 and Appendix A indicates that the large network effect relative to own effect is consistent with either interactions being strong or else low awareness of sheltering opportunities absent interaction with a treating individual.

---

[20] For the individual's ownership in Fig. 7, there is also a discontinuity at 5%, which is, however, not present in the network sample in Fig. 9. This is likely related to the set up of a common E-firm for two individuals' owning 5% is more common within families, and such individuals are dropped from the network sample used in Fig. 9 by us restricting attention to family network links between individuals who do not own shares in the same firm.

Table 4 The effect of crossing 10% ownership by a family network member on E-firm take-up

Range	Everyone			At least 10% share		
	No controls	Flexible	N (Unique N)	No controls	Flexible	N (Unique N)
[0.095, 0.105]	0.034**	− 0.005	1935	0.062**	− 0.005	1117
	(0.017)	(0.028)	(748)	(0.027)	(0.051)	(481)
[0.090, 0.110]	0.029**	0.031	3473	0.045**	0.057	1940
	(0.012)	(0.023)	(1443)	(0.020)	(0.039)	(916)
[0.080, 0.120]	0.019**	0.040**	6799	0.031**	0.065**	3652
	(0.008)	(0.016)	(2963)	(0.014)	(0.027)	(1859)
[0.070, 0.130]	0.015**	0.033**	10580	0.023**	0.056**	5632
	(0.007)	(0.014)	(4508)	(0.011)	(0.023)	(2846)
[0.050, 0.150]	0.009*	0.028***	18295	0.015*	0.043**	9651
	(0.005)	(0.011)	(7667)	(0.009)	(0.018)	(4799)
…with demographic controls						
[0.095, 0.105]	0.034**	− 0.002	1931	0.056**	− 0.007	1116
	(0.017)	(0.028)	(747)	(0.027)	(0.049)	(481)
[0.090, 0.110]	0.028**	0.034	3467	0.039*	0.053	1939
	(0.012)	(0.023)	(1440)	(0.020)	(0.037)	(916)
[0.080, 0.120]	0.018**	0.039**	6785	0.026*	0.057**	3647
	(0.008)	(0.016)	(2958)	(0.014)	(0.027)	(1856)
[0.070, 0.130]	0.013**	0.033**	10562	0.017	0.048**	5626
	(0.007)	(0.014)	(4502)	(0.011)	(0.023)	(2843)
[0.050, 0.150]	0.008	0.027***	18263	0.011	0.037**	9643
	(0.005)	(0.010)	(7653)	(0.009)	(0.017)	(4795)
Treating network members with dividends in 2004						
[0.095, 0.105]	0.056**	0.050	503	0.114**	0.035	285
	(0.028)	(0.035)	(222)	(0.048)	(0.063)	(137)
[0.090, 0.110]	0.058**	0.061*	924	0.110***	0.098*	504
	(0.024)	(0.034)	(420)	(0.040)	(0.059)	(264)
[0.080, 0.120]	0.032**	0.081***	2017	0.060**	0.148***	1060
	(0.015)	(0.029)	(937)	(0.027)	(0.050)	(579)
[0.070, 0.130]	0.025**	0.058**	3134	0.046**	0.113***	1623
	(0.012)	(0.025)	(1418)	(0.021)	(0.043)	(874)
[0.050, 0.150]	0.024***	0.033*	5822	0.038**	0.066**	3056
	(0.009)	(0.019)	(2527)	(0.015)	(0.033)	(1561)
Treating network members without dividends in 2004						
[0.095, 0.105]	0.026	− 0.036	1432	0.043	− 0.047	832
	(0.020)	(0.036)	(531)	(0.031)	(0.067)	(347)
[0.090, 0.110]	0.020	0.018	2549	0.023	0.037	1436
	(0.014)	(0.028)	(1034)	(0.023)	(0.049)	(660)

Table 4 continued

Range	Everyone			At least 10% share		
	No controls	Flexible	N (Unique N)	No controls	Flexible	N (Unique N)
[0.080, 0.120]	0.013	0.026	4782	0.018	0.037	2592
	(0.010)	(0.020)	(2061)	(0.017)	(0.033)	(1301)
[0.070, 0.130]	0.010	0.025	7446	0.013	0.036	4009
	(0.008)	(0.017)	(3152)	(0.014)	(0.028)	(2008)
[0.050, 0.150]	0.001	0.027**	12473	0.004	0.035*	6595
	(0.006)	(0.013)	(5281)	(0.010)	(0.021)	(1561)

Flexible specifications include linear terms on each side of the threshold. N is the number of network pairs: shareholding (j, k) to treated-individuals i, and *Unique N* is the number of treating individuals j (cf. Sect. 5.1)

*significance at 10% level; **significance at 5% level; ***significance at 1% level

Following up on our previous discussion, we further split the sample by whether the network member received dividends in 2004. The bottom two panels of Table 4 show that for those with family members who received dividends, the effects are of the expected sign and not too sensitive to the size of the window or inclusion of controls.[21] They are becoming significant when the window around the threshold (and sample size) grows and in narrow window when no controls are included (the linear terms in ownership share are generally insignificant). The results for those with family members who have not received dividends are smaller and generally insignificant.[22]

This is consistent with the interpretation of take-up by a family member reflecting the presence of the treatment: since the direct effect on take-up for that group was not detectable, observing an impact on their family members would be surprising.[23]

In Table 5 we split the sample in additional ways. First, we look at those with treating family members with dividends in 2004 and firms with over 1000 shares. In this group, the effect of own eligibility (cf. Eq. 1) was strong and the corresponding results are strong here as well. Then, we split the sample by whether the treated individual itself received dividends in 2004. We find more precise statistical evidence for those who did not receive dividends themselves than for those who did, though the large standard errors do not allow for rejecting the possibility that point estimates are not statistically different. Still, even if the coefficients for those without dividends were similar in absolute value, the base take-up for this group is much lower and hence the impact is

---

[21] See also Online Appendix Figs. A21 and A22.

[22] Online Appendix Table A1 shows the results from the specification that pools network links with and without dividends, but includes a dummy for the network member having received dividends, its interaction with crossing the threshold and ownership share controls restricted to be the same across groups—this restriction strengthens the results.

[23] As we discussed in the context of the model in Appendix A, in principle the treatment may have an effect on family members even when it does not affect the decision of the treated individual itself. In particular, those without dividends may choose not to take-up the shelter but having been given an opportunity to do so may now be in a position to inform others. Although the network results for that subsample are for the most part insignificant, they are consistently positive and fairly stable as the window around the discontinuity point widens.

Table 5 The effect of crossing 10% ownership by a family network member on E-firm take-up-decomposition of response

Range	Everyone			At least 10% share		
	No controls	Flexible	N (Unique N)	No controls	Flexible	N (Unique N)
Treating network members with dividends in 2004 and firms over 1000 share						
[0.095, 0.105]	0.074**	0.013	380	0.142***	−0.018	225
	(0.032)	(0.036)	(*170*)	(0.053)	(0.062)	(*106*)
[0.090, 0.110]	0.065**	0.071**	662	0.127**	0.116*	376
	(0.029)	(0.036)	(*290*)	(0.049)	(0.062)	(*186*)
[0.080, 0.120]	0.049**	0.090***	1365	0.090***	0.161***	747
	(0.020)	(0.034)	(*631*)	(0.034)	(0.058)	(*390*)
[0.070, 0.130]	0.035**	0.078**	2161	0.065**	0.142***	1149
	(0.015)	(0.030)	(*955*)	(0.027)	(0.052)	(*591*)
[0.050, 0.150]	0.033***	0.045*	3990	0.054***	0.090**	2153
	(0.011)	(0.024)	(*1680*)	(0.019)	(0.041)	(*1043*)
Treating network members with dividends in 2004, treated-individuals without dividends						
[0.095, 0.105]	0.069**	0.041	328	0.122**	0.007	186
	(0.032)	(0.037)	(*159*)	(0.051)	(0.059)	(*94*)
[0.090, 0.110]	0.081***	0.059	615	0.141***	0.072	328
	(0.029)	(0.039)	(*316*)	(0.049)	(0.062)	(*184*)
[0.080, 0.120]	0.035**	0.096***	1377	0.065**	0.161***	706
	(0.017)	(0.035)	(*732*)	(0.031)	(0.057)	(*414*)
[0.070, 0.130]	0.031**	0.068**	2180	0.062***	0.114**	1085
	(0.012)	(0.029)	(*1125*)	(0.023)	(0.051)	(*637*)
[0.050, 0.150]	0.028***	0.041*	4067	0.048***	0.081**	2032
	(0.009)	(0.021)	(*2043*)	(0.017)	(0.038)	(*1143*)
Treating network members, treated-individuals with dividends in 2004						
[0.095, 0.105]	0.035	0.076	175	0.135	0.134	99
	(0.049)	(0.065)	(*105*)	(0.087)	(0.132)	(*61*)
[0.090, 0.110]	0.011	0.069	309	0.050	0.170	176
	(0.035)	(0.060)	(*190*)	(0.058)	(0.107)	(*114*)
[0.080, 0.120]	0.031	0.054	640	0.062	0.136*	354
	(0.028)	(0.047)	(*406*)	(0.046)	(0.082)	(*237*)
[0.070, 0.130]	0.020	0.049	954	0.027	0.125*	538
	(0.024)	(0.043)	(*604*)	(0.040)	(0.070)	(*366*)
[0.050, 0.150]	0.018	0.030	1755	0.030	0.055	1024
	(0.018)	(0.035)	(*1075*)	(0.028)	(0.057)	(*664*)

The dependent variable is adoption of E-firm by a treated network member. The table shows the effect of restricting the sample to network members that have over 1000 shares and the decomposition of response by whether treated individual had dividends. N is the number of network pairs: shareholding (j, k) to treated-individuals i, and *Unique N* is the number of treating individuals j (cf. Sect. 5.1)

*significance at 10% level; **significance at 5% level; ***significance at 1% level

Table 6 The effect of crossing 10% ownership by a family network member on E-firm take-up—distance of the date of adoption to 1/1/2006

Group	Everyone			At least 10% share		
	OLS					
	No controls	Flexible	N (Unique N)	No controls	Flexible	N (Unique N)
All	0.273	0.669	18295	0.218	0.960	9651
	(0.261)	(0.606)	(7667)	(0.446)	(1.065)	(4799)
Dividends	0.995*	1.218	5822	1.387	2.592	3056
	(0.512)	(1.198)	(2527)	(0.908)	(2.135)	(1561)
No dividends	− 0.080	0.480	12473	− 0.347	0.357	6595
	(0.322)	(0.686)	(5281)	(0.547)	(1.192)	(3320)
	Tobit					
All	5.323*	16.039**	18295	4.863	14.255**	9651
	(3.179)	(6.606)	(7667)	(3.256)	(6.851)	(4799)
Dividends	15.005***	19.821*	5822	13.374**	23.671**	3056
	(5.384)	(11.225)	(2527)	(5.638)	(12.047)	(1561)
No dividends	0.369	15.053*	12473	0.354	10.981	6595
	(4.033)	(8.014)	(5281)	(4.118)	(8.160)	(3320)

The dependent variable is the number of days from the date of adoption to the 1/1/2006 deadline. The coefficients should be interpreted as the number of days. Results for the full sample and a split by whether treating network members had dividends in 2004. N is the number of network pairs: shareholding (j, k) to treated-individuals i, and *Unique N* is the number of treating individuals j (cf. Sect. 5.1)

*significance at 10% level; **significance at 5% level; ***significance at 1% level

economically much more significant—for example, the estimated effect of 0.04 over the (0.05, 0.15) window for the flexible specification corresponds to roughly doubling the take-up. Thus, a very rough taxonomy of the results may be that treating individuals with most to gain (those with dividends) are most responsive to the 10% threshold incentive, but they stimulate take-up by (treated) individuals who have less potential to gain (those without past dividends) and so perhaps least informed otherwise.

6.4 Effect on timing

In order to further substantiate the presence of network effects, we combine the RD evidence with timing. In Sect. 4, we provided evidence of a strong association between timing of adoption in the network and individual adoption, but though suggestive, these results cannot be interpreted as causal. Here we use our regression discontinuity approach to further corroborate the presence of interesting timing effects. In all the following specifications, we focus on the 0.05 window around the discontinuity point.

In Table 6 we look at the effect on the number of days between January 1, 2006 and when the tax shelter was established, with individuals not establishing the shelter assigned a zero value. We estimate regression discontinuity specifications as in Sect. 6.3, just with the distance of adoption until January 1st as the dependent variable. The OLS specification results in Sect. 4 are positive, but for the most part insignificant

Table 7 The effect of crossing 10% ownership by a family network member on E-firm take-up—timing effects, everyone

Timing of adoption	Δp		$\Delta \ln(p)$		N
	No controls	Flexible	No controls	Flexible	
1–30 days	0.000	0.006	0.010	0.198	18295
	(0.004)	(0.006)	(0.101)	(0.216)	
31–60 days	0.007***	0.014***	0.360***	0.806***	18295
	(0.002)	(0.004)	(0.137)	(0.302)	
61–90 days	0.001	0.002**	0.279	0.755	18295
	(0.001)	(0.001)	(0.239)	(0.516)	
91–120 days	−0.001	−0.000	−0.387	−0.088	18295
	(0.002)	(0.003)	(0.430)	(1.081)	
Network members with dividends in 2004					
1–30 days	0.010*	0.014	0.302*	0.423	5822
	(0.005)	(0.011)	(0.178)	(0.375)	
31–60 days	0.009***	0.010*	0.499**	0.638	5822
	(0.003)	(0.006)	(0.213)	(0.485)	
61–90 days	0.001	0.005**	0.227	1.177	5822
	(0.002)	(0.002)	(0.367)	(0.884)	
91–120 days	−0.001	−0.017	−0.310	−5.907*	5822
	(0.002)	(0.017)	(0.782)	(3.280)	
Network members without dividends in 2004					
1–30 days	−0.005	0.003	−0.127	0.098	12473
	(0.005)	(0.008)	(0.123)	(0.260)	
31–60 days	0.005*	0.016***	0.283	0.905**	12473
	(0.003)	(0.005)	(0.179)	(0.372)	
61–90 days	0.001	0.001	0.297	0.494	12473
	(0.001)	(0.001)	(0.294)	(0.603)	
91–120 days	−0.002	0.002	−0.407	0.735	12473
	(0.002)	(0.002)	(0.553)	(1.192)	

Estimates of the effect of probability of adopting in the given timing window before the reform, as a function of having a network member with over 10% ownership (regression discontinuity, using (0.05, 0.15) range)
*significance at 10% level; **significance at 5% level; ***significance at 1% level

for the full and dividend samples. Since distance to January 1st is effectively censored at zero, these results are biased downward. As an alternative, in the following panel we make the normality assumption and estimate the effect on the date of setting up a shelter via Tobit specification. The Tobit estimates have the same sign as the OLS ones but, consistently with the expected OLS bias, are much larger and statistically significant. The results indicate that having a family member exposed to the 10% rule accelerated take-up of the tax shelter by as much as 20 days; the results are robustly significant for the sample with dividends, smaller for the full sample, and possibly zero (with large standard errors) for those with family members who did not have dividends in 2004.

When exactly did these network effects materialize? In Table 7 and Online Appendix Tables A2–A4 we focus on results for particular periods. We ask whether adoption during various time periods—either separately (1–30, 31–60, 61–90, 91–120 days before the reform) or cumulatively (overall, more than 30, 60 or 90 days before), is explained by network exposure. The results are presented both for the full sample and those with at least 10% ownership.

Focusing on the results for everyone, in Table 7 we report results from probit specifications. The table contains an estimate of the effect of crossing the 10% discontinuity on probability of adopting at the threshold and the effects on log probability to allow for more meaningful comparison across different periods.[24] The effect is strongest in the second month before the deadline and it appears to be there for both those with and without network members who received dividends. The evidence of the effect in the last month is weaker, but it is suggestive of the presence of the effect continuing up until the deadline for those with family members who received dividends. The results for three or four months prior to the reform do not indicate an effect though they are noisy and sometimes counterintuitive, reflecting a small number of individuals taking up early on. Online Appendix Table A2 shows the cumulative effect—impact on adoption by the time of the reform (same as our main specification) and by 30, 60 or 90 days pre-reform. Consistently with month-by-month results in Table 7, cumulative results indicate that the bulk of the effect is already there 30 days before the reform. The results for those with at least 10% ownership (in Appendix tables) are qualitatively similar.

Overall, we conclude that this discontinuity-based evidence on timing adds support for the presence of a causal relationship in adopting tax shelters running through the network; as such it also strengthens our observations from Sect. 4 that timing dimension of the response is important.

7 Conclusions

We considered effects of eligibility to adopt a legal tax sheltering strategy in Norwegian family networks. In a descriptive analysis of the timing of take-up of the tax shelter, we find that early adoption in the family network is correlated with individual take-up later on. Looking at the short term impact reveals significant increase in adoption within the week after a network member set up an E-firm. This is very robust to controls and provides very suggestive evidence that take-up in the network stimulates adoption of the tax shelter.

Relying on a regression discontinuity design in the incentives to adopt, we showed that family members of individuals who had a strong incentive to pursue tax sheltering (and who, in fact, responded accordingly) are more likely to pursue tax avoidance themselves. These patterns are not uniform across different group of individuals. The propensity to adopt at discontinuity is strongest by individuals who are most likely to benefit (as measured by history of capital income) and it is their family networks that are affected. At the same time, there is suggestive evidence that it is those members of

---

[24] We use probit, because in periods distant from the reform probabilities are very close to zero. The results are very robust to using linear probability model instead.

family networks who themselves do not have a strong reason to pursue tax avoidance that respond most strongly. This is consistent with two possibilities: these are either uninformed individuals or they face high cost of adoption relative to benefits and that this cost is reduced by having a family member familiar with the process.

More generally, our results provide one of the first empirically well-identified examples that tax planning is a social phenomenon that is affected by what others do. This highlights the importance of accounting for social interactions in understanding enforcement and tax avoidance behavior. Recent work by Pomeranz (2015) highlights that network incentives matter in the VAT context; in our case, however, there is no compliance spillover that may explain our findings—the strategy is legal and networks are not linked by business interests that could explain correlated behavior. Instead, it is likely that knowledge on the benefits of a tax shelter, reduced costs through free advice from a family member, or norms on the acceptability of utilizing a tax shelter are transmitted within a network. Our evidence of heterogeneous patterns of response points to knowledge and cost as likely channels. More research is called for to explore the external validity of the current spillover results for other networks, other tax avoidance strategies, and in other countries.

Our findings are also related to the literature on optimization frictions. Because different individuals have different networks, they effectively are differentially exposed to planning opportunities (either through knowledge or through costs). As a result, networks are linked to optimization frictions: their importance reveals that individuals do not necessarily react to all theoretical incentives out there absent exposure in the family and that they are behind heterogeneity in behavior that varies with the extent of exposure. The recent literature in public finance stresses the relevance of optimization frictions, often taken as abstract barriers to optimization, and our work points to one possible direction for understanding when, how and why they might be present and vary.

Appendix A: Theoretical framework

We are interested in understanding how adoption of tax avoidance strategies within a network affects the individual shareholder's uptake of such strategies. Consider individuals j and i who are linked in the family network. We are interested in the determinants of the decision to adopt an E-firm, E_i, that we presume is determined by a latent variable $\tilde{E}_i$, $E_i = I(\tilde{E}_i > 0)$ (where I is an indicator function). We assume that $\tilde{E}_i = f_i(K_i, X_i) + \varepsilon_i$ where K_i is the set of endogenous variables that may play a

role in individual interactions (though, to simplify notation, we will be explicit about it only where it matters), X_i is the set of individual characteristics and ε_i is the error term orthogonal to $f_i(\cdot)$. To fix attention, we will refer to K_i as taxpayer's awareness of tax shelter possibility. In general, we assume that $K_i(K_j, K^i_{-j}, X_i)$, where K_j is awareness of taxpayer j and K^i_{-j} is awareness of individuals in the network other than j, but we will mostly consider just two individuals, so that $K_i(K_j, X_i)$ and $K_j(K_i, X_j)$. We also assume that derivatives of f and K are non-negative.

The framework assumes that individuals affect each other through variables K and allows for reciprocal reactions.[25] For our purposes, we do not need to assume symmetry so that functions $K_i(\cdot)$ and $K_j(\cdot)$ may be different. We focus on the interaction between i and j and are agnostic about the role of K_{-j} at this point, but we will return to it below. Shocks to the environment are represented by the effect on X_i. Common shocks could affect both X_i and X_j simultaneously, but in what follows we will focus on tracing out implications of an idiosyncratic shock to individual j, ΔX_j, that corresponds to the source of identification that we explore in our empirical work.

Shocks may potentially have four different qualitative effects. First, they affect awareness of the recipient directly (when it is the shock to individual own environment) when $\frac{\partial K_j}{\partial X_j} \neq 0$. Second, they may affect awareness of others when $\frac{\partial K_i}{\partial K_j} \neq 0$ (with the feedback to the original recipient when $\frac{\partial K_j}{\partial K_i} \neq 0$). Third, the overall impact on awareness matters for sheltering behavior when $\frac{\partial f_j}{\partial K_j} \neq 0$. Fourth, sheltering may be affected by the shock directly without altering interactions with others ($\frac{\partial E}{\partial X_j} \neq 0$).[26]

The total impact of the shock to individual j on the individual itself may be traced out as follows:

$$\frac{\Delta \tilde{E}_j}{\Delta X_j} = \frac{\partial f_j}{\partial K_j} \frac{\Delta K_j}{\Delta X_j} + \frac{\partial f_j}{\partial X_j}, \qquad \frac{\Delta K_j}{\Delta X_j} = \frac{\partial K_j}{\partial K_i} \frac{\Delta K_i}{\Delta X_j} + \frac{\partial K_j}{\partial X_j} + \frac{\partial K_j}{\partial K^j_{-i}} \frac{\Delta K^j_{-i}}{\Delta X_j}$$

The effect on individual i has similar structure except for the lack of direct effects:

$$\frac{\Delta \tilde{E}_i}{\Delta X_j} = \frac{\partial f_i}{\partial K_i} \frac{\Delta K_i}{\Delta X_j}, \qquad \frac{\Delta K_i}{\Delta x_j} = \frac{\partial K_i}{\partial K_j} \frac{\Delta K_j}{\Delta X_j} + \frac{\partial K_i}{\partial K^i_{-j}} \frac{\Delta K^i_{-j}}{\Delta X_j}$$

---

[25] This framework is general enough to accommodate many economically interesting special cases. For example, suppose that X_j represents information of individual j. A shock to x_j may affect individual i when person i and j interact, but there need not be a feedback effect on individual j since person i has no additional information over person j as the result of that shock. This case fits in this framework by allowing K to be two-dimensional, (K_1, K_2) with $\frac{\partial K_{1j}}{\partial X_j} \neq 0$, $\frac{\partial K_{2i}}{\partial K_{1j}} \neq 0$ and $\frac{\partial K_{2j}}{\partial K_{2i}} = \frac{\partial K_{1j}}{\partial K_{2i}} = 0$, for example when $K_{1i}(K_{1j}, K_{2j}, X_i) = g(X_i)$ and $K_{2i}(K_{1j}, K_{2j}, X_i) = h(K_{1j})$ for all i, j so that a signal affects own awareness K_1 and, through this channel, network member's outside knowledge K_2 but there is no feedback from K_2 on others.

[26] Note that separating K from adoption $\tilde{E}$ allows shock to individual j to affect take-up of individual i without affecting propensity of an individual j to take-up—this would be the case whenever $\frac{\partial f_j}{\partial K_j} \frac{\Delta K_j}{\Delta X_j} + \frac{\partial f_j}{\partial X_j} = 0$ (in particular, when both derivatives of f_j are zero) but $\frac{\Delta K_i}{\Delta X_j} \neq 0$. For example, an individual j may (exogenously) learn about sheltering opportunity that is not of interest to her (given all her characteristics X_j) but still pass such information to others.

Without observing awareness directly, an econometrician can estimate reduced form effects $\frac{\Delta \tilde{E}_j}{\Delta X_j}$ and $\frac{\Delta \tilde{E}_i}{\Delta X_j}$.

Remark 1 Suppose that $\frac{\Delta \tilde{E}_i}{\Delta X_j} \neq 0$. It implies then that $\frac{\Delta K_i}{\Delta X_j} \neq 0$, thereby providing evidence of social interactions between individuals.

This observation underlies our basic test. We will attempt to estimate the effect on $\frac{\Delta \tilde{E}_i}{\Delta X_j}$ by using variation in X_j that is (assumed) independent of X_i. The independence assumption is natural in the presence of explicit randomization, for example as in Duflo and Saez ([2003](#)); we are going instead to rely on a regression discontinuity design that also makes such an assumption plausible and appealing.

Remark 2 $\frac{\Delta \tilde{E}_j}{\Delta X_j} \neq 0$ is neither necessary nor sufficient for the presence of social interactions.

Observing that an individual responds to her own incentives is not sufficient to establish presence of interactions. It is also strictly speaking not a necessary condition (see an example in footnote 25—a taxpayer who is exposed to information about a shelter may transmit information to others but not act on it) although it is arguably unlikely.

The formulae for $\frac{\Delta K_j}{\Delta x_j}$ and $\frac{\Delta K_i}{\Delta x_j}$ reflect interactions between those terms but can be combined to obtain:

$$\frac{\Delta K_j}{\Delta x_j} = \left(\frac{\partial K_j}{\partial x_j} + \frac{\partial K_j}{\partial K_{-i}^j} \frac{\Delta K_{-i}^j}{\Delta x_j} + \frac{\partial K_j}{\partial K_i} \frac{\partial K_i}{\partial K_{-j}^i} \frac{\Delta K_{-j}^i}{\Delta x_j} \right) \cdot S^{-1} \tag{4}$$

$$\frac{\Delta K_i}{\Delta x_j} = \frac{\partial K_i}{\partial K_j} \left(\frac{\partial K_j}{\partial x_j} + \frac{\partial K_j}{\partial K_{-i}^j} \frac{\Delta K_{-i}^j}{\Delta x_j} + \frac{1}{\partial K_i / \partial K_j} \frac{\partial K_i}{\partial K_{-j}^i} \frac{\Delta K_{-j}^i}{\Delta x_j} \right) \cdot S^{-1} \tag{5}$$

where $S = 1 - \frac{\partial K_i}{\partial K_j} \frac{\partial K_j}{\partial K_i}$ is the "social multiplier" that measures the magnification of the direct effect (with $\frac{\partial K_i}{\partial K_j} \frac{\partial K_j}{\partial K_i} < 1$ being a necessary condition for stability). To illustrate the logic of this condition, assume for the moment that $\frac{\partial K_i}{\partial K_{-j}^i} \frac{\Delta K_{-j}^i}{\Delta X_j} \approx 0$, that is that the effect of X_j on individuals in the network of i other than j is negligible. In such a case, $\frac{\Delta K_i}{\Delta X_j} = \frac{\partial K_i}{\partial K_j} \frac{\Delta K_j}{\Delta X_j}$ and plugging back into formulas for $\tilde{E}$, yields

$$\beta_j \equiv \frac{\Delta \tilde{E}_j}{\Delta X_j} = \frac{\partial f_j}{\partial X_j} + \frac{\partial f_j}{\partial K_j} \left(\frac{\partial K_j}{\partial X_j} \Big/ S \right) \tag{6}$$

$$\beta_i \equiv \frac{\Delta \tilde{E}_i}{\Delta X_j} = \frac{\partial K_i}{\partial K_j} \cdot \frac{\partial f_i}{\partial K_i} \left(\frac{\partial K_j}{\partial X_j} \Big/ S \right). \tag{7}$$

Individual j responds to a change in X_j due to the direct effect it has on sheltering and due to the effect it has on own awareness K_j. The latter effect is magnified due to

the presence of interaction reflected by term S. Sheltering of person i is not affected by X_j directly, but it is affected through the awareness channel. The exogenous shift in awareness is due to the impact it has on awareness of person j (magnified by the presence of spillover effect S). This shift affects person i's awareness to the extent that interactions are present—$\frac{\partial K_i}{\partial K_j}$—and affects the ultimate decision to the extent that awareness matters for sheltering, $\frac{\partial f_i}{\partial K_i} \neq 0$.

In general, separately identifying the direct impact on sheltering $\frac{\partial f_j}{\partial X_j}$, the impact through increased awareness $\frac{\partial f_j}{\partial K_j}\frac{\Delta K_j}{\Delta X_j}$ and the interaction effect $\frac{\partial K_i}{\partial K_j}$ based on estimates of $\frac{\Delta \tilde{E}_j}{\Delta X_j}$ and $\frac{\Delta \tilde{E}_i}{\Delta X_j}$ is not possible. Note though that

Remark 3 Assuming that $\frac{\partial f_j}{\partial X_j} \geq 0$ and $\frac{\partial K_i}{\partial K_{-j}}\frac{\Delta K_{-j}^i}{\Delta x_j} \approx 0$,[27] we have $\frac{\beta_i}{\beta_j} = \frac{\Delta \tilde{E}_i}{\Delta x_j}\Big/\frac{\Delta \tilde{E}_j}{\Delta x_j} \leq \frac{\partial K_i}{\partial K_j}\frac{\frac{\partial f_i}{\partial K_i}}{\frac{\partial f_j}{\partial K_j}}$ with equality when $\frac{\partial f_j}{\partial X_j} = 0$ and $\frac{\partial K_i}{\partial K_{-j}}\frac{\Delta K_{-j}^i}{\Delta x_j} = 0$.

so that the ratio of the coefficients contains information about the strength of the social interactions.

---

[27] Assuming that $\frac{\partial K_i}{\partial K_{-j}}\frac{\Delta K_{-j}^i}{\Delta x_j} \approx 0$ is restrictive but allows for a major simplifications of formulas 4 and 5. That assumption does not rule out effects on others—indeed, other members of the network of person j can be still affected and influence person i—but it rules out feedback effects of higher than second-order: a shock to person j affecting person i who in turn affects some other person k and recognizing the feedback from person k back to i and j. Imposing additional structure on the model may allow for incorporating these types of effects. The complication in our context is that networks are not disjoint so that modeling equilibrium is considerably less tractable than in the case of, for example, peer effects within a school or neighborhood. We leave addressing this issue for future work. The second aspect of restrictiveness of this assumption is that individuals i and j may (and, indeed, usually will) have common network members so that even if one was willing to rule out higher order feedback effects, some members of K_{-j}^i may not be ignored. This is a conceptually separate issue that may be explicitly addressed by enumerating those individuals in arguments of K_i and K_j. Denoting a set of common network members by M,

$$\frac{\Delta K_j}{\Delta x_j} = S\left(\frac{\partial K_j}{\partial x_j} + \frac{\partial K_j}{\partial K_{-i}^j}\frac{\Delta K_{-i}^j}{\Delta x_j}\right) + \sum_{m \in M} S\left(\frac{\partial K_j}{\partial K_i}\frac{\partial K_i}{\partial K_m} + \frac{\partial K_j}{\partial K_m}\right)\frac{\Delta K_m}{\Delta x_j} \quad (8)$$

$$\frac{\Delta K_i}{\Delta x_j} = S\frac{\partial K_i}{\partial K_j}\left(\frac{\partial K_j}{\partial x_j} + \frac{\partial K_j}{\partial K_{-i}^j}\frac{\Delta K_{-i}^j}{\Delta x_j}\right) + \sum_{m \in M} S\left(\frac{\partial K_i}{\partial K_m} + \frac{\partial K_i}{\partial K_j}\frac{\partial K_i}{\partial K_m}\right)\frac{\Delta K_m}{\Delta x_j} \quad (9)$$

The bracketed terms in the last expressions are symmetric so that it is natural to assume that they are the same on average. Then, these two terms can be written as $\frac{\Delta K_j}{\Delta x_j} = a + M \cdot x$ and $\frac{\Delta K_i}{\Delta x_j} = \frac{\partial K_i}{\partial K_j}a + M \cdot x$ where $x = \mathrm{E}\left[S\left(\frac{\partial K_i}{\partial K_m} + \frac{\partial K_i}{\partial K_j}\frac{\partial K_i}{\partial K_m}\right)\frac{\Delta K_m}{\Delta x_j}\right]$, $a = \left(\frac{\partial K_j}{\partial x_j} + \frac{\partial K_j}{\partial K_{-i}^j}\frac{\Delta K_{-i}^j}{\Delta x_j}\right)$ and M is the number of common network members. Suppose that we were able to observe $\frac{\Delta K_j}{\Delta x_j}$ and $\frac{\Delta K_i}{\Delta x_j}$, then by controlling for M, we can identify x and combine $\frac{\Delta K_j}{\Delta x_j}$ and $\frac{\Delta K_i}{\Delta x_j}$ to recover $\frac{\partial K_i}{\partial K_j}$ as before. Since K is not directly observable, pursuing the same exercise as before using observable sheltering decisions E again requires the assumption that $\frac{\partial f}{\partial x_j} = 0$.

To make progress, assume indeed that $\frac{\partial f_j}{\partial X_j} = 0$.[28] Then, making an additional assumption that $\frac{\partial f_j}{\partial K_j} = \frac{\partial f_i}{\partial K_i}$ (or that the ratio of the two terms is some other known constant), the ratio $\frac{\beta_i}{\beta_j}$ would identify the strength of social interactions as the ratio of the two effects.

The assumption, $\frac{\partial f_j}{\partial K_j} = \frac{\partial f_i}{\partial K_i}$, is restrictive. When the level of awareness for the individuals varies, $K_i \neq K_j$. In fact, in our case, we will contend that individuals who are affected by the shock that we rely on to identify the effect are relatively well-informed anyway while those in their networks are not so that $K_j > K_i$. In that case, we would expect $\frac{\partial f_j}{\partial K_j} \leq \frac{\partial f_i}{\partial K_i}$.

Appendix B: Procedure for removing fractions

As mentioned in text, there is clustering of individuals at particular fractional points and it is not limited just to fractions with small base or equal splitting. For example, there are 662 observations with the share of precisely $1/12 \approx 0.833$ and the total number of observations in interval $[0.083, 0.084)$ is 888 while the number of observation in surrounding 1/1000th intervals are 216 in $[0.082, 0.083)$ and 134 in $[0.084, 0.085)$. Our objective is to apply procedure that would eliminate such bunching in a way that is systematic but at the same time would allows us for keeping as many observations as possible (after all, every shareholding is a fraction with the denominator equal to the total number of shares in a firm). We eliminated shareholdings that are exact fractions with denominators between 1 and 20, 25, 30, 40, 50, 100 and 200. In particular, this removes all shareholding that are multiples of 0.005, and of course the discontinuity point 0.1 itself. We additionally remove points that are within 1 share of a fraction with the denominator of 3, 6, 9 or 10—there is evidence of bunching at that kind as well that is particularly strong for these values and occurs very close to the discontinuity point when the denominator is 9 or 10. The resulting histogram is Fig. 5 and it no longer shows evidence of significant bunching when aggregated to 0.001 intervals. The procedure eliminates a large part of the sample—24,294 out of 47,682 observations in the $(0.05, 0.15)$ interval (with almost 7600 removed observations at 0.1 and another 5600 at other 1/100th multiples in the interval). Results are robust to adjustments involving reasonable expanding or limiting the set of denominators accounted for as long as major discontinuity points are eliminated.

---

[28] Assumption of $\frac{\partial f_j}{\partial X_j} = 0$ effectively eliminates the distinction between K and $\tilde{E}$ by ruling out the possibility that X_j may have an impact on sheltering that is not interacting with behavior of others. In particular, it eliminates the natural kind of heterogeneity where individuals are interacting using some variables K but the strength of their response is determined by the value of X_j. In our context, X_j is likely to have an independent effect because it reflects eligibility for setting up a tax shelter that reduces the cost of acting for a particular individual—this effect is conceptually separate from, for example, increased awareness of the shelter and may influence behavior of a taxpayer without affecting others. If, on the other hand, a taxpayer affects others via the decision to shelter only, the assumption of $\frac{\partial f_j}{\partial X_j} = 0$ would hold.

References

Allingham, M. G., & Sandmo, A. (1972). Income tax evasion: A theoretical analysis. *Journal of Public Economics*, *1*(3–4), 323–338.

Alstadsæter, A. (2007). The Achilles heel of the dual income tax: The Norwegian case. *Finnish Economic Papers*, *20*(1), 5–22.

Alstadsæter, A., & Fjærli, E. (2009). Neutral taxation of shareholder income? Corporate responses to an announced dividend tax. *International Tax and Public Finance*, *16*(4), 571–604.

Alstadsæter, A., Jacob, M., Kopczuk, W., & Telle, K. (2016). *Accounting for business income in measuring top income shares: Integrated accrual approach using individual and firm data from norway*. Working paper 22888, National Bureau of Economic Research.

Alstadsæter, A., Kopczuk, W., & Telle, K. (2014). Are closely-held firms tax shelters? *Tax Policy and the Economy*, *28*, 1–32.

Andreoni, J., Erard, B., & Feinstein, J. S. (1998). Tax compliance. *Journal of Economic Literature*, *36*(2), 818–60.

Bohne, A., & Nimczik, J. S. (2018). *Information frictions and learning dynamics: Evidence from tax avoidance in ecuador*. IZA discussion paper.

Boning, W. C, Guyton, J., Hodge Ronald, II, H., Slemrod, J., & Troiano, U. (February 2018). *Heard it through the grapevine: Direct and network effects of a tax enforcement field experiment*. Working paper 24305, National Bureau of Economic Research.

Chetty, R. (2012). Bounds on elasticities with optimization frictions: A synthesis of micro and macro evidence on labor supply. *Econometrica*, *80*(3), 969–1018.

Chetty, R., Friedman, J. N., Leth-Petersen, S., Nielsen, T. H., & Olsen, T. (2014). Active vs. passive decisions and crowd-out in retirement savings accounts: Evidence from Denmark. *Quarterly Journal of Economics*, *129*(3), 1141–1219.

Chetty, R., Friedman, J. N., Olsen, T., & Pistaferri, L. (2011). Adjustment costs, firm responses, and micro vs. macro labor supply elasticities: evidence from danish tax records. *Quarterly Journal of Economics*, *126*(2), 749–804.

Chetty, R., Looney, A., & Kroft, K. (2009). Salience and taxation: Theory and evidence. *American Economic Review*, *99*(4), 1145–1177.

Currie, J. (2006). The take-up of social benefits. In A. J. Auerbach, D. Card, & J. M. Quigley (Eds.), *Public policy and the income distribution* (pp. 80–148). New York: Russell Sage Foundation.

Dahl, G., Kostøl, A. R., & Mogstad, M. (2014). Family welfare cultures. *Quarterly Journal of Economics*, *4*(1), 1711–1752.

Duflo, E., & Saez, E. (2003). The role of information and social interactions in retirement plan decisions: Evidence from a randomized experiment. *Quarterly Journal of Economics*, *118*(3), 815–842.

Finkelstein, A. (2009). E-ZTax: Tax salience and tax rates. *Quarterly Journal of Economics*, *124*(3), 969–1010.

Frimmel, W., Halla, M., & Paetzold, J. (2018). The intergenerational causal effect of tax evasion: Evidence from the commuter tax allowance in Austria. *Journal of the European Economic Association*. https://doi.org/10.1093/jeea/jvy033.

Gordon, R., & Li, W. (2009). Tax structures in developing countries: Many puzzles and a possible explanation. *Journal of Public Economics*, *93*(7–8), 855–866.

Jones, D. (2012). Inertia and overwithholding: Explaining the prevalence of income tax refunds. *American Economic Journal: Economic Policy*, *4*(1), 158–85.

Kleven, H. J., Knudsen, M., Kreiner, C. T., Pedersen, S., & Saez, E. (2011). Unwilling or unable to cheat? Evidence from a randomized tax audit experiment in Denmark. *Econometrica*, *79*(3), 651–92.

Kleven, H. J., & Waseem, M. (2013). Using notches to uncover optimization frictions and structural elasticities: Theory and evidence from Pakistan. *Quarterly Journal of Economics*, *128*(2), 669–723.

Kopczuk, W., & Slemrod, J. (2006). Putting firms into optimal tax theory. *American Economic Review Papers and Proceedings*, *96*(2), 130–134.

Lindhe, T., Södersten, J., & Öberg, A. (2004). Economic effects of taxing different organizational forms under the nordic dual income tax. *International Tax and Public Finance*, *11*(4), 469–485.

Luttmer, E. F. P., & Singhal, M. (2014). Tax morale. *Journal of Economic Perspectives*, *28*(4), 149–68.

Madrian, B. C., & Shea, D. F. (2001). The power of suggestion: Inertia in 401(k) participation and savings behavior. *Quarterly Journal of Economics*, *116*(4), 1149–1187.

Paetzold, J., & Winner, H. (2016). Taking the high road? Compliance with commuter tax allowances and the role of evasion spillovers. *Journal of Public Economics*, *143*, 1–14.

Pomeranz, D. (2015). No taxation without information: Deterrence and self-enforcement in the value added tax. *American Economic Review*, *105*(8–9), 2539–69.

Saez, E. (2010). Do taxpayers bunch at kink points? *American Economic Journal: Economic Policy*, *2*(3), 180–212.

Slemrod, J., & Yitzhaki, S. (2002). Tax avoidance, evasion and administration. In A. J. Auerbach & M. S. Feldstein (Eds.), *Handbook of public economics* (Vol. 3). Amsterdam/New York: Elsevier/North Holland.

Sørensen, P. (2005). Neutral taxation of shareholder income. *International Tax and Public Finance*, *12*(6), 777–801.

Thoresen, T. O., & Alstadsæter, A. (2010). Shifts in organizational form under a dual income tax system. *FinanzArchiv: Public Finance Analysis*, *66*(4), 384–418.