

NBER WORKING PAPER SERIES

SOCIAL NETWORKS AND TAX AVOIDANCE:
EVIDENCE FROM A WELL-DEFINED NORWEGIAN TAX SHELTER

Annette Alstadsæter
Wojciech Kopczuk
Kjetil Telle

Working Paper 25191
<http://www.nber.org/papers/w25191>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
October 2018

We are grateful to Jim Hines, Henrik Kleven, Juliana Londono-Velez, Aureo de Paula, Marzena Rostek, Karl Scholz, Johannes Spinnewijn, Thor Olav Thoresen and seminar participants at many universities and conferences. All errors are naturally ours. Support from the Research Council of Norway is gratefully acknowledged. All authors were employees of Statistics Norway while working on this paper. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2018 by Annette Alstadsæter, Wojciech Kopczuk, and Kjetil Telle. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Social Networks and Tax Avoidance: Evidence from a Well-Defined Norwegian Tax Shelter
Annette Alstadsæter, Wojciech Kopczuk, and Kjetil Telle
NBER Working Paper No. 25191
October 2018
JEL No. D22,D23,H25,H26,H32

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 2006 reform. 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 ability 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 take up 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.

Annette Alstadsæter
Norwegian University of Life Sciences
annette.alstadsater@nmbu.no

Wojciech Kopczuk
Columbia University
420 West 118th Street, Rm. 1022 IAB
MC 3323
New York, NY 10027
and NBER
wk2110@columbia.edu

Kjetil Telle
Statistics Norway
Research Department
Box 8131 Dep
0033 Oslo, Norway
tel@ssb.no

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 work also recognizes that empirical behavioral responses are sometimes puzzlingly small and inconsistent across different contexts — 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, Chetty et al. (2014) show that only a small number of taxpayers makes active saving decisions, and a large literature shows the importance of default options in retirement programs (Madrian and Shea, 2001, and the literature that followed)¹ and imperfect take-up of social benefits (Currie, 2006).

The objective of this paper is to provide empirical evidence regarding a particular class of explanations for tax-motivated behavior. We are interested in testing whether the decision to pursue tax-minimizing behavior spreads within social networks, and in particular, family networks.² We find evidence that this is so. Our research design leverages the presence of a discontinuity in taxpayer’s eligibility for setting up a particular (legal) tax shelter. 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 on the 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 change 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.

One potential direction for reconciling theory and evidence on non-compliance is to provide a more realistic characterization of the economic environment. This is what recent work on tax compliance has done by pointing out to the importance of third-party reporting (Kleven et al., 2011), attachment to the financial sector (Gordon and Li, 2009) or arms-length transactions (Kopczuk and Slemrod, 2006) as factors limiting the extent of evasion. While this strand of work suggests that administrative environment plays 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 evidence so far on the importance

¹Also, see Duflo and Saez (2003) for evidence about the role of social interactions in retirement planning, and Dahl et al. (2014) for social interactions in the take-up of welfare benefits.

²When analyzing how behavior spreads within social networks, there is a variety of networks to choose from, such as family, colleagues, schools, sports, church attendance, shareholders, accountants, board members, neighbourhoods, etc. In this paper, we focus on a particular and natural choice of an exogenous network: family members.

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.

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 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 evidence about relevance 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.

Empirical work on tax avoidance and evasion faces a lot of challenges due to difficulty in observing the outcomes (pursuing and extent of tax avoidance/evasion). We can sidestep this problem due to 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. The existence of this well-defined measure of tax avoidance allows us to overcome measurement issues and focus instead on determinants of pursuing of this type behavior. The particular reform that we analyze introduced discontinuity in opportunities to set up this shelter that we exploit through the regression discontinuity design. Our data allows for constructing full family networks that we can then use to identify spillover effects.

To our knowledge, the only other (and concurrent to this 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. They study father-child pairs and find that tax evasion runs in the family. Where Frimmel et al. 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 section 3 we describe our data. Section 4 is devoted to the empirical strategy. Our main reduced form results are in Section 5, 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. In Section 6 we show the corroborating effect of timing and conclude that the take up in the network accelerates overall take up. 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.³

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 percent on all personal shareholders’ income from shares, including both dividends and capital gains.⁴ During the transition, the tax on realized capital gains on shares for corporate shareholders was removed effective 3/26/2004.⁵ Subsequently, taxes on dividends for personal shareholders were introduced effective as of 1/1/2006. Hence, pre-2004, corporate capital gains and dividends were treated in the same way as individual ones; as of March 2004 corporate capital gains were privileged relative to individual ones and in 2006 both capital gains and dividends on the corporate level were treated favorably relative to individual ones. As the result, 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 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

³This 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).

⁴The 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. Sørensen (2005) and Alstadsæter and Fjærli (2009) provide more information on the Shareholder Income tax.

⁵This exemption of capital gains from taxation was implemented without warning on March 26, 2004. 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.

et al., 2014, for evidence of this type of tax planning), and arbitrage between personal and corporate taxation. Following the reform, exempting capital gains and dividends for corporate owners creates an additional and very important advantage due to deferral of taxation. From our point of view, the key point is that individuals have a stronger incentive to own firms indirectly after the reform; we of course find prima facie evidence of it being so due to the massive 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 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 that would otherwise be due. 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 his 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 report sent to the company register by Dec. 31, 2005. It turned out that this transition rule was restrictive and relatively few shareholders could utilize it, and 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 of the Transition Rule E, the 90% threshold was reduced to 10% — to qualify, the holding company has to hold at least 10% ownership stake in the transferred corporation. Taking advantage of the Rule required that all shares that an individual owns must be transferred, and that the compensation is in the form of shares in the holding corporation. The transfer or foundation must be reported to the Corporate Register by Dec. 31, 2005. 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.

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 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 1/1/2006. 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 E rule provided an opportunity to convert to indirect ownership without the tax. The main purpose of holding companies set up under Transition E rule 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.⁶

⁶Statistics Norway identified new holding corporations set up under the transition rule E by an existing sector

Figure 1 shows the timing of adoption of firms that we classified as being set up under the transition rule. As the figure demonstrates, setting up holding companies was not uniform over time. 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 Register⁷ contains records of every shareholder of every Norwegian corporation for 2004-2008. Relying on the shareholders register, we are able to identify for each person and firm their holdings and, correspondingly, for each firm its owners, whether they are corporate or individual. Relying on this dataset, we select individual shareholders 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.⁸ In particular, 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. 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 resulting sample consists of 318,818 personal shareholders at the end of 2004.

Using other register information we are able to link other 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

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 he 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 NAC E-code 65.238, and during 2005, 16483 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.

⁷More on this data source in Statistics Norway, Share Statistics, 2006, http://www.ssb.no/english/subjects/11/01/aksjer_en/.

⁸We 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 individual channels is one possible explanation for significant take up for individuals who were not eligible in 2004.

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: 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. In other words, the family network of an individual is assumed to include his direct (parents, children, siblings, spouse) relatives and direct relatives of the direct relatives.⁹

4 Basic Framework

Our core econometric framework consists of two equations — for the individual i and the network member j that may potentially affect her

$$E_j = \alpha_j + \beta_j \cdot X_j + \gamma_j Z_j + \mu_j \quad (1)$$

$$E_i = \alpha_i + \beta_i \cdot X_j + \gamma_i Z_i + \mu_i \quad (2)$$

Equation 1 relates sheltering decision E to one’s own incentives represented by X_j , and controlling for own characteristics Z_j — this is the first stage. Equation 2 relates sheltering decision of an individual to his own characteristics Z_i and some characteristics X_j of the network member. In most cases we will use a dummy variable for setting up a tax shelter as the dependent variable and estimate specifications as linear probability model. 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.

In the 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 (under assumptions leading up to it), provides a way to guide the interpretation of their 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

⁹We will usually exclude the spouse from the network because the relevant unit of observation may be a household rather than an individual but the results are robust to this restriction.

mentioned before, we implement a regression discontinuity design that relies on the feature of the reform that required a newly set-up 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 with no additional adjustments while 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. Crucially, we have information about the exact number of shares that an individual owns in 2004 as well as the total number of 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. The first stage corresponds to equation 1. While, as we stated in Remark 2, 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. The first stage 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. Figure 2 suggests that individuals with ownership stake smaller than 10% as of 2010 were less likely to ultimately take advantage of the Transition E rule. In particular, there is evidence of a (statistically significant) discontinuity at the 10% threshold that we will exploit in our regression discontinuity design. We will return to details involved in constructing this figure below.

Equation 2 is the second stage. The decision of an individual is related to incentives (X_j) of his network member. 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.

There are many characteristics of individual 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. Assuming that the assumptions for validity of RD hold, controlling for such additional characteristics is not necessary for obtaining unbiased estimates of the effect of X_j .

We will investigate heterogeneity of the response by splitting the sample along some dimensions (such as history of reliance on 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 is the effect of *any* interactions. As long as assumptions for the validity of RD hold, the estimates reflect though the average effect of exposure to 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.

4.1 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 more than 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, using instead 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 and it turns out to correspond to a small sample size around the 10% threshold (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 (i.e., 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 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* level. That is, our outcome variable E_j represents whether an individual adopted any e-firm for any of her shareholdings. Hence, in our first stage regression, 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 (second stage), we want to retain the same structure on the treatment level. The discontinuity is defined on the level of the shareholding of the network member, (j, k) . The corresponding treatment affects all individuals i who are related to the person 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 the result, a single shareholding k of person j gives rise to multiple observations for all individuals who are in the same network as j . Conversely, person i gives rise to multiple observations corresponding to links with shareholdings of all her network members. 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 . We address the corresponding dependence by two-way clustering of standard errors on i and j level. The likelihood

of including multiple observations for an individual i should vanish as the window around the threshold goes to zero but the likelihood of including multiple observations corresponding to the network member j does not vanish since the same shareholding with a given value S_j corresponds to multiple observations. Hence, correcting for the dependence on the level of a network member is important even asymptotically.

4.2 Interpretation of the estimated coefficients

As we discussed, the unit of observation for our analysis is the (directed) network relationship and our baseline specifications 1 and 2 control for X_j only, rather than for characteristics of all individuals in the family network. In general, individuals may be influenced by many different network members

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

Suppose that $X_j \perp X_{-j}$ (i.e., that in our RD 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 $\beta_i = E \left[\frac{\partial g}{\partial X_j} \middle| Z_i \right]$ — 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 relationships (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.

5 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. Recall Figure 2 that 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. $[0.10, 0.11)$) for a subsample of individuals that we describe below. As discussed before, 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. The adoption of the transition E firm here is defined as having set up an E-firm corresponding to any shareholding (rather than for the one associated with the observation). 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 below. First, there is an appearance of discontinuity at the threshold but there is also enough variation in the data overall that careful statistical testing is necessary to establish its presence. Second, it is a “fuzzy” RD 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 effects are very likely to be heterogeneous across different groups 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 share above which around 20% of the population adopts (and the data is considerably noisier). There is also some evidence that adoption may be declining at higher ownership levels, far from the threshold. 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.

Smoothness of the distribution

Figure 2 and our regression discontinuity analysis that follows does not utilize the full sample. The full distribution is “lumpy” in many places including the threshold itself,¹⁰ and in order to convincingly employ the regression discontinuity approach, the distribution of individual characteristics should be smooth around the threshold. A closer inspection reveals that bunching is very systematic — it occurs at points that correspond to splitting shares of the firm as exact fractions. 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.¹¹

Hence, we proceed by eliminating exact fractions from the sample as explained in the appendix. The outcome of this trimming procedure in terms of the number of observations is shown on Figure 3. The procedure is necessary to apply the regression discontinuity approach. It introduces a natural limitation for 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

¹⁰Appendix Figure A1 shows the log of the number of observations in the full sample, by 0.1% bins. In the Appendix we also show the analogue of Figure 2 in Figures A2 and A3 using the full sample with no adjustments. Figure A2 shows the likelihood of adopting the transition E rule for the particular shareholding, while Figure A3 shows the likelihood of adopting for any shareholding of the corresponding individual.

¹¹Beyond the number of observations, we found indication of non-smoothness for individual characteristics that we investigate below for the restricted sample — 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 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.

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 RD design.

We proceed with the subsample defined in that way in what follows. Figure 2 that we have discussed before is based on this sample. Appendix Figure A4 shows the same on shareholding level. In this sample, a number of characteristics - age, number of owners in a firm, gender and log of the initial capital all change fairly continuously around the 10% mark for our restricted sample (see Appendix Figures A5, A6, A7 and A8). The conditional mean of age and gender is noisy but this is so mostly away from the threshold. An inspection of the density demonstrates that it 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 the consequence, there are many examples of firm 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 $1/6$ th — some of the irregularities are likely associated with that. 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.

The effect of 10% rule on individual adoption

Figure 4 zooms in to the smaller region $(0, 0.30)$ (using adoption defined on the individual level as in Figure 2), that more clearly displays the 10% threshold (with bins corresponding to 0.01 intervals). It also show 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 2. The baseline regression is a linear probability model of the dummy for taking up an E-firm 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 shareholding level and (of our main interest) the effect on shareholder level. The results indicate that the discontinuity is present and very statistically significant both if adoption is defined for shareholding and when it is defined for an individual.¹² In particular, our preferred estimates (on shareholder level, using larger windows around the threshold) indicate that individuals just

¹²In 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 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.

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.

In the following panel we pursue basic robustness checks by including a set of individual controls that we investigate before — 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 transition E rule is robust, we are primarily interested in it in order to use it as “first stage” and trace its implications in the network. We are more likely to be able to statistically trace such responses if the first stage effect is strong. Hence, 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 also Appendix Figure A14), and there is no discernible effect for the remainder of subsample (Appendix Figure A15). The third panel of Table 2 tests formally that the robust effect is there for those with dividends in 2004 and that the magnitude of the effect is much 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 (consistently with the Appendix Figure A15, there is no robust evidence of an effect for those with no dividends; the results are not reported).

The following panel 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. Figure 5 shows the likelihood of adoption on an individual level for that sample, and we see a jump at the threshold. The last panel of Table 2 shows that the effects are large and robust.¹³

Our final piece of evidence on the individual level relates to how e-firms are set up. 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. Figures 6 and 7 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.

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

¹³Appendix Figure A16 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.

those below 10% share who are otherwise similar do not end up taking up an e-firm. The e-firms stimulated by the 10% rule are single owner ones, with no evidence of crowdout of multiple-owner e-firms, suggesting that setting up an e-firm with others was not the alternative entertained by the population complying with the treatment. 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.

Network effects

We now turn to the network level analysis by analyzing take up of E-firms of an individual (i) as a function of ownership 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 family network member is a separate observation affecting the impacted individual. We limit attention to 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).

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. Figure A17 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 figure 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. Beyond the necessity of imposing this restriction to exploit discontinuity for identification purposes, it 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.¹⁴

Figure 9 shows the discontinuity-based evidence of adoption elsewhere in the network on individual adoption, and top panel of Table 3 shows the corresponding estimates. The estimates of the discontinuity are generally significant and reasonably stable as the window around 10% is adjusted. Zooming in on individuals who own at least 10% ownership share in any firm strengthens the results. 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

¹⁴Appendix Figure A9 shows the number of observations and Appendix Figures A10, A11, A12, A13, show age, number of owners, gender and capital of treated firms with similar conclusions as in the case of individual-level analysis: all these variables appear to sail smoothly through the 10% mark.

and hence may be more strongly affected. At the same, 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. The effects 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.

Following up on our previous discussion, we further split the sample by whether the network member received dividends in 2004. Figure 10 shows suggestive evidence of discontinuity when the family member received dividends — this is the group for which the first stage was strong. At the same time, Figure 11 shows no evidence of a discontinuity for the rest of the sample. The bottom two panels of Table 3 show the corresponding estimates that confirm these impressions. 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. 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, consistently with the impression from the figures). The results for those with family members who have not received dividends are smaller and generally insignificant.¹⁵

That the results for those with family members who have not received dividends are generally insignificant 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, one should not expect that their family members are affected.¹⁶

In Table 4 we split the sample in additional ways. First, we look at those with family members with dividends in 2004 and firms with over 1000 shares. In this group, the first stage 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 much stronger statistical evidence for those who did not receive dividends themselves than for those who did. The coefficients for those without dividends are larger in absolute value despite the lower base and hence are also economically very significant – for example, the estimated effect of 0.04 for the flexible specification corresponds to roughly doubling the take up. A rough taxonomy of the results is that individuals with most to gain (those with dividends) are most responsive to the 10% threshold incentive, but they stimulate take up by individuals who have less potential to gain (those without past dividends) and so perhaps least informed otherwise.

¹⁵Appendix 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.

¹⁶As 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.

6 Effect on timing

In order to further substantiate the presence of network effects, we investigate a different dimension — timing. We will return below to regression discontinuity based evidence but begin by providing suggestive evidence of a strong association between timing of adoption in the network and individual adoption.

Figures 12-15 illustrate the dynamics of setting up tax shelters in the data. Figure 12 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 2005 (ie., those who were able to meet the tight eligibility criteria, one needed 90% ownership stake to be transferred to the e-firm). Exposed individuals end up approximately 6 percentage points more likely to eventually set up a tax shelter. Figure 13 represents the same information as the CDF of the timing of adoption conditional on ultimately setting up the tax shelter and it shows that even conditional on ultimate adoption there are differences in timing — those who have exposed family members adopt earlier than others. There are few adoptions during the early period: as also visible in Figure 1, the timing of adoption is heavily concentrated toward the end of the period. Figures 14 and 15 show analogous exercise but this time splitting the sample according to having a family network adopter prior December 1st, 2005 (the date is arbitrarily selected for illustrative purposes; close to half of all of the ultimate E-firms have been set up by that point). As before, individuals in exposed networks are more likely to ultimately set up a holding company and they do so earlier than those who have no family members who already set up an E firm. Perhaps because of the relatively short period of time left before the deadline, it is a bit harder to make the claim that the potential network effect wears off over time, although Figure 15 appears consistent with the two series converging a few days before the end of December. These patterns cannot be interpreted as causal but they do suggest that there is correlation between adoption by network members in the past and 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.

Table 5 shows the result of regressing E-firm adoption dummy on the indicator 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 panels show 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 graphs discussed before, 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.

These results suggest, perhaps unsurprisingly, that the early adoption in the network is correlated with individual economic circumstances. At the same time, expecting that adoption by a family member before June or even November makes the difference for ultimate adoption may be somewhat of a stretch: the effect may be on timing rather than ultimate adoption especially if members of networks adopting early are likely to adopt in general. 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 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's possible that it's hard to distinguish from the effect of other characteristics correlated with early adoption.

This motivates our subsequent strategy that is more careful about timing. We regress the dummy for taking up an e-firm in a particular week on having somebody in the network take up a week before. Under this strategy, the interest is in the timing of adoption rather than the longer term effect. 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.

Figure 16 shows the results for family network based on simple OLS regressions. This is again a linear probability model and this is a hazard-like context. Week 1 corresponds to the last week before 1/1/2006 (and the right-hand side variable is adoption a week before that) and higher numbers correspond to earlier adoption. The figure shows the baseline effect (constant) 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 eight 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.

Figure 17 compares the baseline estimated coefficients to the coefficients based on a specification with the set of demographic and economic controls as before. Contrary to the prior analysis that used wide range of adoption, the controls have very little impact on the magnitude of the effect. This strengthens the possibility that these estimates have causal interpretation and that they don't simply reflect correlation in characteristics. Finally, figure 18 shows the effect of adoption in the last 3 weeks rather than 1 week. These estimates are smoother and more robustly extending further out, but a bit smaller, suggesting that the recent take up has stronger effect than take up further out

These results are suggestive, but cannot be interpreted as causal. We use our regression discontinuity approach to further corroborate the presence of interesting timing effects. In all the following specification we focus on the 0.05 window around the discontinuity point.

In Table 6 we look at the effect on the number of days before 1/1/2006 when the tax shelter was established, with individuals not establishing the shelter assigned a zero value. The OLS specification results in positive, but for the most part insignificant coefficients for the full and dividend samples. Since timing is a censored variable, 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.

In Appendix Tables A2-A5 we focus on results for particular periods for the full sample and those with at least 10% ownership. Focusing on the results for everyone, in Appendix Table A2 we report results from probit specifications. The table contains an estimate of the effect on probability of adopting at the threshold and the effects on log probability to allow for more meaningful comparison across different periods.¹⁷ 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 and inconsistent across specifications. 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 in this period. Appendix Table A3 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. For the full sample, the results indicate that the bulk of the effect is already there there 30 days before the reform. For those with family members who have received dividends, the effect appears to be continuing until

¹⁷We 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.

the deadline.

7 Conclusions

We considered adoption of a legal tax sheltering strategy in Norwegian family networks. 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. This is further corroborated by the evidence from the timing of responses. 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 its their family networks that are affected. At the same time, it is those members of 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 uniformed 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 phenomena that is affected by what others do. 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 knowledge, reduced costs of planning or norms that need to be transmitted within a network. Our evidence of heterogeneous patterns of response points to knowledge and cost as likely channels.

References

- Allingham, Michael G. and Agnar Sandmo**, “Income Tax Evasion: A Theoretical Analysis,” *Journal of Public Economics*, November 1972, 1 (3-4), 323–338.
- Alstadsæter, Annette, Wojciech Kopczuk, and Kjetil Telle**, “Are Closely-Held Firms Tax Shelters?,” *Tax Policy and the Economy*, 2014, 28, 1–32.
- Alstadsæter, Annette**, “The Achilles Heel of the Dual Income Tax: The Norwegian Case,” *Finnish Economic Papers*, 2007, 20 (1), 5–22.
- **and Erik Fjærli**, “Neutral taxation of shareholder income? Corporate responses to an announced dividend tax,” *International Tax and Public Finance*, 2009, 16 (4), 571–604.
- Andreoni, James, Brian Erard, and Jonathan S. Feinstein**, “Tax Compliance,” *Journal of Economic Literature*, June 1998, 36 (2), 818–60.
- Chetty, Raj**, “Bounds on Elasticities with Optimization Frictions: A Synthesis of Micro and Macro Evidence on Labor Supply,” *Econometrica*, May 2012, 80 (3), 969–1018.
- **, Adam Looney, and Kory Kroft**, “Salience and Taxation: Theory and Evidence,” *American Economic Review*, September 2009, 99 (4), 1145–1177.
- **, John Friedman, Tore Olsen, and Luigi Pistaferri**, “Adjustment Costs, Firm Responses, and Micro vs. Macro Labor Supply Elasticities: Evidence from Danish Tax Records,” *Quarterly Journal of Economics*, 2011, 126 (2), 749–804.
- **, John N. Friedman, Soren Leth-Petersen, Torben Helen Nielsen, and Tore Olsen**, “Active vs. Passive Decisions and Crowd-Out in Retirement Savings Accounts: Evidence from Denmark,” *Quarterly Journal of Economics*, 2014, 129 (3), 1141 – 1219.
- Currie, Janet**, “The Take-Up of Social Benefits,” in Alan J. Auerbach, David Card, and John M. Quigley, eds., *Public Policy and the Income Distribution*, New York: Russell Sage Foundation, 2006, pp. 80–148.
- Dahl, Gordon, Andreas R. Kostøl, and Magne Mogstad**, “Family Welfare Cultures,” *Quarterly Journal of Economics*, 2014, 4 (1), 1711–1752.
- Duflo, Esther and Emmanuel Saez**, “The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence from a Randomized Experiment,” *Quarterly Journal of Economics*, August 2003, 118 (3), 815–842.
- Finkelstein, Amy**, “E-ZTax: Tax Salience and Tax Rates,” *Quarterly Journal of Economics*, August 2009, 124 (3), 969–1010.

- Frimmel, Wolfgang, Martin Halla, and Jörg Paetzold**, “The Intergenerational Causal Effect of Tax Evasion: Evidence from the Commuter Tax Allowance in Austria,” *Journal of the European Economic Association*, 2018, *forthcoming*.
- Gordon, Roger and Wei Li**, “Tax Structures in Developing Countries: Many Puzzles and a Possible Explanation,” *Journal of Public Economics*, August 2009, *93* (7-8), 855–866.
- Jones, Damon**, “Inertia and Overwithholding: Explaining the Prevalence of Income Tax Refunds,” *American Economic Journal: Economic Policy*, February 2012, *4* (1), 158–85.
- Kleven, Henrik J. and Mazhar Waseem**, “Using Notches to Uncover Optimization Frictions and Structural Elasticities: Theory and Evidence from Pakistan,” *Quarterly Journal of Economics*, May 2013, *128* (2), 669–723.
- , **Martin Knudsen, Claus Thustrup Kreiner, Soren Pedersen, and Emmanuel Saez**, “Unwilling or Unable to Cheat? Evidence from a Randomized Tax Audit Experiment in Denmark,” *Econometrica*, May 2011, *79* (3), 651–92.
- Kopczuk, Wojciech and Joel Slemrod**, “Putting Firms into Optimal Tax Theory,” *American Economic Review Papers and Proceedings*, May 2006, *96* (2), 130–134.
- Lindhe, Tobias, Jan Södersten, and Ann Öberg**, “Economic Effects of Taxing Different Organizational Forms under the Nordic Dual Income Tax,” *International Tax and Public Finance*, 2004, *11* (4), 469–485.
- Madrian, Brigitte C. and Dennis F. Shea**, “The Power of Suggestion: Inertia in 401(k) Participation and Savings Behavior,” *Quarterly Journal of Economics*, 2001, *116* (4), p1149–1187.
- Pomeranz, Dina**, “No Taxation Without Information: Deterrence and Self-Enforcement in the Value Added Tax,” *American Economic Review*, 2015, *105* (8-9), 2539–69.
- Saez, Emmanuel**, “Do Taxpayers Bunch at Kink Points?,” *American Economic Journal: Economic Policy*, 2010, *2* (3), 180–212.
- Slemrod, Joel and Shlomo Yitzhaki**, “Tax Avoidance, Evasion and Administration,” in Alan J. Auerbach and Martin S. Feldstein, eds., *Handbook of Public Economics*, Vol. 3, Amsterdam; New York: Elsevier/North Holland, 2002.
- Sørensen, Peter**, “Neutral Taxation of Shareholder Income,” *International Tax and Public Finance*, 2005, *12* (6), 777–801.
- Thoresen, Thor O. and Annette Alstadsæter**, “Shifts in organizational form under a dual income tax system,” *FinanzArchiv: Public Finance Analysis*, 2010, *66* (4), 384–418.

Table 1: Summary statistics

	#Obs	Mean	Std.dev	Min	Max
Male	199581	0.747	0.435	0	1
Age	199581	46.473	13.285	0	95
Age<30	199581	0.077	0.267	0	1
Age 30-40	199581	0.241	0.428	0	1
Age 40-50	199581	0.269	0.443	0	1
Age 50-70	199581	0.357	0.479	0	1
Age>70+	199581	0.047	0.211	0	1
Married/cohabiting with children	199581	0.729	0.444	0	1
Number of children	199581	1.871	1.257	0	16
Immigrant	199581	0.037	0.189	0	1
Urban	199581	0.780	0.414	0	1
University	199581	0.341	0.474	0	1
High school	199581	0.550	0.498	0	1
Business education	199581	0.086	0.281	0	1
Self-employment	199581	0.234	0.424	0	1
Net worth, 2004	199581	1657344.000	18800000.000	0	4e+09
Capital income, 2004	199581	366129.500	6175710.000	0	2.19e+09
Total income, 2004	199581	757245.600	6207851.000	-40200000	2.18e+09
Total dividends in 2004	199581	297202.700	6188591.000	0	2.18e+09
E-firm dummy	199581	0.082	0.274	0	1
E-firm before June 17	199581	0.001	0.026	0	1
E-firm before Nov 1	199581	0.010	0.098	0	1
E-firm before Dec 1	199581	0.032	0.175	0	1
E-firm in family network	199581	0.120	0.325	0	1
E-firm in family before June 17	199581	0.001	0.037	0	1
E-firm in family before Nov 1	199581	0.016	0.127	0	1
E-firm in family before Dec 1	199581	0.050	0.218	0	1

Table 2: The effect of crossing 10% ownership on E-firm take up

Range	Adoption for shareholding			Adoption by an individual		
	No controls	Flexible	N	No controls	Flexible	N
[0.095, 0.105]	0.034* (0.019)	0.145*** (0.044)	1246	0.041* (0.022)	0.155*** (0.047)	1246
[0.090, 0.110]	0.027** (0.013)	0.070** (0.029)	2416	0.035** (0.015)	0.073** (0.032)	2416
[0.080, 0.120]	0.047*** (0.009)	0.033* (0.019)	5009	0.039*** (0.010)	0.041* (0.021)	5009
[0.070, 0.130]	0.052*** (0.007)	0.041*** (0.015)	7669	0.043*** (0.008)	0.039** (0.017)	7669
[0.050, 0.150]	0.055*** (0.005)	0.049*** (0.011)	13388	0.050*** (0.006)	0.040*** (0.013)	13388
Everyone with demographic controls						
[0.095, 0.105]	0.025 (0.019)	0.152*** (0.043)	1246	0.034 (0.021)	0.160*** (0.046)	1246
[0.090, 0.110]	0.026** (0.013)	0.062** (0.029)	2416	0.034** (0.014)	0.064** (0.032)	2416
[0.080, 0.120]	0.048*** (0.009)	0.029 (0.019)	5008	0.041*** (0.010)	0.036* (0.021)	5008
[0.070, 0.130]	0.057*** (0.007)	0.036** (0.015)	7664	0.049*** (0.008)	0.034** (0.017)	7664
[0.050, 0.150]	0.063*** (0.005)	0.045*** (0.011)	13378	0.060*** (0.006)	0.036*** (0.013)	13378
Individuals with dividends in 2004						
[0.095, 0.105]	0.109** (0.046)	0.461*** (0.095)	390	0.118** (0.047)	0.434*** (0.098)	390
[0.090, 0.110]	0.086*** (0.031)	0.223*** (0.066)	767	0.098*** (0.032)	0.220*** (0.068)	767
[0.080, 0.120]	0.070*** (0.020)	0.133*** (0.044)	1613	0.066*** (0.021)	0.135*** (0.045)	1613
[0.070, 0.130]	0.088*** (0.016)	0.087** (0.035)	2450	0.085*** (0.017)	0.082** (0.037)	2450
[0.050, 0.150]	0.084*** (0.012)	0.101*** (0.026)	4415	0.083*** (0.012)	0.096*** (0.027)	4415
Dividends in 2004 and firms with at least 1000 shares						
[0.095, 0.105]	0.138*** (0.053)	0.453*** (0.100)	293	0.141** (0.055)	0.425*** (0.103)	293
[0.090, 0.110]	0.121*** (0.039)	0.240*** (0.072)	528	0.144*** (0.041)	0.227*** (0.075)	528
[0.080, 0.120]	0.112*** (0.027)	0.156*** (0.053)	1072	0.112*** (0.028)	0.170*** (0.055)	1072
[0.070, 0.130]	0.125*** (0.022)	0.132*** (0.044)	1614	0.128*** (0.023)	0.130*** (0.046)	1614
[0.050, 0.150]	0.114*** (0.016)	0.139*** (0.034)	2909	0.121*** (0.017)	0.141*** (0.035)	2909

Notes: Flexible specifications include linear terms on each side of the threshold

Table 3: 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	No controls	Flexible	N
[0.095, 0.105]	0.034** (0.017)	-0.005 (0.028)	1935	0.062** (0.027)	-0.005 (0.051)	1117
[0.090, 0.110]	0.029** (0.012)	0.031 (0.023)	3473	0.045** (0.020)	0.057 (0.039)	1940
[0.080, 0.120]	0.019** (0.008)	0.040** (0.016)	6799	0.031** (0.014)	0.065** (0.027)	3652
[0.070, 0.130]	0.015** (0.007)	0.033** (0.014)	10580	0.023** (0.011)	0.056** (0.023)	5632
[0.050, 0.150]	0.009* (0.005)	0.028*** (0.011)	18295	0.015* (0.009)	0.043** (0.018)	9651
Everyone with demographic controls						
[0.095, 0.105]	0.034** (0.017)	-0.002 (0.028)	1931	0.056** (0.027)	-0.007 (0.049)	1116
[0.090, 0.110]	0.028** (0.012)	0.034 (0.023)	3467	0.039* (0.020)	0.053 (0.037)	1939
[0.080, 0.120]	0.018** (0.008)	0.039** (0.016)	6785	0.026* (0.014)	0.057** (0.027)	3647
[0.070, 0.130]	0.013** (0.007)	0.033** (0.014)	10562	0.017 (0.011)	0.048** (0.023)	5626
[0.050, 0.150]	0.008 (0.005)	0.027*** (0.010)	18263	0.011 (0.009)	0.037** (0.017)	9643
Network members with dividends in 2004						
[0.095, 0.105]	0.056** (0.028)	0.050 (0.035)	503	0.114** (0.048)	0.035 (0.063)	285
[0.090, 0.110]	0.058** (0.024)	0.061* (0.034)	924	0.110*** (0.040)	0.098* (0.059)	504
[0.080, 0.120]	0.032** (0.015)	0.081*** (0.029)	2017	0.060** (0.027)	0.148*** (0.050)	1060
[0.070, 0.130]	0.025** (0.012)	0.058** (0.025)	3134	0.046** (0.021)	0.113*** (0.043)	1623
[0.050, 0.150]	0.024*** (0.009)	0.033* (0.019)	5822	0.038** (0.015)	0.066** (0.033)	3056
Network members without dividends in 2004						
[0.095, 0.105]	0.026 (0.020)	-0.036 (0.036)	1432	0.043 (0.031)	-0.047 (0.067)	832
[0.090, 0.110]	0.020 (0.014)	0.018 (0.028)	2549	0.023 (0.023)	0.037 (0.049)	1436
[0.080, 0.120]	0.013 (0.010)	0.026 (0.020)	4782	0.018 (0.017)	0.037 (0.033)	2592
[0.070, 0.130]	0.010 (0.008)	0.025 (0.017)	7446	0.013 (0.014)	0.036 (0.028)	4009
[0.050, 0.150]	0.001 (0.006)	0.027** (0.013)	12473	0.004 (0.010)	0.035* (0.021)	6595

Notes: Flexible specifications include linear terms on each side of the threshold

Table 4: 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	No controls	Flexible	N
Network members with dividends in 2004 and firms over 1000 share						
[0.095, 0.105]	0.074** (0.032)	0.013 (0.036)	380	0.142*** (0.053)	-0.018 (0.062)	225
[0.090, 0.110]	0.065** (0.029)	0.071** (0.036)	662	0.127** (0.049)	0.116* (0.062)	376
[0.080, 0.120]	0.049** (0.020)	0.090*** (0.034)	1365	0.090*** (0.034)	0.161*** (0.058)	747
[0.070, 0.130]	0.035** (0.015)	0.078** (0.030)	2161	0.065** (0.027)	0.142*** (0.052)	1149
[0.050, 0.150]	0.033*** (0.011)	0.045* (0.024)	3990	0.054*** (0.019)	0.090** (0.041)	2153
Network members with dividends in 2004, individuals without dividends						
[0.095, 0.105]	0.069** (0.032)	0.041 (0.037)	328	0.122** (0.051)	0.007 (0.059)	186
[0.090, 0.110]	0.081*** (0.029)	0.059 (0.039)	615	0.141*** (0.049)	0.072 (0.062)	328
[0.080, 0.120]	0.035** (0.017)	0.096*** (0.035)	1377	0.065** (0.031)	0.161*** (0.057)	706
[0.070, 0.130]	0.031** (0.012)	0.068** (0.029)	2180	0.062*** (0.023)	0.114** (0.051)	1085
[0.050, 0.150]	0.028*** (0.009)	0.041* (0.021)	4067	0.048*** (0.017)	0.081** (0.038)	2032
Network members and individuals with dividends in 2004						
[0.095, 0.105]	0.035 (0.049)	0.076 (0.065)	175	0.135 (0.087)	0.134 (0.132)	99
[0.090, 0.110]	0.011 (0.035)	0.069 (0.060)	309	0.050 (0.058)	0.170 (0.107)	176
[0.080, 0.120]	0.031 (0.028)	0.054 (0.047)	640	0.062 (0.046)	0.136* (0.082)	354
[0.070, 0.130]	0.020 (0.024)	0.049 (0.043)	954	0.027 (0.040)	0.125* (0.070)	538
[0.050, 0.150]	0.018 (0.018)	0.030 (0.035)	1755	0.030 (0.028)	0.055 (0.057)	1024

Notes: The table shows the effect of restricting the sample to network members that have over 1000 shares and the decomposition of response by whether the individual had dividends

Table 5: The effect of early adoption on take up

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

Notes: Demographic controls include gender, immigrant dummy, urban dummy, self-employment status, education dummies, business 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

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.

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

Notes: 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

Figure 1: Timing of setting up holding firms under Transition E rule

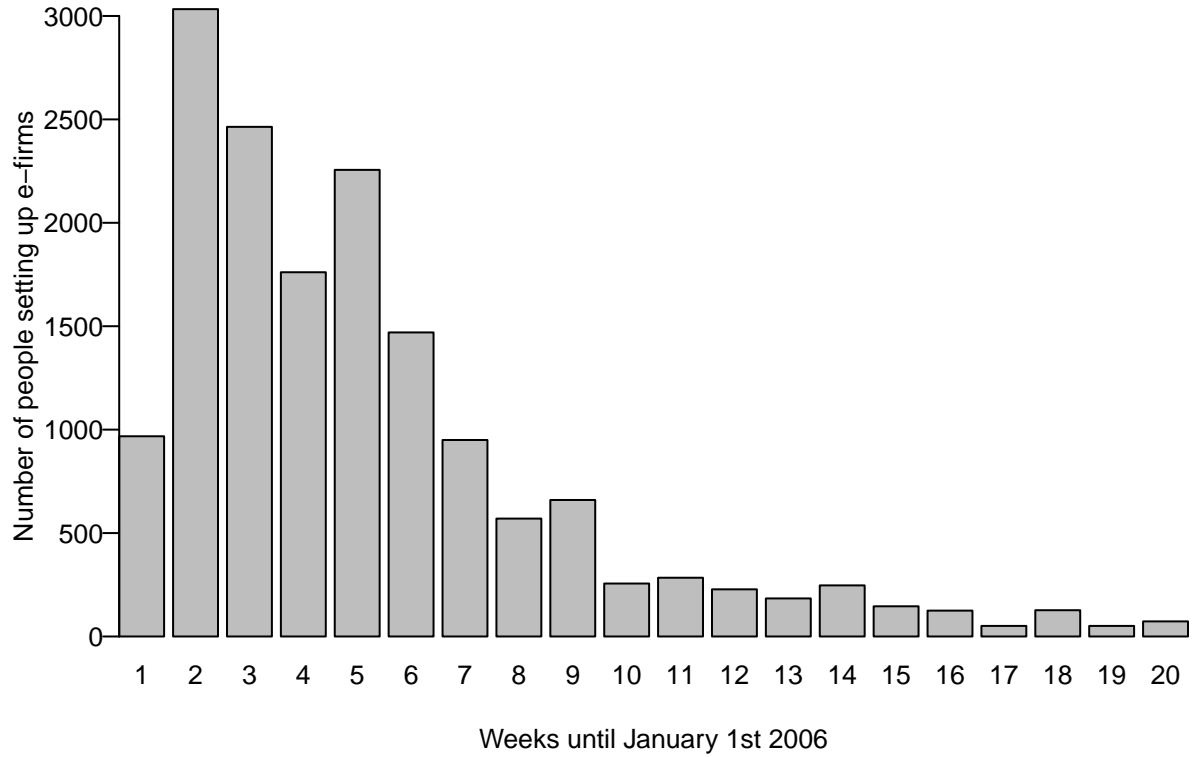


Figure 2: Ownership share in 2004 and probability of ultimate adoption of the Transition E rule

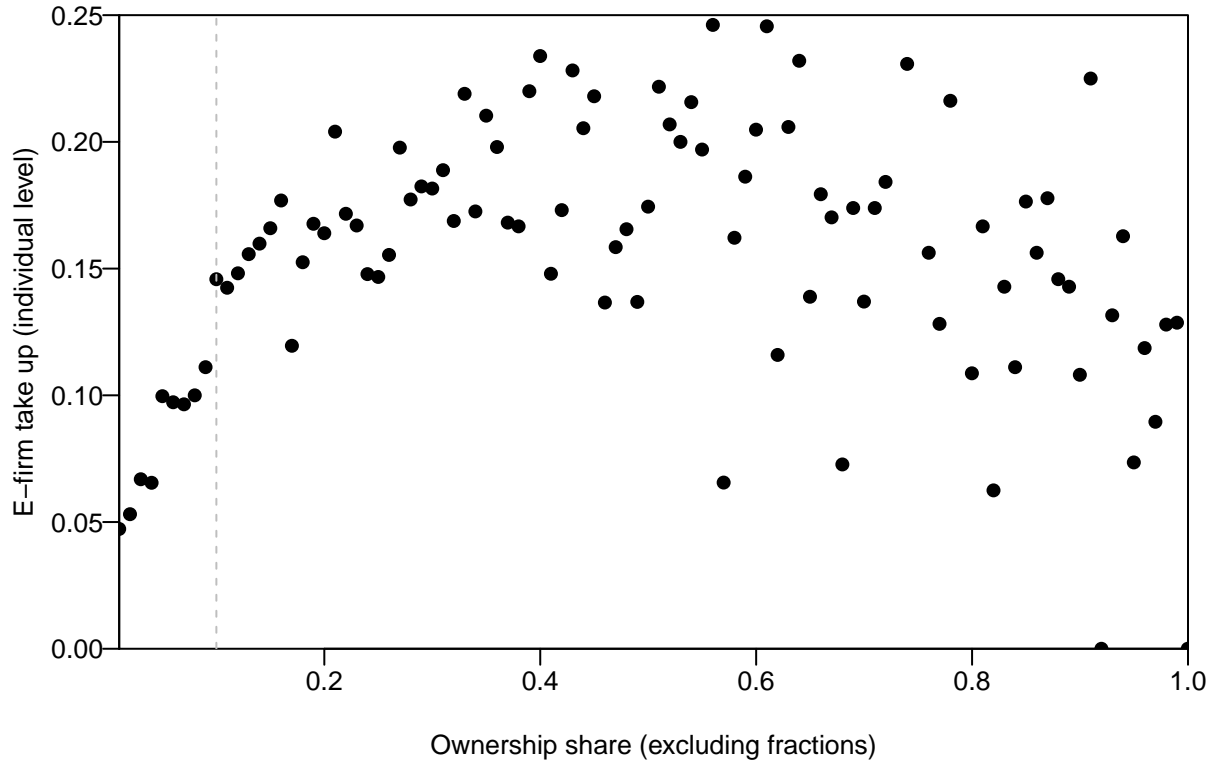


Figure 3: Distribution of ownership shares in 2004 around the 10% threshold

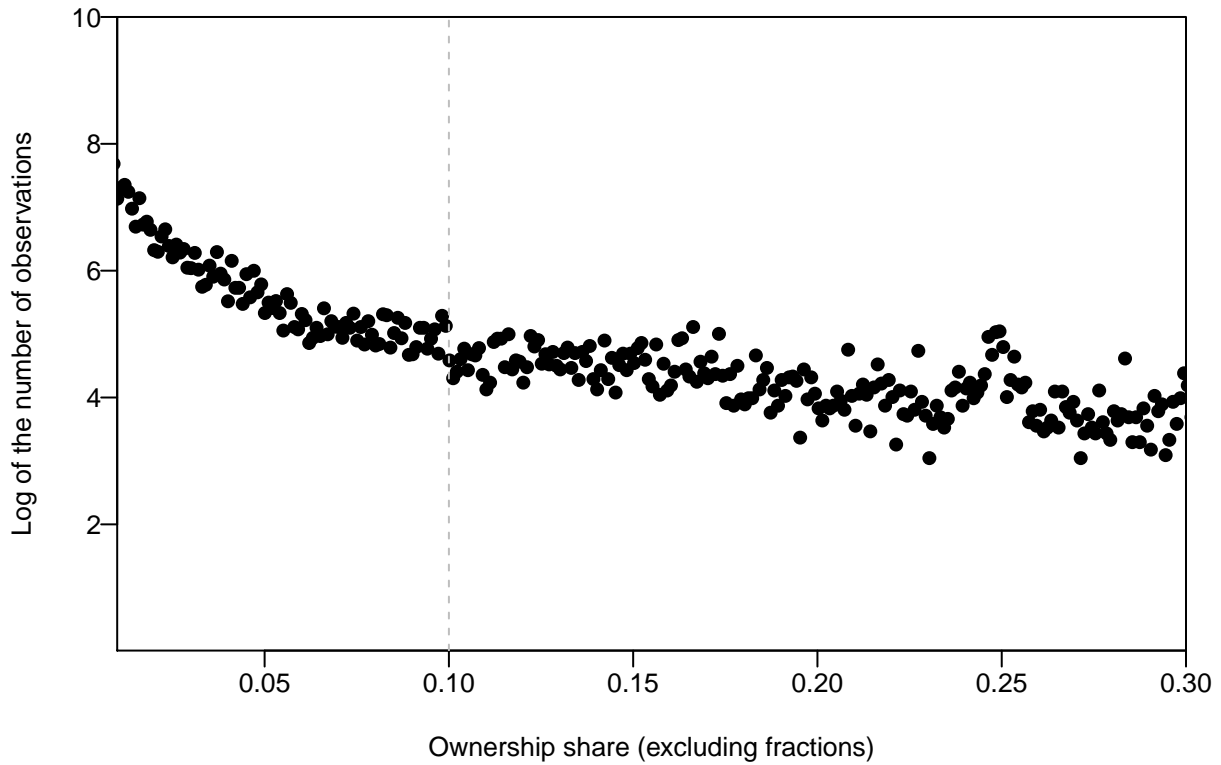


Figure 4: Ownership share in 2004 and probability of ultimate adoption of the Transition E rule — individual level, excluding fractions

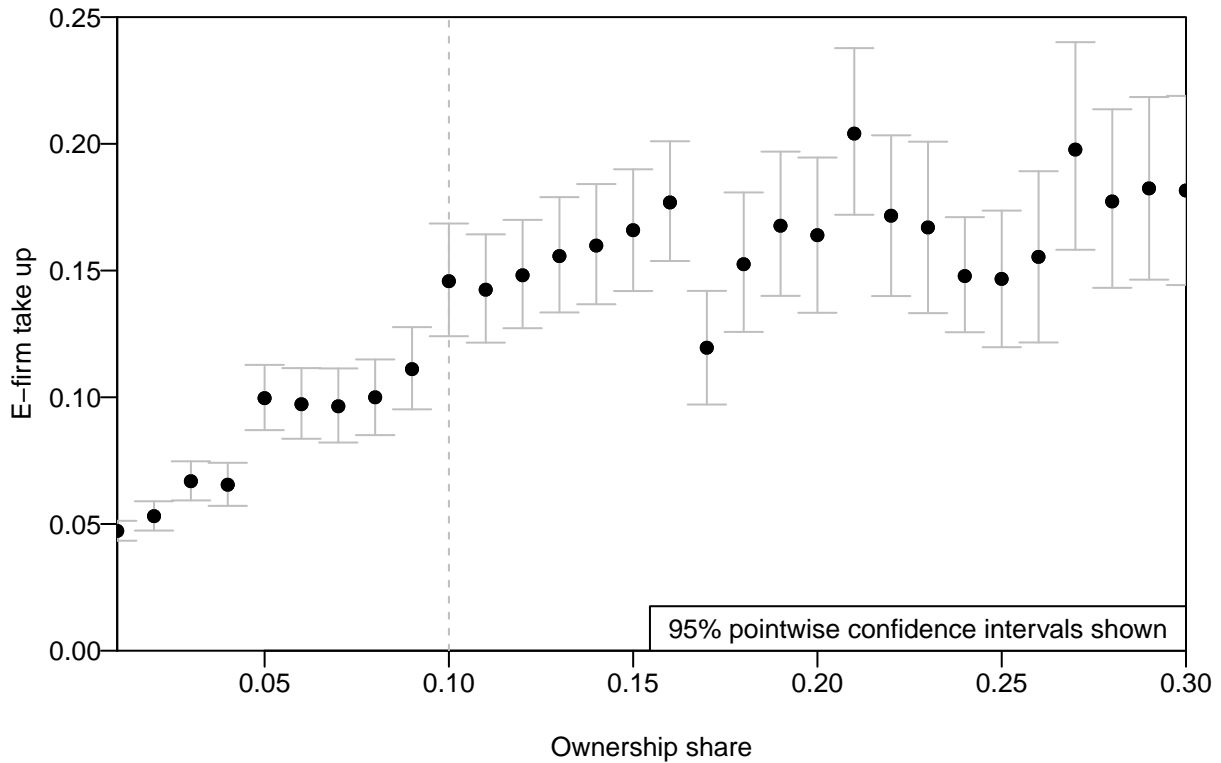


Figure 5: Ownership share in 2004 and probability of ultimate adoption of the Transition E rule — firms with over 1000 shares and dividends realized in 2004

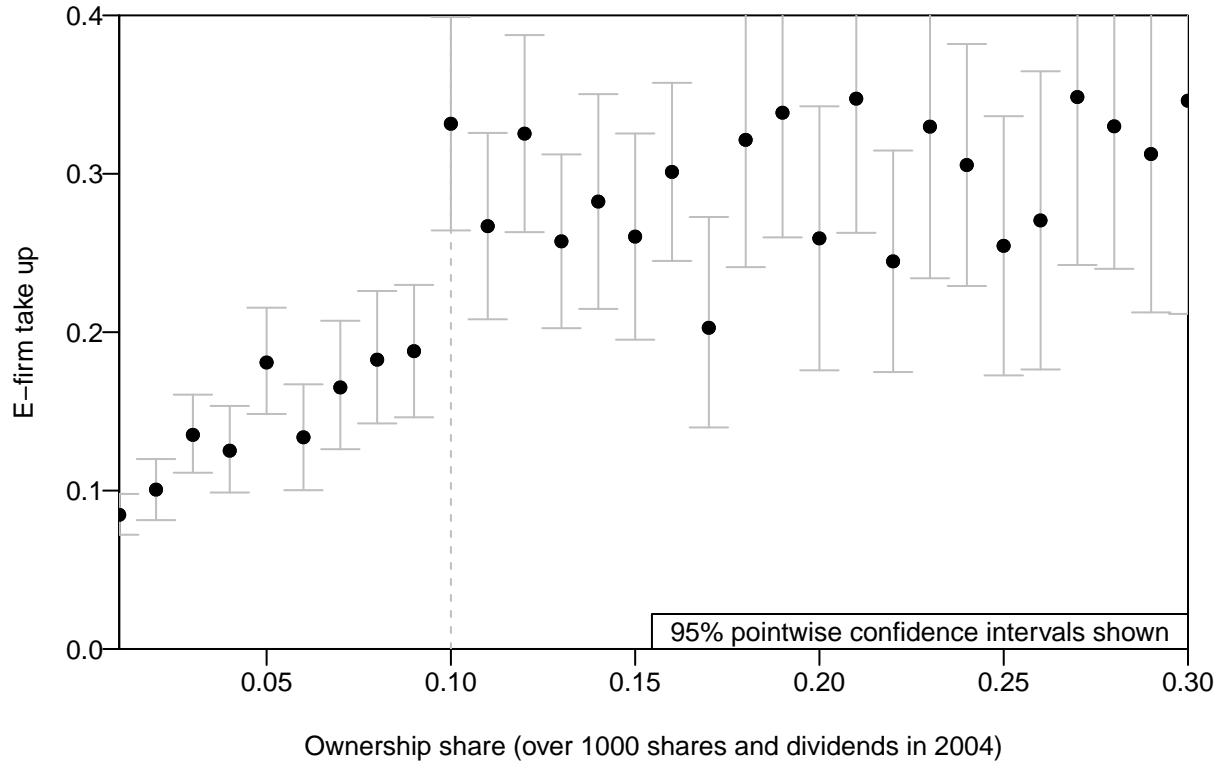


Figure 6: Ownership share in 2004 and single-owner e-firms around the 10% threshold

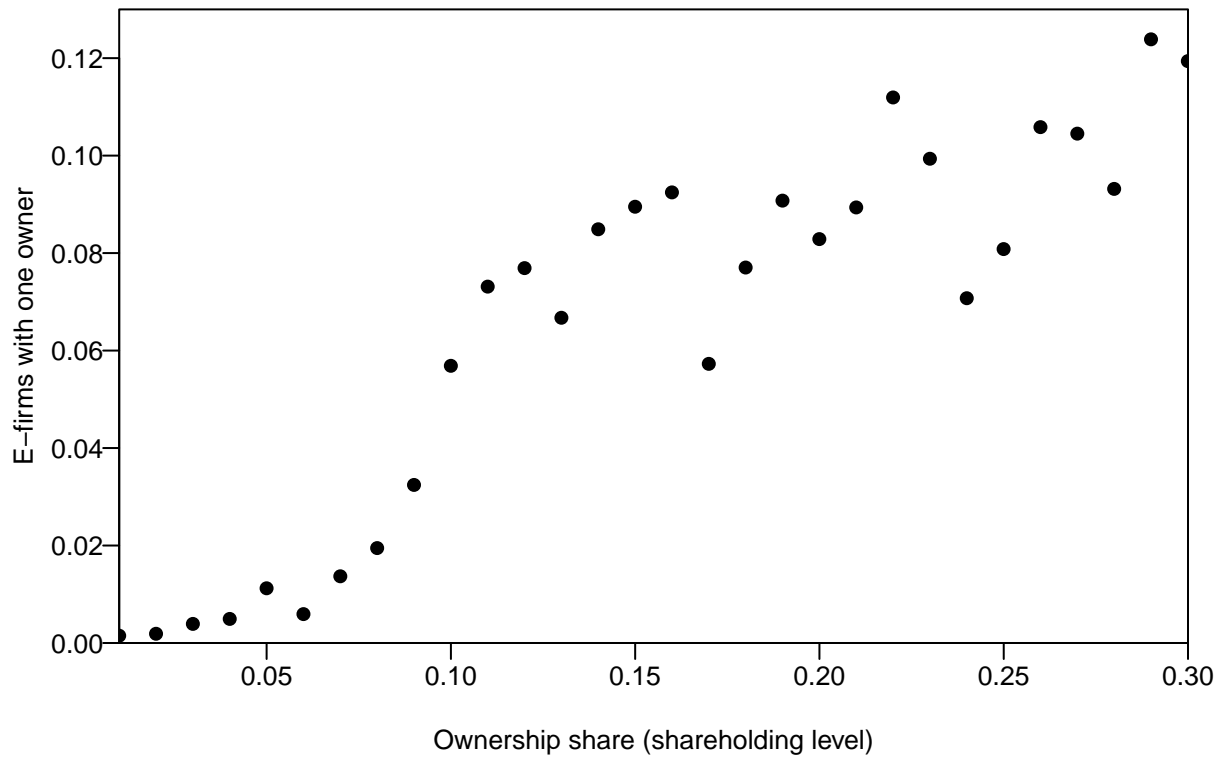


Figure 7: Ownership share in 2004 and multiple owner E-firms, around the 10% threshold

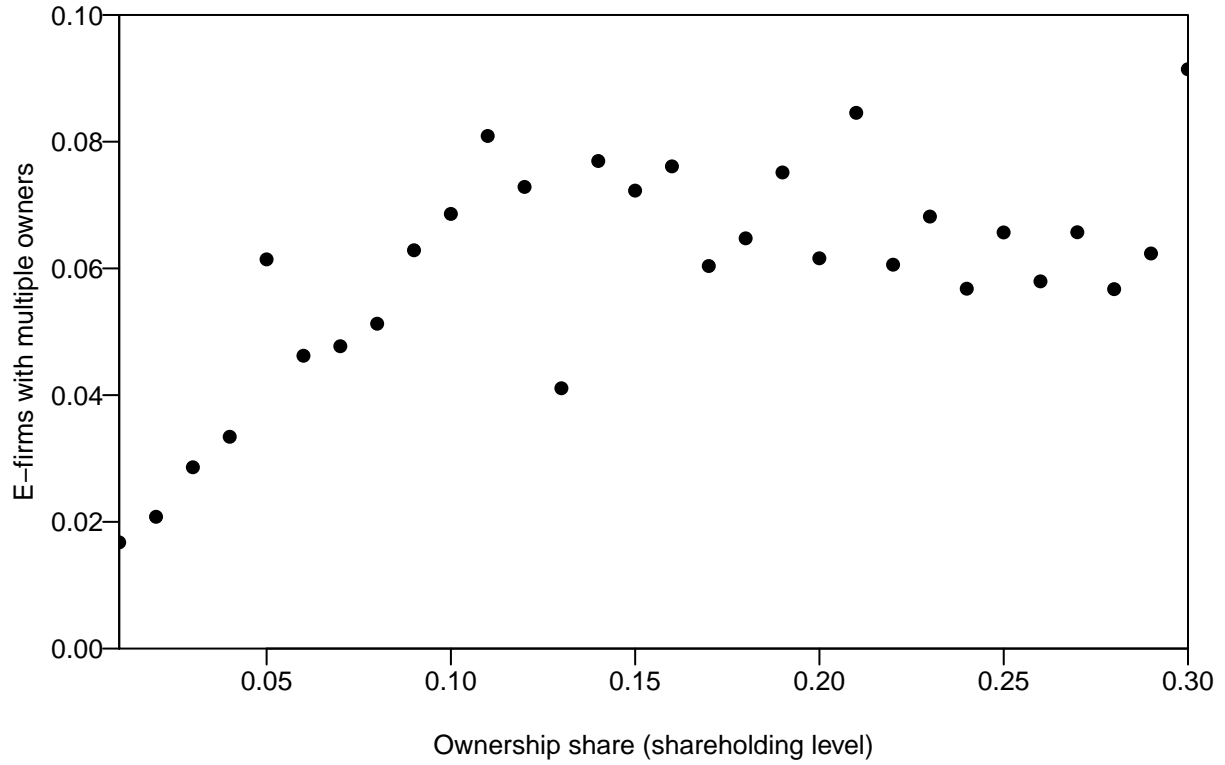


Figure 8: Family member's ownership share in 2004 and number of individuals at least 10% shares — individuals not owning the same firm

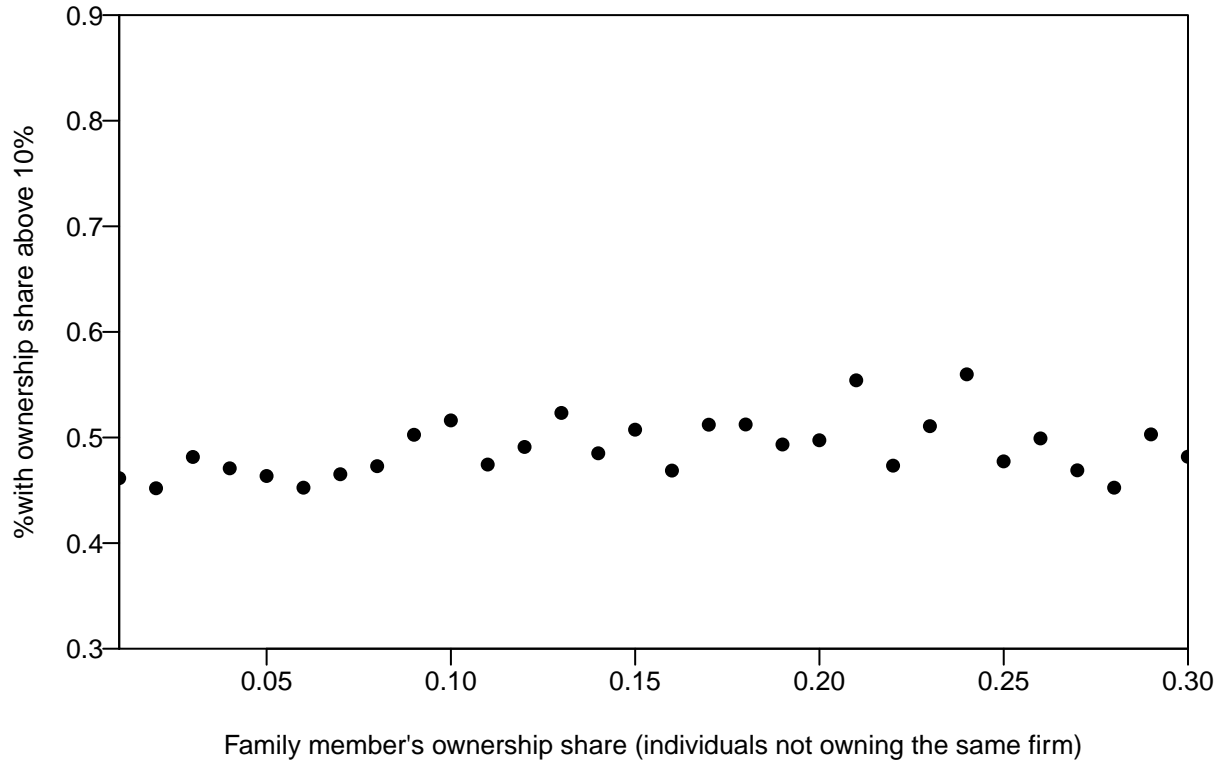


Figure 9: Family member's ownership share in 2004 and ultimate adoption of the Transition E rule

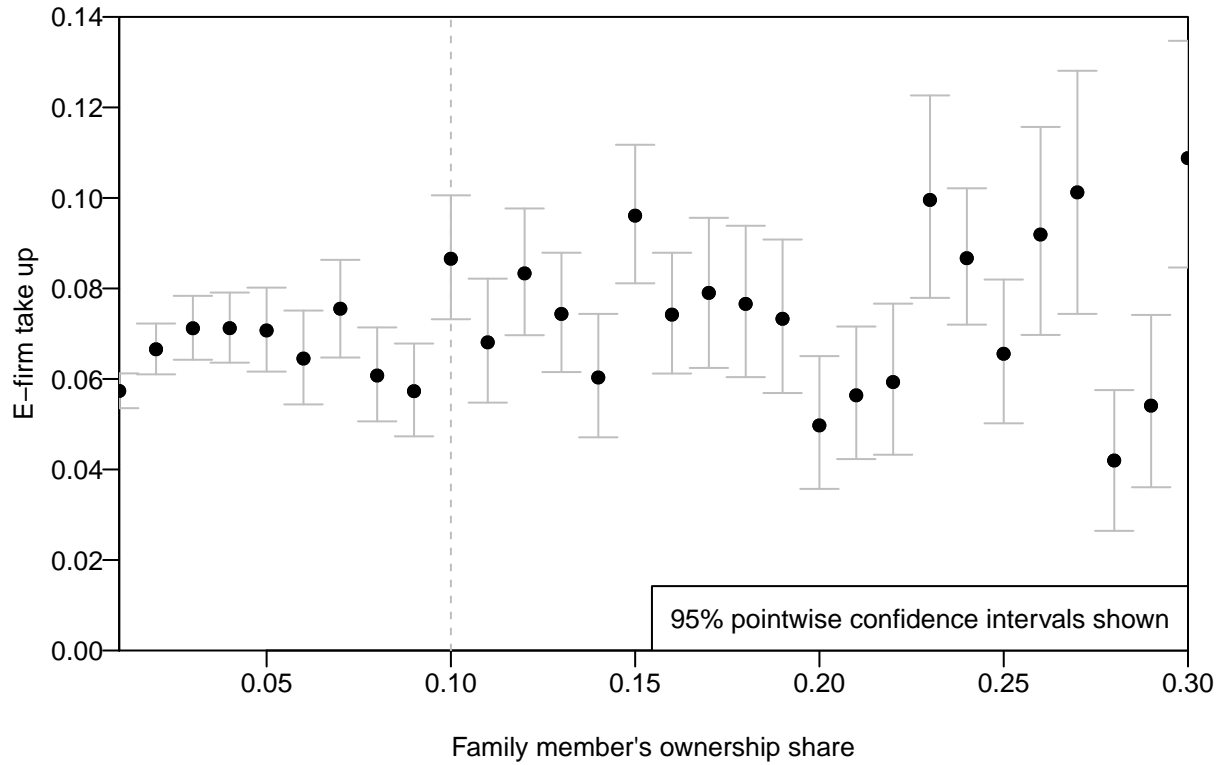


Figure 10: Family member's ownership share in 2004 and ultimate adoption of the Transition E rule — network member's who received dividends in 2004

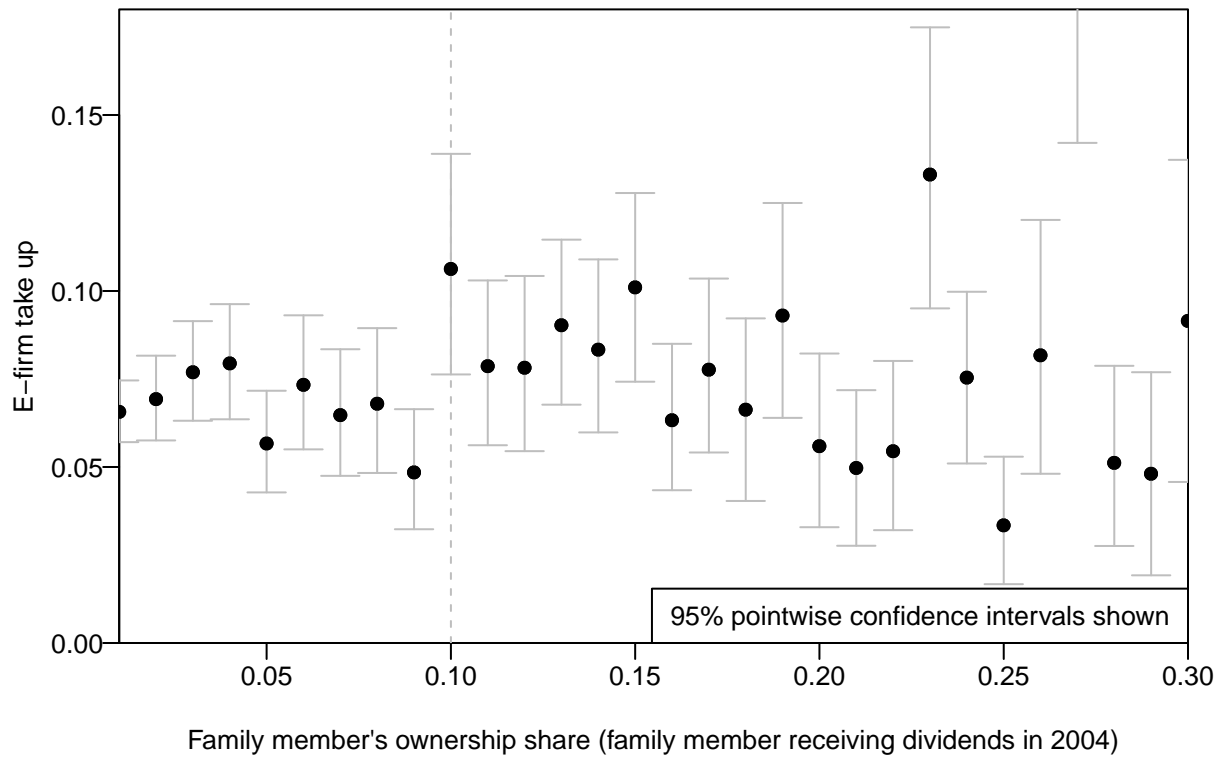


Figure 11: Family member's ownership share in 2004 and ultimate adoption of the Transition E rule — network members who did not receive dividends in 2004

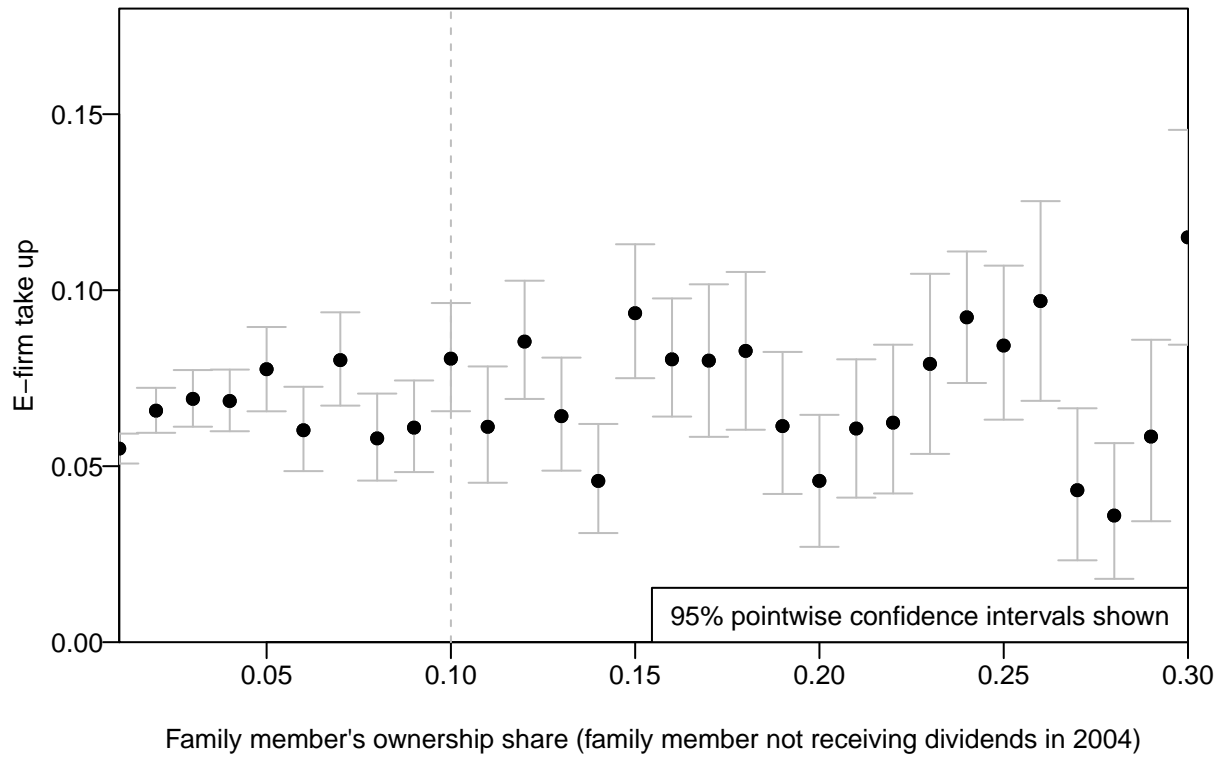


Figure 12: Adoption by individuals with and without a family member adopting before 6/17/2005

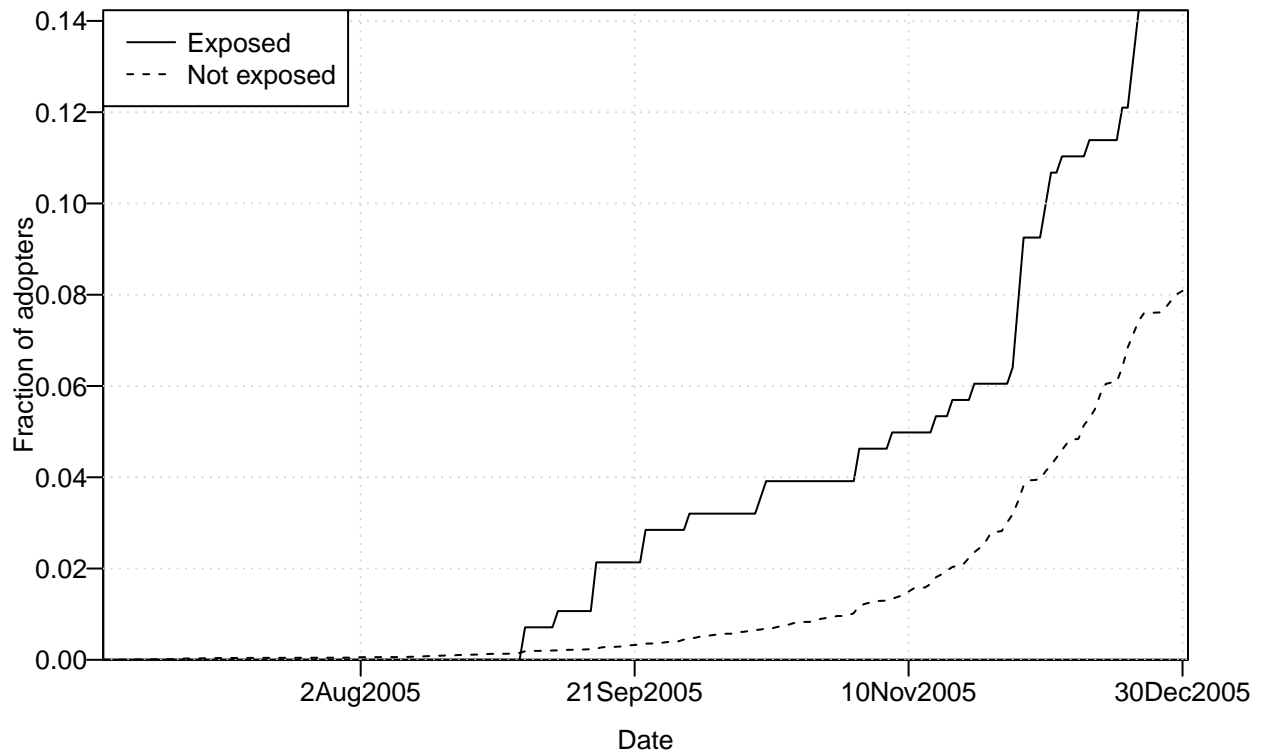


Figure 13: Timing of adoption by individuals with and without a family member adopting before 6/17/2005 (conditional on ultimately adopting)

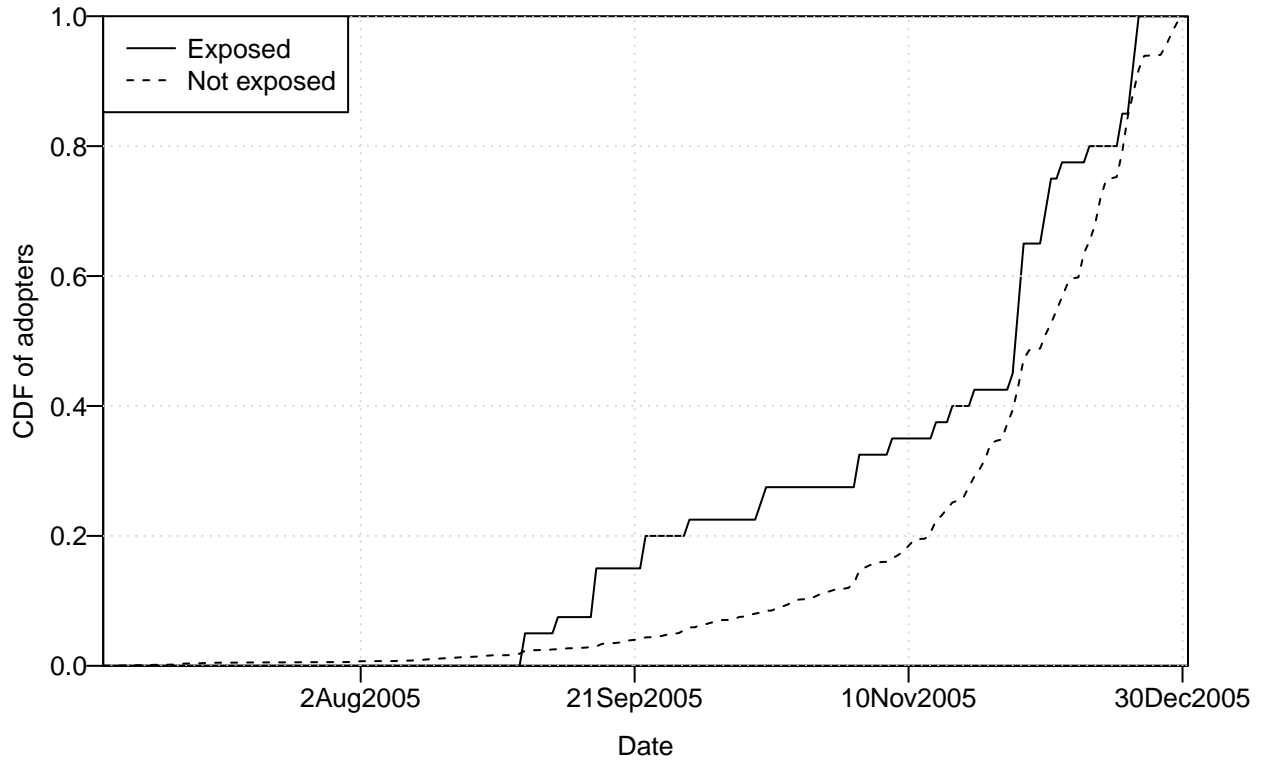


Figure 14: Adoption by individuals with and without family member adopting before 12/1/2005

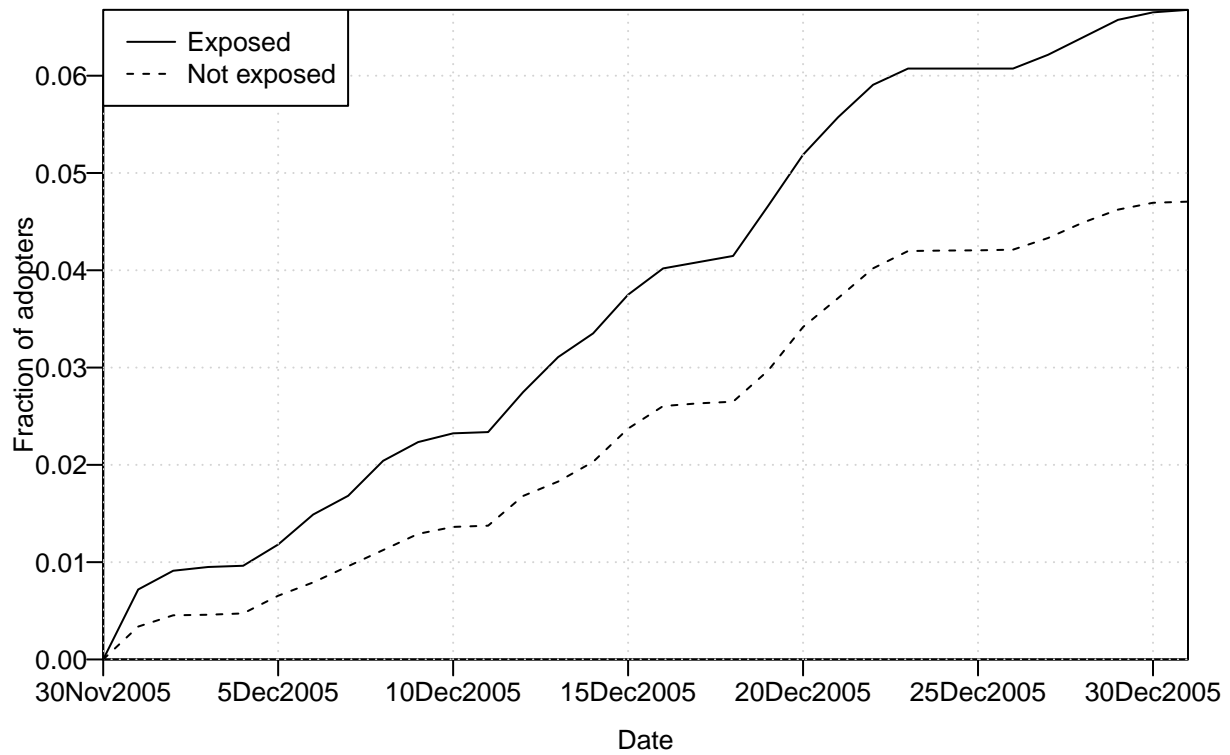


Figure 15: Timing of adoption by individuals with and without family member adopting before 12/1/2005 (conditional on ultimately adopting)

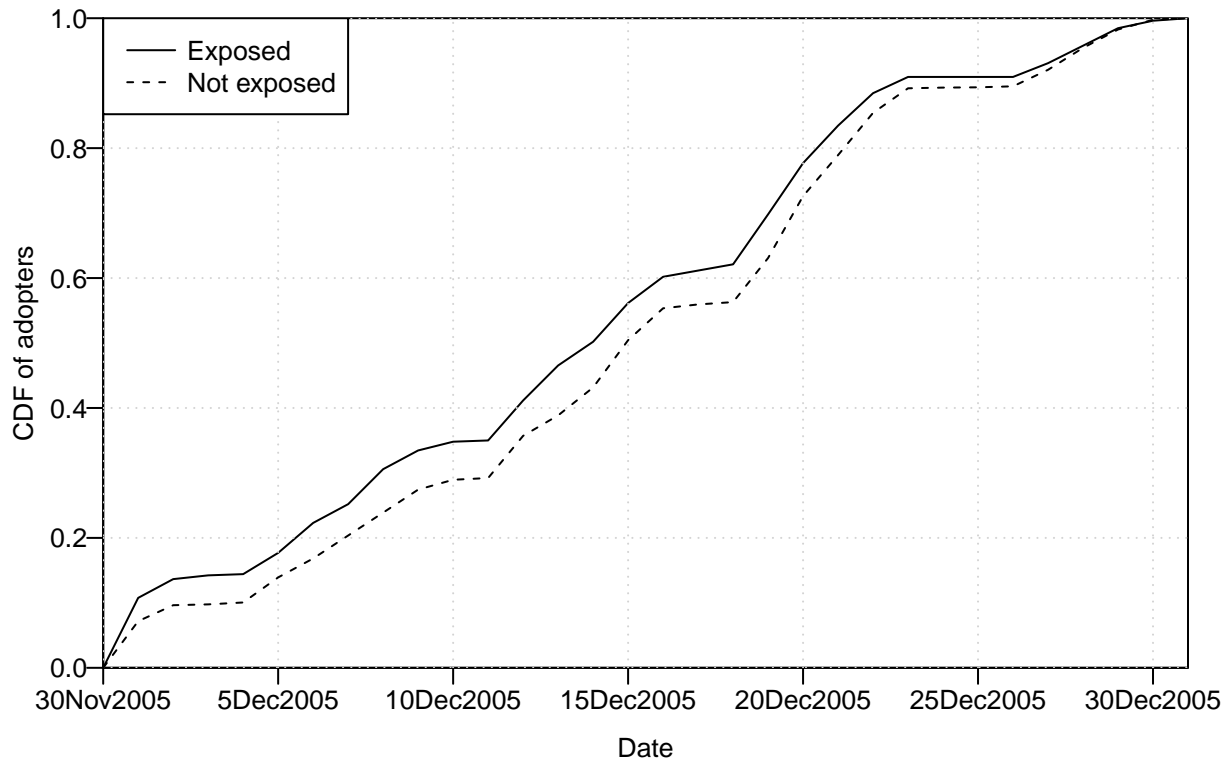


Figure 16: The effect of a family member adopting a week earlier (OLS with no controls)

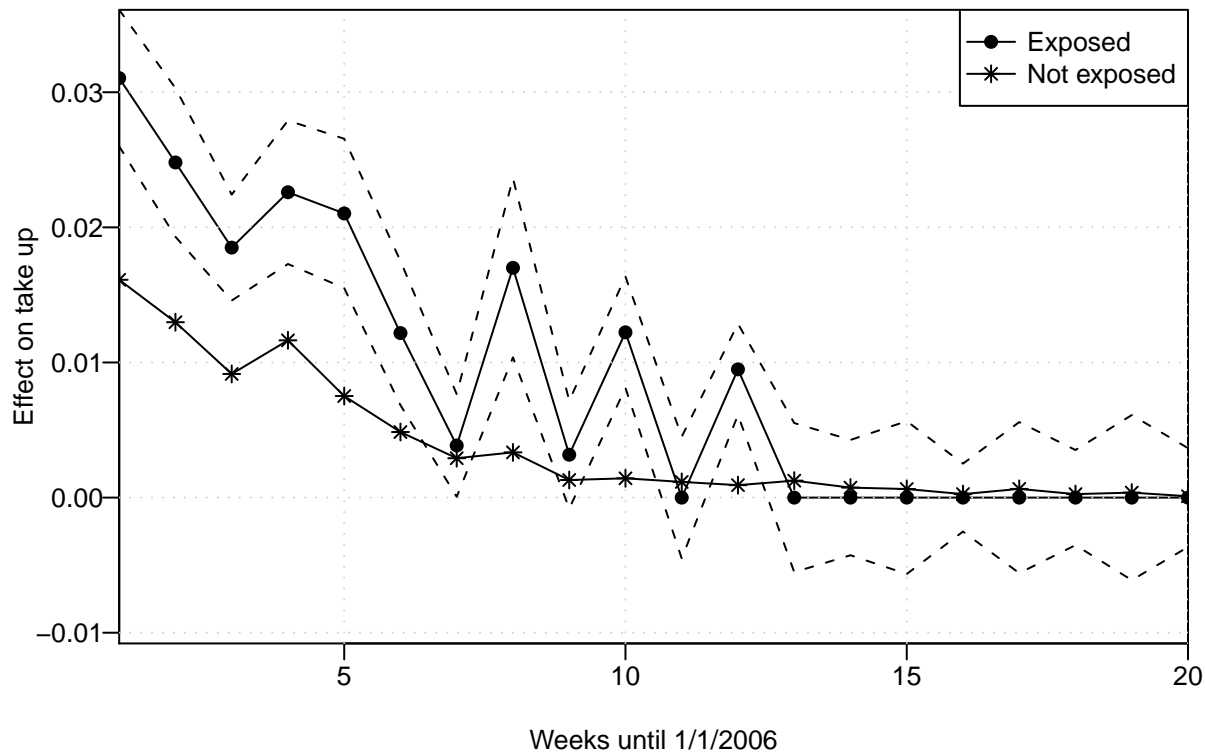


Figure 17: The effect of a family member adopting a week earlier (OLS with controls)

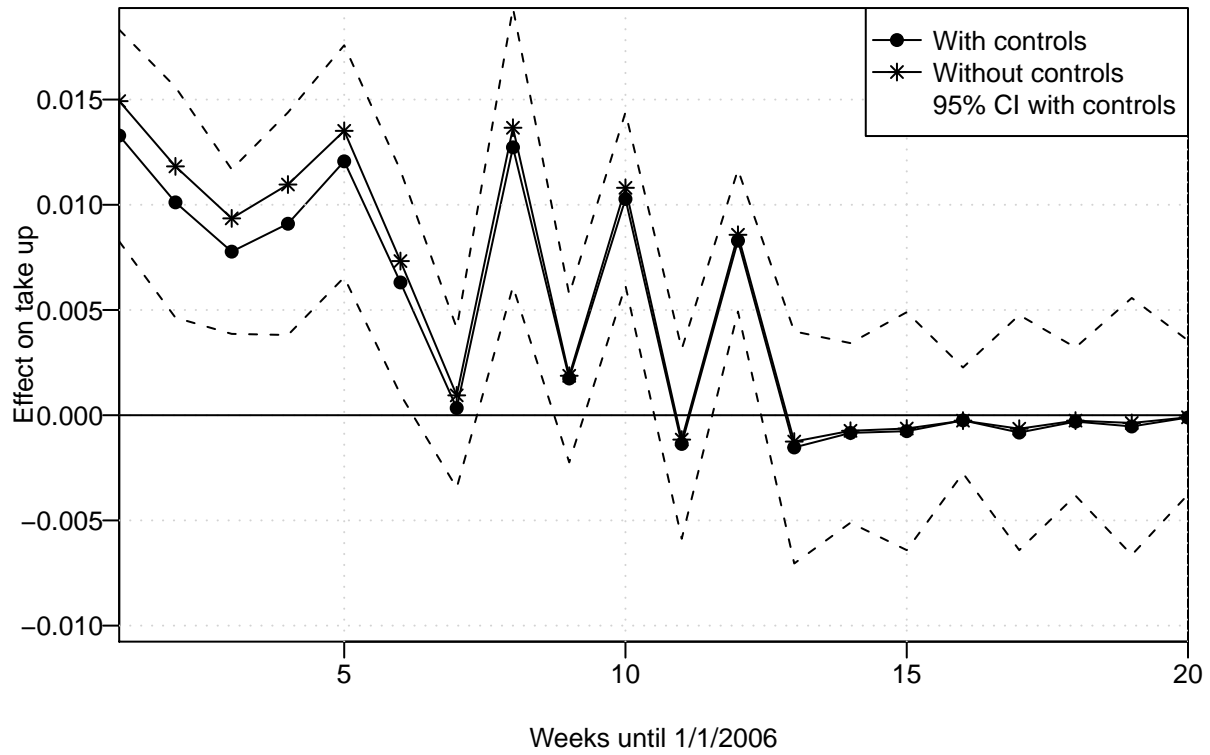
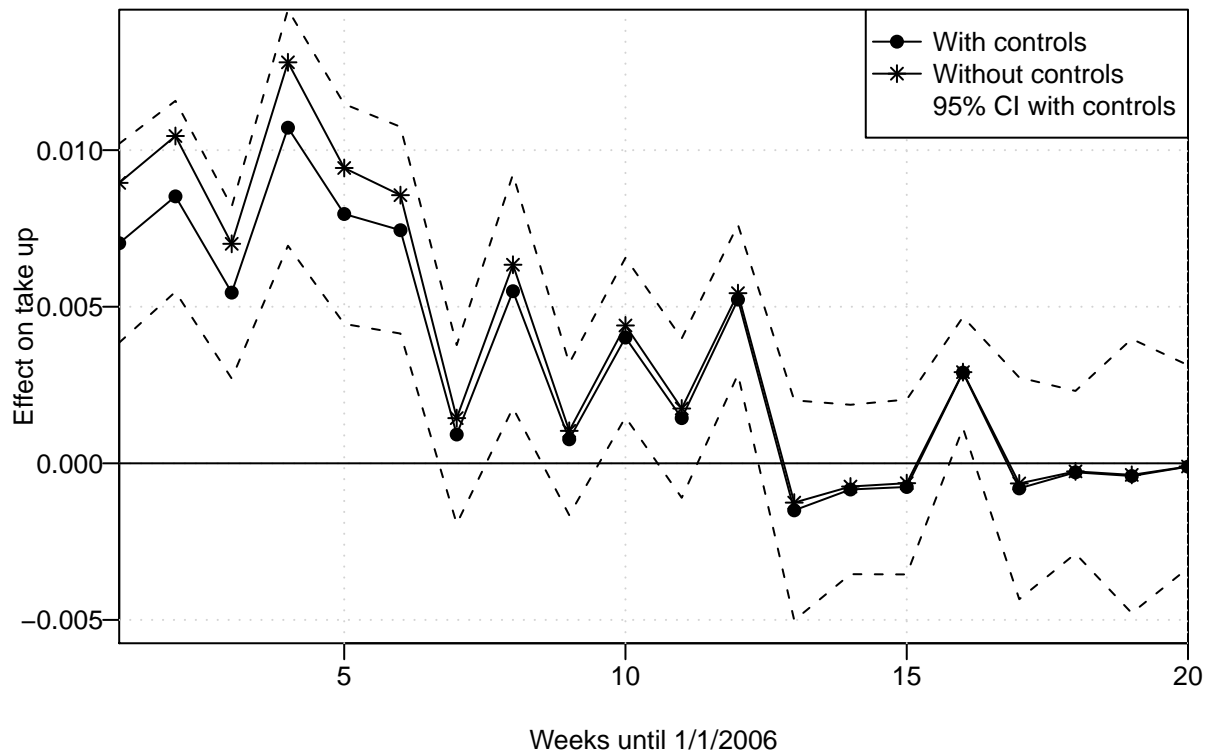


Figure 18: The effect of a family member adopting in the last 3 weeks (OLS with controls)



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 $K_i(K_j, K_{-j}^i, X_i)$, where K_j is awareness of taxpayer j and K_{-j}^i 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.¹⁸ 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 ($\frac{\partial E}{\partial X_j} \neq 0$).¹⁹

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

$$\frac{\Delta \tilde{E}_j}{\Delta X_j} = \frac{\partial f_j}{\partial K_j} \frac{\Delta K_j}{\Delta X_j} + \frac{\partial f_j}{\partial X_j}, \quad \frac{\Delta K_j}{\Delta X_j} = \frac{\partial K_j}{\partial K_i} \frac{\Delta K_i}{\Delta X_j} + \frac{\partial K_j}{\partial X_j} + \frac{\partial K_j}{\partial K_{-j}^j} \frac{\Delta K_{-j}^j}{\Delta X_j}$$

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

$$\frac{\Delta \tilde{E}_i}{\Delta X_j} = \frac{\partial f_i}{\partial K_i} \frac{\Delta K_i}{\Delta X_j}, \quad \frac{\Delta K_i}{\Delta X_j} = \frac{\partial K_i}{\partial K_j} \frac{\Delta K_j}{\Delta X_j} + \frac{\partial K_i}{\partial K_{-j}^i} \frac{\Delta K_{-j}^i}{\Delta X_j}$$

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

¹⁸This 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.

¹⁹Note 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_i 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.

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 his own incentives is not sufficient to establish presence of interactions. It is also strictly speaking not a necessary condition (see an example in footnote 18 — 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:

$$\frac{\Delta K_j}{\Delta x_j} = \left(\frac{\partial K_j}{\partial x_j} + \frac{\partial K_j}{\partial K_{-i}^j} \frac{\Delta K_{-i}^j}{\Delta x_j} + \frac{\partial K_j}{\partial K_i} \frac{\partial K_i}{\partial K_{-j}^i} \frac{\Delta K_{-j}^i}{\Delta x_j} \right) \cdot S^{-1} \quad (3)$$

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

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

$$\beta_j \equiv \frac{\Delta \tilde{E}_j}{\Delta X_j} = \frac{\partial f_j}{\partial X_j} + \frac{\partial f_j}{\partial K_j} \left(\frac{\partial K_j}{\partial X_j} / S \right) \quad (5)$$

$$\beta_i \equiv \frac{\Delta \tilde{E}_i}{\Delta X_j} = \frac{\partial K_i}{\partial K_j} \cdot \frac{\partial f_i}{\partial K_i} \left(\frac{\partial K_j}{\partial X_j} / S \right). \quad (6)$$

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 $\frac{\partial K_i}{\partial K_{-j}^i} \frac{\Delta K_{-j}^i}{\Delta x_j} \approx 0$,²⁰ we have $\frac{\beta_i}{\beta_j} = \frac{\Delta \tilde{E}_i}{\Delta x_j} / \frac{\Delta \tilde{E}_j}{\Delta x_j} \leq \frac{\partial K_i}{\partial K_j} \frac{\frac{\partial f_i}{\partial K_i}}{\frac{\partial f_j}{\partial K_j}}$ with equality

when $\frac{\partial f_j}{\partial X_j} = 0$ and $\frac{\partial K_i}{\partial K_{-j}^i} \frac{\Delta K_{-j}^i}{\Delta x_j} = 0$.

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

²⁰ Assuming that $\frac{\partial K_i}{\partial K_{-j}^i} \frac{\Delta K_{-j}^i}{\Delta x_j} \approx 0$ is restrictive but allows for a major simplifications of formulas 3 and 4. 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 effects. 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 ,

To make progress, assume indeed that $\frac{\partial f_j}{\partial X_j} = 0$.²¹ 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 $\frac{\partial f_j}{\partial K_j} \leq \frac{\partial f_i}{\partial K_i}$.

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 $1/1000$ th 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 Figure 3 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 7,600 removed observations at 0.1 and another 5,600 at other $1/100$ th 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.

$$\frac{\Delta K_j}{\Delta x_j} = S \left(\frac{\partial K_j}{\partial x_j} + \frac{\partial K_j}{\partial K_{-i}^j} \frac{\Delta K_{-i}^j}{\Delta x_j} \right) + \sum_{m \in M} S \left(\frac{\partial K_j}{\partial K_i} \frac{\partial K_i}{\partial K_m} + \frac{\partial K_j}{\partial K_m} \right) \frac{\Delta K_m}{\Delta x_j} \quad (7)$$

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

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 $x = E \left[S \left(\frac{\partial K_i}{\partial K_m} + \frac{\partial K_i}{\partial K_j} \frac{\partial K_i}{\partial K_m} \right) \frac{\Delta K_m}{\Delta x_j} \right]$, $a = \left(\frac{\partial K_j}{\partial x_j} + \frac{\partial K_j}{\partial K_{-i}^j} \frac{\Delta K_{-i}^j}{\Delta x_j} \right)$ 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 $\frac{\partial f}{\partial x_j} = 0$.

²¹Assumption 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.

Online appendix

Table A1: The effect of crossing 10% ownership by a family network member on E-firm take up, heterogeneity by dividend ownership of a network member

Range	Everyone			Above 10%		
	Dividends	No dividends	N	Dividends	No dividends	N
[0.095, 0.105]	0.017 (0.033)	-0.018 (0.032)	1935	0.047 (0.056)	-0.044 (0.059)	1117
[0.090, 0.110]	0.060** (0.028)	0.021 (0.025)	3473	0.119** (0.048)	0.033 (0.043)	1940
[0.080, 0.120]	0.059*** (0.021)	0.036** (0.017)	6799	0.104*** (0.036)	0.056* (0.029)	3652
[0.070, 0.130]	0.045*** (0.017)	0.029** (0.014)	10580	0.080*** (0.030)	0.047** (0.024)	5632
[0.050, 0.150]	0.046*** (0.013)	0.021* (0.011)	18295	0.070*** (0.023)	0.033* (0.018)	9651

Notes: The table shows interaction effects of the treatment dummy with the dummies for the network member having/not having dividends

Table A2: The effect of crossing 10% ownership by a family network member on E-firm take up — timing effects, everyone

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

Notes: 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)

Table A3: The effect of crossing 10% ownership by a family network member on E-firm take up — cumulative timing effects, everyone

Range	Δp		$\Delta \ln(p)$		N
	No controls	Flexible	No controls	Flexible	
Overall	0.008* (0.005)	0.024*** (0.008)	0.127* (0.073)	0.409*** (0.158)	18295
30+ days early	0.008*** (0.003)	0.017*** (0.005)	0.257** (0.109)	0.623** (0.247)	18295
60+ days early	0.001 (0.002)	0.002 (0.004)	0.056 (0.181)	0.197 (0.425)	18295
90+ days early	-0.001 (0.002)	-0.002 (0.004)	-0.149 (0.267)	-0.267 (0.634)	18295
Network members with dividends in 2004					
Overall	0.021*** (0.007)	0.028** (0.013)	0.343*** (0.122)	0.472* (0.266)	5822
30+ days early	0.011*** (0.004)	0.013 (0.008)	0.381** (0.173)	0.514 (0.399)	5822
60+ days early	0.002 (0.003)	0.003 (0.006)	0.150 (0.305)	0.266 (0.693)	5822
90+ days early	0.000 (0.003)	-0.006 (0.008)	0.053 (0.512)	-0.987 (1.053)	5822
Network members without dividends in 2004					
Overall	0.001 (0.006)	0.023** (0.009)	0.019 (0.091)	0.398** (0.192)	12473
30+ days early	0.006 (0.004)	0.019*** (0.006)	0.188 (0.141)	0.692*** (0.304)	12473
60+ days early	0.000 (0.003)	0.002 (0.004)	0.001 (0.232)	0.175 (0.509)	12473
90+ days early	-0.002 (0.003)	-0.000 (0.005)	-0.229 (0.343)	-0.001 (0.734)	12473

Notes: 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)

Table A4: The effect of crossing 10% ownership by a family network member on E-firm take up — timing effects, taxpayers with at least 10% shareholdings

Range	Δp		$\Delta \ln(p)$		N
	No controls	Flexible	No controls	Flexible	
1-30 days	0.004 (0.006)	0.008 (0.012)	0.063 (0.104)	0.156 (0.228)	9651
31-60 days	0.011** (0.004)	0.024*** (0.007)	0.304** (0.138)	0.751** (0.300)	9651
61-90 days	0.002 (0.002)	0.004** (0.002)	0.207 (0.249)	0.808 (0.532)	9651
91-120 days	-0.002 (0.003)	-0.001 (0.006)	-0.356 (0.435)	-0.097 (1.025)	9651
Network members with dividends in 2004					
1-30 days	0.018** (0.009)	0.027 (0.018)	0.317* (0.184)	0.502 (0.402)	3056
31-60 days	0.014** (0.007)	0.019 (0.012)	0.375* (0.217)	0.593 (0.496)	3056
61-90 days	0.003 (0.003)	0.007** (0.003)	0.260 (0.395)	1.792* (0.997)	3056
91-120 days	-0.001 (0.004)	-0.020 (0.022)	-0.152 (0.801)	-5.777* (3.370)	3056
Network members without dividends in 2004					
1-30 days	-0.004 (0.008)	0.000 (0.015)	-0.064 (0.127)	0.004 (0.271)	6595
31-60 days	0.009 (0.006)	0.028*** (0.009)	0.263 (0.178)	0.835** (0.367)	6595
61-90 days	0.001 (0.002)	0.002 (0.003)	0.162 (0.300)	0.316 (0.608)	6595
91-120 days	-0.003 (0.004)	0.004 (0.004)	-0.424 (0.566)	0.647 (1.015)	6595

Notes: 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)

Table A5: The effect of crossing 10% ownership by a family network member on E-firm take up — cumulative timing effects, taxpayers with at least 10% shareholdings

Range	Δp		$\Delta \ln(p)$		N
	No controls	Flexible	No controls	Flexible	
Overall	0.015* (0.008)	0.038*** (0.014)	0.129* (0.073)	0.368** (0.159)	9651
30+ days early	0.011** (0.005)	0.028*** (0.009)	0.198* (0.109)	0.571** (0.244)	9651
60+ days early	-0.000 (0.004)	0.002 (0.007)	-0.024 (0.186)	0.137 (0.425)	9651
90+ days early	-0.003 (0.003)	-0.004 (0.009)	-0.236 (0.276)	-0.329 (0.610)	9651
Network members with dividends in 2004					
Overall	0.034*** (0.012)	0.054** (0.022)	0.309** (0.126)	0.559* (0.291)	3056
30+ days early	0.016* (0.009)	0.025* (0.014)	0.294* (0.179)	0.591 (0.428)	3056
60+ days early	0.002 (0.006)	0.007 (0.008)	0.117 (0.330)	0.550 (0.799)	3056
90+ days early	-0.001 (0.005)	-0.008 (0.014)	-0.058 (0.557)	-0.856 (1.165)	3056
Network members without dividends in 2004					
Overall	0.004 (0.010)	0.032* (0.017)	0.037 (0.090)	0.305 (0.187)	6595
30+ days early	0.008 (0.007)	0.031*** (0.012)	0.145 (0.140)	0.586** (0.290)	6595
60+ days early	-0.002 (0.005)	0.000 (0.009)	-0.098 (0.235)	0.007 (0.478)	6595
90+ days early	-0.004 (0.005)	-0.002 (0.010)	-0.306 (0.355)	-0.126 (0.661)	6595

Notes: 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)

Figure A1: Distribution of ownership shares in 2004 around the 10% threshold

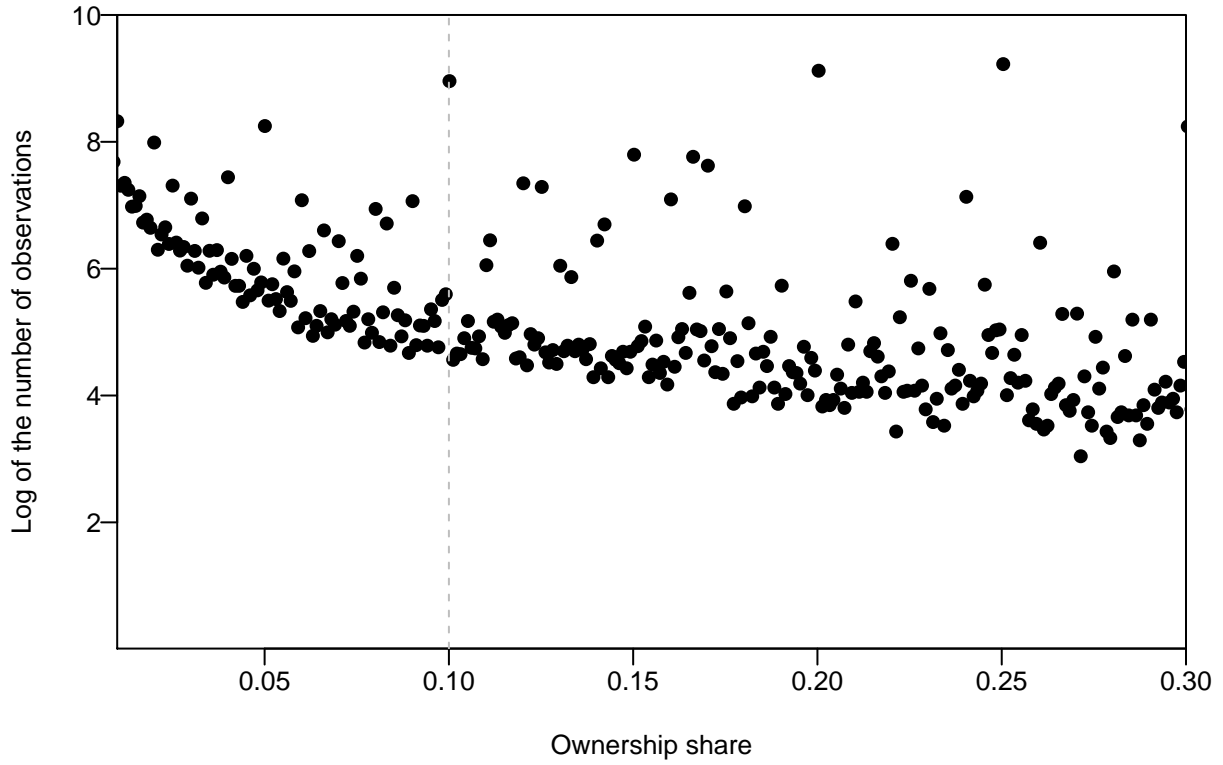


Figure A2: Ownership share in 2004 and probability of ultimate adoption of the Transition E rule — shareholding level, raw data including fractions

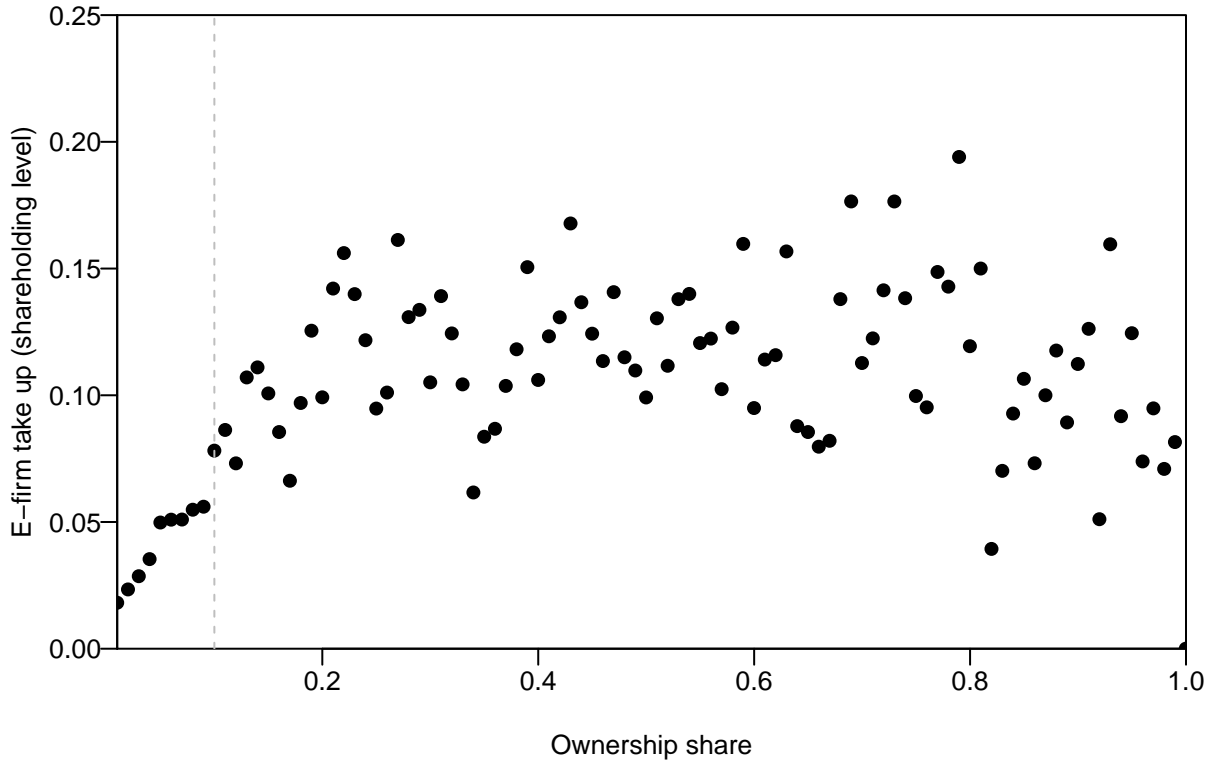


Figure A3: Ownership share in 2004 and probability of ultimate adoption of the Transition E rule
— individual level, raw data including fractions

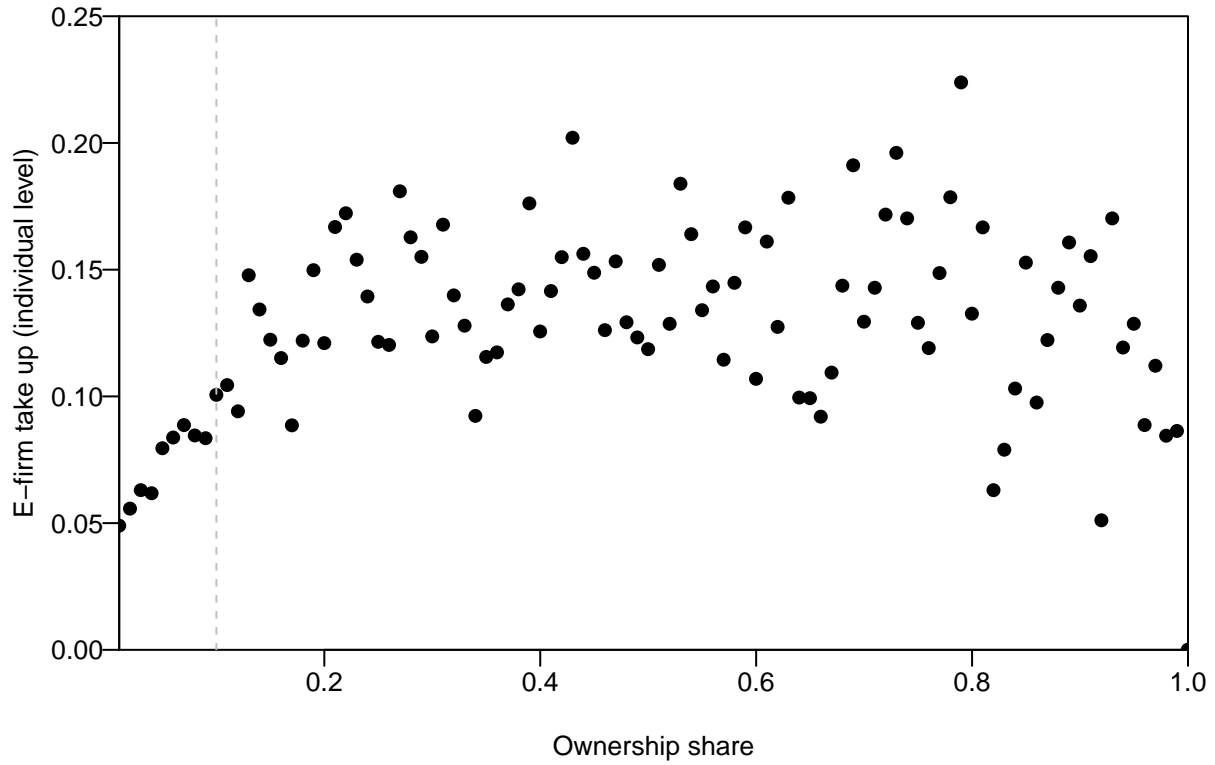


Figure A4: Ownership share in 2004 and probability of ultimate adoption of the Transition E rule
— shareholding level

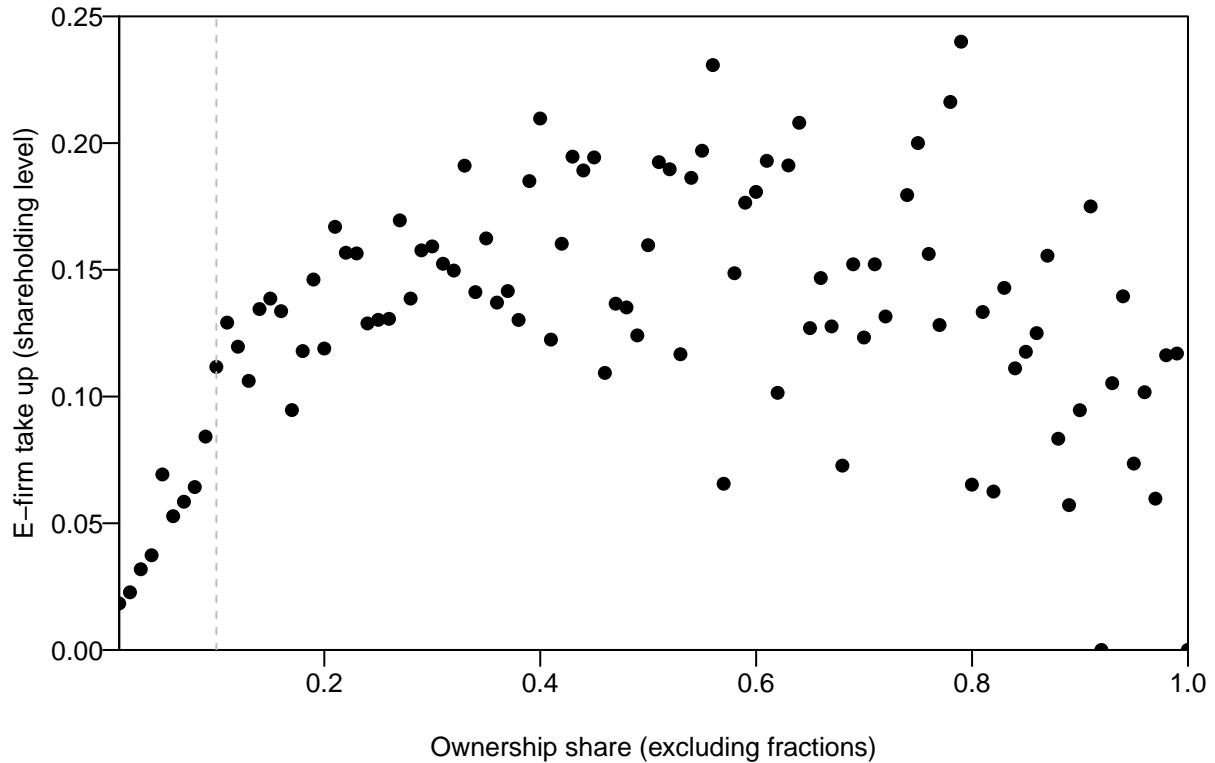


Figure A5: Ownership share in 2004 and number of owners around the 10% threshold

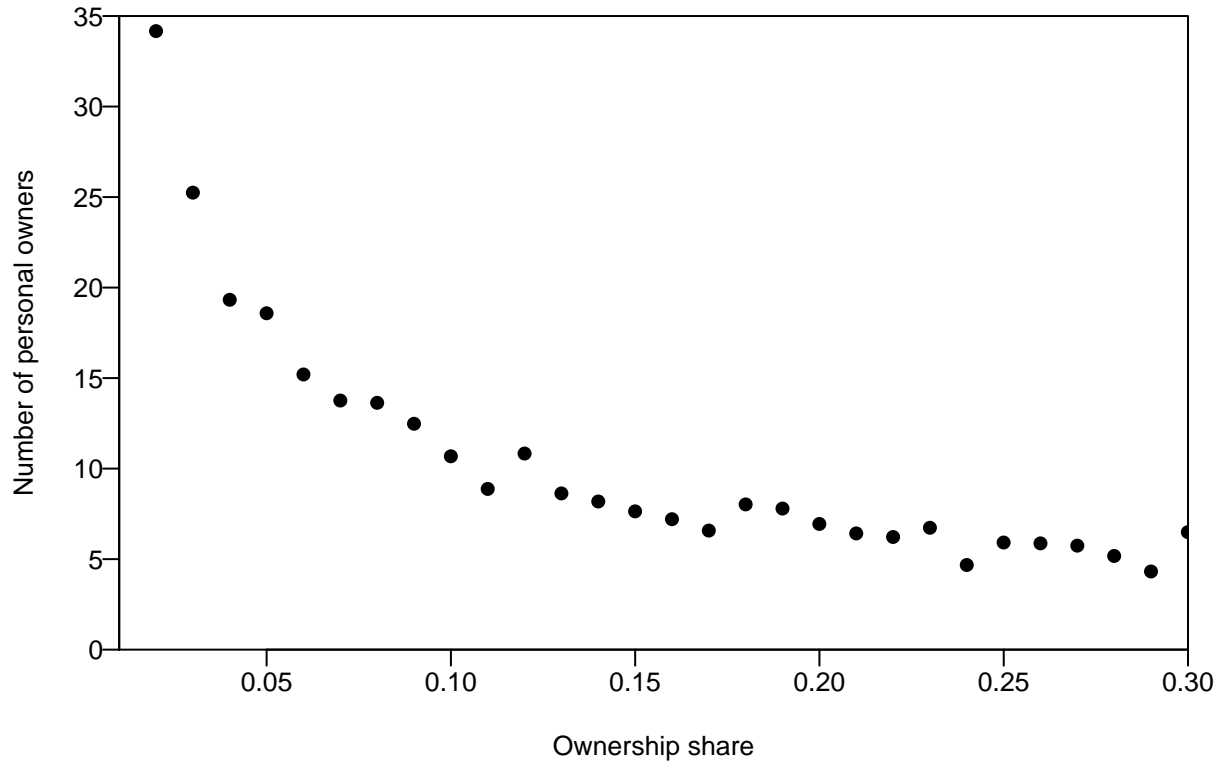


Figure A6: Ownership share in 2004 and log capital around the 10% threshold

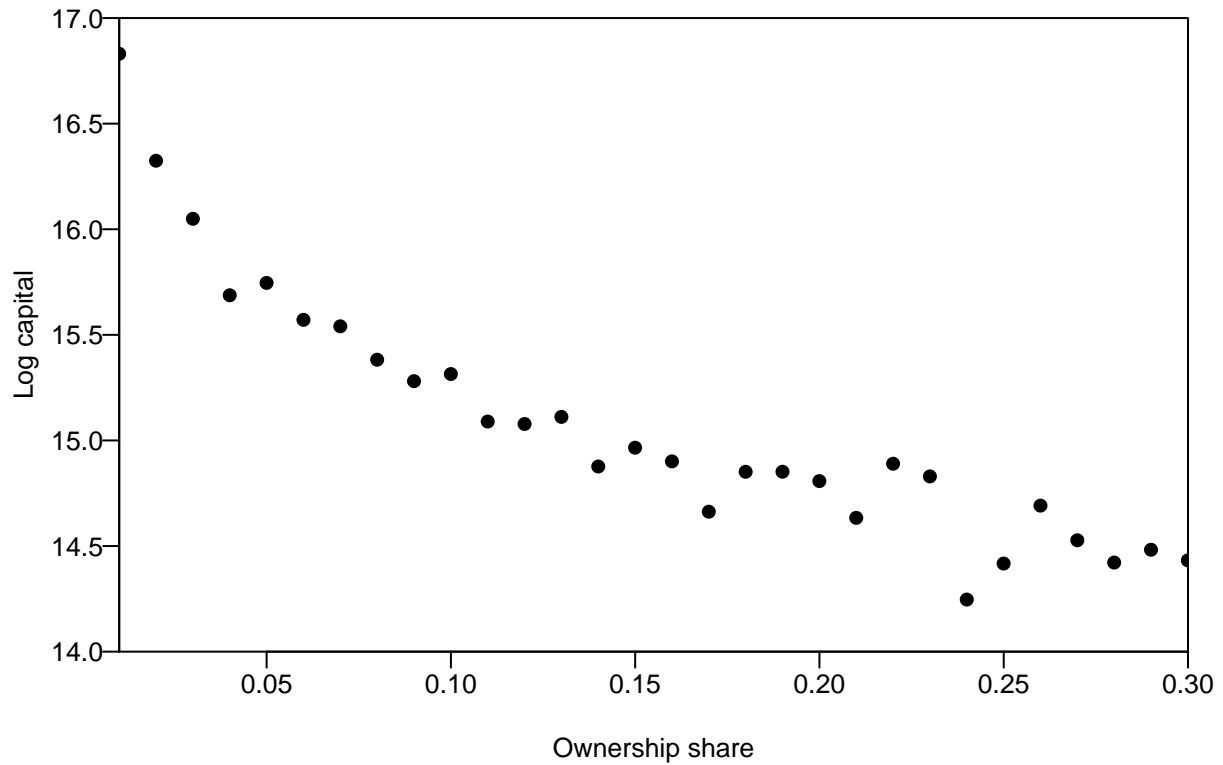


Figure A7: Ownership share in 2004 and gender around the 10% threshold

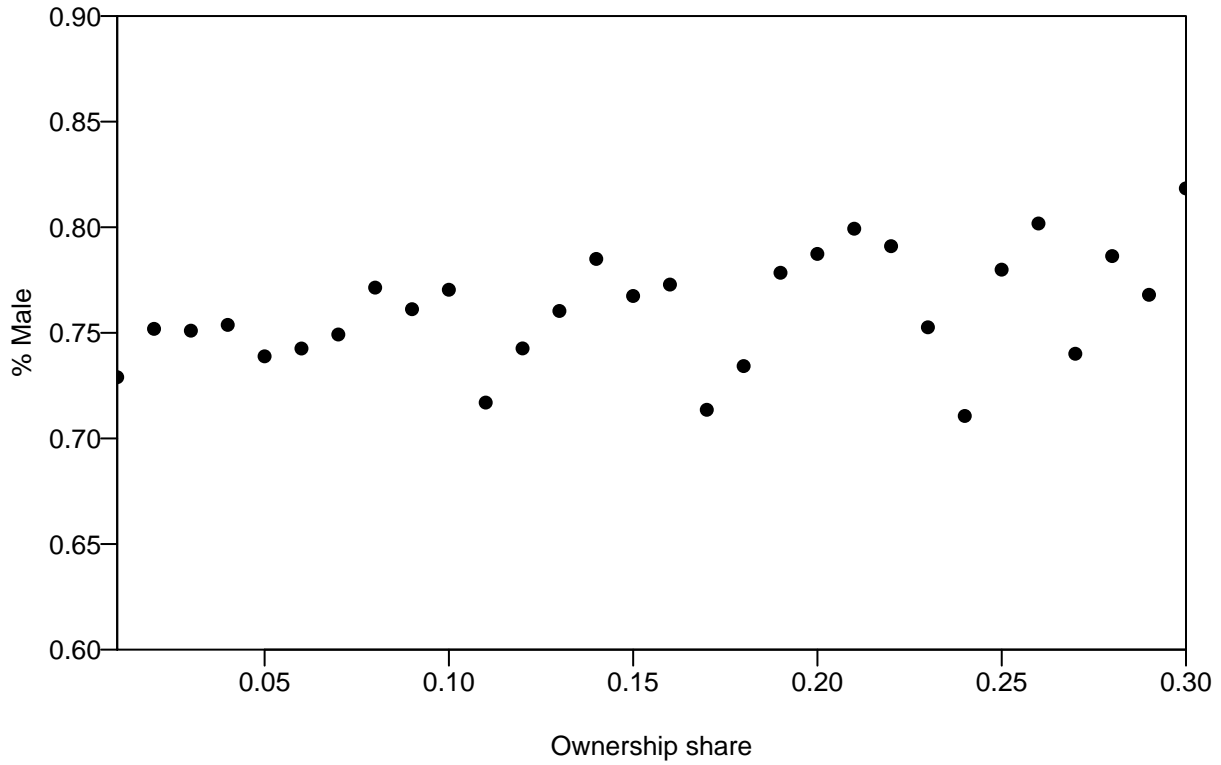


Figure A8: Ownership share in 2004 and average age of shareholders around the 10% threshold

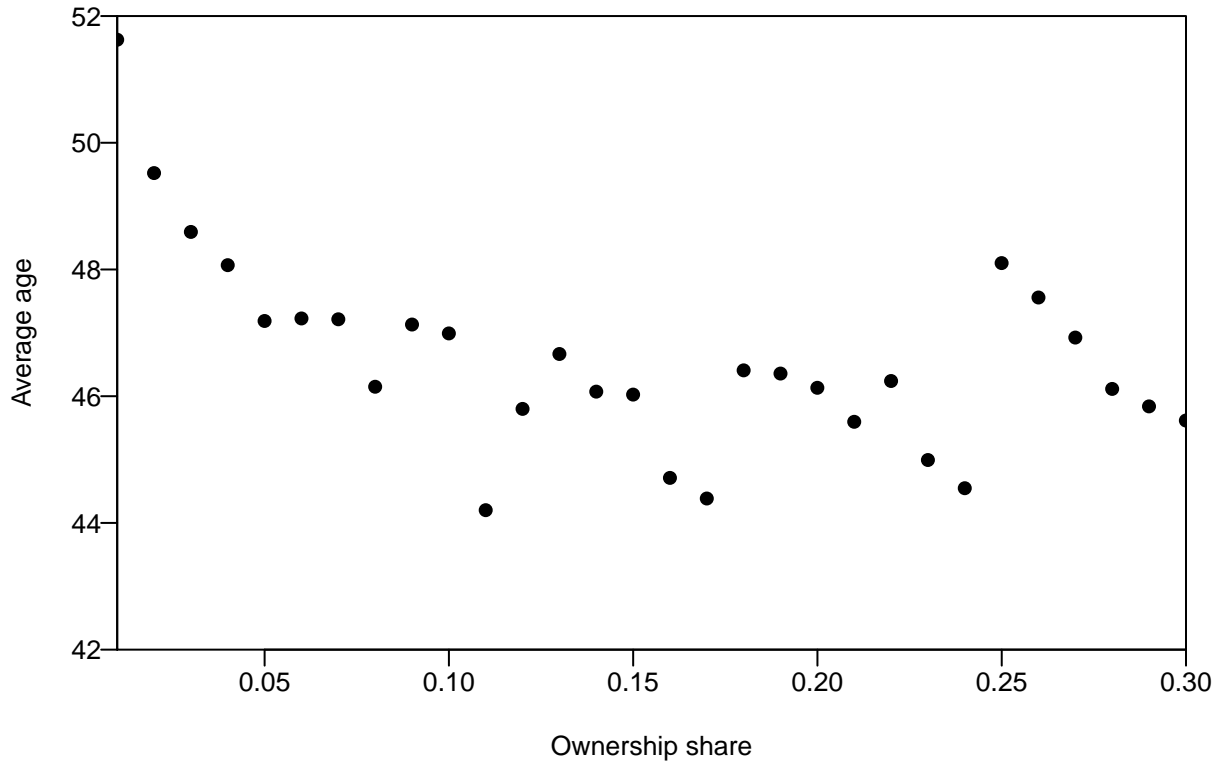


Figure A9: Distribution of family member's ownership shares in 2004 around the 10% threshold

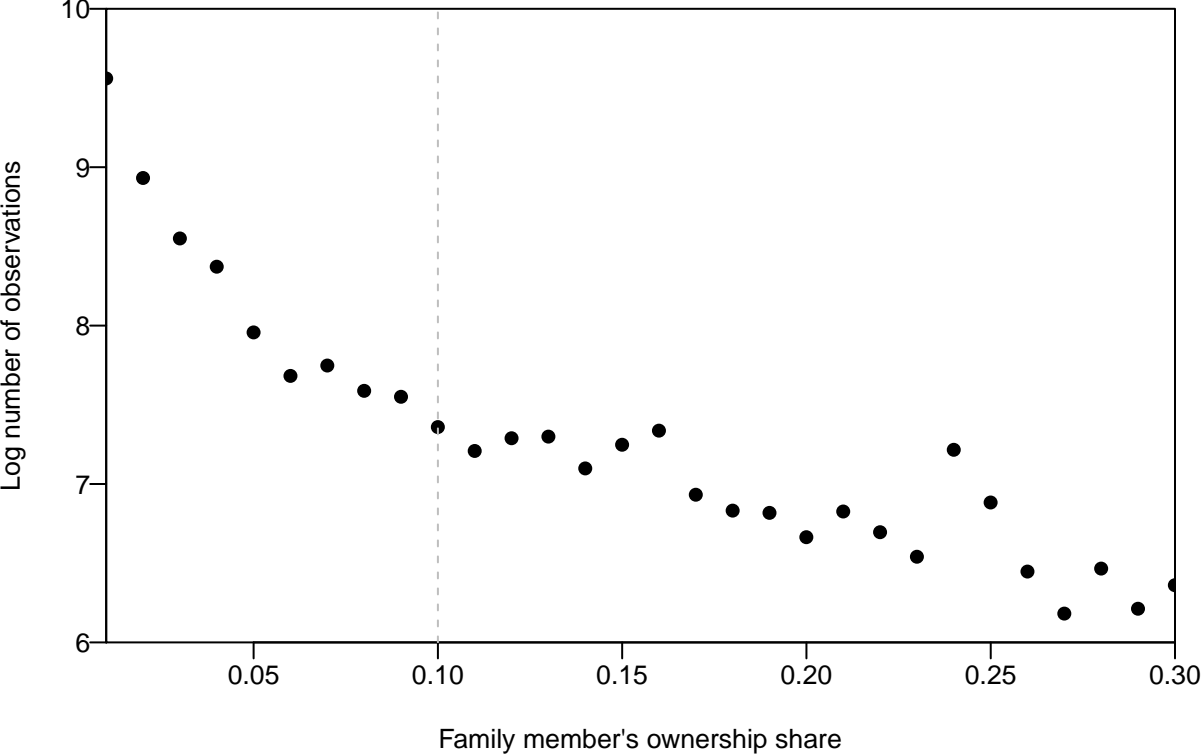


Figure A10: Family member's ownership share in 2004 and number of owners around the 10% threshold

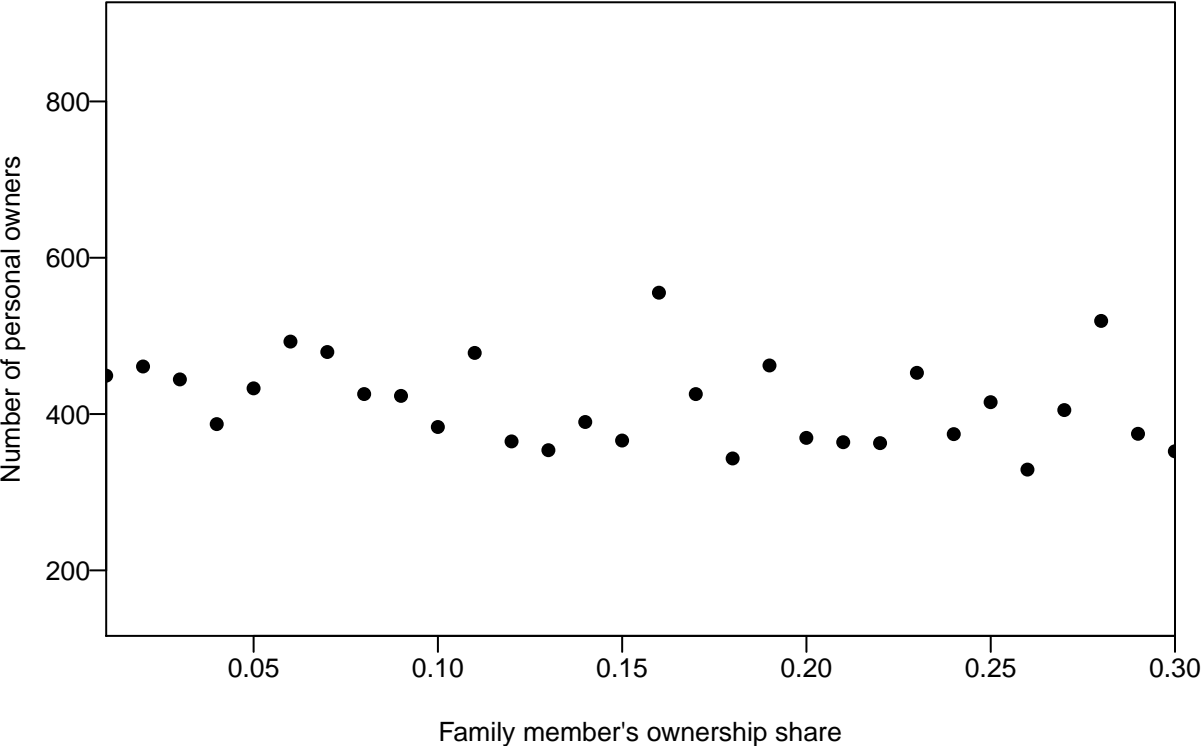


Figure A11: Family member's ownership share in 2004 and log capital around the 10% threshold

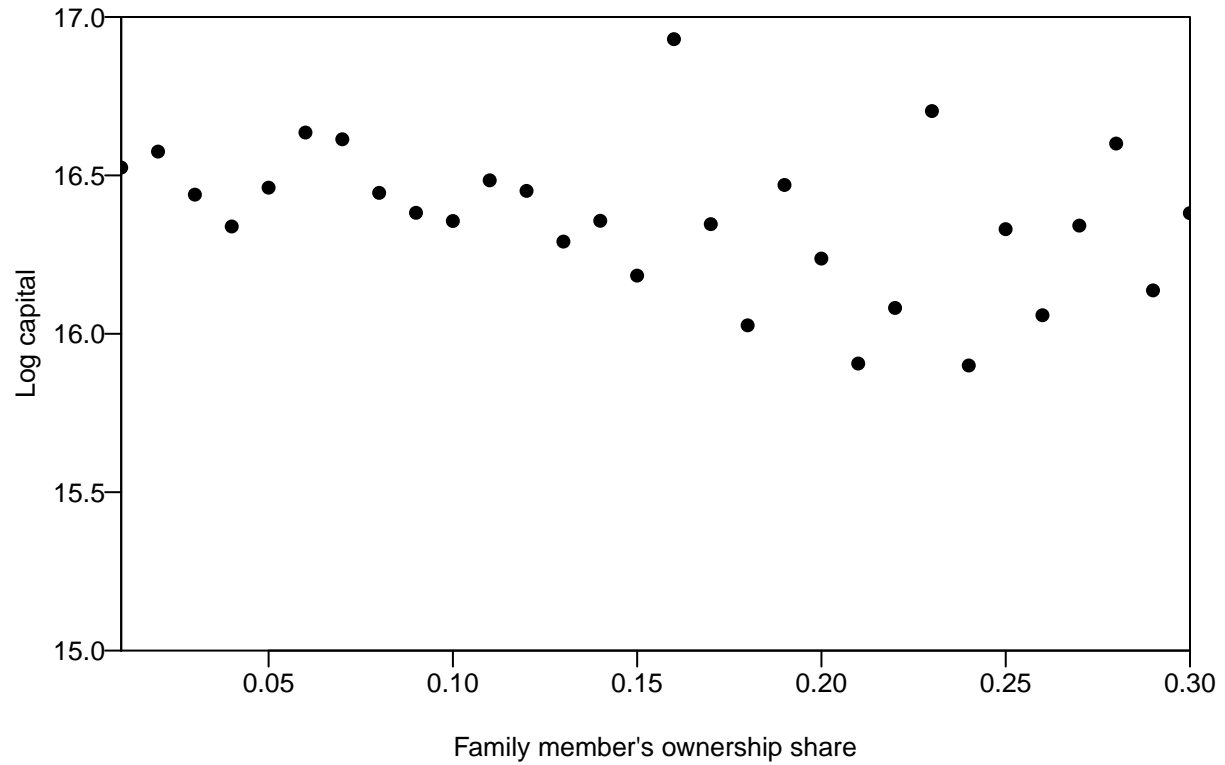


Figure A12: Family member's ownership share in 2004 and average age of shareholders around the 10% threshold

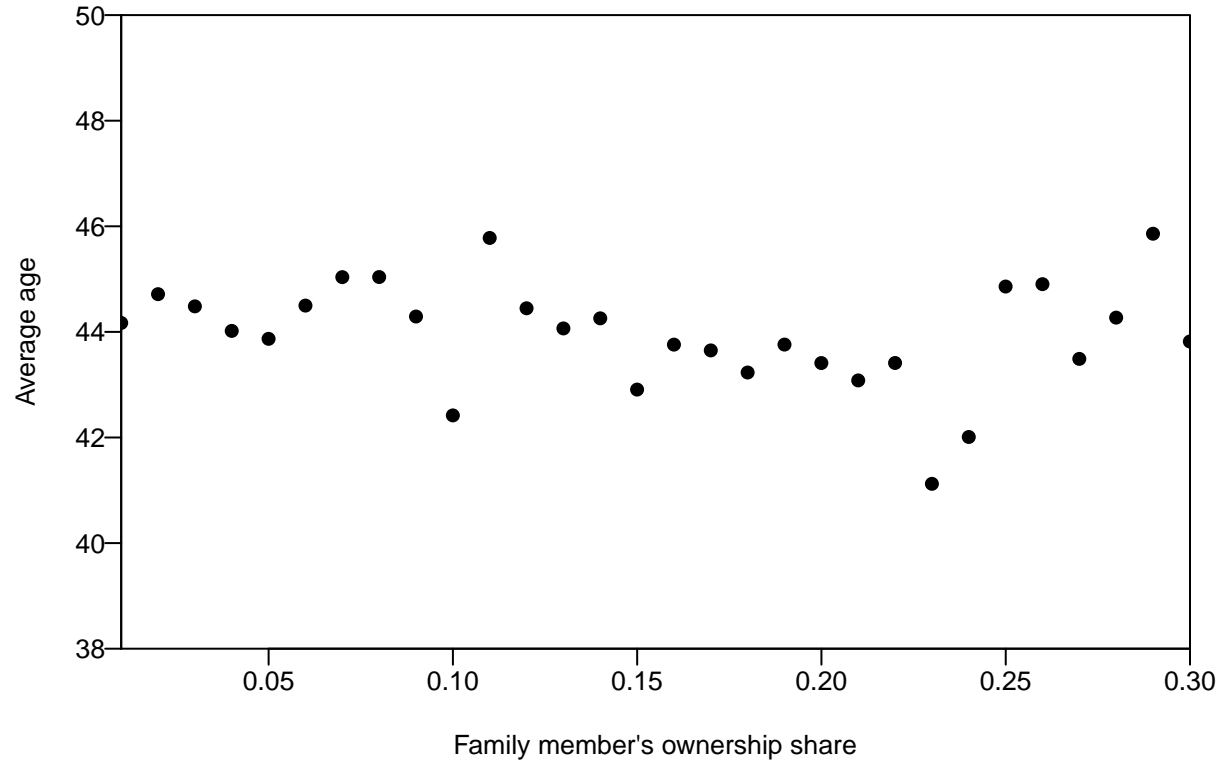


Figure A13: Family member's ownership share in 2004 and gender around the 10% threshold

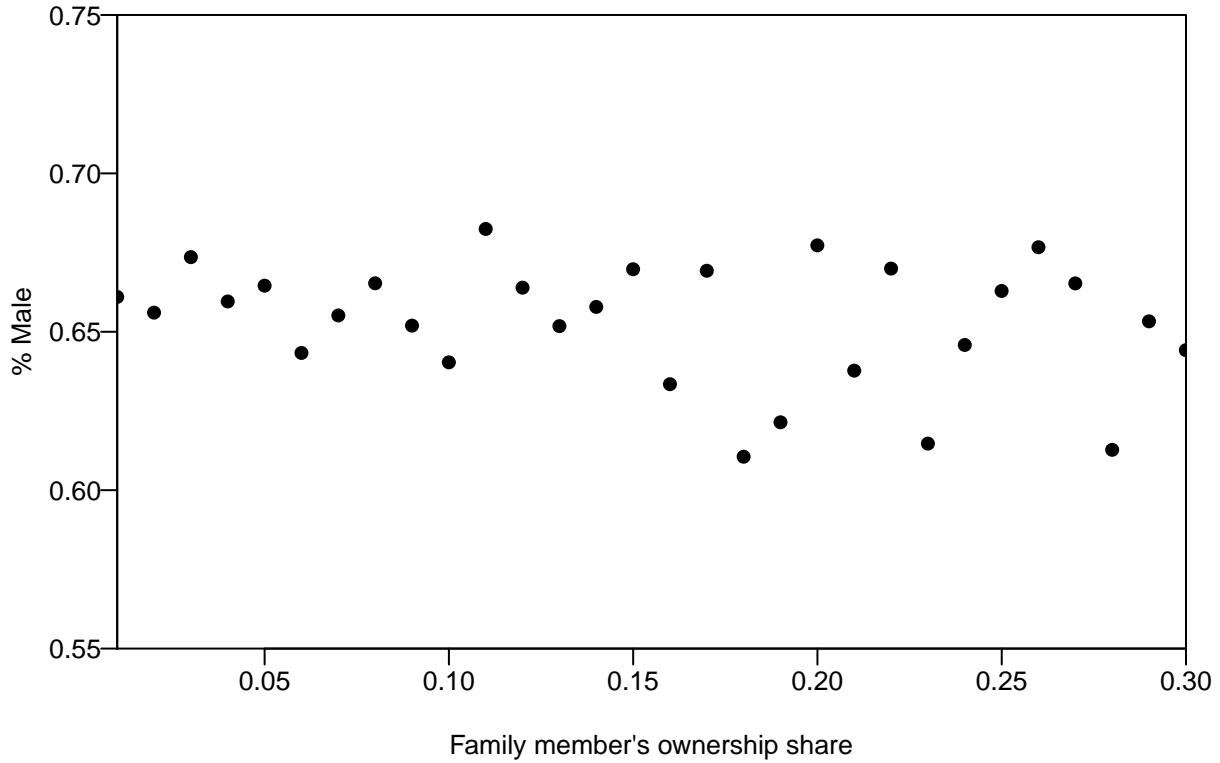


Figure A14: Ownership share in 2004 and probability of ultimate adoption of the Transition E rule — individuals receiving dividends in 2004

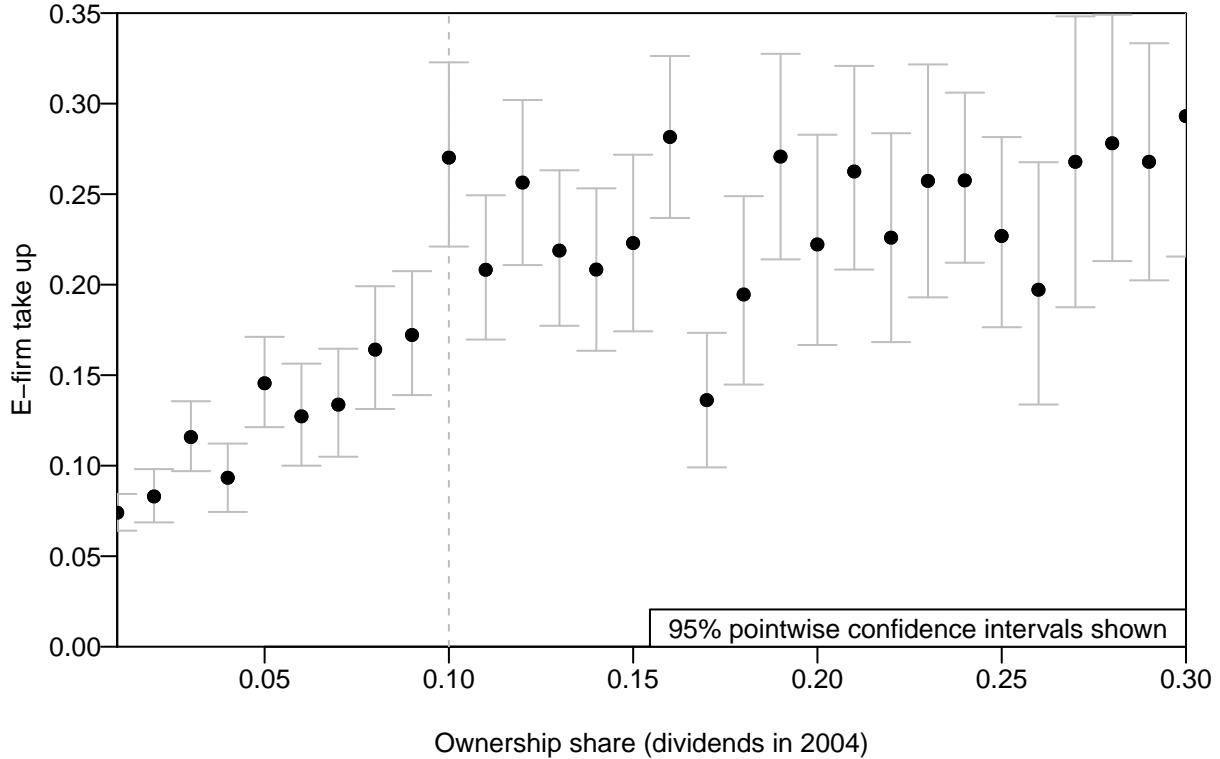


Figure A15: Ownership share in 2004 and probability of ultimate adoption of the Transition E rule
— individuals not receiving dividends in 2004

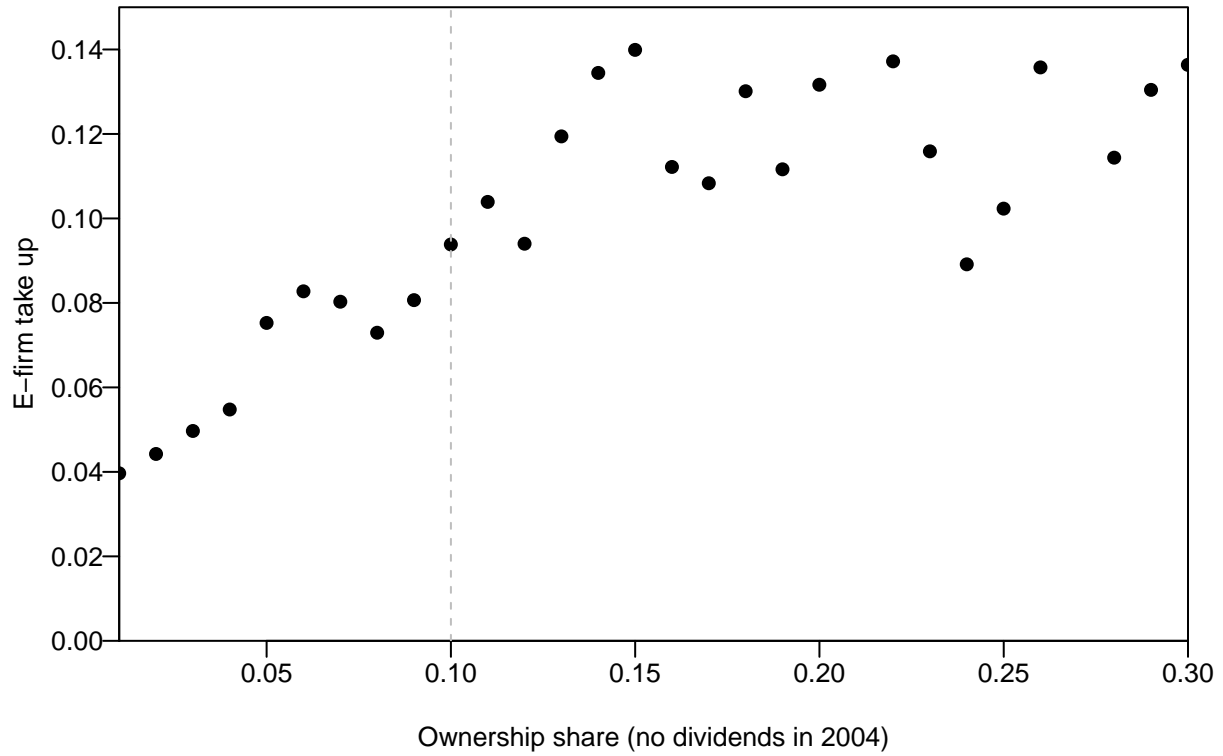


Figure A16: Ownership share in 2004 and probability of ultimate adoption of the Transition E rule
— firms with at most 1000 shares or no dividends realized in 2004

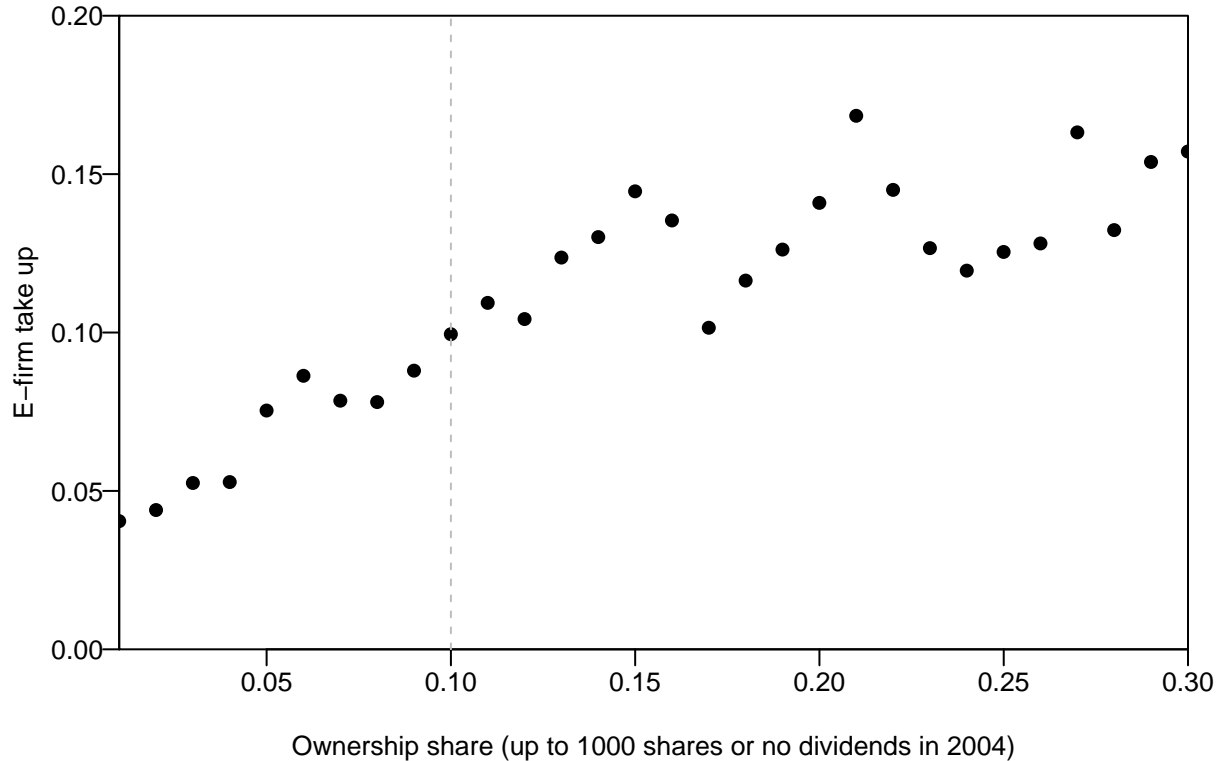


Figure A17: Family member's ownership share in 2004 and number of individuals with at least 10% share

