The impact of international tax information exchange agreements on the use of tax amnesty: evidence from Norway

BY Jonas Andersson, Fred Schroyen and Gaute Torsvik

DISCUSSION PAPER

NHH

Institutt for foretaksøkonomi
Department of Business and Management Science

FOR 12/2019
ISSN: 1500-4066
October 2019

The impact of international tax information exchange agreements on the use of tax amnesty: evidence from Norway*

Jonas Andersson[†], Fred Schroyen[‡] and Gaute Torsvik[§]

30.09.2019

Abstract

In this paper we develop a model for tax amnesty applications in a multi-period setting. One key insight from the model is that applying for amnesty becomes more attractive at the moment when stricter enforcement is announced, even if the implementation of the policy is in the distant future. We use our model to make sense of how international tax information exchange agreements affects voluntary disclosure of wealth and income previously hidden in tax havens. Our data is from Norway. In accordance with the dynamic amnesty model we observe a strong announcement effect of a tax information exchange agreement between Norway and Switzerland and Luxembourg, the two most important tax havens for Norwegian tax evaders. However, the effect levels off very quickly, much faster than our model predicts. We think this is because the initial announcement of the tax agreement exaggerated the risk the agreement imposed to those who had hidden taxable income and wealth in Switzerland. We also estimate and find significant effects of the press releases the Norwegian Tax Authority issues to inform taxpayers about new international tax agreements and the amnesty, or voluntary disclosure, option that exists in the Norwegian tax code.

JEL: H26,H27,K34,C22 and C23

Keywords: Tax Evasion, Tax Amnesty, Tax Information Exchange Agreement

*We are grateful to Lars Aarnes, Karianne Berg, Torhild Henriksen, Paul-Gunnar Larsen, Marcus Zackrisson, Audun Solli (all Norwegian Tax Authority) for making the data available and for help with the interpretation of the data; to Trude Sønvisen (Tax Law Dept, Ministry of Finance) for advice on the interpretation of bilateral tax agreements; to participants of presentations at Statistics Norway (Oslo), the 2nd Belgo-Japanese Workshop on Public Economics (Louvain-la-Neuve, Belgium), Department of Statistics, Uppsala University, Department of Statistics, Athens University of Economics and Business, Skatteforum 2017, the Bergen NoCeT Fall Conference 2017, the Copenhagen workshop on "Tax Evasion, Tax Avoidance, and Inequality" and to Louis Eeckhoudt for helpful comments.

[†]Norwegian School of Economics, Dept of Business and Management Science and NoCeT, email: `jonas.andersson@nhh.no`.

[‡]Norwegian School of Economics, Dept of Economics and NoCeT, email: `fred.schroyen@nhh.no`.

[§]University of Oslo, Dept of Economics and OFS, email: `gaute.torsvik@uio.econ.no`.

1 Introduction

Many countries have a permanent voluntary disclosure (VD) programme, sometimes also called permanent tax amnesty arrangement. Of the 47 countries surveyed by (OECD, 2015), 38 had a permanent VD programme. If tax evaders come forward and voluntarily disclose hidden income or wealth they will have to pay the taxes due and interest charges, but are generally exempted from paying (part of) the penalty tax.

Two types of changes can explain why a person chooses to reverse an evasion decision that was previously deemed optimal: idiosyncratic and systemic shocks. The former refer to changes in attitudes towards risk or towards infringements of social norms. The inheritor of assets hidden in a tax haven may, for example, not share the same risk tolerance or low tax morale as the bequator of those assets. A VD clause in the tax law will allow him to escape from the uncomfortable situation of owning undeclared assets. Another example is a negative income shock making the tax evader more risk averse and hence no longer willing to bear the risk of detection and subsequent penalty. These are idiosyncratic events that ensure a steady demand for VD.

One of the reasons why VD programs have become increasingly relevant over the years are international reforms that have increased the systemic risk associated with tax evasion. People become aware that bank secrecy cannot be shielded from whistleblower-risk.[1] At the same time reforms in the institutional framework regulating the flow of information on incomes and wealth across tax jurisdictions have made such information exchange easier and more efficient. At the level of OECD countries, two sets of reforms are noteworthy. The first, gradually implemented after the financial crisis, relates to the exchange of information *upon request*: the update of article 26 in bilateral double taxation agreements (DTAs for short) empowering tax authorities of the state where the suspected evader resides to request the active assistance of the authorities in the other state–where the evader is suspected to have earned income or parked funds–in the procurement of relevant information on these incomes/funds. In the same vein, many tax information exchange agreements (TIEAs) were concluded with tax haven countries ensuring a similar type of access to information.[2] The second reform, coordinated by the OECD and with full support of the G20, consists in the creation of a legal instrument–the Multilateral Competent Authority Agreement–and the provision of a common standard to regulate and organise the annual *automatic* exchange between signatory states of information on financial accounts held in other states on January 1.[3]

One can therefore assert that these two reforms–the strengthening of information exchange upon request since 2005 and the introduction of automatic information exchange in more recent years–have increased the audit risk for all evaders, and therefore may have shifted the aggregate demand for VD to a higher level. This paper attempts to quantify these two shifts by studying the VD applications

[1] E.g., publicity on the LGT Bank customer file leak in 2008, on the HSBC Swiss-Leak in 2009, on Panama-papers in 2016.

[2] See Keen and Ligthart (2006) for an overview of the development in international information exchange procedures, and Thi and Nikolka (2016) for a more recent update.

[3] The automatic exchange of information between signatory states of the MCAA is preceded (and inspired) by the FATCA agreements which the US government concluded with many states (OECD, 2016, 2012).

from Norwegian citizens during the period 2007-2016. Our data contain all VD applications that were sent to the Norwegian tax authorities in the course of this eleven year period. For each application, we know the exact date, the country in which the income/wealth was hidden and we have information on some characteristics of the applicant.

TIEAs are negotiated by governments and the information that an agreement has been reached, its content, signing date and when it will enter into force, are often announced in government press releases which then dissipate to the general public through the media. Hence, in addition to the data on VD applications, we have gathered information on all press releases on TIEAs from the Norwegian Ministry of Finance, press releases form the Norwegian tax authorities on the existence of the VD clause in the tax law (and the use that people have made of it), as well as media coverage that follow up on these press releases. To make sense of the VD application behaviour, we construct a dynamic decision model where news on future higher detection probabilities may already trigger an application today because of a change in the option value that the evader-status provides. This allows us to distinguish between announcements effects and implementation effects.

Several papers make inference on the effects of TIEAs by measuring compliance using financial statistics. Two recent papers stand out. Johannesen and Zucman (2014) investigate the effect of the conclusion of TIEAs between non-haven and haven countries on the compliance of savers in the former countries using Bank of International Settlement data on bilateral bank deposit holdings. They find a moderate negative effect on the deposit holdings in tax havens that signed information exchange agreements. They find no evidence that these funds are reallocated to the home country and conclude that the G20 tax haven crackdown has mainly resulted in a relocation of deposits to the least compliant tax havens.

The study by Hanlon et al. (2015) consider a particular type of tax evasion: round-trip investments in US bonds and stocks by US residents via a tax haven bank.[4] The authors find two results. The first is that US ordinary and capital gains tax rates positively influence inbound foreign portfolio investment from tax havens (relative to such investment from non-havens), indicating that round tripping is used to evade US taxes. The second is that the conclusion of TIEAs between the US and tax havens, and the specific focus from the OECD on tax havens, have a significant negative impact on inbound foreign portfolio investments from tax havens, indicating that evaders care about the risk of detection. As in Johannesen and Zucman (2014), the second result needs not indicate increased compliance since funds can be relocated to tax havens with which the US does not have a TIEA.

Both these studies use aggregate financial statistics data that include both funds for evasion purposes, but also regular funds. The estimated effects therefore constitute upper bounds on the effects on evasion. In contrast, our paper uses a direct compliance measure – the weekly VD application rate. On the other hand, we have no measure of the stock of evaders. Unlike Hanlon *et al.* (2015), we are not concerned with the decision to become a tax evader (in the spirit of Allingham and Sandmo

[4]US residents set up a sham corporation in a tax haven and open there a bank account in its name. Next, funds are transferred to the account and used to invest in US bonds and equity. Since the immediate owner of the investment is a non US-resident corporation interest income and capital gains are exempted from US tax, and neither does the tax haven tax them.

(1972)). We also ignore how a permanent tax amnesty arrangement influences evasion decisions. This issue has been analysed by Andreoni (1991), in a theoretical paper and by Langenmayr (2017) who also gives empirical support to the idea that a permanent tax amnesty may increase tax evasion.

The overall time pattern of VD applications in Norway mirrors the two waves of TIEAs discussed above. When we use a difference in differences logic to zoom in on the bilateral TIEA agreements Norway signed with Switzerland and Luxembourg we find a strong immediate announcement effect of this agreement. The effect, however, quickly fades away and we observe no impact on VDs from Switzerland and Luxembourg when the agreement is set in force, one and a half year after the announcement. We find a positive effect of the biannual press releases from the Norwegian Tax Authorities, informing the public on the VD opportunity and reminding them about the increased risk of hiding taxable income and wealth abroad. Both our results indicate that announcement and information campaigns are important drivers of tax compliance behaviour.

The paper unfolds as follows. In the next section, we first remind the reader of the Andreoni (1991) model to frame the decision whether or not to apply for amnesty, and then extend this model to a dynamic setting in order to distinguish between announcement effects and implementation effects of policy reforms that increase the likelihood of evasion detection. In section 3, we first describe the permanent tax amnesty rule in Norway and then summarise the application data for the period 2007-2016. Section 4 sketches the institutional aspects of the international exchange of information in tax matters. In section 5, we present the results from the empirical analysis. First we estimate linear reduced form equations. Next we estimate a statistical version of our theoretical model. In section 6 we summarise our results and conclude.

2 Theoretical framework

In this section we develop a theoretical framework that will be useful to structure our thoughts and yield some predictions about the effects of announcements and implementation of future tax enforcement policies. For this purpose we start from Andreoni's (1991) static model and extend it to a dynamic setting.

2.1 The static Andreoni (1991) model

Consider a citizen with a wealth w that is subject to a nominal tax rate t.[5] Preferences over consumption are given by a cardinal utility function $u(\cdot)$ that is increasing and strictly concave (risk aversion), has strictly positive third derivative (prudence), and $\lim_{c\to\infty} u'(c) = \mu$ with $0 \leq \mu < \infty$. Wealth net of due taxes is $w^n = (1-t)w$. Part of this wealth, x, has been underreported to the tax authorities. We do not focus on this underreporting decision, it may have come about by a deliberate choice to evade wealth by the same person, or it may be the result of an inherited wealth that was never declared by the bequeather. The citizen perceives the probability the tax authority will discover

[5]Read income/wealth whenever we write wealth.

the unreported wealth as p. In that case she will be asked to pay the due tax, tx, supplemented with a fine γ per unit of evaded tax. Available consumption is then $w^n - \gamma tx$. On the other hand, if no audit takes place, available consumption is given by $w^n + tx$. Thus a tax evading citizen will have expected utility $(1-p)u(w^n + tx) + pu(w^n - \gamma tx)$. We assume that $(1-p)t - p\gamma t > 0$ so that it was indeed deemed optimal to engage in some evasion.

As in Andreoni (1991), we now introduce a consumption shock ε.[6] If negative, such a shock can be thought of as an unexpected but unavoidable outlay. If positive, it can be regarded as a windfall gain (net of taxes). *Ex post*, expected utility is then

$$U^N(\varepsilon, p) \stackrel{\text{def}}{=} (1-p)u(w^n + tx + \varepsilon) + pu(w^n - \gamma tx + \varepsilon).$$

The alternative to being exposed to the audit risk is to apply for tax amnesty. If the application is successful, the citizen will have to pay the evaded taxes, as well as the fine, a fraction α of which is pardoned as a reward for the voluntary disclosure. Thus in the case of amnesty, *ex post* utility is

$$U^A(\varepsilon) \stackrel{\text{def}}{=} u(w^n - (1-\alpha)\gamma x + \varepsilon).$$

The decision rule for the citizen is to apply for tax amnesty if and only if the immediate utility gain, $\Omega(\varepsilon, p) \stackrel{\text{def}}{=} U^A(\varepsilon) - U^N(\varepsilon, p)$, is positive. It is worth pointing out that without any consumption shock, and if the evasion decision is optimally chosen (thus not 'inherited'), amnesty application is always sub-optimal ($\Omega(0) \leq 0$) since 'by definition' utility under amnesty can never be larger than when remaining honest. In Norway, no penalty is imposed when voluntarily disclosing evasion, so $\alpha = 1$. In the remainder of the theoretical analysis we stick to this case of full pardoning.

We assume that there exists a critical value for ε, $\widehat{\varepsilon}^{stat}(p) \in (\gamma x - w^n, \infty)$, such that $\Omega(\widehat{\varepsilon}^{stat}(p), p) = 0$. In Appendix A, we show that (i) if $u(\cdot)$ has decreasing absolute risk aversion (DARA), then $\Omega_\varepsilon(\widehat{\varepsilon}^{stat}(p), p) < 0$, implying that (because $\Omega(\cdot, p)$ is continuous in ε) $\widehat{\varepsilon}^{stat}(p)$ is unique (Lemma 1); (ii) if $u(\cdot)$ has $-\frac{u''}{u'} < -\frac{u'''}{u''}, -\frac{u''''}{u'''}, -\frac{u'''''}{u''''}$, etc, then $\Omega_\varepsilon(\varepsilon, p) < 0$ for at least all $\varepsilon \in (\gamma x - w^n, \widehat{\varepsilon}^{stat}(p)]$ (Lemma 2); and (iii) $\lim_{\varepsilon \to \infty} \Omega(\varepsilon, p) \leq 0$ (Lemma 3). The assumption under (ii) is stronger than DARA (which amounts to $-\frac{u''}{u'} < -\frac{u'''}{u''}$) but weaker than the property of mixed risk aversion ($-\frac{u''}{u'} < -\frac{u'''}{u''} < -\frac{u''''}{u'''} < -\frac{u'''''}{u''''} < ...$) which is shared by a large class of utility functions.[7] Thus we will assume that $\Omega(\cdot, p)$ is first falling in ε, also when crossing the horizontal axis, reaches a global minimum for some $\varepsilon > \varepsilon^{stat}(p)$, and never crosses the horizontal axis again for $\varepsilon > \widehat{\varepsilon}^{stat}(p)$. The solid line in Figure 1 illustrates the shape when $u(\cdot)$ is the CRRA function. In case $u(\cdot)$ is the linex function (a linear term plus a CARA term) then $\Omega(\cdot, p)$ remains flat after having reached the minimum.

[6]Other idiosyncratic shocks, like shocks to morality and to the perceived detection probability, would also work.

[7]E.g., all Hyperbolic Absolute Risk Aversion (HARA) utility functions (of which CRRA is a special case) are mixed risk averse. Cf. Caballé and Pomansky (1996).

Figure 1. Example of $\Omega(\cdot, p)$, the utility gain from applying for amnesty.

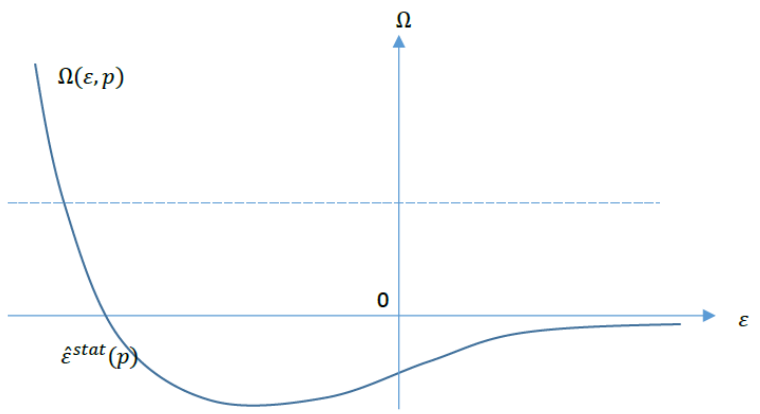


Figure 1. Example of $\Omega(\cdot, p)$, the utility gain from applying for amnesty.

Thus the static model tells us that all citizens with a consumption shock below a critical level $\widehat{\varepsilon}^{stat}(p)$ (implicitly defined by $\Omega(\widehat{\varepsilon}^{stat}(p), p) \equiv 0$) will apply for tax amnesty. If the shocks are i.i.d. according to $F(\cdot)$, a fraction $F(\widehat{\varepsilon}^{stat}(p))$ will apply for amnesty. Increases in the perceived detection probability p shift the Ω-function to the right, resulting in a larger fraction applying for amnesty.

2.2 Dynamic extension of the model

We now extend the model to a dynamic setting where citizens live an infinite number of periods. Period utility is given by $u(\cdot)$ and the period discount factor is $\beta(< 1)$. We abstract from any savings decision and let w stand for the fruits of the orchard that can be consumed each period (non-consumed fruit perishes). The citizen has today honesty status $a = 0$ but considers changing this to status $a = 1$ by applying for amnesty. Honesty status 1 is absorbing, so once having come to terms with the tax authority, one never evades again (e.g, because one fears very close monitoring). To keep the model simple, we assume that tax and penalty liabilities on evaded income expire after one period.[8]

The "present-future"-model In a first instance we analyse the case where the present perceived audit probability is p and the citizen firmly believes that from next period on this probability will rise to $p' \geq p$. We call this the "present-future"-model; later we will inform how the results change when the implementation of the higher audit rate is expected to happen in some arbitrary future period.

An upward revision of beliefs may be triggered by press statements from the tax authorities about TIEAs being signed with tax haven countries where part of the orchard (x) is located. Let the maximal value, as a function of the state variables (a, ε, p, p') be given by $V(\cdot)$, and assume a

[8] See Engel and Hines Jr (1999) for a model where the tax authorities upon an audit also get access to information on evaded incomes in previous periods.

stationary environment. Then by Bellman's principle, we may write

$$V(0,\varepsilon,p,p') = \max_{a\in\{0,1\}} \{a[U^A(\varepsilon) + \beta E_{\widetilde{\varepsilon}} V(1,\widetilde{\varepsilon},p',p')] \tag{2.1}$$
$$+(1-a)[U^N(\varepsilon,p) + (1-p)\beta E_{\widetilde{\varepsilon}} V(0,\widetilde{\varepsilon},p',p') + p\beta E_{\widetilde{\varepsilon}} V(1,\widetilde{\varepsilon},p',p')]\}, \text{ and}$$
$$V(1,\varepsilon,p,p') = U^A(\varepsilon) + \beta E_{\widetilde{\varepsilon}} V(1,\widetilde{\varepsilon},p',p'),$$

where the expectation is over the future consumption shock $\widetilde{\varepsilon}$. By not applying for VD, the evader exposes himself to an audit risk which results in a random continuation utility. It follows immediately from the second expression that $E_{\widetilde{\varepsilon}} V(1,\widetilde{\varepsilon},p,p') = \frac{1}{1-\beta} E_{\widetilde{\varepsilon}} U^A(\widetilde{\varepsilon})$.

In the appendix (Theorem 1) we show that the optimal decision rule is to apply for tax amnesty if and only if the following inequality holds:

$$\begin{array}{rcl} \Omega(\varepsilon,p) & \geq & \dfrac{\beta(1-p)[1-F(\widehat{\varepsilon}(p',p'))]}{1-\beta(1-p')[1-F(\widehat{\varepsilon}(p',p'))]} E_{\widetilde{\varepsilon}>\widehat{\varepsilon}(p',p')}[-\Omega(\widetilde{\varepsilon},p')] \stackrel{\text{def}}{=} \widehat{\Omega}(p,p') \\ & \Updownarrow & \\ \varepsilon & \leq & \widehat{\varepsilon}(p,p'), \end{array} \tag{2.2}$$

where the critical level $\widehat{\varepsilon}(p,p')$ is implicitly defined by letting (2.2) hold with equality.[9] Thus the probability of an application is $F(\widehat{\varepsilon}(p,p'))$.

Condition (2.2) says that it is optimal to apply if the immediate gain of doing so exceeds the present discounted value of the expected gain of remaining an evader and still having the future option to change status whenever the future draw of $\widetilde{\varepsilon}$ is low enough.[10] Any future draw exceeding $\widehat{\varepsilon}(p',p')$ warrants not applying for amnesty. This critical shock level is defined by letting (2.2) hold with equality but with p replaced by p'. In terms of Figure 1, the *rhs* of (2.2) corresponds to the dashed horizontal line and the intersection with the Ω-curve defines $\widehat{\varepsilon}(p,p')$.

A first result, of mainly theoretical interest, is that when evaders anticipate a stable audit policy ($p' = p$), they are less likely to apply for amnesty than in the static setting. The reason is simple: remaining an evader gives an option value and with full pardoning this value is always positive.

Proposition 1. *When $p' = p$ then $\widehat{\varepsilon}(p,p') < \widehat{\varepsilon}^{stat}(p)$.*

(Proof in Appendix A.)

Our second and main result is that the critical consumption shock level increases in both the perceived current and future detection probability. Thus we have

Proposition 2. *When $p' > p$, then $\widehat{\varepsilon}(p,p) < \widehat{\varepsilon}(p,p') < \widehat{\varepsilon}(p',p')$.*

[9] The *rhs* of (2.2) involves the critical value $\widehat{\varepsilon}(p',p')$. This value can be obtained by solving (2.2) for $p = p'$.

[10] The expected benefit of not applying in a typical future period is $[1-F(\widehat{\varepsilon}(p',p'))] \times E_{\widetilde{\varepsilon}>\widehat{\varepsilon}(p',p')}[-\Omega(\widetilde{\varepsilon},p')] + F(\widehat{\varepsilon}(p',p')) \times 0$. The factor $\frac{\beta}{1-\beta(1-p')[1-F(\widehat{\varepsilon}(p',p'))]}$ discounts this typical benefit to the present (since it is $\beta \sum_{t=0}^{\infty} (\beta(1-p')[1-F(\widehat{\varepsilon})])^t$, where $[(1-p')(1-F(\widehat{\varepsilon}))]^t$ is the probability of still having evader status t periods in the future. Finally, $(1-p)$ is the probability of surviving to the future as an evader.

(This proposition is a special case of Proposition 3 to be stated shortly.)

The anticipation of a higher future audit probability reduces the option value of remaining an evader and thus makes VD application more likely today. Once the tougher audit policy is implemented, e.g., because a tax information exchange agreement is in force, this likelihood gets even higher because also the immediate return to an amnesty application increases. The announcement effect corresponds to a downward shift of the horizontal line in Figure 1. The implementation effect produces a further downward shift of that line (because p enters the discount factor–a higher perceived chance of detection this period makes it less likely to survive as an evader), as well as a rightward shift of the Ω-curve.

The general model So far, implementation of the higher audit rate was expected to happen in the next period. We now present the results when implementation is expected to happen in $T+1$ periods from today. This can be denoted as $(p_{[T]}, p')$ meaning that the current audit rate applies for T periods and then rises to p' from period $T+1$ onwards. Let us denote $\widehat{\Omega}(p_{[T]}, p')$ as the present discounted option value of remaining an evader and $\widehat{\varepsilon}(p_{[T]}, p)$ as the critical ε that equates the immediate benefit of applying for amnesty with this option value. Thus $\widehat{\Omega}(p_{[T]}, p') \overset{\text{def}}{=} \Omega(\widehat{\varepsilon}(p_{[T]}, p), p)$. In the "present-future"-model, we had $\widehat{\Omega}(p_{[1]}, p') = \widehat{\Omega}(p, p')$.

Then we can write $\widehat{\Omega}(p_{[T]}, p')$ recursively as

$$\widehat{\Omega}(p_{[T]}, p') = \beta(1-p) \int_{\widetilde{\varepsilon} \geq \widehat{\varepsilon}(p_{[T-1]}, p)} \left[\widehat{\Omega}(p_{[T-1]}, p') - \Omega(\widetilde{\varepsilon}, p) \right] \mathrm{d}F(\widetilde{\varepsilon}) \tag{2.3}$$

where $\widehat{\Omega}(p_{[1]}, p')$ is as defined in (2.2) and $\widehat{\Omega}(p_{[0]}, p')$ is as defined in (2.2) with p replaced by p'. For all draws of next period's shock $\widetilde{\varepsilon}$ exceeding $\widehat{\varepsilon}(p_{[T-1]}, p)$, it is optimal to remain an evader in that period, resulting in a utility gain given by the square bracket term. The expected present discounted value of this option value is the *rhs* of (2.3). It is only optimal to apply for amnesty today if the immediate gain of doing so exceeds (2.3).

It can then be shown that as the number of periods before implementation increases, the critical ε-value below which applying for amnesty today is optimal declines:

Proposition 3. *$\widehat{\varepsilon}(p_{[T]}, p') < \widehat{\varepsilon}(p_{[T-1]}, p')$ for $T \geq 1$.*

(Proof in Appendix A.)

Thus we obtain the intuitive result that the more remote the implementation date in the future, the lower the likelihood of application today. As the implementation date approaches, it become more likely to apply because the option value declines. Upon implementation, the application rate gets an extra boost, because the survival probability on the *rhs* of (2.3) falls from $1-p$ to $1-p'$ (see the *rhs* of ((2.2) with $p = p'$)).

Before the announcement, the application rate is $F(\widehat{\varepsilon}(p_{[\infty]}, p'))$.[11] Upon announcement that the

[11]I.e., it is as if an announcement of a rise in p' occurring not before an infinite periods in the future.

higher audit rate p' will be implemented $T+1$ periods from now, the application rate jumps up to $F(\widehat{\varepsilon}(p_{[T]}, p'))$. As the implementation day approaches, the application rate monotonically increases, to make another jump at implementation. Evaders may perceive the timing of the implementation differently. For those who believe it is in the remote future, the announcement will hardly affect their inclination to apply for amnesty. To others, the announcement may give the impression that implementation of a higher detection rate is around the corner; their application rate may raise significantly. If new information arises that moderates the perception of a speedy implementation, then the application rate among those still evading may fall.

Proposition 3 is a statement about the application rate. If the number of potential applicants remains constant, then the behavior of this rate should translate into a similar behaviour in the number of applications. This qualification, however, is unlikely to hold. The pool of potential applicants decreases over time because (i) people apply for tax amnesty and (ii) a fraction of the non-applicants is caught and fined. On the other hand, there may be entrance of new evaders. However, because tax information exchange agreements increase in number and scope, it is likely that new entrants come from the more risk tolerant part of the population. If risk tolerance is distributed skewly to the right they may not outnumber the flow of exists from the pool. We therefore conjecture that the pool of potential applicants shrinks over time. Thus in terms of the number of applications, a strong announcement effect of a future increase in the detection probability may weaken the implementation effect, just because the latter applies on a significantly reduced stock of potential applicants. It is, therefore, important to keep the announcement effect in mind when looking at the data: ignoring the anticipatory responses may downwardly bias estimates of the effects of TIAEs, especially if a difference in differences methodology is used to estimate reform effects. This point is also stressed by Blundell et al. (2011) in the context of announced welfare reforms.

3 Voluntary correction of tax returns in Norway

This section provides a brief outline of the Norwegian VD programme and presents descriptive statistics on its use.

3.1 The rule

Since 1987, The Norwegian Tax Assessment Law has had a paragraph opening up for tax amnesty, although that wording is never used: "The supplementary tax due when supplying the tax authorities with false information can be calculated at a lower rate [...] or waived if a tax payer voluntarily corrects or supplements information that was given or used earlier, such that it allows for a correct tax assessment. This does not apply when the correction [of information] can be regarded as triggered by monitoring that has been or will be activated, or by information which the tax authorities have received from others" (Tax Assessment Law, §10-3 2.c). Hence the official Norwegian name *voluntary correction* rather than amnesty. Since 2010, the wording "can be calculated at a lower rate [...] or waived" has been replaced with "is waived". In praxis, the supplementary tax due before 2010 was 1% of the due tax. The normal supplementary tax upon audit is 30% of taxes due (60% for serious offenses).

3.2 The use of the VD rule

Between January 2007 and December 2016, 2309 persons applied for VD, on average 4.4 per week. 58% of these applicants reside in Oslo or the neighboring county Akershus. 1602 (69.4%) application files have been concluded by April 2017. Of these, 991 resulted in a positive adjustment of either taxable income or wealth.[12] The concluded applications resulted in around 58bn NOK (around 6.5bn €) of extra taxable wealth and around 2.3bn NOK (250 mn €) of extra taxable income. Conditional on a positive adjustment, the median income and wealth adjustments are 467.000 NOK and 15.2mn NOK, respectively. The maximal income and wealth adjustment amounted to 305mn NOK and 3.04bn NOK, respectively.

The large frequency of zero adjustments for both income and wealth is precisely due to the fact that VD applications are voluntary *corrections*. If a correction does not give rise to a change in taxable income/wealth, then the VD will not result in a higher income/wealth tax liability. In principle, an investor who took up a loan with a Swiss bank to buy stocks that subsequently plummeted in value, could apply for VD with his net taxable wealth being adjusted downwards; the adjustment in tax liability, however, can never be negative.

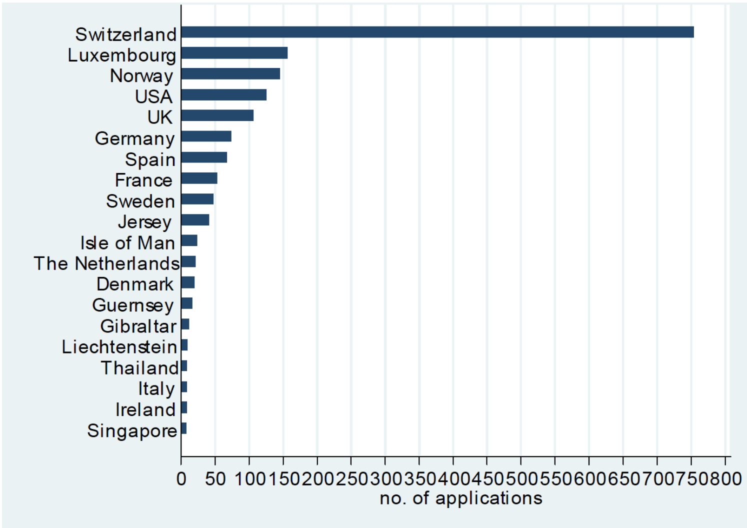
For 1932 out of 2309 VD applications (83.7%) our data set has information on the tax haven used. For these remainder 84% of cases, 75 distinct tax havens were mentioned. In most cases, the applicant made use of a single tax haven. The 20 most 'popular' tax havens are presented in Figure 2. These are the havens with 10 or more VD applications in the data set. The first place is taken by Switzerland, followed by Luxembourg, Norway and the US.

4 Tax treaties and exchange of information

Around 2006, Norway had bilateral tax agreements with around 80 countries. Most of these agreements were based on the model tax convention of the OECD or the United Nations, and had an article regulating the exchange of information (in most cases, article 26 or 27). Paragraph 1 of this article says that "[T]he competent authorities of the Contracting States shall exchange such information as is necessary for carrying out the provisions of this Convention." The 2nd paragraph specifies that the

[12]Two remarks are in order. (1) These figures are the amounts by which *taxable* wealth/income increase due to VD. Thus if w_0 is net reported wealth/income prior to VD, and w_{vd} is additional wealth/income revealed through VD, then the increase in the tax base is $\max\{0, w_0 + w_{vd}\} - \max\{0, w_0\}$. Because net wealth/income prior to VD can be negative (e.g., because the imputed value of one's residence falls short of the mortgage, or because labour earnings fall short of significant losses on the stock exchange) the wealth/income that is revealed through a VD application need not result in an increase in taxable wealth/income. (2) These figures are accumulated taxable wealth/incomes for up to 10 years, since by law tax authorities cannot change tax files older than 10 years. Thus if a 1mn NOK wealth was kept hidden for 5 years, and the citizen had a net wealth of -1mn NOK for the first year, but positive taxable wealth for the 4 remaining years, then the figure that shows up in our data is 4mn NOK.

Figure 2. The most important source countries for Norwegian VD cases



Note: The graph shows the total number of VD applications from different source countries over the period 2006 - 2017. Only countries with 10 or more VD applications are included.

obtained information shall be treated as secret in line with domestic law of the contracting state and paragraph 3 imposes certain limitations to the main rule of §1, in favour of the requested state, by allowing not to go beyond its own internal laws and administrative practices in putting information at the disposal of the requesting state. In particular, it says that paragraph 1 cannot be interpreted as imposing on a state the obligation to supply information that would entail the disclosure of business or commercial secrets. In the tax agreement with Switzerland, the term "banking secrets" is explicitly mentioned.

Revisions of the OECD's Model Tax Convention in 2005 and 2010 resulted in an updated article on information exchange to remove doubts as to its proper interpretation. But these revisions also added two new paragraphs. Paragraph 4, according to which contracting states must use their information gathering measures, was added to deal explicitly with the obligation to exchange information in situations where the requested information is not needed by the requested state. And paragraph 5 stipulates that a contracting state shall not decline to supply information to a treaty partner solely because the information is held by a bank or other financial institution. Thus, paragraph 5 overrides paragraph 3 to the extent that paragraph 3 would otherwise permit a requested contracting state to decline to supply information on grounds of bank secrecy. Although several countries made reservations with the revisions, most countries adopted them by updating existing bilateral tax agreements through a protocol or when negotiating new ones.

In 2002, the OECD also released a model agreement on the exchange of information on tax matters, to be used on a bilateral or multilateral basis by countries that have not entered or have no desire to enter a bilateral tax agreement. The model agreement pertained to the exchange of information upon request, and in 2015 the OECD also released a model agreement to regulate the automatic exchange of information.

In 2006 the Nordic Council of Ministers instructed a steering group to coordinate the negotiations with tax havens to establish TIEAs with tax havens on behalf of six Nordic countries.[13] The first TIEA was signed with Isle of Man in late October 2007. When in 2009, the OECD urges tax havens to sign at least 12 TIEAs, many of the tax havens found it useful to start negotiations with the Nordic countries since this would immediately result in six bilateral TIEAs. As a result, by 2011, 38 TIEAs were signed with tax havens. The project was concluded on November 2015 when the TIEA with the United Arab Emirates was signed. By then Norway and its Nordic neighbours had signed TIEAs with 45 jurisdictions, including all those identified as tax haven by the OECD.

In addition to make information exchange upon request more effective by providing model agreements and urging tax havens to sign TIEAs framed according to the model agreements, the OECD has also played a coordinating role in the process of facilitating the *automatic* exchange of tax information. Automatic exchange of tax information (AEI) is not new from a Nordic perspective. Since the Nordic Convention on Mutual Administrative Assistance in Tax Matters was signed on 7 December 1989, the Nordic countries yearly exchange information about incomes and wealth earned or held by residents of another Nordic state with that state.

Both the Multilateral Convention on Mutual Administrative Assistance in Tax Matters[14] and Article 26 of the OECD 2015 Model Tax Convention allow for the possibility of automatic exchange of tax information. In 2010 US Congress adopted the Foreign Account Tax Compliance Act (FATCA) requiring US persons including those living outside the US to file yearly reports on their non-US financial accounts. It also requires all non-US (foreign) financial institutions to report the assets and identities of US persons to the US Treasury Dept. The act was internationally implemented by intergovernmental agreements where foreign governments committed to collect the required information from financial institutions and transmit it to the US Treasury, in return for receiving the same information about their residents from the US. Norway and the US signed the FATCA on 14 April 2013 which came into force on January 24, 2014. Similar FATCA agreements between the US and France, Germany, Spain, Italy and the UK, made the finance ministers of these latter countries to announce their intention to establish similar exchange of financial account information amongst themselves. In September 2013 the OECD was asked to develop a single global standard for automatic exchange of information, and further political support came at the OECD Ministerial Council meeting in Paris on May 6-7 2014 where 47 countries, including Luxembourg, Switzerland and Singapore signed the Declaration on Automatic exchange of Information in Tax Matters.

The standard for exchanging information was approved by the OECD Council on July 15 2014. On October 29 2014, at the Berlin meeting of the Global Forum on Transparency and Exchange of Information for Tax Purposes, 51 jurisdictions including Luxembourg and Singapore signed the Multilateral Competent Authority Agreement to legally exchange information according to the OECD

[13] Although, for constitutional reasons, the TIEAs could only be entered on a bilateral basis.

[14] This convention was open for signing since 25 January 1988. It was amended by a protocol in 2010 to align it to the international standard of exchange of information upon request.

standard. Switzerland signed the agreement three weeks later (November 19 2014). A year later, the number of signatories was 74 and it raised to 101 by December 22 2016. It is important to note that this does not imply the automatic exchange of information between all 5050 pairs of jurisdictions. It is only when two jurisdictions have each other on their list of partners they wish to exchange information with that a bilateral relationship is established.[15] E.g., in a press statement on October 8 2014, the Swiss Federal Council states that Switzerland will in the first place enter into bilateral exchange relationships with the EU countries and the US (FATCA). Regarding other countries, "consideration will be given to countries with which there are close economic and political ties and which provide their taxpayers with *sufficient scope for regularization* and which are considered to be important and promising in terms of their market potential for Switzerland's financial industry (Swiss Federal Swiss-Federal-Counsil, 2014). In a press release on July 6 2016, the Swiss Federal Council announced that Switzerland will adopt the AEI standard with eight countries, including Norway. However, it was already clear in January 2016 that negotiations between Norway and Switzerland about AEI were successfully concluded, as announced in a press statement released by the Norwegian Ministry of Finance on January 20 2016 that the Minister of Finance and the Swiss Ambassador have signed a declaration of intention to automatically exchange financial account information.

5 Empirical analysis

We address two empirical questions related to voluntary disclosure of previously undeclared wealth and income. The first is whether TIEAs between Norway and tax havens had an effect on amnesty applications from these countries, and especially whether there is an announcement effect as predicted by our dynamic tax amnesty model. We use a linear difference in differences model to address this question, exploiting the fact that Norway signed bilateral information exchange agreements with other countries at different dates. We also develop and use a statistical model, more in line with our theoretical model, allowing for a decreasing pool of potential users of the voluntary disclosure program.

The second question we address is whether information campaigns about audit risks and the VD option provided by the Tax Assessment Law had an effect on the use of voluntary disclosure. The tax administration regularly sends out press releases where they inform Norwegian tax payers (i) about the (increasing) risk of holding undeclared wealth and income in tax havens, and (ii) about the possibility to avoid penalty taxes by voluntarily disclosing previously hidden wealth and income. These information campaigns affect all tax payers who have hidden income and/or wealth abroad, hence there is no natural comparison group to use to estimate the effect of these interventions and we use event study methods to estimate their impact.

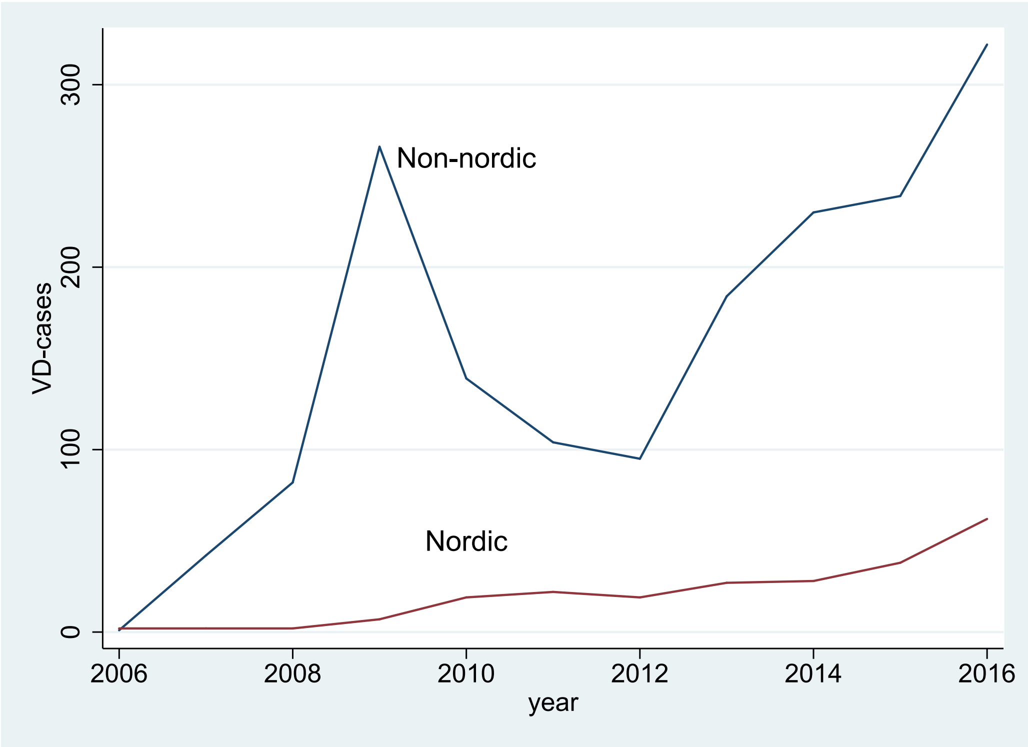
[15] By December 22 2016, 1300 bilateral relationships have been established.

5.1 The effect of TIEAs.

5.1.1 Graphical evidence

There are two waves of TIEAs in the period covered by our data. The first made exchange of information more operational by updating old style "Article 26" in bilateral tax agreements, or by signing explicit agreements to exchange information "upon request". For Norway, this wave started to develop early 2008 and reached its momentum in 2012. The second wave started in the spring of 2013, with the signing of the FATCA agreement with the US, enabling automatic exchange of information on financial accounts held by Norwegian citizens in the US. This agreement came into force in January 2014. At the end of 2018 Norway has automatic exchange of financial information with 58 other jurisdictions, among these Switzerland and Luxembourg.

Figure 3 depicts yearly VD-applications over the entire time-period of our data.



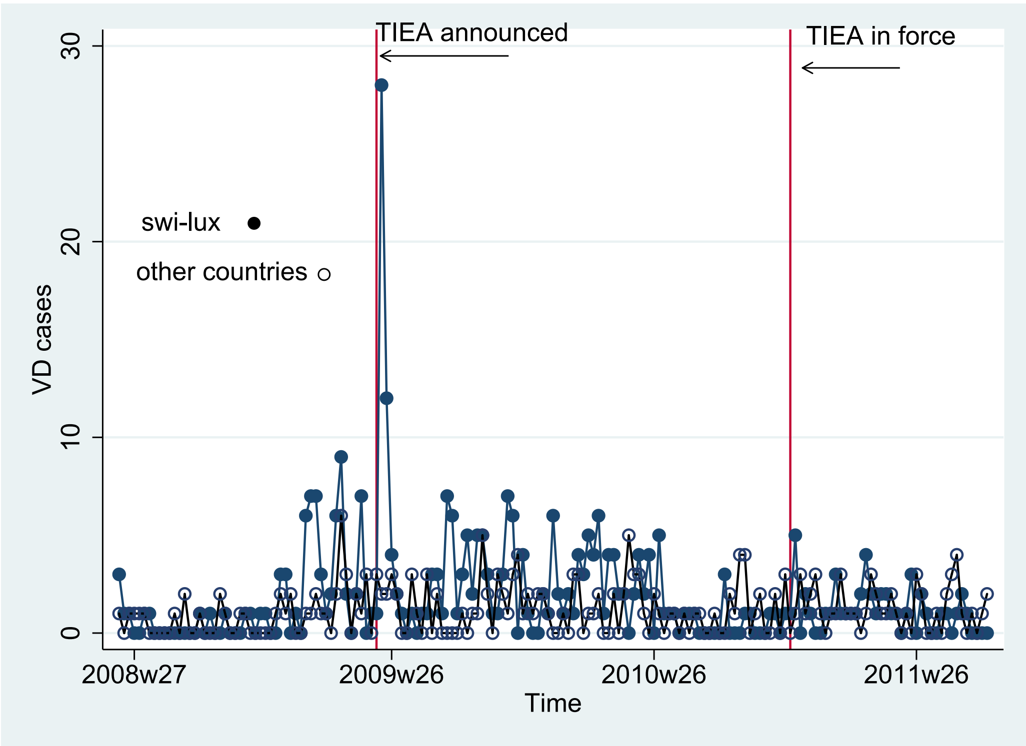
Note: We plot yearly VD applications from Norwegian taxpayers who had wealth or income hidden either in Nordic countries or in non-Nordic countries.

Figure 3. The time profile of VD applications

The distribution of VD-applications from non-Nordic countries have a bimodal shape, mirroring the two waves of TIEAs described above. Among the Nordic countries there have been automatic exchange of tax information over the entire period covered in this graph. The time trend of VD-applications from the Nordic countries has a different pattern, a more monotone increase in VD-applications over time, probably reflecting increased awareness of the VD option over time.

The patterns in Figure 3 support – at an aggregate level – the conjecture that TIEAs increase the perceived audit probability and the risk of hiding taxable wealth or income in a foreign country. To examine the causal effect of the announcement and implementation of TIEAs on VD requires a more focused study, zooming in on the relevant dates. It is especially the first wave of TIEAs, operationalizing information exchange upon request, that have clearly defined dates for the announcement and implementation of different exchange agreements. Information about the introduction of automatic exchange of information has had a more gradual dissemination. It is therefore hard to pin down some specific dates when potential VD applicants should update their information that there would be automatic exchange of tax relevant information from tax havens at some future date.

For the first wave of TIEAs the central date is the 17th of June 2009. The dominant source country for VD-applications is – by a large margin – Switzerland and the second most important source country is Luxembourg (confer Figure 2). Figure 4 plots two time series, the *swi-lux* series gives the number (per week) of voluntary disclosure of wealth and income that has been hidden in Switzerland and Luxembourg, the other time-series depicts the number of VD-applications from all other countries that did not change their TIEA status with Norway over the period 2008 until 2012.



Note: This figure is based on weekly data. We aggregate all weekly VDs from Switzerland and Luxembourg and compare this series with the aggregate VDs from other countries that did not change their TIEA status with Norway during this period.

Figure 4. VD-applications from Switzerland-Luxembourg against VD-applications from other countries

During this time period Norway entered an agreement with Switzerland and Luxembourg that obliged their banks to exchange tax relevant information “upon request”. The agreement was vigorously announced in a press release by the Norwegian Ministry of Finance (on the 17th of June). This press release was widely published in the Norwegian media. The information exchange agreements between Switzerland and Luxembourg were set in force at the commence of 2011.

The relevant dates are marked in Figure 4. The peak just after the press release indicates a very strong announcement effect of the agreement, while there is no visible response around the data when it was set in force at January the 1st of 2011. Figure 4 clearly shows that it is important to take the announcement of information exchange agreements into account to get a reliable estimate of how these agreements affect foreign tax evasion. We use a difference in differences estimator to assess the impact of the agreement that was announced in June 2009, first in a simple linear framework, thereafter in a statistical model closer to our theoretical framework.

A linear model

We don't need formal statistical analyses to conclude that the announcement of the TIEA agreements between Switzerland, Luxembourg and Norway had a strong immediate impact on tax amnesty applications from these two tax havens, a look at Figure 4 suffices. A more interesting, less obvious, question is whether the TIEAs with Switzerland and Luxembourg had a positive effect on VD-cases also when we consider a longer time interval after the agreement was announced, or set in force. To address this question we report estimation results based on a comparison of the two time series depicted in Figure 4.

The linear model estimates the equation

$$A_{i,t} = \alpha + \beta SWILUX_i + \gamma TIEA_t + \delta SWILUX_i \times TIEA_t + \varepsilon_{i,t} \tag{5.1}$$

$A_{i,t}$ measures the number of VD-applications coming from source country i at time t. Just as in Figure 4 we collapse our data into two time series, one counting voluntary disclosure of wealth and income that has been hidden in Switzerland and/or Luxembourg and one counting applications from all other countries that did not change their TIEA status with Norway for the relevant period. This means that the SWILUX dummy is equal to 1 for Switzerland and Luxembourg and 0 for other countries. The TIEA dummy is equal to 0 before either announcement date (17th of June) or the in-force date (1st of January 2011) and 1 thereafter.

The linear model above does not take into account that over time an increase in VD-applications reduces the number of potential VD-applicants later on (the stock of persons that are hiding wealth or income abroad is decreasing). If there is an announcement effect of a TIEA on VD-applications, the reduction in the stock of potential applicants will contribute to an underestimation of the implementation effect of a previously announced TIEA. We therefore also estimate a model that take this stock effect explicitly into account.

A model with a shrinking stock of potential VD applicants

Our theory model predicts that for each person at risk (who still holds untaxed wealth in a foreign country) there will be an increase in the probability that he or she will apply for VD after the TIEA is announced and when it is set in force. To test these predictions we derive and estimate a model where the probability, rather than the number of application, is modeled, and the stock of potential applicants decreases over time as more and more tax payers use the possibility of VD.

Let N_{it} be the size of the population of individuals with hidden wealth in country $i = 1, 2, ..., H$ at time $t = 1, 2, ..., T$. At time t, $A_{i,t}$ of them apply for tax amnesty. We now assume that the number of applicants is distributed according to

$$\begin{cases} (A_{i,t}|N_{i,t}) \sim Normal(N_{i,t}(\mathbf{x}'_{i,t}\boldsymbol{\beta}), \sigma^2) \\ N_{i,t+1} = N_{i,t} - A_{i,t} \end{cases} \tag{5.2}$$

where $\mathbf{x}_{i,t}$ is a vector of variables for group i at time t. The term $\mathbf{x}'_{i,t}\boldsymbol{\beta}$ is an approximation of the propensity to apply for VD. In the next period of time, $t + 1$, the stock of potential applicants has decreased to $N_{i,t} - A_{i,t}$ [16].

Model (5.2) can be rewritten as

$$A_{i,t} = N_{i,t}(\boldsymbol{x}'_{i,t}\boldsymbol{\beta}) + \varepsilon_{i,t}, \tag{5.3}$$

where $\varepsilon_{i,t} \sim Normal(0, \sigma^2)$.

An estimation procedure for this model is described in Appendix A. It is based on iterated ordinary least squares where, given $\boldsymbol{\beta}$, the starting sizes of the stock, $N_{i,1}$, $i = 1, ..., H$, is estimated. In the second step these starting sizes are held fixed and $\boldsymbol{\beta}$ is estimated. This is repeated until convergence.

Results

The variable TIEA in model (5.1) is represented in one of two ways in the empirical analysis; either by a dummy, ANN, which is zero before and one after the announcement of the TIEA or the dummy INFOR, which is zero before and one after the TIEA is put into force.

All models in this section have been estimated on a window of two years of data, in total 102 weeks, around the announcement or implementation date for the TIEA with Switzerland and Luxembourg[17]. The first two columns of Table 1 report the results from the linear model. The *announcement* variable is 0 before the 17th of June 2009 and 1 after. The interaction between ANN and SWILUX captures the difference in differences estimate - and under the assumption that the development in VD-applications from other countries captures the counterfactual development for Switzerland and Luxembourg – the causal effect of the announcement of the TIEA. Likewise for the INFOR variable that switches to 1 when the agreement was legally in force.

[16] Of these, $N_{i,t} - A_{i,t}$ potential applicants, we assume that there are a negligible number who are caught in audits

[17] Estimation of the models have been performed in R, R-Core-Team (2018)

The estimates in column 1a show that the reform, on average, generated 1.42 additional VD-applications from Switzerland and Luxembourg per week in the year that followed the announcement (the diff-in-diff effect). Column 1b shows the estimates of the shrinking stock model. The corresponding TIEA effects are more difficult to interpret because that model considers the conditional mean of the propensity to apply for VD, not the absolute number of applications. At the time of the announcement, the coefficient for SWILUX × ANN, 0.0211, can be translated into a marginal effect of 3.90 additional VD-applications.[18] The large deviation from the effect in the linear model is explained by the fact that the linear model gives an average effect over the entire time period while the shrinking stock model yields the effect at a particular point in time. However, we can compute the implied average diff-in-diff effect from the shrinking stock model. We do this by taking for each group of countries the difference between the average number of predicted applications before and after the announcements, and next taking the difference between the two groups. The shrinking stock model yields an estimate of 1.42 extra VDs from Switzerland/Luxembourg after the announcement of the TIEA, an estimate that is identical to that of the linear model.

Using approximately the same time span before and after the in-force date, we find a negative but insignificant estimate. For the in-force case, the shrinking stock model predicts average diff-in-diff effect of -0.44, compared to -0.50 for the linear model.[19]

[18]The computation is made by multiplying the estimated $N_{\text{SWILUX},t}$ at the time of the announcement with the estimated coefficient: $(260 - 75) \times 0.0211 \approx 3.90$ additional VD-applications from Switzerland and Luxembourg.

[19]N_1 is the estimate of the 2008 population of taxpayers eligible for VD. This estimate would be unbiased if (i) we include all the data from 2008 until 2017 and (ii) our probability model for the propensity to use VD was correctly specified. We estimate that the number of taxpayers with hidden wealth or income in Switzerland or Luxembourg on June 17, 2008, is 259. We know that this number is far too low, since by 2017 there have been more than 900 VD cases from those two countries. The main reason why our estimate is too low is because we use data only until 2012. Our aim is not to find a valid estimate of the stock of hidden wealth, but rather to use a framework, a simple generalisation of the linear model, to account for a shrinking stock of hidden wealth.

Table 1. The linear and the shrinking stock model

	Fixed effects model		**Shrinking stock model**	
	Announce (1a)	In force (2a)	Announce (1b)	In force (2b)
Intercept			0.0001	0.0142***
			(0.0029)	(0.0036)
Ann	0.1592		-0.0019	
	(1.0228)		(0.0086)	
Swilux × Ann	1.4246*		0.0211**	
	(0.8082)		(0.0106)	
Infor		1.2194***		0.0140**
		(0.3654)		(0.0058)
Infor × Swilux		-0.5000		-0.0054
		(0.3699)		(0.0049)
Week	6.9679***	-4.8142***	0.0209	-0.0161**
	(2.1239)	(1.6086)	(0.0185)	(0.0054)
Weeksq	-6.2383***	2.3148*	0.0101	
	(2.9515)	(1.2929)	(0.0328)	
Sigma			2.3832	1.2533
			(0.1180)	(0.0620)
N1 Other			145.0360	135.0887
			(36.8879)	(34.4725)
N1 Swilux			259.2883	208.2693
			(19.2274)	(60.6266)
Mse	6.1340	1.5945	6.0048	1.6338
Observations	204	208	204	208

Note: All models, with the exception of (2b) are estimated with a quadratic trend. Model (2b) is estimated with a linear trend. Standard errors in parenthesis. (1) uses data one year before and after the announcement (17th of June 2009), (2) uses data one year before and after the date when the agreement was set in force (January 1st 2011). The variable WEEK is normalized to $t/102$ where $t = 1, 2..102$ in $1a$ and $2a$ and as $t/104$ where $t = 1, 2..104$ in $1b$ and $2b$. For the linear model robust, and for the approximated binomial asymptotic, standard errors are presented in parentheses * p<0.10, ** p<0.05, *** p<0.01.

To summarize, we observe (from Figure 4) a large immediate impact of the announcement of the TIEA between Switzerland-Luxembourg and Norway. Using the development in VDs from other countries (not signing TIEAs with Norway in that period) as a counterfactual we estimate that effect to be around 1.4 additional VDs per week in the year following the announcement. The announcement effect is in accordance with our model.

Our theoretical model also predicts an upward jump in VDs when the agreement is set into force. We do not observe this in the data. The strong announcement effect and the absence of an "in force" effect could be explained by the differences in publicity of these events. The agreement was vigorously announced. On the 17th of June 2009, the Norwegian finance minister declared in a press release, picked up by the main news-paper in Norway, that with this agreement it was no longer possible for Norwegians to hide financial assets in Swiss banks; "the era of Switzerland as a tax haven is over".

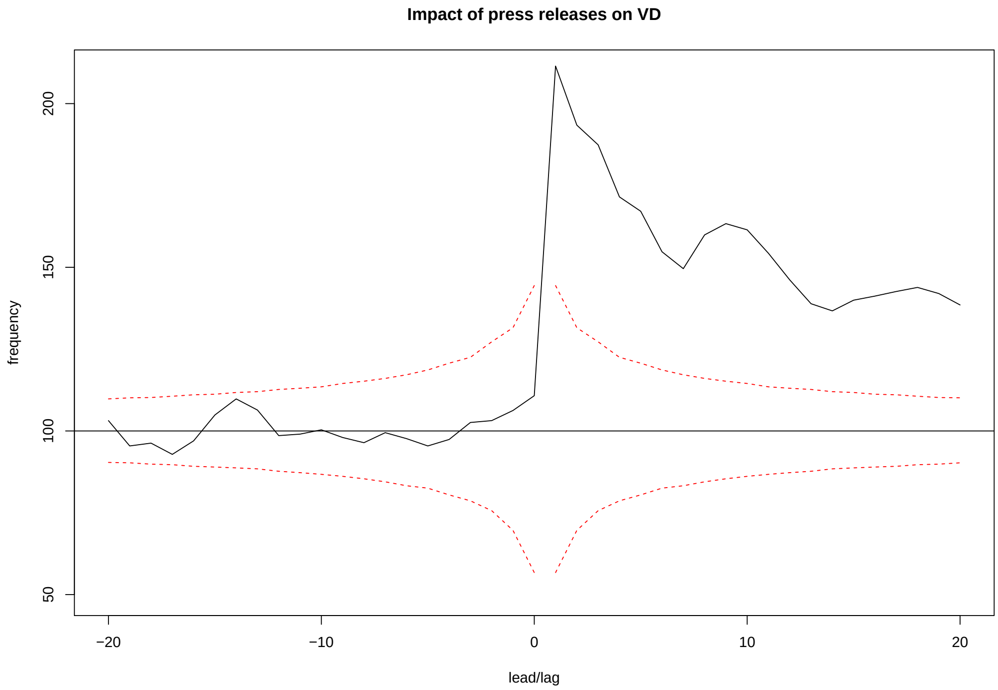
This announcement was an exaggeration, and in the following weeks there appeared newspaper articles pointing out that retrieving financial information from Swiss banks would not be as easy as indicated by the finance minister at the announcement. Maybe the forceful announcement of the TIEA with Switzerland and Luxembourg, both in loudness and content, led potential VD applicants to immediately adjust the audit probability sharply upwards, towards one. When more nuanced information became available, those still evading adjusted the audit probability downwards again. This, and also the fact that there was absolutely no media coverage of the event when the agreement was set in force, could explain a strong immediate announcement effect and no significant effect of making the agreement operational.

5.2 Information campaigns

In this section we consider the development in the aggregate time series of VD-applications and we estimate the impact of events that could affect individual taxpayer's inclination to make use of the VD clause. We are especially interested in the question whether policy campaigns to spread information about the VD-option had an effect on the use of this alternative.

Starting in 2010, Tax Norway regularly issues a press release informing the public about taxpayers use of the voluntary disclosure program. Between beginning of 2010 and mid-2016 we observed 1615 amnesty applications, i.e., on average, 0.68 applications per day. In the same period there are 28 press releases; 315 weeks without any press release, 23 weeks with one press release and 1 week with two and three press releases.

To investigate the effect of these press releases, we developed a graphical technique to investigate if they had a significant impact on the frequency of voluntary disclosure. The underlying idea is that if the press releases have no effect, the VD-frequency around press releases should not, systematically, deviate from the frequency at other time points. We measure the frequency as the daily average number of VD-applications during the window of k=1,2,. . . ,20 days before and after a press release. The method is explained in Appendix C.



Note: We realign the dates for the different press releases and plot the aggregate number of VD cases 20 days before and after the press release. The dashed lines indicate the 95 % confidence interval for the aggregate development in VD.

Figure 5. Impact plot for the cumulative effect of press releases.

As can be seen in the Figure 5, the frequency of voluntary disclosures are systematically above the average frequency after the press releases. This is indeed a clearly systematic pattern showing that the days of and after a press release are by no means ordinary days in terms of the frequency of voluntary disclosures. The analysis is exploratory in nature since no controls are used. As an example, some of the press releases come in the period of due date for tax returns (month of April), something which in itself is a potential reminder for tax payers having undeclared assets or income.

In order to control for additional factors we fit a zero-inflated Poisson (ZIP) regression to the time series of VDs. The model can be written as

$$P(A_t = a_t|\boldsymbol{x}_t) = \pi I(A_t = 0) + (1-\pi)p_{Po}(A_t = a_t|\boldsymbol{x}_t),$$

where p_{Po} is the probability mass function for a Poisson distribution and $I(A_t = 0)$ is one for $A_t = 0$ and zero otherwise, and π is the probability of zero applications in a given period. In line with standard Poisson regression, the ZIP-regression is specified so that the expectation of the p_{Po}-distribution, μ_t, is

$$\mu_t = \exp(\boldsymbol{x}_t'\boldsymbol{\beta}).$$

The probability π could be modelled as a function of explanatory variables; however, this is not pursued in this paper. [20]

The reason for using a zero-inflated distribution is that the fraction of days with no voluntary disclosure is 62%, one disclosure 21% and there is a mean of 0.68 disclosures per day. This set of facts is irreconcilable with a Poisson-distribution. Figure 6 in Appendix D clearly shows that the zero-inflated Poisson distribution is a more realistic fit to the data.

The motivation given is based on the unconditional distribution of voluntary disclosures and not the distribution conditional on explanatory variables. However, there is no reason to believe that explanatory variables would make the standard Poisson distribution fit any better.

In addition to study the relationship on daily data we investigate if there are more persistent effects by aggregating the data to a weekly frequency. The results are shown in Table 2.

[20] To perform the estimation, we use the R-package pscl, see Zeileis et al. (2008). For the HAC, heteroscedastic and autocorrelation corrected robust standard errors, the package sandwich is used (Zeileis, 2004, 2006).

Table 2. The effect of press releases on VD

	Daily data	**Weekly data**
Press Release	0.372*	0.1975**
	(0.216)	(0.0883)
Press Release[-1]	0.0033	0.0289
	(0.1564)	(0.1235)
Press Release[-2]	0.4773***	0.0926
	(0.1638)	(0.1435)
Press Release[-3]	−0.2371**	0.1661
	(0.1033)	(0.1789)
Press Release[-4]	0.3825	0.1493*
	(0.3046)	(0.0783)
Press Release[-5]	−0.1660	0.1257
	(0.5110)	(0.1369)
Press Release[-6]	−0.0705	
	(0.1548)	
Press Release[-7]	0.3493**	
	(0.1426)	
Tax return	0.6865***	0.5861**
	(0.1383)	(0.1980)
Week-end	−1.7571***	
	(0.1300)	
Constant	0.2026***	1.4726***
	(0.0534)	(0.0871)
$\log(\pi/1-\pi)$	-0.6311***	-2.6697***
	(0.1017)	(0.2639)
R^2	0.22	0.44
Observations	2373	340

Note: ZIP-regression of VDs on press releases from Tax Norway. We use data from 01.01.2010 until 30.06.2016. The lags of the press releases are in days for the daily data and in weeks for the weekly data. We control for weekends when using daily data and the tax return period, i.e. the month of April. We have also estimated the model with year dummies or quadratic time trends and that had only negligible effects on the coefficients reported here. Robust HAC standard errors are presented in parentheses * p<0.10, ** p<0.05, *** p<0.01.

As expected based on the impact plot, the immediate effect of a press release on the frequency of voluntary disclosures is positive. In the following, the marginal effects presented are all conditional on the event $A_t > 0$. Everything else equal, the frequency of VD is 45% higher ($\exp(0.373) = 1.45$) than on a day without a press release. This effect is significant at the 10%-level. In the week after a press release there is only one more day that has a significant marginal effect. The two-day after marginal effect of a press release on VD is significant and corresponds to an increase of 61%. The marginal effect three days after is significant and negative; a decrease of 22%. During a weekend the number of VD per day is only, on average, 17% of the frequency during a weekday, everything else equal. In April, the period when the deadline for annual tax returns is approaching, the daily frequency of VD is 98% higher than during other days. The estimate of π in the model is 0.35, which might look small in comparison to the fraction of days without any VDs in the data, 0.62. However, since zero can also occur within the Poisson model, the ZIP-model implies a probability of zero which is $\pi + (1 - \pi)p_{Po}(A_t = 0|\mathbf{x}_t)$ which is larger than π, how much larger depends on $\mathbf{x}_t$.

In order to study the more long-term effects we aggregate the data to weekly frequencies and perform the same analysis as we did for the daily data, including 5 (weekly) lags. The results show that during weeks with a press release, the frequency of VD is, on average, 22% higher than in weeks without, everything else equal and the effect is significant. The increased frequency four weeks after is significant and 16%. Thus, there is an indication of an effect of a press release which is delayed three to four week. In the tax return period the weekly frequency of VDs is estimated to be 80% higher than in other parts of the year, everything else equal. The estimate of π in the weekly model is 0.06.

6 Conclusion

We develop a dynamic model of tax amnesty with a key insight that applying for amnesty becomes more attractive at the moment when stricter enforcement is announced, even if the implementation of the policy is in the distant future. We use this model to sharpen our understanding of how international tax information exchange agreements affect voluntary disclosure of wealth and income, previously hidden in tax havens. Our data is from Norway. In accordance with the dynamic amnesty model we observe a strong announcement effect of a tax information exchange agreement between Norway and Switzerland and Luxembourg, the by far two most important tax havens for Norwegian tax evaders. However, the effect levels off very quickly, and much faster than our model predicts. We think this is because the initial announcement of the tax agreement exaggerated the risk the agreement imposed to those who had hidden taxable income and wealth in Switzerland and Luxembourg.

We also test the effect of press releases from the Norwegian tax authorities, reminding individuals that there is a voluntary disclosure section in the Norwegian tax law and that increasing international transparency makes it more and more risky to hide money abroad. We find a positive effect of these press releases, of such a magnitude that it surely covers their costs. Our data does not enable us to estimate the effect of the OECD initiated reform of automatic exchange of tax relevant information. That is left for future work.

References

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

Andreoni, J. (1991). The desirability of a permanent tax amnesty. *Journal of Public Economics 45*(2), 143–159.

Blundell, R. W., M. Francesconi, and W. Van der Klaauw (2011). Anatomy of welfare reform evaluation: announcement and implementation effects. Technical report.

Caballé, J. and A. Pomansky (1996). Mixed risk aversion. *Journal of Economic Theory 71*(2), 485–513.

Engel, E. M. and J. R. Hines Jr (1999). Understanding tax evasion dynamics. Technical report, National Bureau of Economic Research.

Hanlon, M., E. L. Maydew, and J. R. Thornock (2015). Taking the long way home: Us tax evasion and offshore investments in us equity and debt markets. *The Journal of Finance 70*(1), 257–287.

Johannesen, N. and G. Zucman (2014). The end of bank secrecy? an evaluation of the g20 tax haven crackdown. *American Economic Journal: Economic Policy 6*(1), 65–91.

Keen, M. and J. E. Ligthart (2006). Information sharing and international taxation: a primer. *International Tax and Public Finance 13*(1), 81–110.

Langenmayr, D. (2017). Voluntary disclosure of evaded taxes—increasing revenue, or increasing incentives to evade? *Journal of Public Economics 151*, 110–125.

OECD (2012). Automatic exchange of information. what it is, how it works, benefits, what remains to be done. Technical report.

OECD (2015). Update on voluntary disclosure programmes: A pathway to tax compliance. Technical report, OECD: Paris.

OECD (2016). Automatic exchange of financial account information–background information brief. Technical report, OECD:Paris.

R-Core-Team (2018). R: A language and environment for statistical computing; 2015.

Swiss-Federal-Counsil (2014). Automatic exchange of information in tax matters: Federal council adopts negotiation mandates with partner states.

Thi, H. H. N. and T. Nikolka (2016). The development of tax transparency in oecd countries. *CESifo DICE Report 14*(4), 85–89.

Zeileis, A. (2004). Econometric computing with hc and hac covariance matrix estimators. *Journal of Statistical Software 11*, 1–17.

Zeileis, A. (2006). Object-oriented computation of sandwich estimators. *Journal of Statistical Software 16*(9), 1–16.

Zeileis, A., C. Kleiber, and S. Jackman (2008). Regression models for count data in r. *Journal of Statistical Software 27*(8), 1–25.

A Theoretical results

Lemma 1. *Suppose there is a critical value for the shock* $\widehat{\varepsilon}^{stat}(p)$ *for which* $\Omega(\widehat{\varepsilon}^{stat}(p),p)\equiv 0$. *Under decreasing absolute risk aversion,* $\Omega_\varepsilon(\varepsilon,p)|_{\varepsilon=\widehat{\varepsilon}^{stat}}<0$ *and the critical value is unique with* $\Omega(\varepsilon,p)<(>)0$ *for all* $\varepsilon<(>)\widehat{\varepsilon}^{stat}(p)$.

Proof

In the static model, define $\widehat{\varepsilon}^{stat}(p)$ as $\Omega(\widehat{\varepsilon}^{stat}(p),p)\equiv 0$. Hence

$$u(w^n+\widehat{\varepsilon}^{stat})=Eu(\widetilde{c}+\widehat{\varepsilon}^{stat}), \tag{A.1}$$

where $\widetilde{c}=(w^n+tx,w^n-\gamma tx;1-p,p)$, the gamble that a tax evader faces. By definition of the risk premium (Π),

$$Eu(\widetilde{c}+\widehat{\varepsilon}^{stat})=u(E\widetilde{c}+\widehat{\varepsilon}^{stat}-\Pi), \tag{A.2}$$

and by definition of the prudence premium (Ψ) for the gamble that the evader faces,

$$Eu'(\widetilde{c}+\widehat{\varepsilon}^{stat})=u'(E\widetilde{c}+\widehat{\varepsilon}^{stat}-\Psi). \tag{A.3}$$

We are interested in the sign of $\Omega_\varepsilon(\varepsilon,p)|_{\varepsilon=\widehat{\varepsilon}^{stat}}$:

$$\Omega_\varepsilon(\varepsilon,p)|_{\varepsilon=\widehat{\varepsilon}^{stat}}=u'(w^n-(1-\alpha)\gamma x+\widehat{\varepsilon}^{stat})-Eu'(\widetilde{c}+\widehat{\varepsilon}^{stat}). \tag{A.4}$$

From (A.1) and (A.2), it follows that $u'(w^n-(1-\alpha)\gamma x+\widehat{\varepsilon}^{stat})=u'(E\widetilde{c}+\widehat{\varepsilon}^{stat}-\Pi)$. Then using (A.3), we can rewrite (A.4) as

$$\Omega_\varepsilon(\varepsilon,p)|_{\varepsilon=\widehat{\varepsilon}^{stat}}=u'(E\widetilde{c}+\widehat{\varepsilon}^{stat}-\Pi)-u'(E\widetilde{c}+\widehat{\varepsilon}^{stat}-\Psi).$$

Hence $\Omega_\varepsilon(\varepsilon,p)|_{\varepsilon=\widehat{\varepsilon}^{stat}}\gtrless 0$ if and only if $\Pi\gtrless\Psi$. Since the risk (prudence) premium is proportional to the coefficient of absolute risk aversion R_a (absolute prudence, P_a), we have $\Omega_\varepsilon(\varepsilon,p)|_{\varepsilon=\widehat{\varepsilon}^{stat}}\gtrless 0$ if and only if $R_a\gtrless P_a$ (both evaluated at consumption level $E\widetilde{c}+\widehat{\varepsilon}^{stat}$). Under risk aversion, the coefficient of absolute risk aversion falls in consumption if and only R_a exceeds P_a. Hence DARA is necessary and sufficient for $\Omega_\varepsilon(\varepsilon,p)|_{\varepsilon=\widehat{\varepsilon}^{stat}}<0$.

Suppose now that there exists a second level for ε, $\widehat{\widehat{\varepsilon}}^{stat}$, such that $\Omega(\widehat{\widehat{\varepsilon}}^{stat},p)=0$. Then by the previous argument, also $\Omega_\varepsilon(\varepsilon,p)|_{\varepsilon=\widehat{\widehat{\varepsilon}}^{stat}}<0$. But since $\Omega(\cdot,p)$ is continuous in ε, there should be a third critical value at which $\Omega_\varepsilon>0$, which is not possible. Hence, there will exists a unique $\widehat{\varepsilon}^{stat}(p)$ for which $\Omega(\widehat{\varepsilon}^{stat}(p),p)\equiv 0$. For all $\varepsilon<(>)\widehat{\varepsilon}^{stat}(p)$, $\Omega(\varepsilon,p)>(<)0$.■

Lemma 2. *If $u(\cdot)$ is completely monotone*[21] *and belongs to the class of utility functions satisfying*

$$A_1(c) = -\frac{u^{(2)}(c)}{u^{(1)}(c)} < A_n = -\frac{u^{(n+1)}(c)}{u^{(n)}(c)} \text{ for all } c \text{ and } n \geq 2.$$

then $\Omega(\cdot, p)$ is monotonically decreasing in ε over $(\gamma tx - w^n, \widehat{\varepsilon}^{stat}]$.

Proof

To simplify notation, let $\Omega(\varepsilon) = \Omega(\varepsilon, p)$. Take a Taylor expansion of $\Omega'(\varepsilon)$ around $\widehat{\varepsilon}^{stat}$:

$$\begin{aligned}\Omega'(\varepsilon) &= \Omega'(\widehat{\varepsilon}^{stat}) + \Omega''(\widehat{\varepsilon}^{stat})(\varepsilon - \widehat{\varepsilon}^{stat}) + \frac{1}{2}\Omega'''(\widehat{\varepsilon}^{stat})(\varepsilon - \widehat{\varepsilon}^{stat})^2 + \frac{1}{6}\Omega''''(\widehat{\varepsilon}^{stat})(\varepsilon - \widehat{\varepsilon}^{stat})^3 + ... \\ &= \sum_{n=1}^{\infty} \frac{1}{(n-1)!}\Omega^{(n)}(\widehat{\varepsilon}^{stat})(\varepsilon - \widehat{\varepsilon}^{stat})^{n-1}.\end{aligned}$$

By definition: $\Omega(\widehat{\varepsilon}^{stat}) = 0$, i.e., $u(w^n + \widehat{\varepsilon}^{stat}) = Eu(\widetilde{c} + \widehat{\varepsilon}^{stat}) = u(E\widetilde{c} + \widehat{\varepsilon}^{stat} - \Pi)$ where Π is the risk premium. Notational convention: let $\Pi = \Pi_1$. Therefore $u^{(n)}(w^n + \widehat{\varepsilon}^{stat}) = u^{(n)}(E\widetilde{c} + \widehat{\varepsilon}^{stat} - \Pi_1)$.

By definition, $Eu^{(n)}(\widetilde{c} + \widehat{\varepsilon}^{stat}) = u^{(n)}(E\widetilde{c} + \widehat{\varepsilon}^{stat} - \Pi_{n+1})$ (thus Π_2 corresponds to the prudence premium, Π_3 to the temperance premium, Π_4 to the edginess premium, etc).

Then

$$\Omega^{(n)}(\widehat{\varepsilon}^{stat}) = u^{(n)}(E\widetilde{c} + \widehat{\varepsilon}^{stat} - \Pi_1) - u^{(n)}(E\widetilde{c} + \widehat{\varepsilon}^{stat} - \Pi_{n+1}) \gtrless 0$$
$$\Updownarrow \; u(\cdot) \text{ completely monotone}$$
$$\Pi_1 \gtrless \Pi_{n+1} \text{ if } n \text{ is odd and } \Pi_1 \lesseqgtr \Pi_{n+1} \text{ if } n \text{ is even,}$$
$$\Updownarrow$$
$$A_1 \gtrless A_{n+1} \text{ if } n \text{ is odd, and } A_1 \lesseqgtr A_{n+1} \text{ if } n \text{ is even,}$$

so that $\text{sign}\Big(\Omega^{(n)}(\widehat{\varepsilon}^{stat})\Big) = \text{sign}(A_1 - A_{n+1}) \times (-1)^{n+1}$.

If $u(\cdot)$ belongs to class with $A_1 < A_2, A_3, ...$ then $\text{sign}\Big(\Omega^{(n)}(\widehat{\varepsilon}^{stat})\Big) = (-1)^n$. If $\varepsilon < \widehat{\varepsilon}^{stat}$, then $\text{sign}[(\varepsilon - \widehat{\varepsilon}^{stat})^{n-1}] = (-1)^{n-1}$. Therefore $\text{sign}[\Omega^{(n)}(\widehat{\varepsilon}^{stat})(\varepsilon - \widehat{\varepsilon}^{stat})^{n-1}] = (-1)$. It then follows that $\Omega'(\varepsilon) < 0$ for all $\varepsilon < \widehat{\varepsilon}^{stat}$. ■

Lemma 3. $\lim_{\varepsilon \to \infty} \Omega(\varepsilon, p) \leq 0$.

Proof

Taking a first order Taylor expansion of $pu(w^n + tx + \varepsilon)$ and $(1-p)u(w^n - \gamma tx + \varepsilon)$ around $w^n + \varepsilon$, gives

$$pu(w^n + tx + \varepsilon) + (1-p)u(w^n - \gamma tx + \varepsilon) = u(w^n + \varepsilon) + [(1-p)u'(\overline{c}(\varepsilon))tx - pu'(\underline{c}(\varepsilon))\gamma tx$$

[21] A function is completely monotone if it is increasing and the signs of its derivatives are constant and alternate in sign.

where $\overline{c}(\varepsilon)$ $[\underline{c}(\varepsilon)]$ is some consumption level between $w^n + \varepsilon$ and $w^n + \varepsilon + tx$ $[w^n + \varepsilon - \gamma tx]$. Then

$$\lim_{\varepsilon \to \infty} \Omega(\varepsilon, p) = -\lim_{\varepsilon \to \infty} [(1-p)u'(\overline{c}(\varepsilon))tx - pu'(\underline{c}(\varepsilon))\gamma tx]$$

By assumption $\lim_{c \to \infty} u'(c) = \mu$ with $\mu \in [0, \infty)$. Also by assumption, before the realisation of the income shock, the tax payer considers evasion to be optimal: $(1-p)t - p\gamma t > 0$. Therefore $\lim_{\varepsilon \to \infty} \Omega(\varepsilon, p) = -[(1-p) - p\gamma]tx\mu \leq 0$..

Corollary 1. *If $u(\cdot)$ is completely monotone and belongs to the class of utility functions with $A_1 < A_n$ (all $n \geq 2$) then $\Omega(\cdot, p)$ is monotonically decreasing in ε over an interval strictly including $(\gamma tx - w^n, \widehat{\varepsilon}^{stat}]$, and does not exceeds zero on $(\widehat{\varepsilon}^{stat}, \infty)$.*

Lemma 4. *The region of ε-values for which applying for tax amnesty is optimal is a connected subset of $(-\infty, \widehat{\varepsilon}^{stat})$.*

Proof

The dynamic programming problem (2.1) is an optimal stopping problem and can be rewritten as

$$V(0, \varepsilon, p, p') - V(1, \varepsilon, p, p') = \max\{0, U^N(\varepsilon, p) - U^A(\varepsilon) + (1-p)\beta[E_{\widetilde{\varepsilon}}V(0, \widetilde{\varepsilon}, p', p') - E_{\widetilde{\varepsilon}}V(1, \widetilde{\varepsilon}, p', p')]\} \tag{A.5}$$

Define $\Delta(\varepsilon, p, p') \stackrel{\text{def}}{=} V(0, \varepsilon, p, p') - V(1, \varepsilon, p, p')$ then

$$\Delta(\varepsilon, p, p') = \max\{0, -\Omega(\varepsilon, p) + (1-p)\beta E_{\widetilde{\varepsilon}}\Delta(\widetilde{\varepsilon}, p', p')\} \tag{A.6}$$

It follows immediately that $\Delta(\varepsilon, p, p')$ and therefore $E_{\widetilde{\varepsilon}}\Delta(\widetilde{\varepsilon}, p', p')$ are positive. From Corollary 1, $-\Omega(\varepsilon, p)$ is increasing in ε on $(-\infty, \widehat{\varepsilon}^{stat}]$. Therefore, the solution to (A.6), $\Delta(\varepsilon, p, p')$, is non-decreasing in ε on $(-\infty, \widehat{\varepsilon}^{stat}]$. We know that $-\Delta(\varepsilon, p, p') < 0$ iff $\varepsilon < \widehat{\varepsilon}^{stat}$. Suppose that there is a critical ε-value $\widehat{\varepsilon}(p, p')$ for which the 2nd argument in max{} exceeds $\widehat{\varepsilon}^{stat}$. Then there must exist an ε-value, $\widehat{\widehat{\varepsilon}} \in [\widehat{\varepsilon}^{stat}, \widehat{\varepsilon}(p, p')]$ such that $-\Omega(\widehat{\widehat{\varepsilon}}, p) + (1-p)\beta E_{\widetilde{\varepsilon}}\Delta(\widetilde{\varepsilon}, p', p') < 0$. But this contradicts that $\Omega(\widehat{\widehat{\varepsilon}}, p) < 0$. Hence, all critical values for which the 2nd *rhs* of (A.6) are zero must belong to $(-\infty, \widehat{\varepsilon}^{stat}]$. But since $-\Omega(\varepsilon, p)$ is increasing in ε on $(-\infty, \widehat{\varepsilon}^{stat}]$ there can only be one such critical value. ■

Theorem 1. *The optimal decision rule for the dynamic problem is given by (2.2)*

Proof

If it is optimal to apply, then from the first expression in the maximisation problem we get that

$$\begin{aligned} V(0, \varepsilon, p, p') &= U^A(\varepsilon) + \frac{\beta}{1-\beta}E_{\widetilde{\varepsilon}}U^H(\widetilde{\varepsilon}) \\ &\geq U^N(\varepsilon, p) + (1-p)\beta E_{\widetilde{\varepsilon}}V(0, \widetilde{\varepsilon}, p', p') + p\frac{\beta}{1-\beta}E_{\widetilde{\varepsilon}}U^H(\widetilde{\varepsilon}) \end{aligned} \tag{A.7}$$

The tax payer will then apply for amnesty when

$$\Omega(\varepsilon,p) \geq (1-p)\beta\left[E_{\widetilde{\varepsilon}}V(0,\widetilde{\varepsilon},p',p') - \frac{1}{1-\beta}E_{\widetilde{\varepsilon}}U^H(\widetilde{\varepsilon})\right], \tag{A.8}$$

By Lemma 4 (A.8) with equality has one solution $\widehat{\varepsilon}(p,p') < \widehat{\varepsilon}^{stat}(p)$. Consider now the first equality in (A.7), take conditional expectations and replace p by p':

$$E_{\widetilde{\varepsilon}\leq\widehat{\varepsilon}(p',p')}V(0,\widetilde{\varepsilon},p',p') = E_{\widetilde{\varepsilon}\leq\widehat{\varepsilon}(p',p')}U^A(\widetilde{\varepsilon}) + \frac{\beta}{1-\beta}E_{\widetilde{\varepsilon}}U^A(\widetilde{\varepsilon}). \tag{A.9}$$

Next, assume that it is not optimal to apply for amnesty today. Then

$$\begin{gathered} U^A(\varepsilon) + \beta E_{\widetilde{\varepsilon}}V(1,\widetilde{\varepsilon}) \leq U^N(\varepsilon,p) + \\ (1-p)\beta E_{\widetilde{\varepsilon}}V(0,\widetilde{\varepsilon},p',p') + p\beta E_{\widetilde{\varepsilon}}V(1,\widetilde{\varepsilon},p',p') = V(0,\varepsilon,p,p'), \text{ or} \\ U^A(\varepsilon) + \frac{\beta}{1-\beta}E_{\widetilde{\varepsilon}}U^A(\widetilde{\varepsilon}) \leq U^N(\varepsilon,p) + \\ (1-p)\beta E_{\widetilde{\varepsilon}}V(0,\widetilde{\varepsilon},p',p') + p\frac{\beta}{1-\beta}E_{\widetilde{\varepsilon}}U^A(\widetilde{\varepsilon}) = V(0,\varepsilon,p,p'). \end{gathered} \tag{A.10}$$

Thus no application takes place if

$$\begin{array}{rcl} \Omega(\varepsilon,p) & \leq & (1-p)\beta\left[E_{\widetilde{\varepsilon}}V(0,\widetilde{\varepsilon},p',p') - \frac{1}{1-\beta}E_{\widetilde{\varepsilon}}U^A(\widetilde{\varepsilon})\right] \\ & \Updownarrow & \\ \varepsilon & \geq & \widehat{\varepsilon}(p,p') \end{array}$$

Then consider the equality in (A.10), take conditional expectations and replace p by p' :

$$\begin{aligned} E_{\widetilde{\varepsilon}\geq\widehat{\varepsilon}(p',p')}V(0,\varepsilon,p',p') &= E_{\widetilde{\varepsilon}\geq\widehat{\varepsilon}(p',p')}U^N(\varepsilon,p') + (1-p')\beta E_{\widetilde{\varepsilon}}V(0,\widetilde{\varepsilon},p',p') \\ &\quad + p'\frac{\beta}{1-\beta}E_{\widetilde{\varepsilon}}U^A(\widetilde{\varepsilon}). \end{aligned} \tag{A.11}$$

Using (A.9) and (A.11), the unconditional expectation of the maximal utility of an evader is then obtained as

$$
\begin{aligned}
& E_{\widetilde{\varepsilon}} V(0,\widetilde{\varepsilon},p',p') = \\
& E_{\widetilde{\varepsilon}\leq\widehat{\varepsilon}(p',p')} V(0,\widetilde{\varepsilon},p',p') \Pr(\widetilde{\varepsilon}\leq\widehat{\varepsilon}(p',p')) + E_{\widetilde{\varepsilon}\geq\widehat{\varepsilon}(p,p')} V(0,\widetilde{\varepsilon},p',p') \Pr(\widetilde{\varepsilon}\geq\widehat{\varepsilon}(p',p')) \\
& = E_{\widetilde{\varepsilon}\leq\widehat{\varepsilon}(p,p')} U^A(\widetilde{\varepsilon},q) \Pr(\widetilde{\varepsilon}\leq\widehat{\varepsilon}(p',p')) + \frac{\beta}{1-\beta} E_{\widetilde{\varepsilon}} U^A(\widetilde{\varepsilon}) \Pr(\widetilde{\varepsilon}\leq\widehat{\varepsilon}(p',p')) \\
& +E_{\widetilde{\varepsilon}\geq\widehat{\varepsilon}(p,p')} U^N(\widetilde{\varepsilon},p') \Pr(\widetilde{\varepsilon}\geq\widehat{\varepsilon}(p',p')) \\
& +(1-p')\beta E_{\widetilde{\varepsilon}} V(0,\widetilde{\varepsilon},p',p') \Pr(\widetilde{\varepsilon}\geq\widehat{\varepsilon}(p',p')) \\
& +p'\frac{\beta}{1-\beta} E_{\widetilde{\varepsilon}} U^A(\widetilde{\varepsilon}) \Pr(\widetilde{\varepsilon}\geq\widehat{\varepsilon}(p',p'))
\end{aligned}
$$

Solving for $E_{\widetilde{\varepsilon}} V(0,\widetilde{\varepsilon},p',p')$ gives

$$
\begin{aligned}
& E_{\widetilde{\varepsilon}} V(0,\widetilde{\varepsilon},p',p') = \frac{1}{1-(1-p')\beta\Pr(\widetilde{\varepsilon}\geq\widehat{\varepsilon}(p',p'))} \times \\
& \Big\{ E_{\widetilde{\varepsilon}\leq\widehat{\varepsilon}(p',p')} U^A(\widetilde{\varepsilon}) \Pr(\widetilde{\varepsilon}\leq\widehat{\varepsilon}(p',p')) + \frac{\beta}{1-\beta} E_{\widetilde{\varepsilon}} U^A(\widetilde{\varepsilon}) \Pr(\widetilde{\varepsilon}\leq\widehat{\varepsilon}(p',p')) \\
& +E_{\widetilde{\varepsilon}\geq\widehat{\varepsilon}(p',p')} U^N(\widetilde{\varepsilon},p') \Pr(\widetilde{\varepsilon}\geq\widehat{\varepsilon}(p',p')) + p'\frac{\beta}{1-\beta} E_{\widetilde{\varepsilon}} U^A(\widetilde{\varepsilon}) \Pr(\widetilde{\varepsilon}\geq\widehat{\varepsilon}(p',p')) \Big\}.
\end{aligned}
$$

Now go back to inequality (A.8) and make use of the just obtained expression for $E_{\widetilde{\varepsilon}} V(0,\widetilde{\varepsilon},p',p')$:

$$
\begin{aligned}
\Omega(\varepsilon,p) & \geq \beta\,(1-p) \left[E_{\widetilde{\varepsilon}} V(0,\widetilde{\varepsilon},p',p') - \frac{1}{1-\beta} E_{\widetilde{\varepsilon}} U^A(\widetilde{\varepsilon}) \right] \\
& = \frac{\beta\,(1-p)}{1-(1-p')\beta\Pr(\widetilde{\varepsilon}\geq\widehat{\varepsilon}(p',p'))} \times \\
& \Big[E_{\widetilde{\varepsilon}\leq\widehat{\varepsilon}(p',p')} U^A(\widetilde{\varepsilon}) \Pr(\widetilde{\varepsilon}\leq\widehat{\varepsilon}(p',p')) + \frac{\beta\Pr(\widetilde{\varepsilon}\leq\widehat{\varepsilon}(p',p'))}{1-\beta} E_{\widetilde{\varepsilon}} U^A(\widetilde{\varepsilon}) \\
& +E_{\widetilde{\varepsilon}\geq\widehat{\varepsilon}(p',p')} U^N(\widetilde{\varepsilon},p') \Pr(\widetilde{\varepsilon}\geq\widehat{\varepsilon}(p',p')) + p'\frac{\beta}{1-\beta} E_{\widetilde{\varepsilon}} U^A(\widetilde{\varepsilon}) \Pr(\widetilde{\varepsilon}\geq\widehat{\varepsilon}(p',p')) \Big] \\
& -\frac{(1-p)\beta}{1-\beta} E_{\widetilde{\varepsilon}} U^A(\widetilde{\varepsilon}),
\end{aligned}
$$

or

$$
\begin{aligned}
\Omega(\varepsilon,p) & \geq \frac{\beta(1-p)}{1-\beta(1-p')\Pr(\widetilde{\varepsilon}\geq\widehat{\varepsilon}(p',p'))} \times \\
& \Big[E_{\widetilde{\varepsilon}\leq\widehat{\varepsilon}(p',p')} U^A(\widetilde{\varepsilon}) \Pr(\widetilde{\varepsilon}\leq\widehat{\varepsilon}(p',p')) + E_{\widetilde{\varepsilon}>\widehat{\varepsilon}(p',p')} U^N(\widetilde{\varepsilon},p') \Pr(\widetilde{\varepsilon}\geq\widehat{\varepsilon}(p',p')) \Big] \\
& +\frac{\beta(1-p)}{1-\beta} \left[\frac{\beta[\Pr(\widetilde{\varepsilon}\leq\widehat{\varepsilon}(p',p')) + p'\Pr(\widetilde{\varepsilon}\geq\widehat{\varepsilon}(p',p'))}{1-\beta(1-p')\Pr(\widetilde{\varepsilon}\geq\widehat{\varepsilon}(p',p'))} - 1 \right] E_{\widetilde{\varepsilon}} U^A(\widetilde{\varepsilon}) \\
& = \frac{\beta(1-p)}{1-\beta(1-p')\Pr(\widetilde{\varepsilon}\geq\widehat{\varepsilon}(p',p'))} \times \Big[E_{\widetilde{\varepsilon}\leq\widehat{\varepsilon}(p',p')} U^A(\widetilde{\varepsilon}) \Pr(\widetilde{\varepsilon}\leq\widehat{\varepsilon}(p',p')) \\
& +E_{\widetilde{\varepsilon}\geq\widehat{\varepsilon}(p',p')} U^N(\widetilde{\varepsilon},p') \Pr(\widetilde{\varepsilon}\geq\widehat{\varepsilon}(p',p')) - E_{\widetilde{\varepsilon}} U^A(\widetilde{\varepsilon}) \Big],
\end{aligned}
$$

or

$$\Omega(\varepsilon,p) \geq \frac{\beta(1-p)}{1-\beta(1-p')\Pr(\widetilde{\varepsilon}\geq\widehat{\varepsilon}(p',p'))}\times \left[E_{\widetilde{\varepsilon}\geq\widehat{\varepsilon}(p',p')}[U^N(\widetilde{\varepsilon},p')-U^A(\widetilde{\varepsilon})\right]\Pr(\widetilde{\varepsilon}\geq\widehat{\varepsilon}(p',p'))\},$$

or

$$\Omega(\varepsilon,p) \geq \frac{\beta(1-p)\Pr(\widetilde{\varepsilon}\geq\widehat{\varepsilon}(p',p'))}{1-\beta(1-p')\Pr(\widetilde{\varepsilon}\geq\widehat{\varepsilon}(p',p'))}\left[E_{\widetilde{\varepsilon}\geq\widehat{\varepsilon}(p',p')}[-\Omega(\widetilde{\varepsilon},p')]\right]$$

or

$$\begin{aligned}\Omega(\varepsilon,p) &\geq \frac{\beta(1-p)[1-F(\widehat{\varepsilon}(p',p'))]}{1-\beta(1-p')[1-F(\widehat{\varepsilon}(p',p'))]}\left[E_{\widetilde{\varepsilon}\geq\widehat{\varepsilon}(p',p')}[-\Omega(\widetilde{\varepsilon},p')]\right] \\ &\Updownarrow \\ \varepsilon &\leq \widehat{\varepsilon}(p,p') \quad \blacksquare\end{aligned}$$

Proof of Proposition 1.

By definition of $\widehat{\varepsilon}(p,p)$, (2.2) holds with equality for this shock value:

$$\Omega(\widehat{\varepsilon}(p,p),p) = \frac{\beta(1-p)[1-F(\widehat{\varepsilon}(p,p))]}{1-\beta(1-p')[1-F(\widehat{\varepsilon}(p,p))]}\left[E_{\widetilde{\varepsilon}>\widehat{\varepsilon}(p,p)}[-\Omega(\widetilde{\varepsilon},p))\right]\}. \tag{A.12}$$

Suppose the statement is false, i.e., $\widehat{\varepsilon}(p,p)\geq\widehat{\varepsilon}_{stat}(p)$. Because $\Omega(\widehat{\varepsilon}_{stat}(p),p)=0$ and because $\widehat{\varepsilon}(p,p)\geq\widehat{\varepsilon}_{stat}(p)$, it follows that $\Omega(\widehat{\varepsilon}(p,p),p)<0$. Hence, the *lhs* of (A.12) is negative. Consider a future shock draw $\widetilde{\varepsilon}>\widehat{\varepsilon}(p,p)$. Because $\Omega(\widetilde{\varepsilon},p)<0$ we get $E_{\widetilde{\varepsilon}>\widehat{\varepsilon}(p,p)}[-\Omega(\widetilde{\varepsilon},p))>0$. Hence the *rhs* of (A.12) is positive, a contradiction. ■

Proof of proposition 3

The proof is by induction. We first show that the result holds for $T=1$ and $T=2$. We then show that holds for $T\geq 3$.

(i) Proof that $\widehat{\varepsilon}(p_{[1]},p')<\widehat{\varepsilon}(p_{[0]},p')$.

By definition,

$$\Omega(\widehat{\varepsilon}(p_{[0]},p'),p') = (1-p')\int_{\widehat{\varepsilon}(p_{[0]},p')}^{\infty}[\Omega(\widehat{\varepsilon}(p_{[0]},p'),p')-\Omega(\widetilde{\varepsilon},p')]\mathrm{d}F(\widetilde{\varepsilon}) \tag{A.13}$$

$$\Omega(\widehat{\varepsilon}(p_{[1]},p'),p) = (1-p)\int_{\widehat{\varepsilon}(p_{[0]},p')}^{\infty}[\Omega(\widehat{\varepsilon}(p_{[0]},p'),p')-\Omega(\widetilde{\varepsilon},p')]\mathrm{d}F(\widetilde{\varepsilon}) \tag{A.14}$$

Since $\widehat{\varepsilon}(p_{[0]},p')<\widehat{\varepsilon}^{stat}(p')$ (Proposition 1), $\Omega(\widehat{\varepsilon}(p_{[0]},p'),p')\geq\Omega(\widetilde{\varepsilon},p')$ for all $\widetilde{\varepsilon}>\widehat{\varepsilon}(p_{[0]},p')$ (Corollary 1) so that the integral is positive and the *rhs* of (A.14) exceeds the *rhs* of (A.13). Therefore $\Omega(\widehat{\varepsilon}(p_{[0]},p'),p')<\Omega(\widehat{\varepsilon}(p_{[1]},p'),p)$. Because $\Omega(\cdot)$ is falling in ε and increasing in p, it is necessary that $\widehat{\varepsilon}(p_{[0]},p')<\widehat{\varepsilon}(p_{[1]},p')$.

(ii) Proof that $\widehat{\varepsilon}(p_{[2]}, p') < \widehat{\varepsilon}(p_{[1]}, p')$.
By definition,

$$\Omega(\widehat{\varepsilon}(p_{[2]}, p'), p) = (1-p) \int_{\widehat{\varepsilon}(p_{[1]}, p')}^{\infty} [\Omega(\widehat{\varepsilon}(p_{[1]}, p'), p) - \Omega(\widetilde{\varepsilon}, p)] \mathrm{d}F(\widetilde{\varepsilon}). \tag{A.15}$$

Since $\Omega(\cdot)$ is increasing in p, $\Omega(\cdot, p') > \Omega(\cdot, p)$. Therefore

$$\Omega(\widehat{\varepsilon}(p_{[0]}, p'), p') - \Omega(\widetilde{\varepsilon}, p) > \Omega(\widehat{\varepsilon}(p_{[0]}, p'), p') - \Omega(\widetilde{\varepsilon}, p') \text{ for all } \widetilde{\varepsilon}.$$

Earlier, it was shown that $\Omega(\widehat{\varepsilon}(p_{[0]}, p'), p') < \Omega(\widehat{\varepsilon}(p_{[1]}, p'), p)$. As a result

$$\Omega(\widehat{\varepsilon}(p_{[1]}, p'), p) - \Omega(\widetilde{\varepsilon}, p) > \Omega(\widehat{\varepsilon}(p_{[0]}, p'), p') - \Omega(\widetilde{\varepsilon}, p') \text{ for all } \widetilde{\varepsilon}.$$

Therefore

$$\int_{\widehat{\varepsilon}(p_{[0]}, p')}^{\infty} [\Omega(\widehat{\varepsilon}(p_{[1]}, p'), p) - \Omega(\widetilde{\varepsilon}, p)] \mathrm{d}F(\widetilde{\varepsilon}) > \int_{\widehat{\varepsilon}(p_{[0]}, p')}^{\infty} [\Omega(\widehat{\varepsilon}(p_{[0]}, p'), p') - \Omega(\widetilde{\varepsilon}, p')] \mathrm{d}F(\widetilde{\varepsilon}).$$

Moreover, $\Omega(\widetilde{\varepsilon}, p) < \Omega(\widehat{\varepsilon}(p_{[1]}, p'), p)$ for $\widetilde{\varepsilon} \in (\widehat{\varepsilon}(p_{[1]}, p'), \widehat{\varepsilon}(p_{[0]}, p'))$ because $\Omega_{\varepsilon}(\cdot) < 0$ for $\widehat{\varepsilon}(p_{[1]}, p') \leq \widetilde{\varepsilon} \leq \widehat{\varepsilon}(p_{[0]}, p') < \widehat{\varepsilon}^{stat}(p')$. As a result $\int_{\widehat{\varepsilon}(p_{[1]}, p')}^{\widehat{\varepsilon}(p_{[0]}, p')} [\Omega(\widehat{\varepsilon}(p_{[1]}, p'), p) - \Omega(\widetilde{\varepsilon}, p)] \mathrm{d}F(\widetilde{\varepsilon}) > 0$ and therefore

$$\int_{\widehat{\varepsilon}(p_{[1]}, p')}^{\infty} [\Omega(\widehat{\varepsilon}(p_{[1]}, p'), p) - \Omega(\widetilde{\varepsilon}, p)] \mathrm{d}F(\widetilde{\varepsilon}) > \int_{\widehat{\varepsilon}(p_{[0]}, p')}^{\infty} [\Omega(\widehat{\varepsilon}(p_{[0]}, p'), p') - \Omega(\widetilde{\varepsilon}, p')] \mathrm{d}F(\widetilde{\varepsilon}).$$

Thus the *rhs* of (A.15) exceeds the *rhs* of (A.14) and $\Omega_{\varepsilon}(\cdot) < 0$ for $\varepsilon < \widehat{\varepsilon}^{stat}(p)$ implies that$\widehat{\varepsilon}(p_{[2]}, p') < \widehat{\varepsilon}(p_{[1]}, p')$.

(iii) Finally, consider

$$\begin{aligned}
\Omega(\widehat{\varepsilon}(p_{[T]}, p'), p) &= (1-p) \int_{\widehat{\varepsilon}(p_{[T-1]}, p')}^{\infty} [\Omega(\widehat{\varepsilon}(p_{[T-1]}, p'), p) - \Omega(\widetilde{\varepsilon}, p)] \mathrm{d}F(\widetilde{\varepsilon}) && \text{(A.16)} \\
\Omega(\widehat{\varepsilon}(p_{[T-1]}, p'), p) &= (1-p) \int_{\widehat{\varepsilon}(p_{[T-2]}, p')}^{\infty} [\Omega(\widehat{\varepsilon}(p_{[T-2]}, p'), p) - \Omega(\widetilde{\varepsilon}, p)] \mathrm{d}F(\widetilde{\varepsilon}) && \text{(A.17)}
\end{aligned}$$

for $T \geq 3$. Consider the integral $\int_{\widehat{\varepsilon}}^{\infty} [\Omega(\widehat{\varepsilon}, p) - \Omega(\widetilde{\varepsilon}, p)] \mathrm{d}F(\widetilde{\varepsilon})$. Since $\Omega_{\varepsilon}(\widehat{\varepsilon}, p) < 0$ for $\widehat{\varepsilon} < \widehat{\varepsilon}^{stat}(p)$, this integral is decreasing in $\widehat{\varepsilon}$ for $\widehat{\varepsilon} < \widehat{\varepsilon}^{stat}(p)$:

$$\frac{\mathrm{d}}{\mathrm{d}\widehat{\varepsilon}} \int_{\widehat{\varepsilon}}^{\infty} [\Omega(\widehat{\varepsilon}, p) - \Omega(\widetilde{\varepsilon}, p)] \mathrm{d}F(\widetilde{\varepsilon}) = (1-p) \Omega_{\varepsilon}(\widehat{\varepsilon}, p) [1 - F(\widehat{\varepsilon})] < 0.$$

Therefore the *rhs* of (A.16) is larger than the *rhs* of (A.17) if $\widehat{\varepsilon}(p_{[T-1]} < \widehat{\varepsilon}(p_{[T-2]}, p')$. As a result, since $\Omega_{\varepsilon}(\cdot) < 0$,

$$\widehat{\varepsilon}(p_{[T]}, p') < \widehat{\varepsilon}(p_{[T-1]}, p') \text{ if } \widehat{\varepsilon}(p_{[T-1]} < \widehat{\varepsilon}(p_{[T-2]}, p'). \tag{A.18}$$

Since it was earlier on shown that the last inequality holds for $T = 3$, (A.18) shows that $\widehat{\varepsilon}(p_{[T]}, p') < \widehat{\varepsilon}(p_{[T-1]}, p')$ holds for all $T \geq 3$. ■

Proof of proposition 2

Proposition 3 implies that $\widehat{\varepsilon}(p_{[\infty]}, p') < \widehat{\varepsilon}(p_{[1]}, p') < \widehat{\varepsilon}(p_{[0]}, p')$, which is the statement in Proposition 2 for the "present-future"-version of the model. ■

B Estimation

If we denote the observations $\{a_{i,t}\}$ for $i = 1, 2, ..., n$ and $t = 1, 2, ..., T$, the likelihood of the model can be written

$$lnL(\boldsymbol{\beta}, \boldsymbol{N}_1) = \sum_{i=1}^{n} \sum_{t=1}^{T} \ln P(A_{i,t} = a_{i,t}|N_{i,t})] \tag{B.1}$$

where

$$N_{i,t+1} = N_{i,t} - a_{i,t} \tag{B.2}$$

The maximum likelihood estimator for the parameters is obtained by a numerical method. In the following two sub-sections we will explain how this is done in some detail.

A Gaussian simplification

To estimate the model (5.2) we proceed with the following steps.

1. Choose starting values for $\boldsymbol{N}_1 = (N_{1,1}, N_{2,1}, ..., N_{n,1})$
2. Estimate $\boldsymbol{\beta}$ given $\boldsymbol{N}_1$.
3. Estimate $\boldsymbol{N}_1$ given the estimated $\boldsymbol{\beta}$.
4. Repeat 2-3 until convergence.
5. To ensure that we have obtained the maximum likelihood estimator we maximize the likelihood numerically and obtain the hessian in the maximum, from which we can compute asymptotic standard errors.

Estimate $\boldsymbol{\beta}$ given $\boldsymbol{N}_1$

When $N_{i,1}$ is known, so is $N_{i,t}$ since it can be written

$$N_{i,t} = N_{i,1} - \sum_{k=1}^{t-1} A_{i,t} \tag{B.3}$$

We can therefore estimate $\boldsymbol{\beta}$ by running a linear regression of $A_{i,t}$ on $N_{i,t}\boldsymbol{x}_{it}$.

Estimate $\boldsymbol{N}_1$ given $\boldsymbol{\beta}$

Given $\boldsymbol{\beta}$ and using (B.3), we can rewrite model (5.3) as

$$A_{i,t} + \boldsymbol{x}'_{i,t}\boldsymbol{\beta} \sum_{k=1}^{t-1} A_{i,k} = N_{i,1}\boldsymbol{x}'_{i,t}\boldsymbol{\beta} + \varepsilon_{i,t} \tag{B.4}$$

or, defining

$$u_{i,t} = A_{i,t} + (\boldsymbol{x}'_{i,t}\boldsymbol{\beta}) \sum_{k=1}^{t-1} A_{i,k}$$

and

$$v_{i,t} = \boldsymbol{x}'_{i,t}\boldsymbol{\beta}$$

and obtain the simple regression without intercept

$$u_{i,t} = N_{i,t} v_{i,t} + \varepsilon_{i,t} \tag{B.5}$$

which can be run separately for $i = 1, 2, ..., n$.

C Impact plots

We observe two sequences of time points measured in days. These are $d_1, d_2, ..., d_m$, representing the dates where a tax-payer has made contact with Tax Norway in connection with an amnesty application, and $\boldsymbol{p} = \{p_1, ..., p_n\}$ which are dates of press releases from the Ministry of Finance (or Tax Norway). The period of interest is January, 1, 2010 until December, 31, 2016, a period containing 2373 days. In this time period we observed 1615 amnesty applications, i.e., on average, 0.68 applications per day. In order to investigate how press releases affect the frequency of contacts, the following is done. First, we compute how many contacts where done within one day of the press releases.

$$y_1 = \sum_{i=1}^{m} I(d_i \in \boldsymbol{p})$$

where

$$I(d_i \in \boldsymbol{p}) = \begin{cases} 1 \text{ if } d_i \in \boldsymbol{p} \\ 0 \text{ otherwise.} \end{cases}$$

. y_2 represents the number of contacts done within two days of the press release

$$y_2 = \sum_{i=1}^{m} I(d_i \in \boldsymbol{p}) + \sum_{i=1}^{m} I(d_i \in \boldsymbol{p} + 1)$$

where $\boldsymbol{p} + 1$ denotes "one day after the press release". This is a number for a period of length two days. Therefore, we divide it by 2 in order to get an daily average.

$$D_2 = \frac{\sum_{i=1}^{m} I(d_i \in \boldsymbol{p}) + \sum_{i=1}^{m} I(d_i \in \boldsymbol{p} + 1)}{2}$$

This makes it comparable to the expected number of contacts 26.3. In general

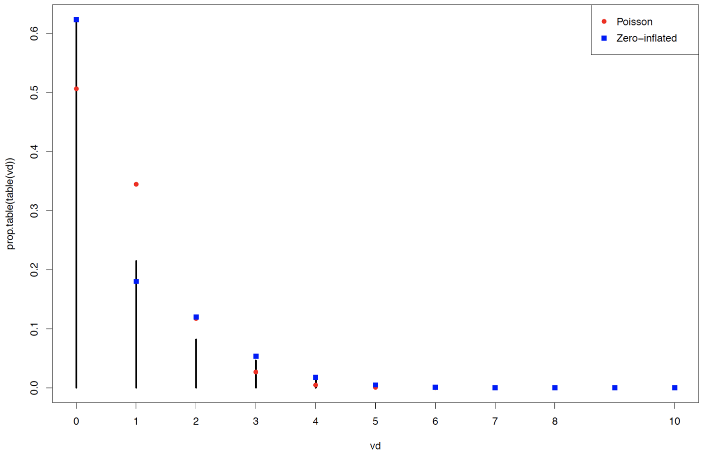
$$D_k = \frac{\sum_{l=0}^{k-1} \sum_{i=1}^{m} I(d_i \in \boldsymbol{p} + k)}{k}$$

represents the number of contacts within k days of a press release, normalized to a daily average. Finally we normalize D_k such that the typical value before press releases is 100.

In order to quantify the uncertainty of A_k we compute point-wise 5% critical values for the test that $A_k = \lambda$, under the assumption that y_k is $Po(\lambda k)$-distributed where λ is the expected number of events, in our case VD-applications, per day before the press releases.

D The zero-inflated Poisson distribution

Figure 6. Comparing poisson with the zero-inflated poisson



Note: The (unconditional) empirical distribution of VD compared to the fitted Poisson distribution and the zero-inflated Poisson distribution. The vertical lines represent the empirical distribution.

NHH

NORGES HANDELSHØYSKOLE
Norwegian School of Economics

Helleveien 30
NO-5045 Bergen
Norway

T +47 55 95 90 00
E nhh.postmottak@nhh.no
W www.nhh.no