Estimating Dark Figures of Crime

Nicole Bellert*

Andrea Günster†

Damian Kozbur‡

January 5, 2026

Preliminary Version

– please do not quote without permission of the authors–

Abstract

The dark figure of crime confounds the true prevalence with enforcement, since detection and crime are observed only in conjunction. We introduce a novel Maximum Likelihood estimator that simultaneously estimates probabilities of two latent outcomes - crime and enforcement - using instrumental variables. Identification relies on exogenous instruments excluded from shifting the probability of crime or enforcement occurring to disentangle the selective sample from its population. In a corporate cybercrime application in two Swiss cantons, the reported monthly rates (0.03% and 0.01%) severely mask the true prevalence; estimated rates are 2.4% to 17.9%, with dark rates exceeding 99%.

Keywords: Cybercrime, Dark Figure, Simultaneous Equation Modeling, Maximum Likelihood

JEL Codes: C35, C36, K42, L96

*Nicole Bellert (University of Applied Sciences Zürich (ZHAW), Institute for Wealth & Asset Management, Gertrudstrasse 8, 8401 Winterthur, Switzerland; bell@zhaw.ch; University of Zurich (UZH), Institute for Informatics, Binzmühlestrasse 14, 8050 Zürich, Switzerland.

†Andrea Günster (University of Applied Sciences Zürich (ZHAW), Institute of Business Information Technology, Theaterstrasse 17, 8400 Wintherthur, Switzerland; gues@zhaw.ch).

‡Damian Kozbur (University of Zurich, Department of Economics, Schönberggasse 1, 8001 Zurich, Switzerland; damian.kozbur@econ.uzh.ch).

In an earlier version, this paper was entitled "Estimating Dark Figures: An Application to Cybercrime." We sincerely thank Raphael Arnold, Dominik Balogh, Dominik Boos, Martin Carree, Jérémy Decerle, David Jaggi, Eduardas Lazebnyi, Nadine Marti, Roman Meyer, Alexander Posth, Luca Persia, Peter Schwedner, Mark Weibel and Felix Wullschleger for their valuable feedback and insights. Data was kindly provided by the Cantonal Police Bureaus (Kantonspolizei) St Gallen and the Police of Zug. We are also grateful for comments by the participants of the 2025 MSc Management and Law International Conference in Luzern (with special thanks to Peter Münch and Arnaud Raynouard for their thoughtful discussions), the 2025 SSES Annual Congress (Zurich), the 2025 EARIE (Valencia) and the 2025 Institute of Wealth and Asset Management Research Colloquium (Zurich University of Applied Science (ZHAW)). This research was supported by Innosuisse *Quantifying Illegal Activity: Estimating Dark Rates and Predicting Offenses* (https://www.zhaw.ch/de/forschung/projekt/73547) and *Dark Rates in Crime* (https://www.zhaw.ch/en/research/project/73003).

Ganz unbekannt war ja sein Prozess nicht, wenn es auch noch nicht ganz klar war, wer davon wusste und wieviel.[1]

1 Introduction

The estimation of true criminal activity remains one of the most persistent and fundamental challenges in empirical economics, law and criminology. The core empirical challenge lies in the fact that official statistics and observed recorded cases (R) capture only the light figure of crime (C) — a potentially highly selective, non-random sample of an underlying criminal population. As Biderman and Reiss [1967] argue, "...in exploring the dark figure of crime, the primary question is not how much of it becomes revealed, but rather what will be the selective properties of any particular innovation for its illumination." The resulting gap between the true prevalence of offenses and observed incidents is known as the Dark Figure of Crime (DFC) (Quételet [1832]), and quantifying it is essential to accurately measure social harm, formulate effective policies, and evaluate enforcement (E). The second central empirical obstacle when identifying the DFC, next to selection bias, is the simultaneity problem: a criminal incident appears in official records (R) only if two independent, latent events occur concurrently — a criminal act (C) and its enforcement (E) (i.e., detection, prosecution, or reporting). The observed record is the conjunction $R = (C \text{ and } E)$. Shifts in observed or recorded crime rates may, therefore, reflect changes in crime itself, or changes in detection technology, victim reporting propensity, or law enforcement effort, making it impossible to disentangle the true phenomena from its revelation. As Levitt [1998] notes "the debate over the validity of reported crime statistics is almost as old as reported crime statistics themselves," yet confounding factors remain unresolved.

We address this fundamental identification challenge by proposing a novel structural econometric method based on Maximum Likelihood Estimation (MLE), which simultaneously estimates the probabilities of two latent outcomes (C and E). Estimation is operationalized by specifying a likelihood function where the probability of a recorded incident, $Pr(R)$, is modeled as a product of two probits, one for $Pr(C)$ and one for $Pr(E)$. Identification of this system hinges on the crucial availability of exogenous instrumental variables (IVs). These instruments must be strictly excluded from shifting one probability (e.g., $Pr(C)$) while affecting the other (e.g., $Pr(E)$), thereby providing the exclusion restrictions necessary to separate the contributions of C and E to the observed record R. This structural approach directly addresses the selection and simultaneity biases that plague existing estimation methods, like Capture-Recapture (CR) (Ormosi [2014]), Detection Controlled Estimation (DCE) (Feinstein [1990], Foros [2004]), Difference-In-Difference (Diff-in-Diff) settings (Sovinsky [2022], Heim et al. [2022]), Hazard Rate (HR) estimation (Bryant and Eckard [1991], Levenstein and Suslow [2008]) or structural modeling (Craig [1987], Cornwell and Trumbull [1994]).

Section 2 situates our contribution within related work in criminology, economics, empirical legal studies, epidemiology, and finance, where researchers attempt to infer population sizes from selective samples. Section 3 develops our estimator. In Section 4, we demonstrate the favorable statistical properties of this estimator in finite samples using simulated data, describing also how to address the bi-convex nature of the product of two probits to ultimately find the global optimum of the minimization. Section 5 shows the results of our methodology when run on a comprehensive and unique administrative data of reported cybercrime incidents experienced by companies in the Swiss Cantons of St. Gallen and Zug between 2016 and 2023, 2024. Cybercrime

[1] His trial was not entirely unknown, even if it was not yet entirely clear who knew about it and how much. (Kafka [1925], Der Prozess, Page 130).

presents an acute challenge for estimation due to strong corporate disincentives for reporting and long detection lags, exacerbating the DFC problem. The observed monthly reporting rates (0.03% and 0.01%) severely mask the effective prevalence of cybercrime which we estimate at the company level to be 9.8% to 17.9% and 2.4% to 10.8% in St. Gallen and Zug, respectively. This yields dark rates of 99% in both Cantons. The stability of our latent prevalence estimates across multiple instrument specifications provides initial empirical support for our identification strategy. However, as Orsagh [1973] laments "it is undoubtedly true that the quality of the data available for empirical investigation is unusually poor" when trying to quantify the relation between crime and enforcement.[2] Section 6 concludes.

2 Related Work

The DFC was first introduced by Quételet [1832] and refers to the number of unreported or unregistered criminal cases. The Dark Rate of Crime (DRC) is the percentage of unknown cases in the population of all cases of crime (C): $DRC = DFC/C$. The Light Figure of Crime (LFC) in observable official crime statistics and cases is the difference between all crime and its dark part ($C = DFC + LFC$). It is a potentially unrepresentative sample of criminal activity (Biderman and Reiss [1967]). The extent and reasons for unrecorded crime have been widely discussed in various contexts (Duffee et al. [2000], MacDonald [2001], Lynch and Addington [2006], Loftin and McDowall [2010], Mosher et al. [2010], Biderman and Lynch [2012]). The willingness to report a crime case depends on various factors, including individual characteristics of the victim (Schneider et al. [1975]), economic factors (MacDonald [2001]), the gravity and type of the crime (Kääriäinen and Sirén [2011]) or opportunity costs (Tarling and Morris [2010]). Some types of crime may remain unknown even by the victim (cybercrime (IBM Corporation [2023])) or have no individual victim to report it (e.g., collusion, tax evasion, money laundering, bribery). Identification of the DFC represents a fundamental challenge in criminology and legal enforcement, as it distorts statistics on the magnitude of illegal or criminal activity, thus hindering the effective allocation of resources for law enforcement, prevention, deterrence, and evaluation of legal effectiveness (Skogan [1977]).

Traditional approaches to explain or quantify the DFC analyze discrepancies between official police records and victimization reports (Messner [1984]). Concerned about possible bias in OLS estimates, Orsagh [1973] applies the simultaneous equations from the classical demand and supply model to crime and sanctions for the number of reported crime cases in the 58 Californian counties in the year 1960. He finds a significant bias in OLS versus two-stage least squares estimation. Nagin [1978] warns of possible identification problems in the relation between crime and police clearance rates. He estimates a simultaneous model for the relation between the risk of imprisonment and crime, controlling for imprisonment duration and using prison capacity as IV, for cross-sectional crime data of 47 US states in 1960. In contrast, Wolpin [1980] estimates a single equation approach for time-series robbery cases between 1955 and 1971 in California, England and Japan. He relies on quite strong assumptions, e.g., that the changes in the crime rate do not affect the strength of deterrence. Cornwell and Trumbull [1994] aim to address heterogeneity and simultaneity issues by controlling for county-specific characteristics in a panel dataset of North Carolina counties with arrest and offense data. In addition to two-stage modeling, Wheeler et al. [2011] address spatial dependencies including average crime rates in neighboring areas as control variable. However, all these approaches do not separate crime from its enforcement.

[2] Orsagh [1973], Page 354.

Craig [1987] explicitly models enforcement (the allocation of police resources) and victim reporting in a simultaneous model with four equations. He estimates the success of crime deterrence measures for Baltimore neighborhoods in 1972. Comparing police recorded crimes from the Baltimore police department with reported and unreported cases in the US National Crime Victimization Survey (NCVS), he finds a significant reporting bias. To improve the measurement of reporting bias, Levitt [1998] controls for the variation of police resources per capita in the NCVS. In another approach, he assumes nearly 100% recording rates for murder cases and uses them as a reference to estimate the rate of unrecorded cases in other subject areas. Also focusing on murder cases, but for the reason that they are much more costly than other types of crime, Chalfin and McCrary [2018] analyze the relationship between police resources and violent crime rates. They address measurement errors in police force data estimating a GMM with two different data sources of police staffing numbers, with crime data for 242 US cities between 1960 and 2010.

Combining different police surveillance with victim behavior and socio-economic characteristics in geographic areas, Buil-Gil et al. [2021] present a geographic map showing dark figures of crime based on estimates of small areas. Based on the Crime Survey for England and Wales from 2011 to 2017, they find larger dark figures of crime in suburban areas than in big cities. However, questionnaire based survey data may suffer itself from various forms of bias, in particular selection bias in survey participation (Haverkamp [2020]), but also measurement error from misremembering or misclassification, respondent fatigue, and interviewer effects (Pina-Sánchez et al. [2023], Fé [2024]). More fundamentally, these approaches assume the difference between surveys and official records represents the dark figure, but this assumption fails when survey responses themselves are systematically biased - precisely the problem in contexts like corporate cybercrime where firms have strategic incentives to under-report even in confidential surveys.

Recognizing these fundamental identification challenges, recent work has turned to partial identification approaches that provide bounds rather than point estimates under weaker assumptions. Manski and Pepper [2013] apply partial identification to estimate the deterrent effect of capital punishment on homicides in all US states, comparing the country-wide abolishment in 1975 with state-individual legalization status in 1977. They define the legalization status as treatment effect. In addition, they estimate the Difference-in-Difference (Diff-In-Diff) to compare changes in homicide rate over time in states with and without treatment. Building on Manski and Pepper [2013] in applying partial identification to criminology, Fé [2024] builds a partial identification framework to derive upper and lower bounds for victims misreporting crime incidents in the Crime Survey for England and Wales between 2012 and 2020. His contribution acknowledges non-identification of latent crime rates without strong assumptions.

DCE estimates the percentage of undetected offenses for a sample of investigated companies (Feinstein [1990]). In applications to nuclear power plants (Feinstein [1989]) and tax evasion (Feinstein [1991]), each investigated company is paired with its responsible investigating agency. Feinstein [1990] derives an MLE function based on two probit equations, one estimating the probability of committing an offense and the other estimating the probability of detection in the case of an offense, controlling for characteristics of company and investigating agency. Although the approach has similarities to ours, it is only applicable if the sample is identifiable. As Feinstein [1990] (Page 245) points out, "the DCE decomposes a single datum, detected violations, into two disjoint behavioral categories, violation and detection, and it is not initially clear whether this decomposition can be performed uniquely". It also strictly requires subjects to be investigated, which is information most often missing from enforcement records or even a non-existence prerequisite in enforcement

regimes. Feinstein [1990] therefore only touches on problems with non-random sampling defined in Heckman [1979]. Foley et al. [2019] identify illegal activities of bitcoin users in measuring the share of transaction with other illegal users. They adapt DCE for instrumental variables that either affect only the participation in illegal activity, or its detection, running the model on the population of all bitcoin trades. Their method is most closely related to our identification strategy, but does not account for repeated recording and requires a review process identifying illegal activity. The latter is most often not applicable in real data or enforcement regimes more generally.

Estimation methods deriving unknown population sizes from (selective) samples originate most often in life sciences; populations of animals in a specific geographic area, prevalence and incidence of viruses, and diseases for a population of potential hosts. The main type of econometric specification are HR models in survival analyses developed in biomedical studies. Bryant and Eckard [1991] use HR to estimate the detection rates of cartels uncovered and prosecuted by the Department of Justice (DoJ) for price-fixing conspiracies in the 1980s. Based on the duration of the cartel, they estimate the probability of birth and death, which when combined provides the probability of a cartel being alive.[3] Bryant and Eckard [1991] find the 13-17% detection probability in a given year for a sample of eventually detected US cartels between 1961 and 1988. Combe et al. [2008] estimate 12.9-13.3% for EU cartels from 1969 to 2007. The estimated probability of death (which is equal to one divided by the average duration (Harrington and Wei [2017])), may not be representative of the entire (unknown) population (Bryant and Eckard [1991], Harrington and Chang [2009], Davies and Ormosi [2012], and Ormosi [2014]). Other critical features of the econometric models explaining the duration of cartels (HR models) used so far are the underlying assumptions of (i) constant and (time) independent rates of cartel birth, death, and detection (Harrington and Chang [2009] and Bos and Harrington [2010]), (ii) homogeneous firms and industries, and (iii) complete cartels (Günster [2010]). Diff-in-Diff studies aim for, but do not quantify a reduction in collusive activity effectively, ignoring selection biases as well (Miller [2009], Harrington and Chang [2015], Hellwig and Hüschelrath [2018], Heim et al. [2022], and Sovinsky [2022]).

Their effect on cartel formation, duration and deterrence is investigated by Motta and Polo [2003], Chen and Harrington [2007], Harrington [2008], Spagnolo [2008], Miller [2009], Gärtner and Zhou [2012], Harrington [2013], Chen and Rey [2013], Duso et al. [2014], Harrington and Chang [2015], Hellwig and Hüschelrath [2018], Heim et al. [2022] and Sovinsky [2022]. Although Diff-in-Diff studies do not quantify a reduction in collusive activity, studies relying solely on samples of detected cartels base the analysis on information from convictions, ignoring selection biases. Theoretical studies establish theoretically that firms can use potentially very strategically leniency applications (selection bias) (Motta and Polo [2003], Harrington [2008], Spagnolo [2008], Harrington [2013], Chen and Rey [2013], Harrington and Chang [2015]).

CR is the second most prominent method to establish the number of cartels alive, originally developed in biology (Ormosi [2014]). It compares the proportion of (re)captured animals, but also cartels, of all captured animals in a defined area relative to the total area over time.[4] Rivest and Baillargeon [2014] provide a widely

[3]The most basic model is $\lambda(t) = \lim_{dt\to 0} Pr(t \leq T < t + dt | T \geq t)/dt = f(t)/S(t) = -S'(t)/S(t)$ where t represents time and S is the hazard function; the hazard function can also be represented as a cumulative hazard function $\Lambda(t) = -\log S(t)$.

[4]The capture histories are fitted using a Poisson regression with MLE (McCrea and Morgan [2015]). The most simple representation of Capture-Recapture is based on Amstrup et al. [2005]: $n = mc/r$, where n represents population size, m the number of animals captured and marked, c the total number of captures during the second visit, and r the number of recaptures on the second visit. For ecological studies, which often involve temporal processes, the assumption of closure (no births, deaths, or migration) was relaxed, leading to open population models such as the Cormack-Jolly-Seber model (Rivest and Baillargeon [2014]). For epidemiological applications, which typically collate individuals from different lists, a seminal advancement was the introduction of log-linear models (Fienberg [1972]). These models specify expected cell counts in a log-linear form, allowing for interactions between surveys and forming the basis of most Multiple Systems Estimation applied to epidemiological data.

used statistical R package for closed and open populations of animals. Chan et al. [2021] extend it for non-overlapping lists of illegal immigrants in the US. They derive a multiple system estimate for victims of human trafficking appearing in different victimization surveys during the same time period.

Any of the described estimation methods might suffer from sample selection biases that have not been adequately addressed (Heckman [1979]). Bellert et al. [2023] show how HR and CR models do not derive unbiased and consistent estimators because samples of illegal activity might not be representative of its population. All methods discussed do not adhere to random sampling. CR inspires our approach as we explicitly control for being previously recorded. We expand upon CR with structural equation modeling using instrumental variables, while Foley et al. [2019] expand DCE with IVs. We also explicitly model treatment (crime) and potential selection (discovery), which is an analogy to estimating Local Average Treatment Effect (LATE) (Imbens and Angrist [1994]). Our method expands LATE by deriving latent outcome variables of crime and its revelation, which are affected by treatment and compliance. This process simultaneously identifies unobserved probabilities of crime (C) and enforcement (E). It also addresses selection and simultaneity biases by using instrumental variables to identify groups of compliers who act according to their group. These subjects comply, they take the treatment when assigned to treatment and abstain when assigned to the control group. It nets the effect of a treatment on compliers neglecting non-compliers. Our method relies on a combination of several methodologies (i.e., CR, LATE, simultaneous equation modeling including IVs) to address the main problem of previous attempts; neglecting selection (i.e., CR, DCE, HR), measuring indirect effects (i.e., Diff-In-Diff) or bounds (i.e., partial identification).

Cybercrime presents particularly acute challenges for dark figure estimation. Unlike traditional crimes, cyberattacks often go undetected for extended periods (IBM Corporation [2023] reports 207 days average detection time), and even detected incidents face strong disincentives for reporting due to reputational concerns. In a survey by Keeper Security [2023], 48% of the participants stated that the incidents were not reported to the authorities and 41% that the incidents were not even reported internally. August et al. [2024] investigate the conditions that motivate companies to report cyber incidents and disclose information about cybersecurity. Cybercrime statistics are particularly affected by over- or under-reporting (Anderson et al. [2013]). The number of reported cybercrime cases is steadily increasing due to digital innovation and a shift of activities to the digital space (Wu et al. [2023] gives a general overview of recent studies). Cloud services, mobile devices, and social networks create new vulnerabilities, with online property crime averaging around 50% of worldwide property crime (Anderson et al. [2019]). Global damage estimates range from $1 trillion (McAfee [2020]) to $3 trillion (Herjavec [2019]) annually. Prevalence estimates use victim surveys (Reep-van den Bergh and Junger [2018], Accenture [2019], Hiscox [2022], ISACA [2023]) and predictive Hawkes and Machine Learning models (Bessy-Roland et al. [2021], Elluri et al. [2023]). The US DoJ estimates that only 15% of cyber incidents are reported (US Department of Justice [2015]). Applying our estimation method to Swiss cybercrime, we estimate that the true monthly prevalence rates are 9.8% to 17.9% and 2.4% to 10.8% for cyber incidents on companies in the Swiss cantons of St Gallen and Zug, while the monthly reporting rates are only 0.03% and 0.01% respectively, implying monthly detection rates of only 0.10% to 0.17% and 0.28% to 0.43%.

3 Estimating Dark Figures

The basis for law enforcement is the sample of observed recorded cases (R), with a recording $R = 1$ for every observation. However, for an illegal offense to be recorded, two conditions must be met

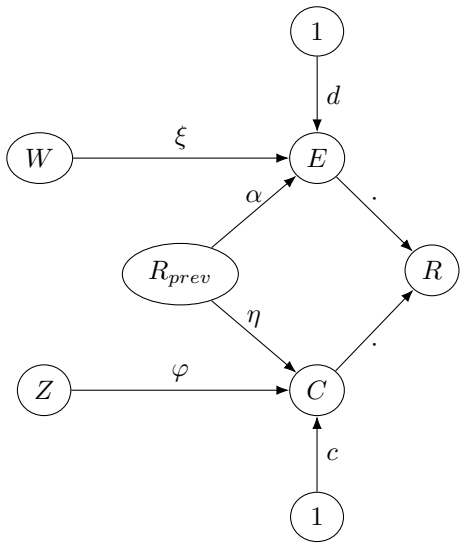
- a criminal incident takes place (C)
- there is a possibility of enforcement (E)

Observing a criminal incident (R) in reality is the conjunction of the incident occurring (C) and being detected, prosecuted, and convicted (enforcement (E) of crime (C)), which is $R = C \cdot E$. Consequently, we explicitly model the observation of a crime (R) as the multiplication of the two latent outcomes of the crime (C) and its enforcement (E), which we simultaneously estimate.

Denote an incident recorded by the same individual as R, then R_{prev} is the same subject involved in an offense recorded in a previous period. Both recording variables are binary; individuals record an incident or not. Both R_{prev} and R are observed. The interval between R_{prev} and R is defined and depends on the concrete application. With the assumption that former participation has an (unknown) influence on present events, we let R_{prev} affect both C and E. The observed variable R takes the value 1 for an individual in a certain period of time if at least one illegal event is recorded and 0 otherwise. In contrast, Feinstein [1990] explicitly models Bayes [1958] using Heckman [1979], but therefore requires whether an entity is investigated. Instead, we run our method on the entire population of potential perpetrators or victims without identification of selection. Note also that we do not require making assumptions about the sign or magnitude of the relations between instruments and previously reporting on the latent variables we aim to estimate. These will be the outcome of the estimate. We exclusively require instruments to be exogenous to the latent variable that they should not affect. We do not assume structure here. However, we assume that there are no Type II errors, that is, no falsely recorded incidents.

Traditional methods fail because they only account for reported cases ($R = 1$, $C = 1$, $E = 1$), and not for the large number of unobserved cases ($R = 0$ and either $C = 0$ or $E = 0$). In the dual-process logic used in the simultaneous two-probit model, a recorded observation is $R = C \cdot E$, which requires the conjunction of both latent variables. On the one hand, individuals are subject to crime with a probability $P(C)$ consisting of a conditional mean of crime (c), an instrument affecting the probability of crime occurring (Z), the impact of the instrument (φ) and the impact of having been previously recorded (η), with $c, \varphi, \eta \in \mathbb{R}$. On the other hand, individuals are subject to the enforcement of crime with a probability $P(E)$ consisting of a conditional mean of the enforcement of the crime (d), an instrument that affects the probability of the enforcement occurring (W), the impact of the instrument (ξ) and having been subject to previous recordings (α), with $d, \xi, \alpha \in \mathbb{R}$. The probability of crime enforcement must be independent of the probability of committing or experiencing a crime and *vice versa* (i.e., strict exogeneity); instrumental variables must identify either the likelihood of a crime or its enforcement to disentangle the sample from its population. The instrumental variable describing the probability of crime enforcement might consist of legislation (law creation), the judiciary (law interpretation), and actual enforcement (law execution).[5] Criminal activity depends on behavioral factors, gains and punishment, and the legal system in turn. In Section 5, we show concrete examples of instruments that affect the probability of being subject to cybercrime or the propensity to report a cyber incident at the firm level in Switzerland.

[5] Becker [1968] was the first of very many to analyze the interplay between crime, punishment and the cost for society. His contributions offer guidance on instrumental variable choice determining the enforcement of crime.



This Directed Acyclic Graph describes the relations between offenses, their identification and their reporting.
R (R_{prev}): Recorded (Previously)
Z: IV Affecting Crime
W: IV Affecting Crime Enforcement
E: Crime Enforcement
C: Crime

Figure 1: DGP of Crime and its Enforcement

The Directed Acyclic Graph (DAG) in Figure 1 describes the relationship of the Data Generating Process (DGP) of illegal offenses and their enforcement (Pearl [1995]). The weighted directed arrows indicate the order and direction of the relation between the observed variables (Z, W, R_{prev}, R) and the unobserved variables (E, C). The size of the weights is unknown and is observed as a vector of unobserved parameters $\theta = [c, \varphi, \eta, d, \xi, \alpha]$, which we want to estimate. The core relationship we model remains $\Pr(R = 1) = \Pr(C = 1) \cdot \Pr(E = 1)$, which we specify as Equations 1 to 3 to summarize the DGP, with $v, e \sim \mathcal{N}(0, 1)$:

$$C = \mathbb{I}\{v < c + \varphi \cdot Z + \eta \cdot R_{prev}\} \tag{1}$$

$$E = \mathbb{I}\{e < d + \xi \cdot W + \alpha \cdot R_{prev}\} \tag{2}$$

$$C \cdot E = R \tag{3}$$

The estimated models provide the weights and thus derive the direction and magnitude of the relationships between observed variables (Z, W, R_{prev}, R) and unobserved latent variables to be identified (E, C). In addition, we derive the coefficients (φ, η, ξ, α) and the constants (c, d) by estimating the system of equations. Having been recorded in a previous stage (R_{prev}) is observed and essential to isolate C from E, assigning their respective shares (Equations 1 and 2). To simultaneously estimate the system of Equations 1 to 3, we derive the maximum likelihood function in Equation 4 with the cumulative distribution function of the standard normal distribution (Φ) as link function. The latter transforms the models of latent variables into probabilities. Given that the latent variable C is binary, taking the expectation $\mathbf{E}[C]$ over the population is exactly equivalent to calculating the probability of a crime occurring ($\Pr(C = 1)$). Consequently, the system of equations yields the estimated mean latent prevalence of crime, often referred to as the expected crime rate, or alternatively, the probability that a crime occurred. Based on Equations 1 to 3, we estimate the likelihood of the values of unknown parameters given the observed variables (Equation 4). Minimizing the negative log-likelihood delivers the estimated parameter vector $\hat{\theta}$ (Equation 5); the directions and weights of the DGP (θ =[c, φ, η, d, ξ, α])

shown in Figure 1:

$$\mathscr{L}(\theta) = \prod_{obs} \; (\Phi(c + \varphi \cdot Z + \eta \cdot R_{prev}) \tag{4}$$
$$\cdot \Phi(d + \xi \cdot W + \alpha \cdot R_{prev}))^{\mathbb{I}\{R=1\}}$$
$$\cdot (1 - \Phi(c + \varphi \cdot Z + \eta \cdot R_{prev})$$
$$\cdot \Phi(d + \xi \cdot W + \alpha \cdot R_{prev}))^{\mathbb{I}\{R=0\}}$$
$$\hat{\theta} = \underset{\theta}{\operatorname{argmin}} \quad -\ln \mathscr{L}(\theta) \tag{5}$$

The MLE objective function $\mathscr{L}(\theta)$ in Equation 4 is the product of the two probit functions shown in Equations 1 and 2. The left hand side of both equations is unknown (Crime and its Enforcement), but the product of both is observed (Recorded Crime). This product results in a non-linear objective function without closed form solution. Therefore, we solve the MLE using numerical optimization. This is done iteratively, starting from an initially arbitrary vector $\hat{\theta}_1$, seeking the convergence of a sequence $\{\hat{\theta}_n\}$. In the resulting estimates, the score function should be zero on average, with the log-likelihood gradient being $S(\hat{\theta}) = \partial logL(\hat{\theta}, data)/\partial\hat{\theta} \approx 0$. To ensure a global minimum, the Hessian in the resulting estimates has to be a positive definite non-singular matrix, with all Eigenvalues above zero, to compute standard errors and confidence intervals from its inverse.

MLE enjoys statistical consistency bounds when specified well and convexity conditions apply. It delivers consistent, asymptotically unbiased and asymptotically efficient estimates (Davidson and MacKinnon [2004]). These conditions include continuity of the normal Cumulative Distribution Function (CDF), identification of parameters (i.e., different parameters lead to different distributions), and compactness of the parameter space. Our objective function is suitable for optimization with the Newman-Raphson method. We choose the Broyden, Fletcher, Goldfarb and Shanno (BFGS) algorithm to derive $\hat{\theta}$ (Broyden [1970], Fletcher [1970], Goldfarb [1970], Shanno [1970]). BFGS is a robust numerical approach for non-linear unconstrained optimization problems. As a quasi-Newton method, it approximates the Hessian iteratively instead of calculating it.[6] With the estimated parameters in $\hat{\theta}$, we reconstruct the estimated probabilities of $\hat{E}$ and $\hat{C}$ for each observation, with C being our latent variable of interest. The average of $\hat{C}$ gives the estimated percentage of observed and unobserved crimes. Subtracting the number of observed incidents from the number of estimated ones yields the DFC.

While the objective function is locally convex near the maximum likelihood estimate as shown by a positive definite Hessian, the overall likelihood surface is non-convex due to the product structure of the probabilities and the non-linear probit link function. Fixing the parameters of one probit, the remaining problem is quasi-convex. This means that the optimization problem satisfies a property that is weaker than bi-convexity. It requires careful optimization with multiple starting values to ensure that the global optimum is found. The number of optimization rounds with different initial parameter values depends on the concrete application. The global optimum is determined from the rounds with the converging algorithm. In the resulting local optimum, the gradient must be near zero and the hessian must be non-singular. Several equal solutions show the same smallest objective value. They represent the global minimum, if all alternative solutions have substantially higher objective values and also higher condition numbers. The condition number κ of the hessian matrix is

[6] For robustness checks, we estimate parameters for the simulated data additionally with both the Nelder-Mead (Nelder and Mead [1965]) and Conjugate Gradient algorithms (Fletcher and Reeves [1964]). All of them produce similar results in the simulation (available on request), with BFGS being the fastest, which is important for applications with a large number of observations (Nocedal and Wright [2006]).

a measure of ill conditioning. If the Hessian is positive definite, κ can be calculated as the ratio of its largest to its smallest Eigenvalue (Todd [1950]). We use it in addition to the objective value to compare different optimization results.

For enforcement, the relevance and exclusion restriction assumptions regarding the instrumental variables have to hold. We assume that the instrumental variable Z is relevant and affects crime (Equation 1), but must not affect enforcement. In analogy, we assume that instrument W is relevant and affects E (Equation 2), but must not affect C. If C and E were observed, we could check the rank condition to ensure instrument relevance. Further, we could check the order condition to test the exclusion restriction; number of omitted exogenous variables $\geq$ number of endogenous variables. These two assumptions cannot be tested, because the left-hand side of Equations 1 and 2 is unknown. The assumptions rely on the underlying economic theory. To assess the plausibility of these restrictions, we calculate the Lagrange Multiplier statistics (LM), a hypothesis test for parameter restrictions violations. For both exclusions restrictions we estimate the unrestricted model allowing cross-effects. First, we add the instrument used in the second equation to the first equation, changing Equation 1 to Equation 6. Second, we add the instrument used in the first equation to the second equation, changing Equation 2 to Equation 7 and optimize them separately.

$$C = \mathbb{I}\{v < c + \varphi \cdot Z + \rho_C \cdot W + \eta \cdot R_{prev}\} \tag{6}$$

$$E = \mathbb{I}\{e < d + \xi \cdot W + \rho_E \cdot Z + \alpha \cdot R_{prev}\} \tag{7}$$

The LM for each restriction is

$$LM = \nabla_{\theta_u}^T \widehat{V} \nabla_{\theta_u} / n \tag{8}$$

where n is the sample size and ∇_{θ_u} is the gradient of the objective function optimized for the unrestricted vector $\theta_u = [\theta, \rho]$. $\widehat{V}$ is the estimated asymptotic variance covariance matrix of the estimated restricted parameter vector $\hat{\theta}$. Since the LM converges to a χ^2 distribution with 1 degree of freedom (one restricted parameter), the critical value for the 5% Level is $z = 6.63$ (Breusch and Pagan [1980]). If $LM < z$, we cannot reject the Null Hypothesis that the additional parameter ρ is zero, which provides evidence for the exclusion restriction.

4 Simulating Dark Figures

To evaluate and demonstrate the performance of our method, we simulate data for a population of potential subjects of crime of one million ($n = 1'000'000$ observations), which are in one of three possible states: (i) not subject to crime and, therefore, by definition not observed ($C = 0$, $R = 0$), (ii) subject to crime and observed ($C = 1$, $R = 1$), (iii) subject to crime but not observed ($C = 1$, $R = 0$). Note that the last state constitutes the dark figure which is unobserved in reality. Note also that this excludes the possibility of Type II Errors ($C = 0$, $R = 1$), which possibly exist in law enforcement. The DAG in Figure 1 describes the DGP. Some individuals have previously experienced crimes ($R_{prev} = 1$). In addition, we arbitrarily set the following parameters at $\theta = [c = .4, \varphi = .3, \eta = -.2, d = .5, \xi = .4, \alpha = .2]$. The effect η of R_{prev} on C is set negative, assuming that individuals take preventive measures after being subject to crime. For each observation, we let the instruments Z and W independently take values $\in \{0, 1, 2\}$, following a Binomial Distribution with two

trials and probability 0.5. R_{prev} is simulated as $R_{prev} = \mathbb{I}\{u < 0\}$, with $u \sim \mathcal{N}(0, 1)$. Inserting the parameters and simulated variables in Equations 1 and 2 generates values for C and E. Equation 3 generates values for R. This DGP delivers ground-truth data for n observations with recorded incident, unrecorded incident, or without incident (Figure 1).

We optimize the objective function (Equation 4) on the synthetic data. To find the global optimum, we run 20 optimization rounds with randomly set initial parameters and select converging solutions with the lowest objective value. Appendix Table 16 lists all results sorted by objective value. The nine results with the same lowest value (up to 7 digits) have equal estimated coefficients (up to 2 digits) and condition number around 695. All other results have larger objective values and a singular Hessian. Their condition number is either very large ($> 4e^{17}$) or not defined. For the result with the lowest objective value, Table 1 shows in columns 3 to 6 the estimated coefficient, its standard deviation and the resulting Confidence Interval (CI) at 95%. Column 2 shows the simulation set parameters to facilitate comparison. Rows 1 to 3 of Table 1 provide the estimation results for the probit that estimates the effect of the instrument Z and was recorded in a previous period (R_{prev}) on crime (C). Rows 4 to 6 show the results for the probit estimating the effect of the instrument W and R_{prev} on E (Figure 1).

Table 1: Crime Simulated and Estimated Parameters in Simulation

	Simulation	Estimate	SE	95% CI	
Constant on Crime ($\hat{c}$)	0.4	0.4223	0.0163	(0.3903	0.4543)
IV on Crime ($\hat{\varphi}$)	0.3	0.3049	0.0060	(0.2932	0.3166)
R_{prev} on Crime ($\hat{\eta}$)	-0.2	-0.2094	0.0113	(-0.2314	-0.1873)
Constant on Enforcement ($\hat{d}$)	0.5	0.4759	0.0152	(0.4462	0.5056)
IV on Enforcement ($\hat{\xi}$)	0.4	0.3878	0.0122	(0.3638	0.4117)
R_{prev} on Enforcement ($\hat{\alpha}$)	0.2	0.2030	0.0139	(0.1757	0.2303)
$\hat{E}$ (Enforcement estimated on $\hat{\theta}$)	0.82958	0.8229	0.0738	(0.6829	0.9271)
$\hat{C}$ (Crime estimated on $\hat{\theta}$)	0.72050	0.7273	0.0781	(0.5843	0.8490)
N		1′000′000			

This table shows for one million observations ($n = 1'000'000$) the parameter values set in the simulation, the estimated parameters, their standard errors and CI.

Table 1 confirms that our novel method describes the DGP of dark figure estimation well (Figure 1). The estimated coefficients are within two standard deviations (67%) at most. In addition to the estimated parameters, Table 1 shows the simulated values for enforcement (E) and crime (C) in the last two rows. Applying Equations 1 and 2 to every observation in the simulated data provides the distribution of probabilities for E and C. The mean of the simulation and the estimate of E are 83.0% and 82.3%, respectively. For the probability of C the means are even closer (72.1% and 72.7%). Table 2 shows the condition number (κ) for the resulting hessian matrix. For both exclusion restriction assumptions, we calculate the LM statistics for the unrestricted model. Both LM_C and LM_E are below the critical value of $z = 6.63$ (Section 3), with corresponding p-values. We cannot reject the Null Hypothesis that an additional parameter in the unrestricted models is zero, giving evidence for the exclusion restrictions. The confidence intervals and condition numbers decrease (i) in increasing numbers of observations (consistency), (ii) in higher rates of recorded cases (R) and (iii) more points of support of the instrumental variables (W and Z). Appendix Figure 4 shows the distribution of all estimated parameters for 1'000 simulations.[7]

[7]Simulation code written in R is publicly available at Bellert [2025] to facilitate replication and application by other researchers. We provide a description (README) on how to reproduce the simulation and amend it, to generate basic statistics, estimate the hidden population, and calculate the dark rate. Also in Bellert [2025], we provide an Online Appendix with additional results and

Table 2: Evaluation of Results (Simulation)

Condition Number κ	LM_C	p_C	LM_E	p_E
694.72	0.7315	0.3924	8e-8	0.9998

This table shows for the estimates in Table 1 the condition number and the Lagrange Multiplier statistics with p-values for the unrestricted Equations 6 and 7.

5 Dark Rates of Cybercrime in the Cantons St Gallen and Zug

We estimate the percentage of unreported cybercrime incidents on companies, using a sample of reported incidents in the Swiss Cantons St Gallen (2016-2023) and Zug (2016-2024). Companies are potential victims of cybercrime attacks (C) and we define the incentive for the company to report as enforcement of crime (E); the higher the incentive to report translates into more recordings of crime independent of whether more crime was committed.

Cybercrime is a novel area of enforcement with the Budapest Cybercrime Convention (Council of Europe [2001]), signed by Switzerland only in 2001, defining cybercrime as *"action directed against the confidentiality, integrity and availability of computer systems, networks and computer data, as well as the misuse of such systems, networks and data."*[8] In Switzerland, digital or cybercrime enters the country-wide criminal statistics records for the first time in 2020. Cantons St. Gallen and Zug publish the first statistic on cybercrime cases in 2020 and 2017, respectively. We focus only on cyber incidents affecting companies, which cover mostly cyber fraud, but also all other categories. In Switzerland, offenses with a digital crime component per year increased significantly by 80% from 24'398 recorded cases in 2020 to 43'839 in 2023. The major part of cybercrime concerns economic cybercrime, with a share of 84% in 2020 increasing to 92% in 2023 (Swiss Federal Statistical Office [2024b]). This rise is shown to be representative for phenomena internationally, with the dark rate of cybercrime described as *"not all digital offenses are reported to the police, even if they were correctly identified by the victims or company"* (Swiss Federal Office of Justice [2024]).

The frequency and severity of cybercrime depend on socioeconomic, political, and technological factors, as well as technology standards, data handling, and prevention methods (Solano and Reinoso Peinado [2017]). Studies find positive correlations between company size and cyberattack frequency (Paoli et al. [2018], Isenhardt et al. [2022]). Additional cybercrime drivers include remote work arrangements (FINMA [2020a]), technological innovations such as GPT-3 (ETCISO [2024]) or cryptocurrencies (August et al. [2025]), economic crises such as COVID-19 (Naidoo [2020], Panda Security [2020], FINMA [2020b]) and geopolitical events such as the Russian attack on Ukraine (National Cybersecurity Centre [2023], CyberPeace Institute [2024]). Table 3 lists possible IV (Z) that influence the probability of cybercrime incidents in Swiss companies.

Factors that affect the incentive to report a cybercrime incident (enforcement E) include government regulations and facilitation measures. In 2016, the EU established the Directive for the Security of Network and Information Systems ($NIS1$), including a reporting obligation in case of a cyber incident for operators of essen-

figures.

[8] We rely on the definition provided by the Swiss Federal Statistical Office. The term cybercrime covers all offenses with a digital component, i.e., committed in telecommunication (networks), particularly on the internet (Swiss Federal Statistical Office [2021]). Swiss Federal Statistical Office [2023b] defines 29 different offenses with 33 possible types of modus operandi as cybercrime. Offense types are defined by the respective Articles in Swiss Criminal Code (SCC), with fraud (Article 146) accounting for the majority of cases (53% in 2023), followed by computer fraud (Article 147) with 17%, money laundering (Article 305bis) with 9%, pornography (Article 197) with 6% and the unauthorized obtaining of data (Article 143) with 4%. Modus operandi can be categorized in economic cybercrime (92.4% in 2023), cyber sexual crimes (6%) and cyber damage to reputation and unfair practices (1.7%) (Swiss Federal Statistical Office [2023a]).

Table 3: Determinants of Cyber Crime (C)

Instrument Z	Start (to End) Date	Description
Covid	2020-03-16 (to 2022-04-01)	Covid-19 Lockdown / Home Office Recommendation (CH)
Ukraine	2022-02-24	Russian Attack on Ukraine
ChatGPT	2022-11-30	Introduction ChatGPT (GPT-3)
GPT-4	2023-03-14	Introduction GPT-4
N_Emp	2025	Number of Employees
Mid_Sized	2025	Number of Employees ≥ 50
Large_Sized	2025	Number of Employees ≥ 250
Retail	2025	Retail Trade (NACE 47)
Online	2025	Retail via Mail Order Houses or Internet (NACE 47.91)

The variable *Covid* is 1 for all months between the start of the official Covid-19 lockdown in Switzerland (Swiss Federal Council [2020]) and the final end of the home office recommendation (Swiss Federal Council [2022a]). All other shocks in Z are 1 for all time periods after the shocks, and 0 before. The number of employees (*N_Emp*) could be used as an approximation of the firm size. However, we use categories of firm size instead, to not combine continuous with binary variables. We include information on the economic activities of the company as instruments. Retail firms (1 if NACE Code starts with 47, else 0) and online stores (1 if NACE Code is 4791, else 0) (European Commission [2008]) are overrepresented in our specific sample, as they are particularly often victims of cyber fraud (Table 5 and 6, and Appendix Table 31. The variables we included are in italics.

tial services such as electricity, transport, health, drinking water, financial services and digital service providers (European Union [2016a]). $NIS1$ was adopted by the Member States in May 2018.[9]

In Switzerland, the Data Protection Act (DSG) came into force in 1992 and regulates the processing of personal data. Only since the revised version in September 2023 does it contain Article 24, which obliges companies processing personal data of natural persons to report breaches of data security to the Federal Data Protection and Information Commissioner (FDPIC) (Swiss Federal Council [2023a]).[10]

Liability companies (AG) are subject to certain disclosure requirements to their shareholders (Article 697 of the Swiss Code of Obligations).[11] Companies incorporated on the Swiss Stock Exchange (SIX) are required to disclose price sensitive information (SIX Exchange Regulation AG [2024]). According to Mathys [2021], this might include cyber incidents.[12]

Independent of reporting obligations, all persons and legal entities in Switzerland affected by a cyber incident are encouraged to report the incident (National Cybersecurity Centre [2024]). Reporting cyber incidents does not necessarily lead to its prosecution, requiring the involvement of a law enforcement agency. To bring charges, it is additionally necessary to file a complaint with the cantonal police bureau (Kantonspolizei (KAPO)). In Switzerland, the police are organized on a decentralized basis. The police forces in the individual cantons use different software systems and are independently responsible for recording, analyzing and documenting the complaints.[13] Table 4 lists possible IV (W) that affect the proneness of companies in Switzerland to report a cyber incident.

There is a third category of factors that influence both cybercrime and reporting, such as corporate gover-

[9]Extending the scope of $NIS1$ to medium-sized enterprises and new sectors like food, chemistry and electronics and harmonizing reporting obligations, NIS2 came into force in 2022 and had to be adopted by the member states until October 2024 (European Union [2022]). Both NIS affect Swiss companies when they engage in the EEA.

[10]In the European Union, the European General Data Protection Regulation (DSGVO) entered into force in May 2018 with the aim of protecting data of natural persons (European Union [2016b]). In Switzerland, it affects all companies that trade within the European Economic Area (EEA). Article 33 DSGVO requires Swiss companies that process personal data of natural persons to report breaches of data security to the supervisory authority of each EU Member State whenever a person is affected by a data breach (Federal Data Protection and Information Commissioner [2018]). In addition, the Swiss Parliament decided in September 2023 to introduce a reporting obligation for cyberattacks on critical infrastructure as amendment to the Information Security Act (ISA); enacted in April 2025 (Federal Office for Cybersecurity [2025]).

[11]In addition, Article 29 (2) FINMA [2024] requires companies under FINMA supervision to report any *"incident that is of substantial importance to the supervision,"* affecting financial institutions, insurance companies and stock exchange trading platforms.

[12]In comparison, the US Securities and Exchange Commission (SEC) implemented stricter disclosure rules in July 2023, obliging publicly traded companies to disclose any material cybersecurity breach (US Securities and Exchange Commission [2023]).

[13]To facilitate complaints, several cantonal polices offer the online tool *Swiss epolice* for minor offenses, including three types of cybercrime (Polizeitechnik und Informatik Schweiz [2023]).

Table 4: Determinants of Reporting (E)

Instrument W	Date Entering in Force	Description
NIS1	2018-05-09	EU reporting critical infrastructure
DSGVO	2018-05-25	EU reporting companies
ePolice SG	2023-06-01	SG online crime reporting
ePolice ZG	2023-07-04	ZG online crime reporting
DSG	2023-09-01	CH reporting companies (*DSG* Article 24)
NIS2	2024-10-18	EU strengthening of NIS1
ISA	2025-04-01	CH reporting critical infrastructure
AG	2025	If legal form is Limited Company (AG) 1, else 0
Listed	2025	If company is listed on the stock exchange 1, else 0
Not_Virtual	2025	If no indication for virtual office 1, else 0
Not_Canton	2025	If company is registered in other canton 1, else 0

This table gives an overview of changes in EU and Swiss law affecting the probability that companies report cybercrime incidents in Switzerland. Additionally, we add several firm characteristics as instruments: limited companies (*AG*) and stock market listed firms (*Listed*) are subject to stricter regulatory provisions. Furthermore, we add information on the indication that the company is only registered as virtual office, which is mostly relevant for Canton Zug. A significant number of incidents are reported by companies that are not registered in the canton where they reported, while (by construction of our sample) all companies that never report are registered in the respective canton. Adding being registered in another canton (*Not_Canton*) as an instrument addresses this issue (Tables 5, 6 and Appendix Table 31). The variables we included are in italics.

nance, which may indicate higher reporting propensity while also correlating with greater security investments (Higgs et al. [2016], Amir et al. [2018]). Taking into account time and previous reporting (R_{prev}), we explicitly allow for companies being generally more prone to report due to good governance (Amir et al. [2018]) while also facing a higher probability of being the victim of cybercrime. Our method allows for both factors to be at work simultaneously, accommodating the possibility that firms at high risk invest more in preventive measures (decreasing the probability of a second cyber incident), as well as the possibility that firms with prior cybercrime experience show greater propensity to report subsequent incidents (Kamiya et al. [2021], see also Ormosi [2014] for Capture-Recapture).

We employ incidents of crime, provided by the Cantonal Police Bureau St Gallen and the Police of Zug, which contain a digital component, i.e., the Article SCC joined with modus operandi defining cybercrime by the Federal Statistical Office.[14] We limit the analysis to companies with a Unique Enterprise Identification (UID) (Swiss Federal Council [2022b]), registered in the publicly accessible database Orbis (Moody's Analytics [2024]). As our empirical strategy requires the universe of companies that are potentially and eventually subject to cybercrime, we link the incidents to all companies legally incorporated in both cantons. Also from Orbis, we retrieve the number of employees, leaving out observations with zero registered employees. We also retrieve the legal form of incorporation (liability company (AG), limited liability company (GmbH), sole proprietorship, association, etc. (Swiss Federal Council [2023b])) and if the company is listed on the stock market. We anonymize the data and add indicator variables for additional instruments capturing the external shocks affecting cybercrime (Z) and the reporting thereof (W) (Tables 3 and 4). We also add the variable R_{prev}, which is 1 for a company and month if the company has already reported a previous event in the last 365 days.

Table 5 provides an overview on the sample of cyber incidents on companies reported in St Gallen (Upper Panel) and Zug (Lower Panel). For St Gallen, there are 1'285 reported incidents between 2016 and 2023, an average of 161 incidents per year for a population of, on average, 38'443 registered companies per year, or 0.03% of reported incidents per company and month. In comparison, in Zug, we have only 360 reported incidents

[14]From the Police of Zug, we receive the following list of SCC Articles and modus operandi used for the Swiss Police Crime Statistics (2023); Articles SCC 143, 144bis, 146, 147, 156, 160, 162, 173, 174, 177, 179, 180, 181, 187, 197, 198, 239, 251, 252, 261bis, 305bis, 320, 321, combined with digital modus operandi starting with 6 and modus operandi 5501300.

between 2016 and 2024; an average of only 40 for a population of 36'747 registered companies per year or 0.01% reported cases per company and month. Out of the reported incidents, several (with a maximum of 134) incidents are reported by the same company.[15] Basis for our estimation is the whole population of registered (in Orbis database) companies build as a panel data set over all years and months, which makes a total of 3'690'564 and 3'968'712 observations for St Gallen and Zug, respectively. Summary statistics on this monthly firm level are shown in Appendix Table 31.

The mean number of employees (*N_Emp*) for reported incidents is 928 for St Gallen and 387 for Zug, while on firm level, the average number of employees (as average of the yearly reported values in Orbis) is 283 and 427. The large difference between incident level and firm level in St Gallen is due to one very large firm reporting multiple incidents, skewing the distribution. However, the average company registered in St Gallen and Zug has only 12 and 13 employees, respectively. Most companies registered in St Gallen have only one employee, while most in Zug have four employees. This makes the distribution of *N_Emp* highly skewed with a standard deviation of 542 and 678, respectively. We differentiate companies by size, following the OECD definition (OECD [2024]), with small companies having less than 50 employees, medium-sized companies have 50 to 249 employees and large enterprises employ 250 or more employees. We suspect that companies with many employees tend to experience cyber incidents more often and use *Mid_Sized* and *Large_Sized* as part of our instruments affecting cybercrime.

Of all incidents, 77% and 81% are reported by a limited company (*AG*). However, the share of *AG* in the population of all registered companies is only 34% and 55% in St Gallen and Zug, respectively. In St Gallen, the share of *AG* in the sample of reporting firms (68%) is twice as large as in the population, while in Zug it is only 1.5 times as large (79%). Of all reporting companies in St Gallen, 0.56% are stock market listed firms (*Listed*), compared to 0.38% of reporting companies in Zug. This is a large share compared to all registered firms, where only 0.03% and 0.08% are stock market listed in St Gallen and Zug, respectively. Again, the share of stock market listed firms reporting in St Gallen is much larger compared to the population than in Zug. Companies with corporation type *AG* and *Listed* firms tend to have much more reported incidents. Assuming this is due to stricter reporting rules and disclosure obligations makes both variables a plausible instrumental variable for the incentive to report (enforcement *E*).

On average, 0.0001 incidents are reported per company and month in Canton Zug. This is only around one third of the Canton St Gallen reporting 0.0003 cases per company and month. The rate of companies per capita is more than three times higher in the Canton Zug (0.3 with a total number of inhabitants of 130'000 in 2023) compared to the Canton St Gallen or entire Switzerland more broadly (around 0.08, with a total number of inhabitants in St Gallen of 535'000 in 2023 (Swiss Federal Office of Justice [2024], Swiss Federal Statistical Office [2024a])). However, the number of incorporated companies is almost identical. A main reason explaining the low incident rate, the large rate of companies per capita and the differences in number of employees, share of *AG* and share of *Listed* companies could be the possible existence of virtual offices in Canton Zug. The Canton Zug is known for facilitating company incorporation (Carbó and Regenass [2021], Domizilagentur [2024]). To account for this phenomenon, we add an additional variable *Not_Virtual*, which is 1 if a company is not suspected to be a virtual office, and 0 if the following criterion applies: the company has only one employee and is either registered in an address that is shared with more than 100 other firms, or has "c/o" in its address which is

[15] Table 6 shows summary statistics on firm level for the 539 different reporting companies in St Gallen (Upper Panel) and the 260 reporting companies in Zug (Lower Panel).

shared with at least one more firm. We suspect that only 0.7% of all companies registered in St Gallen are virtual offices, but 18% of all companies in Zug. Confirming our restriction, all companies reporting a cyber incident are *Not_Virtual=1*.

Analyzing our sample of reporting companies, we notice that only 50% and 71% of them are also registered in the canton where they reported an incident. Around 16% and 12% are registered in Canton Zurich, the rest are registered in other cantons with smaller percentages. Of all reporting companies, 7% and 14% have reported in both cantons. Therefore, the population of all companies experiencing cyber events in one canton is the group of all companies that trade in this canton. Unfortunately, this data is not available. We use the population of all registered companies as a proxy for the population of companies exposed to cybercrime. From Orbis we get the statistical classification of economic activities (NACE (European Commission [2008])) of all registered companies. Of all incidents reported in St Gallen and Zug, 42% and 23% are reported by retail companies (*Retail*). A share of 13% and 7% is reported by online stores (*Online*). In contrast, the share of retail companies in the registered population is only 8% and 4% for St Gallen and Zug, respectively, while the share of online stores in the population is only 0.9% and 0.6%.

Dividing the sample of reporting firms into companies registered in the reporting canton and those registered in another canton shows for St Gallen a share of 14% in retail stores and 3% online stores for the former group, but 43% retail and 14% online stores for the latter group. For Zug, we get 6% in retail stores and 2% online stores for companies registered in Zug, versus 39% and 11% for companies registered elsewhere. This suggests that a large proportion of incidents are caused by major retailers, including some online retailers. The group of reporting firms that are registered elsewhere is also more prone to multiple reports by the same firm.[16] Reporting (R) and having reported in a previous time period (R_{prev}) are binary variables being 0 most often. R_{prev} takes the value 1 during the 12 months following a reported incident. Both variables show a large standard deviation relative to their means.

Table 5: Summary Statistics Sample of Cybercrime Incidents

	Mean	Median	SD	Min	Max	Skew	Obs
Year	2019.6023	2019	2.26	2016	2023	0.05	1'285
N_Emp	928.4887	4	9'369.37	1	106'622	11.10	1'285
Mid_Sized	0.2389	0	0.43	0	1	1.22	1'285
Large_Sized	0.0272	0	0.16	0	1	5.80	1'285
Online	0.1323	0	0.34	0	1	2.17	1'285
Retail	0.4163	0	0.49	0	1	0.34	1'285
AG	0.7696	1	0.42	0	1	-1.28	1'285
Listed	0.0086	0	0.09	0	1	10.66	1'285
Not_Virtual	0.0000	0	0	0	0		1'285
Not_Canton	0.7424	1	0.44	0	1	-1.11	1'285

	Mean	Median	SD	Min	Max	Skew	Obs
Year	2020.5444	2021	2.43	2016	2024	-0.14	360
N_Emp	387.1667	4	5'157.63	1	96'793	18.18	360
Mid_Sized	0.2056	0	0.40	0	1	1.45	360
Large_Sized	0.0306	0	0.17	0	1	5.43	360
Online	0.0667	0	0.25	0	1	3.46	360
Retail	0.2333	0	0.42	0	1	1.26	360
AG	0.8056	1	0.40	0	1	-1.54	360
Listed	0.0028	0	0.05	0	1	18.82	360
Not_Virtual	0.0000	0	0	0	0		360
Not_Canton	0.4028	0	0.49	0	1	0.39	360

This table shows the summary statistics on incident level for Canton St Gallen (Upper Panel) and Canton Zug (Lower Panel).

[16] Appendix Tables 29 and 30.

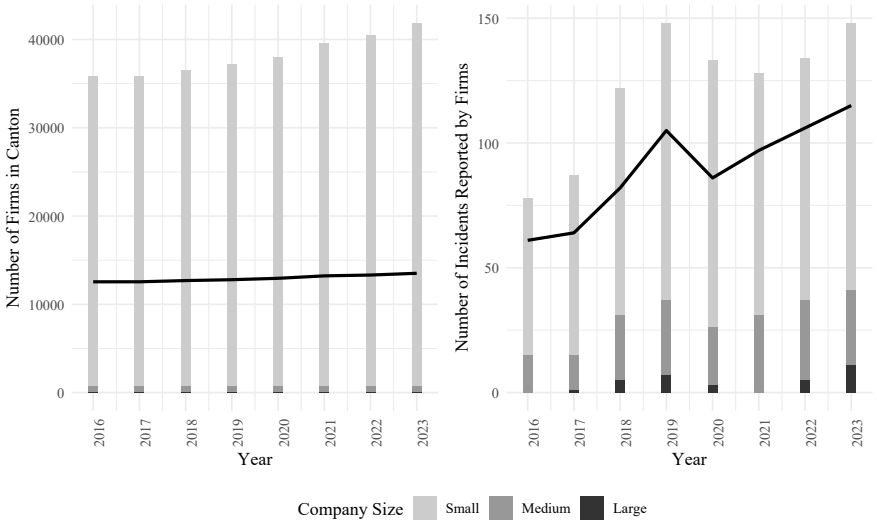


Figure 2: Registered Firms and Reported Cybercrime Cases on Firms in Canton St Gallen
Note: This figure illustrates the difference over time in number of employees and share of *AG* between the population of all registered companies (Left Panel) and the sample of reported incidents (Right Panel) in Canton St Gallen.

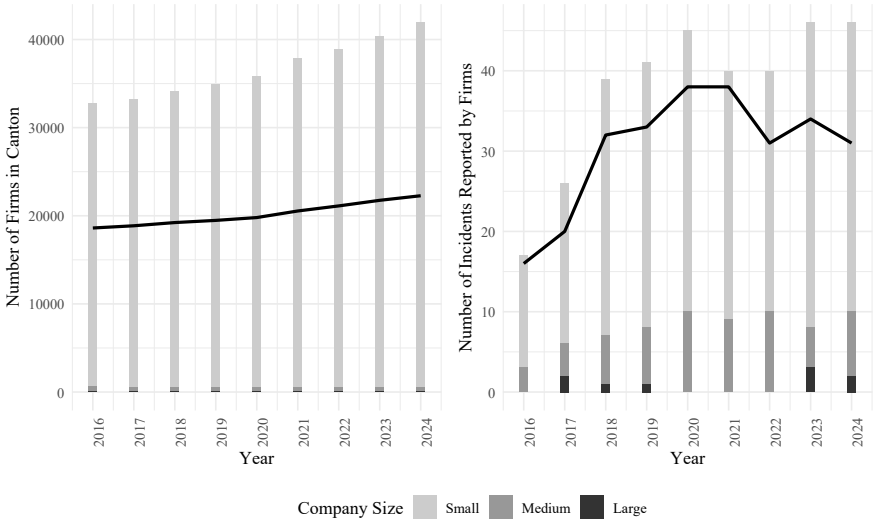


Figure 3: Registered Firms and Reported Cybercrime Cases on Firms in Canton Zug
This figure illustrates the difference over time in number of employees and share of *AG* between the population of all registered companies (Left Panel) and the sample of reported incidents (Right Panel) in Canton Zug.

Figures 2 and 3 illustrate the difference in number of employees and share of *AG* between the population of all registered companies and the sample of reported incidents. The left panel of both figures shows all registered companies in St Gallen from 2016 to 2023 and Zug from 2016 to 2024, respectively. The number of registered companies increases over the sample horizon from 36'759 to 41'916 for St Gallen and from 32'922 to 41'939 for Zug. Differentiating companies by size, most companies in both cantons are small companies with less than 50 employees. The black line indicates the number of companies with the legal form *AG*, which shows a slightly decreasing share of all firms of 0.34 to 0.32 for St Gallen and 0.57 to 0.53 for Zug, respectively. The right panel of both Figures 2 and 3 shows the number of all incidents reported by companies for the same time interval as the left panel. There is a significant increase in reported cases starting with an increase from 78 in 2016 to almost 148 in 2023 for St Gallen and from 17 reported cases in 2016 to 46 in 2024 for Zug. Focusing on company

size and reporting, we observe for both cantons that the share of medium and large enterprises is much higher in the sample of reported cases than in the population of all registered companies. In analogy to the left panel, the black line indicates companies with the legal form *AG*. Although 34% of all Canton St Gallen companies are incorporated as *AG*, the reported sample shows a much higher share of, on average, 77%. In Canton Zug, 55% of all companies have the legal form *AG*, but in the sample of reported cases the share of *AGs* is 81% on average. Figures 2 and 3 show that company size could be a good instrument for the likelihood of a cyber incident, while we use *AG* as instrument for enforcement.

Table 6: Cybercrime Reporting Firms in St Gallen and Zug

	Mean	Median	SD	Min	Max	Skew	Obs
Mean(*N_Emp*)	283.1437	4	4'659.54	1	106'370.5	22.02	539
Mid_Sized	0.1725	0	0.38	0	1.0	1.73	539
Large_Sized	0.0148	0	0.12	0	1.0	8	539
Online	0.0872	0	0.28	0	1	2.92	539
Retail	0.2839	0	0.45	0	1	0.96	539
AG	0.6827	1	0.47	0	1	-0.78	539
Listed	0.0056	0	0.07	0	1	13.25	539
Not_Virtual	0.0000	0	0	0	0.0		539
Not_Canton	0.4991	0	0.50	0	1.0	0	539

	Mean	Median	SD	Min	Max	Skew	Obs
Mean(*N_Emp*)	426.8436	4	6'006.57	1	96'793	15.88	260
Mid_Sized	0.1615	0.0	0.37	0	1	1.83	260
Large_Sized	0.0231	0.0	0.15	0	1	6.32	260
Online	0.0423	0	0.20	0	1	4.52	260
Retail	0.1538	0	0.36	0	1	1.91	260
AG	0.7885	1	0.41	0	1	-1.40	260
Listed	0.0038	0	0.06	0	1	15.94	260
Not_Virtual	0.0000	0.0	0	0	0		260
Not_Canton	0.2885	0.0	0.45	0	1	0.93	260

This table shows the summary statistics on firm level for companies reporting incidents in Canton St Gallen (Upper Panel) and Canton Zug (Lower Panel).

Table 7: Registered Firms (Not Reporting) in St Gallen and Zug

	Mean	Median	SD	Min	Max	Skew	Obs
Mean(*N_Emp*)	6.6692	1	103.25	1	12'110.67	101.79	53'401
Mid_Sized	0.0134	0	0.12	0	1	8.46	53'401
Large_Sized	0.0004	0	0.02	0	1	50.40	53'401
Online	0.0104	0	0.10	0	1	9.64	53'401
Retail	0.0897	0	0.29	0	1	2.87	53'401
AG	0.2885	0	0.45	0	1	0.93	53'401
Listed	0.0002	0	0.01	0	1	73.05	53'401
Not_Virtual	-0.0075	0	0.09	-1	0	-11.39	53'401
Not_Canton	0.0000	0	0	0	0		53'401

	Mean	Median	SD	Min	Max	Skew	Obs
Mean(*N_Emp*)	8.4026	1	385.81	1	74'794.20	163.01	55'139
Mid_Sized	0.0113	0	0.11	0	1	9.24	55'139
Large_Sized	0.0004	0	0.02	0	1	50.03	55'139
Online	0.0061	0	0.08	0	1	12.73	55'139
Retail	0.0416	0	0.20	0	1	4.59	55'139
AG	0.5203	1	0.50	0	1	-0.08	55'139
Listed	0.0006	0	0.02	0	1	40.23	55'139
Not_Virtual	-0.1953	0	0.40	-1	0	-1.54	55'139
Not_Canton	0.0000	0	0	0	0		55'139

This table shows the summary statistics on firm level for companies that do not report any incidents, but are registered in Canton St Gallen (Upper Panel) and Canton Zug (Lower Panel).

To select instrumental variables for the estimations, we rely on economic reasoning to justify the relevance

condition. However, for a final decision on the instruments, we run the estimator to check that the gradient in the estimated parameters is close to zero and the Hessian is non-singular. Between the models that meet the above conditions, we compare the significance and confidence intervals of the estimated coefficients together with the condition number.

We build the instruments for St Gallen and Zug as follows:

$$\begin{aligned} &\text{Model I} \\ &\text{IV on Crime: } Z = Covid + Online + Retail \\ &\text{IV on Enforcement: } W = AG + DSG + Listed \\ &\quad + NIS1 + Not\_Virtual + Not\_Canton \end{aligned} \tag{9}$$

$$\begin{aligned} &\text{Model II} \\ &\text{IV on Crime: } Z = Covid + Large\_Sized \\ &\quad + Mid\_Sized + Online + Retail \\ &\text{IV on Enforcement: } W = AG + DSG + Listed \\ &\quad + NIS1 + Not\_Virtual + Not\_Canton \end{aligned} \tag{10}$$

The instrument for victim of cybercrime (Z) consists of 1 for the periods when Switzerland was mostly in home office (2020-03-16 until 2022-02-17). In addition, we add 1 if the reporting company is in retail trade, plus 1 if it is also an online store. As alternative Model II, we include company size as part of the instrument (1 for companies with at least 50 employees plus 1 for companies with at least 250 employees). The instrument for enforcement (W) consists of the legal form (1 for AG), the listing status (1 for stock market listed firm) and the two policy shock variables EU NIS1 and CH DSG. With the EU regulation entering into force in 2018, it affects 71% and 76% of the sample period observations. The second policy shock (CH DSG) accounts for 3.4% of the sample in St Gallen, as DSG is only enacted at the end of the sample horizon in September 2023, and 15.7% in Zug due to the larger sample period. The introduction of DSG, as well as the later policies NIS 2 and ISA might have a considerable effect on reporting in a later stage. Future research should focus on the introduction of these laws. We add -1 if there is an indication that the company is registered solely as a virtual office, plus 1 if the company is registered in a canton different from the one where the incident was reported.[17]

To derive the number of companies that might have been subject to cybercrime in St Gallen and Zug between 2016 and 2023, 2024, respectively, we run the method on the universe of companies with very few ever reporting an incident. The method shows to be sufficiently flexible to capture this real-world phenomenon. Note that the coefficients show the real-life effects of the instruments and reporting previously on enforcement and on being subject to cybercrime. As in the simulation in Section 3, we run 20 optimization rounds for each model and canton with randomly set initial parameters (Appendix Tables 17 to 20). Although only 25% (Model I and Model II ZG) and 10% (Model II SG) of the runs converge to the best solution found, this solution is consistently superior to the alternative runs, confirming it as the global maximum likelihood. The low convergence rate reflects the complex geometry of the likelihood surface rather than uncertainty about the

[17] A correlation table of observed instruments and reporting variables is provided in Appendix Tables 22 to 25.

optimum.

Tables 8 and 9 provide the 95% confidence intervals for all estimators (Figure 1) for Models I and II. The first three rows of Tables 8 and 9 show the results for the probit estimating the effect of the instrument measuring the likelihood of cybercrime (Z) and reporting in a previous period (R_{prev}) on cyber incidents (C). Rows 4 to 6 show the corresponding results for the effects of the instrument (W) and previously reporting on enforcement (E). The parameter η (the effect of previously reporting on cybercrime) is not significant for Model I on Zug data. This might be due to the low correlations between instrument and cybercrime (Appendix Tables 22 to 25). The estimated coefficients for all other instruments and control variables are significantly different from zero. They also show the same positive direction and nearly identical magnitude across models and cantons.

The last two rows of Tables 8 and 9 list the estimated values for enforcement of a cyber incident ($\hat{E}$) and for the percentage of cybercrime ($\hat{C}$), both calculated with the estimated parameters ($\hat{\theta}$). For Model I, we estimate that 17.85% of all companies incorporated in Canton St Gallen have been subject to a cyber incident on average per month. Given that only 0.03% of all companies per month reported an incident, we find a detection rate for cybercrime of 0.17%, which corresponds to a dark rate of 99.83%. For Canton Zug, we estimate with Model I that 9.79% of all incorporated companies have been subject to a cyber incident per month. Given that only 0.01% of all companies reported an incident, we find that the dark rate of cybercrime is 99.90%. This dark figure is quite considerable, but to be interpreted with caution as not all estimates are significant.

For Model II, we estimate 10.76% and 2.35% of companies are victim of cyber incidents per month in St Gallen and Zug, respectively. This corresponds to a dark rate of 99.72% and 99.57%, respectively. Cantonal police bureaus support this finding with unpublished reports based on questionnaire data.

In Appendix Tables 14 and 15, we show results of the evaluation tests as in Section 4. For the estimator applied to the data of both St Gallen and Zug, the condition number is around 10 times higher than that of the simulated data, signaling lower estimation accuracy, but lower than the inferior solutions with local optima, which have condition numbers between e^5 and e^{19}. However, with LM statistics for the unrestricted models between 31 and 145, we fail to give evidence for the Null Hypothesis that the additional parameter is zero. Consequently, estimates should be regarded with caution. Although we have carefully selected the instruments, extended data including all Swiss cantons and especially cases reported directly to the Swiss National Cybersecurity Centre could greatly improve relevance and exogeneity.

6 Discussion

We propose a novel structural econometric methodology to solve the three persistent empirical problems that plague traditional estimation methods when estimating the dark figure of crime. First, there might be structural error in detection data and official statistics, leading to substantial under-policing and social welfare losses (Levitt [1998], Chalfin and McCrary [2018]); measurement error in crime data constitutes a non-random noise term. Our method therefore proposes and relies on the use of actual recorded incidents and not aggregated data of illegal activity. Second, the light figure of recorded crime is a potentially highly selective, non-random sample of the underlying population of criminal offenses (Biderman and Reiss [1967]). Traditional sample selection methods based on Heckman [1979] do not account for simultaneous selection by an offender to offend, an enforcer to identify, and/or a victim to report the offense. Feinstein [1990]'s DCE solves this problem only partially for law enforcement areas where one knows whether an offender or victim is part of a screening

Table 8: Estimated Parameters Cybercrime Cantons St Gallen and Zug (Model I)

	St Gallen		Zug	
	Lower	Upper	Lower	Upper
Constant on Crime ($\hat{c}$)	(-1.4723	-0.9523)	(-2.3732	-0.5812)
IV on Crime ($\hat{\varphi}$)	(0.5383	0.8148)	(0.2302	0.8351)
R_{prev} on Crime ($\hat{\eta}$)	(1.0380	2.6736)	(-0.0838	2.8188)
Constant on Enforcement ($\hat{d}$)	(-4.0838	-3.8381)	(-4.2484	-3.3839)
IV on Enforcement ($\hat{\xi}$)	(0.5337	0.6078)	(0.3174	0.4504)
R_{prev} on Enforcement ($\hat{\alpha}$)	(0.7421	1.0530)	(0.3530	1.4737)
$\hat{E}$ (Enforcement estimated on $\hat{\theta}$)	(0.0000	0.0024)	(0.0001	0.0039)
$\hat{C}$ (Cybercrime estimated on $\hat{\theta}$)	(0.1127	0.5560)	(0.0698	0.1724)
N	3'690'564		3'968'712	

Z (IV on Crime): *Covid + Online + Retail*
W (IV on Enforcement): *AG + DSG + Listed + NIS1 + Not_Virtual + Not_Canton*

This table shows for Model I the confidence intervals of the estimated parameters in Equations 1 and 2 for St Gallen and Zug. The confidence intervals show the same direction and almost identical magnitude across model and canton. With the exception of φ for Zug (change of sign), all parameters are significant.

Table 9: Estimated Parameters Cybercrime Cantons St Gallen and Zug (Model II)

	St Gallen		Zug	
	Lower	Upper	Lower	Upper
Constant on Crime ($\hat{c}$)	(-1.8050	-1.3647)	(-2.7721	-1.6117)
IV on Crime ($\hat{\varphi}$)	(0.5734	0.7977)	(0.3392	0.6365)
R_{prev} on Crime ($\hat{\eta}$)	(1.3078	2.8928)	(0.4392	2.5059)
Constant on Enforcement ($\hat{d}$)	(-3.9336	-3.6693)	(-3.8600	-3.0343)
IV on Enforcement ($\hat{\xi}$)	(0.5303	0.6034)	(0.3539	0.5058)
R_{prev} on Enforcement ($\hat{\alpha}$)	(0.5856	0.9013)	(0.0244	1.1763)
$\hat{E}$ (Enforcement estimated on $\hat{\theta}$)	(0.0001	0.0038)	(0.0003	0.0155)
$\hat{C}$ (Cybercrime estimated on $\hat{\theta}$)	(0.0565	0.4154)	(0.0142	0.0442)
N	3'690'564		3'968'712	

Z (IV on Crime): *Mid_Sized + Large_Sized + Covid + Online + Retail*
W (IV on Enforcement): *AG + DSG + Listed + NIS1 + Not_Virtual + Not_Canton*

This table shows for Model II the confidence intervals of the estimated parameters in Equations 1 and 2 for St Gallen and Zug. The confidence intervals show the same direction and almost identical magnitude across model and canton. All parameters are significant (no change of sign).

process or has been subject to an investigation. Finally, the simultaneity problem summarizes the empirical core obstacle of a criminal incident appearing in records iff two independent conditions are met concurrently:

- a criminal incident takes place (C)
- there is a possibility of enforcement (E).

We simultaneously run two structural probits with IVs that separate the conjoint $\Pr(R)$ modeled as $R = C \cdot E$. Identification hinges on exogenous IVs that are strictly excluded from shifting $\Pr(C)$ or $\Pr(E)$, thus providing the exclusion restrictions necessary to separate the contributions of C and E to the observed record R. The validity of the method and its assumed DGP in simulated data demonstrate the method's favorable statistical properties in finite samples. The simultaneous estimation requires finding the global optimum of a weakly bi-convex product of two probits. Identification hinges explicitly on the quality of the instruments, which poses a first limitation. The methodology uncovers the true population of subjects exposed to crime, demonstrating reliable convergence even when faced with extremely sparse reporting.

We are happy to see others using our novel method and that our empirical application confirms the critical policy relevance of quantifying the dark figure in the domain of cybercrime. The observed monthly reporting rates (0.03% and 0.01%) in the Cantons of St Gallen and Zug severely mask the true prevalence; estimated true prevalence rates of 2.4% to 17.9% yield dark rates exceeding 99%. The stability of our latent prevalence estimates across multiple instrument specifications provides initial empirical support for the identifying assumptions. However, the Lagrange Multiplier tests suggest that the exclusion restrictions may not hold perfectly, indicating that our point estimates should be interpreted with appropriate caution.

Despite its strong performance in simulations, the current specification has several limitations that guide future research. The model does not adequately address time, which is crucial for accurately incorporating policy shocks as instruments. Explicitly incorporating time could improve the quality and consistency of the instruments. The model may be susceptible to omitted variable bias. Introducing variables covering information sensitivity or R&D intensity may improve the explanation of why a company is subject to crime (C). Including features on profitability, reputation, or corporate governance can strengthen the instruments for enforcement (E). Future research should extend the data to include other Swiss cantons (e.g., ePolice activation) or use the sample of reported incidents from the Federal Office of the Police (Fedpol). Identifying the country of origin of the incident could provide a crucial instrument in explaining the probability of cyber incidents ($\Pr(C)$). Future work with richer instrument sets may further strengthen identification.

References

Accenture. 2019. Accenture Cost of Cybercrime Study 2019.

Amir, Eli, Levi, Shai, and Livne, Tsafrir. 2018. Do Firms Underreport Information on Cyber-Attacks? Evidence from Capital Markets. *Review of Accounting Studies* 23, 3 (Sept. 2018), 1177–1206.

Amstrup, Steven C., McDonald, Trent L., and Manly, Bryan F. 2005. *Handbook of Capture-Recapture Analysis*. Princeton University Press.

Anderson, Ross, Barton, Chris, Boehme, Rainer, Clayton, Richard, Ganan, Carlos, Grasso, Tom, Levi, Michael, Moore, Tyler, and Vasek, Marie. 2019. Measuring the Changing Cost of Cybercrime. (July 2019). https://doi.org/10.17863/CAM.41598

Anderson, Ross, Barton, Chris, Böhme, Rainer, Clayton, Richard, van Eeten, Michel J. G., Levi, Michael, Moore, Tyler, and Savage, Stefan. 2013. Measuring the Cost of Cybercrime. In *The Economics of Information Security and Privacy*, Rainer Böhme (Ed.). Springer, Berlin, Heidelberg, 265–300.

August, Terrence, Dao, Duy, Kim, Kihoon, and Niculescu, Marius Florin. 2025. The Impact of Cryptocurrency on Cybersecurity. *Management Science* 71, 11 (Nov. 2025), 9606–9627.

August, Terrence, Noh, Daehoon, Shamir, Noam, and Shin, Hyoduk. 2024. Cyberattacks, Operational Disruption, and Investment in Resilience Measures. *Management Science* (Dec. 2024). https://doi.org/10.1287/mnsc.2022.00430

Bayes, Thomas. 1958. An Essay Towards Solving a Problem in the Doctrine of Chances. *Biometrika* 45, 3-4 (1958), 296–315.

Becker, Gary. 1968. Crime and Punishment: An Economic Approach. *The Journal of Political Economy* 76 (1968), 169–217.

Bellert, Nicole. 2025. Estimate Dark Figures of Crime. https://github.com/belln1/drcr.

Bellert, Nicole, Günster, Andrea, and Kozbur, Damian. 2023. Simulating Collusion: Challenging Conventional Estimation Methods. In *CRESSE Competition & Regulation European Summer School and Conference 2023, Rhodes; 2023 SSES Annual Congress, Neuchatel*.

Bessy-Roland, Yannick, Boumezoued, Alexandre, and Hillairet, Caroline. 2021. Multivariate Hawkes Process for Cyber Insurance. *Annals of Actuarial Science* 15, 1 (March 2021), 14–39.

Biderman, Albert and Reiss, Albert J. Jr. 1967. On Exploring the "Dark Figure" of Crime. *The Annals of the American Academy of Political and Social Science* 374, 1 (1967), 1–15.

Biderman, Albert D and Lynch, James P. 2012. *Understanding Crime Incidence Statistics: Why the Ucr Diverges from the Ncs*. Springer Science & Business Media.

Bos, Iwan and Harrington, Joseph E. 2010. Endogenous Cartel Formation with Heterogeneous Firms. *The RAND Journal of Economics* 41, 1 (2010), 92–117.

Breusch, Trevor S. and Pagan, Adrian R. 1980. The Lagrange Multiplier Test and Its Applications to Model Specification in Econometrics. *The Review of Economic Studies* 47, 1 (1980), 239–253.

Broyden, Charles G. 1970. The Convergence of a Class of Double-rank Minimization Algorithms 1. General Considerations. *IMA Journal of Applied Mathematics* 6, 1 (March 1970), 76–90.

Bryant, Peter G. and Eckard, Edwin. 1991. Price Fixing: The Probability of Getting Caught. *Review of Economics and Statistics* 73 (1991), 531–540.

Buil-Gil, David, Medina, Juanjo, and Shlomo, Natalie. 2021. Measuring the Dark Figure of Crime in Geographic Areas: Small Area Estimation from the Crime Survey for England and Wales. *The British Journal of Criminology* 61, 2 (March 2021), 364–388.

Carbó, Adrià Budry and Regenass, Romeo. 2021. Zug - ein Offshore-Eldorado für Briefkastenfirmen. https://www.publiceye.ch/de/themen/korruption/die-schweiz-ein-offshore-paradies/zug.

Chalfin, Aaron and McCrary, Justin. 2018. Are U.S. Cities Underpoliced? Theory and Evidence. *The Review of Economics and Statistics* 100, 1 (March 2018), 167–186.

Chan, Lax, Silverman, Bernard W., and Vincent, Kyle. 2021. Multiple Systems Estimation for Sparse Capture Data: Inferential Challenges When There Are Non-Overlapping Lists. *J. Amer. Statist. Assoc.* 116, 535 (July 2021), 1297–1306. arXiv:1902.05156 [stat]

Chen, Joe and Harrington, Joseph E. 2007. The Impact of the Corporate Leniency Program on Cartel Formation and the Cartel Price Path. *Contributions to Economic Analysis* 282 (2007), 59–80.

Chen, Zhijun and Rey, Patrick. 2013. On the Design of Leniency Programs. *The Journal of Law and Economics* 56, 4 (Nov. 2013), 917–957. https://doi.org/10.1086/674011

Combe, Emmanuel, Monnier, Constance, and Legal, Renaud. 2008. Cartels: The Probability of Getting Caught in the European Union. *Available at SSRN 1015061* (2008). https://doi.org/10.2139

Cornwell, Christopher and Trumbull, William N. 1994. Estimating the Economic Model of Crime with Panel Data. *The Review of Economics and Statistics* 76, 2 (1994), 360–366. https://doi.org/10.2307/2109893 arXiv:2109893

Council of Europe. 2001. Convention on Cybercrime, ETS No 185.

Craig, Steven G. 1987. The Deterrent Impact of Police: An Examination of a Locally Provided Public Service. *Journal of Urban Economics* 21, 3 (May 1987), 298–311.

CyberPeace Institute. 2024. Timeline of Cyberattacks and Operations. https://cyberconflicts.cyberpeaceinstitute.org/threats/timeline.

Davidson, James and MacKinnon, James G. 2004. *Econometric Theory and Methods.* New York, Oxford University Press.

Davies, Stephen W. and Ormosi, Peter L. 2012. A Comparative Assessment of Methodologies Used to Evaluate Competition Policy. *Journal of Competition Law and Economics* 00 (2012), 1–35.

Domizilagentur, GmbH. 2024. Virtual Office vs. Letterbox Company in Switzerland, Zug and Zurich. https://www.domizilagentur.ch/en/virtual-office-vs-letterbox-company-in-switzerland-zug-and-zurich/.

Duffee, David, McDowall, David, Mazerolle, Lorraine Green, and Mastrofski, Stephen D. 2000. Measurement and Analysis of Crime and Justice: An Introductory Essay. *Measurement and Analysis of Crime and Justice* 4 (2000), 1–31.

Duso, Tomaso, Röller, Lars-Hendrik, and Seldeslachts, Jo. 2014. Collusion Through Joint R&D: An Empirical Assessment. *The Review of Economics and Statistics* 96, 2 (May 2014), 349–370.

Elluri, Lavanya, Mandalapu, V., Vyas, P., and Roy, Nirmalya. 2023. Recent Advancements in Machine Learning for Cybercrime Prediction. *Journal of Computer Information Systems* 65 (2023), 249–263.

ETCISO. 2024. Ai Rise Will Lead to Increase in Cyberattacks - Et Ciso.

European Commission (Ed.). 2008. *NACE Rev. 2: Statistical Classification of Economic Activities in the European Community*. Publications Office, Luxembourg.

European Union. 2016a. Directive 2016/1148 of the European Parliament and of the Council (NIS1).

European Union. 2016b. Regulation 2016/679 of the European Parliament and of the Council (General Data Protection Regulation).

European Union. 2022. Directive 2022/2555 of the European Parliament and of the Council (NIS2).

Fé, Eduardo. 2024. Partial Identification of the Dark Figure of Crime with Survey Data Under Misreporting Errors. *Journal of Quantitative Criminology* (2024).

Federal Data Protection and Information Commissioner. 2018. Die EU-Datenschutzgrundverordnung und ihre Auswirkungen auf die Schweiz. https://www.kmu.admin.ch/kmu/de/home/fakten-trends/digitalisierung/datenschutz/eu-regelung-zum-datenschutz.html.

Federal Office for Cybersecurity. 2025. Legal Basis for the Reporting Obligation. https://www.ncsc.admin.ch/ncsc/en/home/meldepflicht/gesetzliche-grundlagen-mp.html.

Feinstein, Jonathan S. 1989. The Safety Regulation of U.S. Nuclear Power Plants: Violations, Inspections, and Abnormal Occurrences. *Journal of Political Economy* 97, 1 (1989), 115–154. arXiv:1831057

Feinstein, Jonathan S. 1990. Detection Controlled Estimation. *The Journal of Law and Economics* 33, 1 (April 1990), 233–276.

Feinstein, Jonathan S. 1991. An Econometric Analysis of Income Tax Evasion and Its Detection. *The RAND Journal of Economics* 22, 1 (1991), 14–35. https://doi.org/10.2307/2601005 arXiv:2601005

Fienberg, Stephen E. 1972. The Multiple Recapture Census for Closed Populations and Incomplete 2k Contingency Tables. *Biometrika* 59, 3 (1972), 591–603.

FINMA. 2020a. Aufsichtsmitteilung 05/2020: Meldepflicht von Cyber-Attacken Gemäss Art. 29 Abs. 2 FINMAG. Eidgenössische Finanzmarktaufsicht.

FINMA. 2020b. New Cyber Supervisory Approach and Guidance. Eidgenössische Finanzmarktaufsicht. https://www.finma.ch/en/documentation/dossier/dossier-cyberrisiken/new-cyber-supervisory-approach-and-guidance/.

FINMA. 2024. Authorisation to Operate on the Financial Market. Eidgenössische Finanzmarktaufsicht. https://www.finma.ch/en/authorisation/authorisation-overview/.

Fletcher, Roger. 1970. A New Approach to Variable Metric Algorithms. *Comput. J.* 13, 3 (Jan. 1970), 317–322. https://doi.org/10.1093/comjnl/13.3.317

Fletcher, R. and Reeves, C. M. 1964. Function Minimization by Conjugate Gradients. *Comput. J.* 7, 2 (Jan. 1964), 149–154.

Foley, Sean, Karlsen, Jonathan R, and Putniņš, Tālis J. 2019. Sex, Drugs, and Bitcoin: How Much Illegal Activity Is Financed through Cryptocurrencies? *The Review of Financial Studies* 32, 5 (May 2019), 1798–1853.

Foros, O. 2004. Strategic Investments with Spillovers, Vertical Integration and Foreclosure in the Broadband Access Market. *International Journal of Industrial Organization* 22, 1 (2004), 1–24.

Gärtner, D. L. and Zhou, Jun. 2012. *Delays in Leniency Application: Is There Really a Race to the Enforcer's Door?* Discussion Paper 2012-044. Tilburg University, Tilburg Law and Economic Center.

Goldfarb, Donald. 1970. A Family of Variable-Metric Methods Derived by Variational Means. *Math. Comp.* 24, 109 (1970), 23–26. https://doi.org/10.1090/S0025-5718-1970-0258249-6

Günster, Andrea. 2010. *On European Antitrust Enforcement.* Ph. D. Dissertation. University of Maastricht.

Harrington, Joseph E. 2008. Optimal Corporate Leniency Programs. *The Journal of Industrial Economics* 56, 2 (2008), 215–246.

Harrington, Joseph E. 2013. Corporate Leniency Programs When Firms Have Private Information: The Push of Prosecution and the Pull of Pre-emption. *The Journal of Industrial Economics* 61, 1 (March 2013), 1–27.

Harrington, Joseph E. and Chang, Myong-Hun. 2009. Modeling the Birth and Death of Cartels with an Application to Evaluating Competition Policy. *Journal of the European Economic Association* 7, 6 (2009), 1400–1435.

Harrington, Joseph E. and Chang, Myong-Hun. 2015. When Can We Expect a Corporate Leniency Program to Result in Fewer Cartels? *The Journal of Law & Economics* 58, 2 (2015), 417–449.

Harrington, Joseph E. and Wei, Yanhao. 2017. What Can the Duration of Discovered Cartels Tell Us About the Duration of All Cartels? *The Economic Journal* 127, 604 (2017), 1977–2005.

Haverkamp, Rita. 2020. An Overview of the Research on the Dark Figure of Crime in Germany. Concept, Methods and Development. *SIAK-Journal - Journal for Police Science and Practice* 10 (2020), 39–53. https://doi.org/10.7396/IE_2020_D

Heckman, James J. 1979. Sample Selection Bias as a Specification Error. *Econometrica: Journal of the econometric society* (1979), 153–161.

Heim, Sven, Hüschelrath, Kai, Laitenberger, Ulrich, and Spiegel, Yossi. 2022. The Anticompetitive Effect of Minority Share Acquisitions: Evidence from the Introduction of National Leniency Programs. *American Economic Journal: Microeconomics* 14, 1 (2022), 366–410.

Hellwig, Michael and Hüschelrath, Kai. 2018. When Do Firms Leave Cartels? Determinants and the Impact on Cartel Survival. *International Review of Law and Economics* 54 (June 2018), 68–84.

Herjavec. 2019. The 2019/2020 Official Annual Cybersecurity Jobs Report - Herjavec Group.

Higgs, Julia L., Pinsker, Robert E., Smith, Thomas J., and Young, George R. 2016. The Relationship between Board-Level Technology Committees and Reported Security Breaches. *Journal of Information Systems* 30, 3 (Sept. 2016), 79–98. https://doi.org/10.2308/isys-51402

Hiscox. 2022. *Cyber Readiness Report 2022 | Hiscox*. Technical Report.

IBM Corporation. 2023. IBM Cost of a Data Breach Report 2023 - Bericht über die Kosten einer Datenschutzverletzung 2023. (2023).

Imbens, Guido W. and Angrist, Joshua D. 1994. Identification and Estimation of Local Average Treatment Effects. *Econometrica : journal of the Econometric Society* 62, 2 (1994), 467–475.

ISACA. 2023. State of Cybersecurity 2023 Report.

Isenhardt, Anna, Frey, Louise Emily, and Hostettler, Ueli. 2022. Befragung Zur Sicherheit in Unternehmen Bezüglich Digitaler Und Physischer Angriffe: Auswertungsbericht Zuhanden Des Verbands Swissmem. (2022).

Kääriäinen, Juha and Sirén, Reino. 2011. Trust in the Police, Generalized Trust and Reporting Crime. *European Journal of Criminology* 8, 1 (2011), 65–81.

Kafka, Franz. 1925. *Der Prozeß* (1. aufl. 2015, unbearb. orig.-ausgabe ed.). Damnick, Braunschweig.

Kamiya, Shinichi, Kang, Jun-Koo, Kim, Jungmin, Milidonis, Andreas, and Stulz, René M. 2021. Risk Management, Firm Reputation, and the Impact of Successful Cyberattacks on Target Firms. *Journal of Financial Economics* 139, 3 (March 2021), 719–749. https://doi.org/10.1016/j.jfineco.2019.05.019

Keeper Security. 2023. Keeper Security Releases Cybersecurity Disasters Survey: Incident Reporting & Disclosure.

Levenstein, Margaret C and Suslow, Valerie Y. 2008. International Cartels. *Issues in Competition Law and Policy* 2. Jg. (2008), 1107–1126.

Levitt, Steven D. 1998. The Relationship Between Crime Reporting and Police: Implications for the Use of Uniform Crime Reports. *Journal of Quantitative Criminology* 14, 1 (March 1998), 61–81. https://doi.org/10.1023/A:1023096425367

Loftin, Colin and McDowall, David. 2010. The Use of Official Records to Measure Crime and Delinquency. *Journal of Quantitative Criminology* 26, 4 (Dec. 2010), 527–532. https://doi.org/10.1007/s10940-010-9120-8

Lynch, James P and Addington, Lynn A. 2006. *Understanding Crime Statistics: Revisiting the Divergence of the Ncvs and the Ucr*. Cambridge University Press.

MacDonald, Ziggy. 2001. Revisiting the Dark Figure : A Microeconometric Analysis of the Under-reporting of Property Crime and Its Implications. *The British Journal of Criminology* 41, 1 (Jan. 2001), 127–149. https://doi.org/10.1093/bjc/41.1.127

Manski, Charles F. and Pepper, John V. 2013. Deterrence and the Death Penalty: Partial Identification Analysis Using Repeated Cross Sections. *Journal of Quantitative Criminology* 29, 1 (March 2013), 123–141. https://doi.org/10.1007/s10940-012-9172-z

Mathys, Roland. 2021. Eine Checkliste aus rechtlicher Sicht. Gastbeitrag der Rechtskommission von swissICT. https://www.computerworld.ch/security/best-practice/checkliste-rechtlicher-sicht-2670114.html.

McAfee. 2020. McAfee Press Release: The Hidden Costs of Cybercrime Report 2020.

McCrea, Rachel and Morgan, Byron. 2015. *Analysis of Capture-Recapture Data*. Chapman and Hall.

Messner, Steven. 1984. The "Dark Figure" and Composite Indexes of Crime: Some Empirical Explorations of Alternative Data Sources. *Journal of Criminal Justice* 12 (1984), 435–444. https://doi.org/10.1016/0047-2352(84)90091-6

Miller, Nathan H. 2009. Strategic Leniency and Cartel Enforcement. *American Economic Review* 99, 3 (2009), 750–768.

Moody's Analytics. 2024. Orbis Bureau van Dijk. https://login.bvdinfo.com/R1/Orbis.

Mosher, Clayton J, Miethe, Terance D, and Hart, Timothy C. 2010. *The Mismeasure of Crime*. Sage.

Motta, Massimo and Polo, Michele. 2003. Leniency Programs and Cartel Prosecution. *International Journal of Industrial Organization* 21, 3 (2003), 347–379.

Nagin, Daniel. 1978. Crime Rates, Sanction Levels, and Constraints on Prison Population. *Law & Society Review* 12, 3 (April 1978), 341–366. https://doi.org/10.2307/3053284

Naidoo, Rennie. 2020. A Multi-Level Influence Model of Covid-19 Themed Cybercrime. *European Journal of Information Systems* 29, 3 (May 2020), 306–321. https://doi.org/10.1080/0960085X.2020.1771222

National Cybersecurity Centre. 2023. *Semi-Annual Report 2023/1*. Technical Report.

National Cybersecurity Centre. 2024. NCSC Report. https://www.report.ncsc.admin.ch/en/.

Nelder, J. A. and Mead, R. 1965. A Simplex Method for Function Minimization. *Comput. J.* 7, 4 (Jan. 1965), 308–313.

Nocedal, Jorge and Wright, Stephen J. 2006. Quadratic Programming. In *Numerical Optimization*. Springer.

OECD. 2024. Enterprises by Business Size. https://www.oecd.org/en/data/indicators/enterprises-by-business-size.html.

Ormosi, Peter. 2014. A Tip of the Iceberg? The Probability of Catching Cartels. *Journal of Applied Econometrics* 29, 4 (2014), 549–566.

Orsagh, Thomas. 1973. Crime, Sanctions and Scientific Explanation. *The Journal of Criminal Law and Criminology (1973-)* 64, 3 (Sept. 1973), 354. https://doi.org/10.2307/1142379 arXiv:1142379

Panda Security. 2020. 43 COVID-19 Cybersecurity Statistics.

Paoli, Letizia, Visschers, Jonas, and Verstraete, Cedric. 2018. The Impact of Cybercrime on Businesses: A Novel Conceptual Framework and Its Application to Belgium. *Crime, Law and Social Change* 70, 4 (Nov. 2018), 397–420. https://doi.org/10.1007/s10611-018-9774-y

Pearl, Judea. 1995. Causal Diagrams for Empirical Research. *Biometrika* 82, 4 (1995), 669–688. https://doi.org/10.1093/biomet/82.4.669

Pina-Sánchez, Jose, Buil-Gil, David, Brunton-Smith, Ian, and Cernat, Alexandru. 2023. The Impact of Measurement Error in Regression Models Using Police Recorded Crime Rates. *Journal of Quantitative Criminology* 39, 4 (2023), 975–1002. https://doi.org/10.1007/s10940-022-09557-6

Polizeitechnik und Informatik Schweiz, PTI. 2023. Suisse ePolice – Der Schweizer Online-Polizeiposten. https://www.suisse-epolice.ch.

Quételet, Adolphe. 1832. Recherches Sur Le Penchant Au Crime Aux Différens Âges. *Mémoires de l'Académie royale de Belgique* 7, 1 (1832), 1–87. https://doi.org/10.3406/marb.1832.2744

Reep-van den Bergh, Carin M. M. and Junger, Marianne. 2018. Victims of Cybercrime in Europe: A Review of Victim Surveys. *Crime Science* 7 (2018). https://doi.org/10.1186/s40163-018-0079-3

Rivest, Louis-Paul and Baillargeon, Sophie. 2014. Capture–Recapture Methods for Estimating the Size of a Population. https://doi.org/10.1201/b16597-19

Schneider, Anne Larason, Wilson, LA, and Burcart, Janie. 1975. *Role of Attitudes in Decisions to Report Crimes to the Police*. Oregon Research Institute Eugene, Oregon.

Shanno, David. F. 1970. Conditioning of Quasi-Newton Methods for Function Minimization. *Math. Comp.* 24, 111 (1970), 647–656. https://doi.org/10.1090/S0025-5718-1970-0274029-X

SIX Exchange Regulation AG. 2024. Listing Rules. *SIX Exchange Regulation* (2024).

Skogan, Wesley G. 1977. Dimensions of the Dark Figure of Unreported Crime. *Crime & Delinquency* 23, 1 (Jan. 1977), 41–50. https://doi.org/10.1177/001112877702300104

Solano, Pablo Casais and Reinoso Peinado, Antonio José. 2017. Socio-Economic Factors in Cybercrime: Statistical Study of the Relation Between Socio-Economic Factors and Cybercrime. In *2017 International Conference On Cyber Situational Awareness, Data Analytics And Assessment (Cyber SA)*. 1–4. https://doi.org/10.1109/CyberSA.2017.8073392

Sovinsky, Michelle. 2022. Do Research Joint Ventures Serve a Collusive Function? *Journal of the European Economic Association* 20, 1 (2022), 430–475.

Spagnolo, G. 2008. Leniency and Whistleblowers in Antitrust. In *Buccirossi P.(Ed.), Handbook of Antitrust Economics*. The MIT Press, Cambridge Massachusetts.

Swiss Federal Council. 2020. Coronavirus: Federal Council Declares 'Extraordinary Situation' and Introduces More Stringent Measures. https://www.news.admin.ch/en/nsb?id=78454.

Swiss Federal Council. 2022a. Coronavirus: Federal Council Lifts Requirements to Quarantine and to Work from Home and Launches Consultation on a Widespread Easing of Measures. https://www.news.admin.ch/en/nsb?id=87041.

Swiss Federal Council. 2022b. Sme Portal for Small and Medium-Sized Enterprises: UID Register. https://www.kmu.admin.ch/kmu/en/home/savoir-pratique/creation-pme/creation-d_entreprise/registre_ide.html.

Swiss Federal Council. 2023a. Federal Law: Federal Act on Data Protection (FADP). https://www.fedlex.admin.ch/eli/cc/2022/491/de.

Swiss Federal Council. 2023b. Sme Portal for Small and Medium-Sized Enterprises: Legal Form: The Limited Company. https://www.kmu.admin.ch/kmu/en/home/concrete-know-how/setting-up-sme/les-differentes-formes-juridiques.html.

Swiss Federal Office of Justice. 2024. Handelsregister-Statistik.

Swiss Federal Statistical Office. 2021. Polizeiliche Kriminalstatistik (PKS) 2020.

Swiss Federal Statistical Office. 2023a. Polizeiliche Kriminalstatistik (PKS) 2023.

Swiss Federal Statistical Office. 2023b. Schema der Tatvorgehen digitale Kriminalität.

Swiss Federal Statistical Office. 2024a. Demografische Bilanz nach Kanton - 1971-2023. https://www.bfs.admin.ch/asset/de/32208093.

Swiss Federal Statistical Office. 2024b. Digital Crime.

Tarling, Roger and Morris, Katie. 2010. Reporting Crime to the Police. *The British Journal of Criminology* 50, 3 (2010), 474–490.

Todd, John. 1950. The Condition of a Certain Matrix. *Mathematical Proceedings of the Cambridge Philosophical Society* 46, 1 (Jan. 1950), 116–118.

US Department of Justice. 2015. Financial Fraud Crime Victims.

US Securities and Exchange Commission. 2023. Cybersecurity Risk Management, Strategy, Governance, and Incident Disclosure.

Wheeler, Sarah A., Round, David K., and Wilson, John K. 2011. The Relationship Between Crime and Electronic Gaming Expenditure: Evidence from Victoria, Australia. *Journal of Quantitative Criminology* 27, 3 (2011), 315–338. arXiv:23883825

Wolpin, Kenneth I. 1980. A Time Series–Cross Section Analysis of International Variation in Crime and Punishment. *The Review of Economics and Statistics* 62, 3 (Aug. 1980), 417. https://doi.org/10.2307/1927109 arXiv:1927109

Wu, Ling, Peng, Qiong, and Lembke, Michael. 2023. Research Trends in Cybercrime and Cybersecurity: A Review Based on Web of Science Core Collection Database. *International Journal of Cybersecurity Intelligence & Cybercrime* 6, 1 (March 2023), 5–28. https://doi.org/10.52306/2578-3289.1154

A Additional Estimated Models

Table 10: Estimated Parameters Cybercrime Cantons St Gallen and Zug (Model III)

	St Gallen		Zug	
	Lower	Upper	Lower	Upper
Constant on Crime ($\hat{c}$)	(-3.4539	-2.8455)	(-4.0010	-3.7799)
IV on Crime ($\hat{\varphi}$)	(0.2088	0.2781)	(0.1266	0.1655)
R_{prev} on Crime ($\hat{\eta}$)	(2.1848	3.2685)	(1.0008	2.1779)
Constant on Enforcement ($\hat{d}$)	(-2.3977	-1.4127)	(-1.4390	-0.6813)
IV on Enforcement ($\hat{\xi}$)	(0.1379	0.2407)	(0.2966	1.4059)
R_{prev} on Enforcement ($\hat{\alpha}$)	(-0.4373	0.5471)	(-1.8563	0.3711)
$\hat{E}$ (Enforcement estimated on $\hat{\theta}$)	(0.0284	0.0634)	(0.1445	0.9324)
$\hat{C}$ (Cybercrime estimated on $\hat{\theta}$)	(0.0008	0.0206)	(0.0001	0.0004)
N	3'690'564		3'968'712	

Z (IV on Crime): Ln(*N_Emp*) + *Covid* + *Online* + *Retail*
W (IV on Enforcement): *AG* + *DSG* + *Listed* + *NIS1* + *Not_Virtual*

This table shows for an additional Model III the confidence intervals of the estimated parameters in Equations 1 and 2 for St Gallen and Zug. The confidence intervals show the same direction and almost identical magnitude across model and canton. With the exception of α for Zug (change of sign), all parameters are significant.

Table 11: Estimated Parameters Cybercrime Cantons St Gallen and Zug (Model IV)

	St Gallen		Zug	
	Lower	Upper	Lower	Upper
Constant on Crime ($\hat{c}$)	(-2.2151	-1.8087)	(-3.1929	-2.1612)
IV on Crime ($\hat{\varphi}$)	(0.2707	0.3682)	(0.1570	0.2763)
R_{prev} on Crime ($\hat{\eta}$)	(1.5235	3.7571)	(0.9748	3.6283)
Constant on Enforcement ($\hat{d}$)	(-3.7372	-3.4322)	(-3.5717	-2.5880)
IV on Enforcement ($\hat{\xi}$)	(0.4552	0.5278)	(0.2931	0.4467)
R_{prev} on Enforcement ($\hat{\alpha}$)	(0.5090	0.8510)	(-0.4517	0.8486)
$\hat{E}$ (Enforcement estimated on $\hat{\theta}$)	(0.0002	0.0046)	(0.0010	0.0244)
$\hat{C}$ (Cybercrime estimated on $\hat{\theta}$)	(0.0221	0.2889)	(0.0037	0.0305)
N	3'690'564		3'968'712	

Z (IV on Crime): Ln(*N_Emp*) + *Covid* + *Online* + *Retail*
W (IV on Enforcement): *AG* + *DSG* + *Listed* + *NIS1* + *Not_Virtual* + *Not_Canton*

This table shows for an additional Model IV the confidence intervals of the estimated parameters in Equations 1 and 2 for St Gallen and Zug. The confidence intervals show the same direction and almost identical magnitude across model and canton. With the exception of α for Zug (change of sign), all parameters are significant.

Table 12: Estimated Parameters Cybercrime Cantons St Gallen and Zug (Model V)

	St Gallen		Zug	
	Lower	Upper	Lower	Upper
Constant on Crime ($\hat{c}$)	(-2.4941	-1.6498)	(-3.7464	-3.4873)
IV on Crime ($\hat{\varphi}$)	(0.3470	0.5467)	(0.1385	0.2448)
R_{prev} on Crime ($\hat{\eta}$)	(1.4511	2.7542)	(1.2198	1.8599)
Constant on Enforcement ($\hat{d}$)	(-3.2544	-2.5938)	(-1.8626	-1.2700)
IV on Enforcement ($\hat{\xi}$)	(0.2552	0.3509)	(0.5287	1.3147)
R_{prev} on Enforcement ($\hat{\alpha}$)	(0.5499	1.2171)	(-1.0193	1.1251)
$\hat{E}$ (Enforcement estimated on $\hat{\theta}$)	(0.0017	0.0102)	(0.0586	0.8847)
$\hat{C}$ (Cybercrime estimated on $\hat{\theta}$)	(0.0191	0.1193)	(0.0001	0.0003)
N	3'690'564		3'968'712	

Z (IV on Crime): *Covid + Online + Retail*
W (IV on Enforcement): *AG + DSG + Listed + NIS1 + Not_Virtual*

This table shows for an additional Model V the confidence intervals of the estimated parameters in Equations 1 and 2 for St Gallen and Zug. The confidence intervals show the same direction and almost identical magnitude across model and canton. With the exception of α for Zug (change of sign), all parameters are significant.

B Additional Results

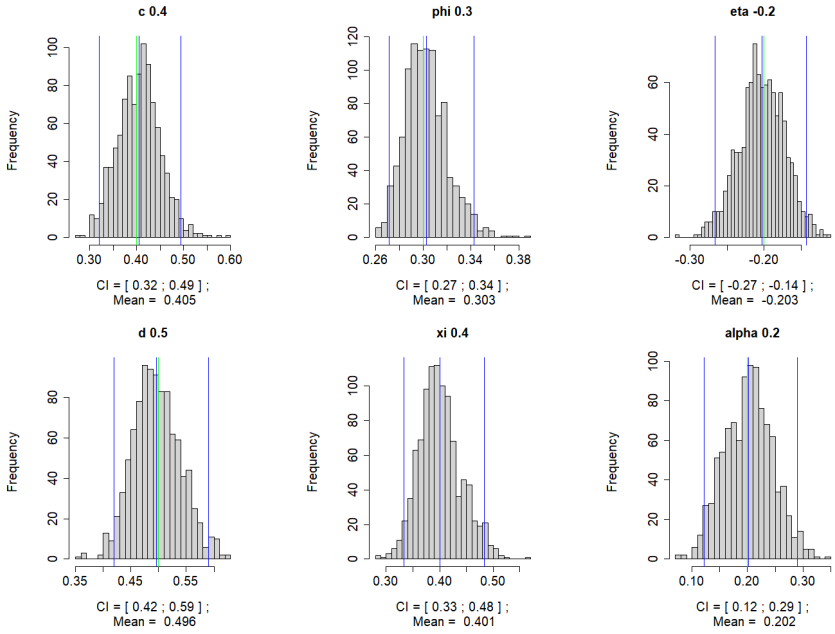


Figure 4: Bootstrapping estimated Parameters for 1000 Simulations
Note: This figure shows the distribution of all estimated parameters in 1000 simulations with 100'000 observations each. We show the confidence intervals below in each subplot, with the lower and upper bounds (thin blue vertical lines). The estimated mean is in the middle of the distribution (blue vertical line in the middle). The simulated (true) value is the green line.

Table 13: Parameter Estimations Cantons St Gallen and Zug

	St Gallen		Zug	
	Model I	Model II	Model I	Model II
Constant on Crime ($\hat{c}$)	-1.2123^{***}	-1.5848^{***}	-1.4772^{**}	-2.1919^{***}
	(0.1327)	(0.1123)	(0.4571)	(0.2960)
IV on Crime ($\hat{\varphi}$)	0.6765^{***}	0.6856^{***}	0.5327^{***}	0.4879^{***}
	(0.0705)	(0.0572)	(0.1543)	(0.0758)
R_{prev} on Crime ($\hat{\eta}$)	1.8558^{***}	2.1003^{***}	1.3675	1.4725^{**}
	(0.4172)	(0.4043)	(0.7405)	(0.5272)
Constant on Enforcement ($\hat{d}$)	-3.9609^{***}	-3.8014^{***}	-3.8162^{***}	-3.4472^{***}
	(0.0627)	(0.0674)	(0.2205)	(0.2106)
IV on Enforcement ($\hat{\xi}$)	0.5708^{***}	0.5668^{***}	0.3839^{***}	0.4299^{***}
	(0.0189)	(0.0187)	(0.0339)	(0.0387)
R_{prev} on Enforcement ($\hat{\alpha}$)	0.8975^{***}	0.7435^{***}	0.9133^{**}	0.6003^{*}
	(0.0793)	(0.0805)	(0.2859)	(0.2939)
$\hat{E}$ (Enforcement Estimated on $\hat{\theta}$)	0.0011	0.0016	0.0008	0.0033
	(0.0034)	(0.0039)	(0.0013)	(0.0040)
$\hat{C}$ (Cybercrime Estimated on $\hat{\theta}$)	0.1785	0.1076	0.0979	0.0235
	(0.1108)	(0.0917)	(0.0549)	(0.0218)
Observations	3′690′564	3′690′564	3′968′712	3′968′712

Note: Significance at the 1%, 5%, and 10% level is indicated by ***, **, and *, respectively.
Model I
Z (IV on Crime): *Covid + Online + Retail*
W (IV on Enforcement): *AG + DSG + Listed + NIS1 + Not_Virtual + Not_Canton*
Model II
Z (IV on Crime): *Covid + Mid_Sized + Large_Sized + Online + Retail*
W (IV on Enforcement): *AG + DSG + Listed + NIS1 + Not_Virtual + Not_Canton*
(Variable Values and Range see Table 31)

Table 14: Evaluation of Results (Model I)

Canton	κ	LM_C	p_C	LM_E	p_E
SG	7484.806	105.8112	0	93.33228	0
ZG	13505.54	31.40451	0	28.92275	1e-07

Z: *Covid + Online + Retail*
W: *AG + DSG + Listed + NIS1 + Not_Virtual Not_Canton*

This table shows for the estimates in Table 13 the condition number and the Lagrange Multiplier statistics with p-values for the unrestricted Equations 6 and 7.

Table 15: Evaluation Values (Model II)

Canton	κ	LM_C	p_C	LM_E	p_E
SG	6650.83	116.9504	0	84.80163	0
ZG	6080.90	145.2108	0	1e-7	0.9998

Z: *Covid + Mid_Sized + Large_Sized + Online + Retail*
W: *AG + DSG + Listed + NIS1 + Not_Virtual Not_Canton*

This table shows for the estimates in Table 13 the condition number and the Lagrange Multiplier statistics with p-values for the unrestricted Equations 6 and 7.

C Multiple Starting Points

Table 16: Different Initial Parameters Simulation

c	φ	η	d	ξ	α	Value	κ	$\bar{se}$	c_0	φ_0	η_0	d_0	ξ_0	α_0
0.422	0.305	-0.209	0.476	0.388	0.203	661165.7	6.95e+02	0.012	-1.121	-0.460	3.117	0.141	0.259	3.430
0.422	0.305	-0.209	0.476	0.388	0.203	661165.7	6.95e+02	0.012	0.922	-2.530	-1.374	-0.891	2.448	0.720
0.422	0.305	-0.209	0.476	0.388	0.203	661165.7	6.95e+02	0.012	0.802	0.221	-1.112	3.574	0.996	-3.933
0.422	0.305	-0.209	0.476	0.388	0.203	661165.7	6.95e+02	0.012	1.403	-0.946	-2.136	-0.436	-2.052	-1.458
0.422	0.305	-0.209	0.476	0.388	0.203	661165.7	6.95e+02	0.012	-1.250	-3.373	1.676	0.307	-2.276	2.508
0.422	0.305	-0.209	0.476	0.388	0.203	661165.7	6.95e+02	0.012	0.853	-0.590	1.790	1.756	1.643	1.377
0.422	0.305	-0.209	0.476	0.388	0.203	661165.7	6.98e+02	0.012	1.108	-0.124	-0.612	-0.761	-1.389	-0.416
0.422	0.305	-0.209	0.476	0.388	0.203	661165.7	6.95e+02	0.012	-2.531	4.338	2.416	-2.246	-0.806	-0.933
0.423	0.305	-0.210	0.475	0.387	0.203	661165.7	6.96e+02	0.012	1.560	-0.167	0.507	-0.057	-0.086	2.737
0.126	0.226	-0.068	0.905	32.269	0.143	662500.9	Inf	0.000	-0.452	3.033	-3.098	1.169	0.248	0.432
0.126	0.226	-0.068	0.905	8.173	0.143	662500.9	Inf	0.000	0.759	-1.005	-0.666	-2.037	-2.144	0.607
0.336	0.242	-0.351	0.596	0.498	8.001	663259.3	Inf	0.000	0.896	0.106	1.845	4.100	-0.982	-4.618
0.132	0.220	-0.126	0.906	9.665	6.236	664102.3	Inf	0.000	2.011	-1.418	-1.376	2.051	-0.570	-2.441
0.057	0.218	0.010	5.050	11.995	-4.001	665215.9	Inf	0.000	0.363	-0.278	0.012	0.771	-0.741	1.289
0.734	0.565	7.743	0.289	0.198	-0.264	665883.9	1.07e+18	0.000	-0.441	0.664	2.194	0.870	-0.652	2.298
0.062	0.213	-0.049	32.312	40.133	18.303	666804.3	Inf	0.000	1.987	1.097	0.477	-1.256	2.721	-1.201
0.062	0.213	-0.049	47.892	-18.403	-4.708	666804.3	3.78e+26	0.000	4.375	3.065	-0.471	-2.053	-1.421	0.514
0.062	0.213	-0.049	25.213	-5.708	-8.357	666804.3	3.33e+24	0.000	-0.493	-0.695	-1.903	-0.090	-1.570	-3.336
6.649	3.065	1.559	0.090	0.184	-0.050	668541.3	4.00e+17	0.000	-0.760	1.838	-1.151	1.216	-3.236	-0.111
46.643	-17.849	-2.765	0.090	0.184	-0.050	668541.3	Inf	0.000	1.039	0.602	0.211	-1.281	-1.699	-2.048

This table shows for the simulation 20 optimization runs with different starting points. The resulting estimates are in columns 1 to 6. The respective objective value, condition number κ and mean standard errors are in columns 7 to 9. The randomly set initial parameter values are shown in columns 10 to 15. The results are sorted by objective value and condition number. Confidence intervals of the best result are shown in Table 1.

Table 17: Different Initial Parameter St Gallen (Model I)

c	φ	η	d	ξ	α	Value	κ	$\bar{se}$	c_0	φ_0	η_0	d_0	ξ_0	α_0
-1.212	0.677	1.856	-3.961	0.571	0.898	6902.927	7.48e+03	0.13	-1.121	-0.460	3.117	0.141	0.259	3.430
-1.212	0.677	1.856	-3.961	0.571	0.898	6902.927	7.48e+03	0.13	0.922	-2.530	-1.374	-0.891	2.448	0.720
-1.212	0.677	1.856	-3.961	0.571	0.898	6902.927	7.48e+03	0.13	0.802	0.221	-1.112	3.574	0.996	-3.933
-1.212	0.677	1.856	-3.961	0.571	0.898	6902.927	7.49e+03	0.13	1.403	-0.946	-2.136	-0.436	-2.052	-1.458
-1.170	0.709	4.356	-3.973	0.564	0.848	6904.291	7.35e+05		-1.250	-3.373	1.676	0.307	-2.276	2.508
-1.170	0.709	11.961	-3.973	0.564	0.848	6904.295	3.17e+18	0.00	0.853	-0.590	1.790	1.756	1.643	1.377
-1.170	0.709	7.117	-3.973	0.564	0.848	6904.295	9.54e+10		1.108	-0.124	-0.612	-0.761	-1.389	-0.416
-1.170	0.709	7.092	-3.973	0.564	0.848	6904.295	1.91e+11		-2.531	4.338	2.416	-2.246	-0.806	-0.933
-1.170	0.709	159.033	-3.973	0.564	0.848	6904.295	2.04e+18	0.00	1.560	-0.167	0.507	-0.057	-0.086	2.737
-0.276	17.433	2.072	-4.251	0.536	1.206	6936.584	1.74e+18	0.00	-0.452	3.033	-3.098	1.169	0.248	0.432
-0.276	6.913	3.428	-4.250	0.535	1.199	6936.642	1.09e+10		0.759	-1.005	-0.666	-2.037	-2.144	0.607
-0.276	4.813	107.814	-4.250	0.535	1.199	6936.644	1.29e+18	0.00	0.896	0.106	1.845	4.100	-0.982	-4.618
-0.276	32.754	14.960	-4.250	0.535	1.199	6936.644	6.04e+18	0.00	2.011	-1.418	-1.376	2.051	-0.570	-2.441
-2.013	0.280	-0.010	-3.698	0.771	9.100	6949.593	3.21e+09		0.363	-0.278	0.012	0.771	-0.741	1.289
-2.013	0.280	-0.010	-3.698	0.771	9.192	6949.593	5.05e+09		-0.441	0.664	2.194	0.870	-0.652	2.298
-2.013	0.280	-0.010	-3.698	0.771	20.178	6949.593	Inf	0.00	1.987	1.097	0.477	-1.256	2.721	-1.201
-2.013	0.280	-0.010	-3.698	0.771	54.504	6949.593	Inf	0.00	4.375	3.065	-0.471	-2.053	-1.421	0.514
-2.013	0.280	-0.010	-3.698	0.771	12.082	6949.593	Inf	0.00	-0.493	-0.695	-1.903	-0.090	-1.570	-3.336
-3.684	0.233	1.707	11.025	29.198	102.537	7544.701	Inf	0.00	-0.760	1.838	-1.151	1.216	-3.236	-0.111
-3.685	0.233	1.707	114.193	43.808	36.576	7545.091	Inf	0.00	1.039	0.602	0.211	-1.281	-1.699	-2.048

Z: *Covid + Online + Retail*

W: *AG + DSG + Listed + NIS1 + Not_Virtual Not_Canton*

This table shows for St Gallen (Model I) 20 optimization runs with different starting points. The resulting estimates are in columns 1 to 6. The respective objective value, condition number κ and mean standard errors are in columns 7 to 9. The randomly set initial parameter values are shown in columns 10 to 15. The results are sorted by objective value and condition number. Confidence intervals of the best result are shown in Table 8.

Table 18: Different Initial Parameter Zug (Model I)

c	φ	η	d	ξ	α	Value	κ	$\bar{se}$	c_0	φ_0	η_0	d_0	ξ_0	α_0
-1.477	0.533	1.367	-3.816	0.384	0.913	3145.263	1.35e+04	0.315	-1.121	-0.460	3.117	0.141	0.259	3.430
-1.477	0.533	1.367	-3.816	0.384	0.913	3145.263	1.35e+04	0.315	0.922	-2.530	-1.374	-0.891	2.448	0.720
-1.477	0.533	1.367	-3.816	0.384	0.913	3145.263	1.35e+04	0.315	0.802	0.221	-1.112	3.574	0.996	-3.933
-1.477	0.533	1.367	-3.816	0.384	0.913	3145.263	1.35e+04	0.315	1.403	-0.946	-2.136	-0.436	-2.052	-1.458
-1.514	0.521	1.387	-3.797	0.385	0.898	3145.268	1.28e+04	0.303	-1.250	-3.373	1.676	0.307	-2.276	2.508
-0.261	17.208	4.828	-4.223	0.344	1.197	3152.185	2.12e+18	0.000	0.853	-0.590	1.790	1.756	1.643	1.377
-2.093	0.344	-0.360	-3.452	0.426	40.910	3159.725	Inf	0.000	1.108	-0.124	-0.612	-0.761	-1.389	-0.416
-2.093	0.344	-0.361	-3.452	0.426	38.431	3159.725	Inf	0.000	-2.531	4.338	2.416	-2.246	-0.806	-0.933
-2.093	0.344	-0.361	-3.452	0.426	87.312	3159.725	Inf	0.000	1.560	-0.167	0.507	-0.057	-0.086	2.737
-2.093	0.344	-0.361	-3.452	0.426	24.588	3159.725	Inf	0.000	-0.452	3.033	-3.098	1.169	0.248	0.432
-2.093	0.344	-0.361	-3.452	0.426	79.309	3159.725	Inf	0.000	0.759	-1.005	-0.666	-2.037	-2.144	0.607
-2.093	0.344	-0.361	-3.452	0.426	57.065	3159.725	Inf	0.000	0.896	0.106	1.845	4.100	-0.982	-4.618
-2.093	0.344	-0.361	-3.452	0.426	41.746	3159.725	Inf	0.000	2.011	-1.418	-1.376	2.051	-0.570	-2.441
8.007	5.749	20.020	-4.305	0.301	1.386	3178.338	Inf	0.000	0.363	-0.278	0.012	0.771	-0.741	1.289
-3.813	0.166	1.509	-1.022	29.107	-8.313	3271.378	Inf	0.000	-0.441	0.664	2.194	0.870	-0.652	2.298
-3.813	0.167	1.494	-1.022	10.304	6.636	3273.136	Inf	0.000	1.987	1.097	0.477	-1.256	2.721	-1.201
-3.813	0.167	1.494	-1.022	144.463	12.040	3273.136	Inf	0.000	4.375	3.065	-0.471	-2.053	-1.421	0.514
-3.851	0.174	1.604	20.317	24.583	-45.485	3294.026	3.40e+10		-0.493	-0.695	-1.903	-0.090	-1.570	-3.336
-3.853	0.177	1.541	14.008	18.778	-18.672	3297.374	8.55e+10	1049	-0.760	1.838	-1.151	1.216	-3.236	-0.111
-3.858	0.181	1.529	10.164	3.652	-1.650	3303.203	Inf	0.000	1.039	0.602	0.211	-1.281	-1.699	-2.048

Z: *Covid + Mid_Sized + Large_Sized + Online + Retail*
W: *AG + DSG + Listed + NIS1 + Not_Virtual Not_Canton*

This table shows for Zug (Model I) 20 optimization runs with different starting points. The resulting estimates are in columns 1 to 6. The respective objective value, condition number κ and mean standard errors are in columns 7 to 9. The randomly set initial parameter values are shown in columns 10 to 15. The results are sorted by objective value and condition number. Confidence intervals of the best result are shown in Table 8.

Table 19: Different Initial Parameter St Gallen (Model II)

c	φ	η	d	ξ	α	Value	κ	$\bar{se}$	c_0	φ_0	η_0	d_0	ξ_0	α_0
-1.585	0.686	2.100	-3.801	0.567	0.743	6829.762	6.65e+03	0.123	-1.121	-0.460	3.117	0.141	0.259	3.430
-1.585	0.686	2.098	-3.801	0.567	0.744	6829.762	6.62e+03	0.123	0.922	-2.530	-1.374	-0.891	2.448	0.720
-1.553	0.713	35.371	-3.818	0.562	0.697	6831.324	1.74e+18	0.000	0.802	0.221	-1.112	3.574	0.996	-3.933
-1.553	0.713	17.927	-3.818	0.562	0.697	6831.324	7.87e+17	0.000	1.403	-0.946	-2.136	-0.436	-2.052	-1.458
-1.553	0.713	11.274	-3.818	0.562	0.697	6831.324	4.22e+19	0.000	-1.250	-3.373	1.676	0.307	-2.276	2.508
-1.553	0.713	146.155	-3.818	0.562	0.697	6831.324	8.33e+17	0.000	0.853	-0.590	1.790	1.756	1.643	1.377
-2.289	0.322	0.276	-3.520	0.794	2.439	6851.527	2.12e+03	0.107	1.108	-0.124	-0.612	-0.761	-1.389	-0.416
-2.309	0.321	0.147	-3.526	0.809	8.671	6872.163	1.07e+09		-2.531	4.338	2.416	-2.246	-0.806	-0.933
-2.309	0.321	0.147	-3.526	0.809	16.010	6872.163	Inf	0.000	1.560	-0.167	0.507	-0.057	-0.086	2.737
-2.309	0.321	0.147	-3.526	0.809	50.658	6872.163	Inf	0.000	-0.452	3.033	-3.098	1.169	0.248	0.432
-2.309	0.321	0.147	-3.526	0.809	202.076	6872.163	Inf	0.000	0.759	-1.005	-0.666	-2.037	-2.144	0.607
-2.309	0.321	0.147	-3.526	0.809	20.871	6872.163	Inf	0.000	0.896	0.106	1.845	4.100	-0.982	-4.618
-2.309	0.321	0.147	-3.526	0.809	10.818	6872.163	Inf	0.000	2.011	-1.418	-1.376	2.051	-0.570	-2.441
-2.309	0.321	0.147	-3.526	0.809	12.413	6872.163	Inf	0.000	0.363	-0.278	0.012	0.771	-0.741	1.289
-2.309	0.321	0.147	-3.526	0.809	52.685	6872.163	Inf	0.000	-0.441	0.664	2.194	0.870	-0.652	2.298
-0.506	6.322	2.211	-4.227	0.533	1.187	6896.476	1.79e+09		1.987	1.097	0.477	-1.256	2.721	-1.201
-0.506	12.427	15.543	-4.227	0.533	1.181	6896.549	1.48e+19	0.000	4.375	3.065	-0.471	-2.053	-1.421	0.514
5.448	33.236	0.224	-4.342	0.506	1.367	7006.910	2.02e+19	0.000	-0.493	-0.695	-1.903	-0.090	-1.570	-3.336
-3.738	0.295	1.611	16.370	30.059	36.271	7416.776	Inf	0.000	-0.760	1.838	-1.151	1.216	-3.236	-0.111
-3.738	0.295	1.612	401.311	52.476	56.489	7417.099	Inf	0.000	1.039	0.602	0.211	-1.281	-1.699	-2.048

Z: *Covid + Mid_Sized + Large_Sized + Online + Retail*

W: *AG + DSG + Listed + NIS1 + Not_Virtual Not_Canton*

This table shows for St Gallen (Model II) 20 optimization runs with different starting points. The resulting estimates are in columns 1 to 6. The respective objective value, condition number κ and mean standard errors are in columns 7 to 9. The randomly set initial parameter values are shown in columns 10 to 15. The results are sorted by objective value and condition number. Confidence intervals of the best result are shown in Table 9.

Table 20: Different Initial Parameter Zug (Model II)

c	φ	η	d	ξ	α	Value	κ	$\bar{se}$	c_0	φ_0	η_0	d_0	ξ_0	α_0
-2.192	0.488	1.473	-3.447	0.430	0.600	3105.311	6.08e+03	0.240	-1.121	-0.460	3.117	0.141	0.259	3.430
-2.192	0.488	1.472	-3.447	0.430	0.600	3105.311	6.08e+03	0.240	0.922	-2.530	-1.374	-0.891	2.448	0.720
-2.192	0.488	1.472	-3.447	0.430	0.600	3105.311	6.08e+03	0.240	0.802	0.221	-1.112	3.574	0.996	-3.933
-2.246	0.477	1.498	-3.409	0.435	0.565	3105.328	6.06e+03	0.245	1.403	-0.946	-2.136	-0.436	-2.052	-1.458
-2.618	0.370	0.015	-3.108	0.490	57.206	3118.793	Inf	0.000	-1.250	-3.373	1.676	0.307	-2.276	2.508
-2.618	0.370	0.015	-3.108	0.490	53.480	3118.793	Inf	0.000	0.853	-0.590	1.790	1.756	1.643	1.377
-2.618	0.370	0.015	-3.108	0.490	12.804	3118.793	Inf	0.000	1.108	-0.124	-0.612	-0.761	-1.389	-0.416
-2.618	0.370	0.015	-3.108	0.490	28.350	3118.793	Inf	0.000	-2.531	4.338	2.416	-2.246	-0.806	-0.933
-2.618	0.370	0.015	-3.108	0.490	29.276	3118.793	Inf	0.000	1.560	-0.167	0.507	-0.057	-0.086	2.737
-2.618	0.370	0.015	-3.108	0.490	79.254	3118.793	Inf	0.000	-0.452	3.033	-3.098	1.169	0.248	0.432
-2.618	0.370	0.015	-3.108	0.490	38.298	3118.793	Inf	0.000	0.759	-1.005	-0.666	-2.037	-2.144	0.607
-0.474	14.185	0.861	-4.208	0.350	1.232	3133.934	4.61e+18	0.000	0.896	0.106	1.845	4.100	-0.982	-4.618
-0.476	10.459	21.097	-4.210	0.351	1.165	3134.916	1.38e+18	0.000	2.011	-1.418	-1.376	2.051	-0.570	-2.441
-3.864	0.251	1.419	-0.973	6.830	-2.102	3227.574	8.82e+05		0.363	-0.278	0.012	0.771	-0.741	1.289
-3.865	0.252	1.406	-0.972	10.409	5.973	3228.900	Inf	0.000	-0.441	0.664	2.194	0.870	-0.652	2.298
-3.865	0.252	1.406	-0.972	12.632	49.145	3228.900	Inf	0.000	1.987	1.097	0.477	-1.256	2.721	-1.201
-3.902	0.260	1.447	3.576	166.363	-20.942	3250.866	Inf	0.000	4.375	3.065	-0.471	-2.053	-1.421	0.514
-3.902	0.260	1.447	6.537	107.889	-35.693	3250.872	Inf	0.000	-0.493	-0.695	-1.903	-0.090	-1.570	-3.336
-3.903	0.261	1.434	19.245	30.700	-3.741	3252.170	Inf	0.000	-0.760	1.838	-1.151	1.216	-3.236	-0.111
-3.907	0.263	1.436	33.857	18.235	11.453	3255.747	Inf	0.000	1.039	0.602	0.211	-1.281	-1.699	-2.048

Z: *Covid + Mid_Sized + Large_Sized + Online + Retail*

W: *AG + DSG + Listed + NIS1 + Not_Virtual Not_Canton*

This table shows for Zug (Model II) 20 optimization runs with different starting points. The resulting estimates are in columns 1 to 6. The respective objective value, condition number κ and mean standard errors are in columns 7 to 9. The randomly set initial parameter values are shown in columns 10 to 15. The results are sorted by objective value and condition number. Confidence intervals of the best result are shown in Table 9.

D Correlation Tables

Table 21: Correlation Matrix Simulation

	R_{prev}	R	Z	W
R_{prev} (Recorded earlier Period)	1			
R (Recorded)	-0.0193	1		
Z (IV affecting Crime)	-0.0013	0.1179	1	
W (IV affecting Enforcement)	0.0025	0.1021	0.0008	1

This table shows the correlation matrix of the simulated values for 1 million observations (Section 4). It indicates that there is no collinearity between Z and W, which requires strict exogeneity in theory. By construction, R correlates with W and Z in the simulation. Note that in our simulation, no observed variables experience a high level of correlation.

Table 22: Correlation Cybercrime Canton St Gallen (Model I)

	R_{prev}	R	Z	W
R_{prev} (Recorded earlier Period)	1			
R (Recorded)	0.1113	1		
Z (IV affecting Crime)	0.0378	0.0139	1	
W (IV affecting Enforcement)	0.0703	0.0269	0.1385	1

This table shows the correlation between the observed variables for cybercrime used for estimation in canton St Gallen (Model I): policy shock for cybercrime (W), external shock for reporting (Z), reporting in earlier time period (R_{prev}) and reporting (R).

Table 23: Correlation Cybercrime Canton St Gallen (Model II)

	R_{prev}	R	Z	W
R_{prev} (Recorded earlier Period)	1			
R (Recorded)	0.1113	1		
Z (IV affecting Crime)	0.0558	0.0208	1	
W (IV affecting Enforcement)	0.0703	0.0269	0.1569	1

This table shows the correlation between the observed variables for cybercrime used for estimation in canton St Gallen (Model II): policy shock for cybercrime (W), external shock for reporting (Z), reporting in earlier time period (R_{prev}) and reporting (R).

Table 24: Correlation Cybercrime Canton Zug (Model I)

	R_{prev}	R	Z	W
R_{prev} (Recorded earlier Period)	1			
R (Recorded)	0.0427	1		
Z (IV affecting Crime)	0.0158	0.0050	1	
W (IV affecting Enforcement)	0.0261	0.0097	0.0259	1

This table shows the correlation between the observed variables for cybercrime used for estimation in canton Zug (Model I): policy shock for cybercrime (W), external shock for reporting (Z), reporting in earlier time period (R_{prev}) and reporting (R).

Table 25: Correlation Cybercrime Canton Zug (Model II)

	R_{prev}	R	Z	W
R_{prev} (Recorded earlier Period)	1			
R (Recorded)	0.0427	1		
Z (IV affecting Crime)	0.0270	0.0090	1	
W (IV affecting Enforcement)	0.0261	0.0097	0.041	1

This table shows the correlation between the observed variables for cybercrime used for estimation in canton Zug (Model II): policy shock for cybercrime (W), external shock for reporting (Z), reporting in earlier time period (R_{prev}) and reporting (R).

Table 26: Correlation Matrix Cybercrime Canton SG

	R_{prev}	R	*N_Emp*	*Covid*	*Online*	*Retail*	*AG*	*Listed*	*NIS1*	*DSG*	*Not_V*	*Not_C*
R_{prev}	1	0.1113	0.0387	-0.0005	0.0577	0.0461	0.0375	0.0298	0.0054	0.0002	0.0040	0.3443
R	0.1113	1	0.0294	0.0011	0.0235	0.0178	0.0136	0.0100	0.0025	0.0007	0.0014	0.1328
N_Emp	0.0387	0.0294	1	-0.0013	-0.0012	-0.0024	0.0058	0.1400	-0.0010	-0.0004	0.0017	0.0795
Covid	-0.0005	0.0011	-0.0013	1	0.0059	0.0025	-0.0161	-0.0001	0.6230	0.1909	0.0017	-0.0027
Online	0.0577	0.0235	-0.0012	0.0059	1	0.3127	-0.0357	-0.0017	0.0045	0.0028	0.0021	0.1155
Retail	0.0461	0.0178	-0.0024	0.0025	0.3127	1	-0.0881	-0.0054	0.0016	0.0020	0.0171	0.1024
AG	0.0375	0.0136	0.0058	-0.0161	-0.0357	-0.0881	1	0.0220	-0.0111	-0.0057	-0.0390	0.0654
Listed	0.0298	0.0100	0.1400	-0.0001	-0.0017	-0.0054	0.0220	1	0.0000	-0.0001	0.0015	0.0519
NIS1	0.0054	0.0025	-0.0010	0.6230	0.0045	0.0016	-0.0111	0.0000	1	0.1190	0.0020	-0.0016
DSG	0.0002	0.0007	-0.0004	0.1909	0.0028	0.0020	-0.0057	-0.0001	0.1190	1	0.0000	-0.0011
Not_Virtual	0.0040	0.0014	0.0017	0.0017	0.0021	0.0171	-0.0390	0.0015	0.0020	0.0000	1	0.0069
Not_Canton	0.3443	0.1328	0.0795	-0.0027	0.1155	0.1024	0.0654	0.0519	-0.0016	-0.0011	0.0069	1

This table shows the correlation between all used observed variables for cybercrime in Canton St Gallen.

Table 27: Correlation Matrix Cybercrime Canton ZG

	R_{prev}	R	*N_Emp*	*Covid*	*Online*	*Retail*	*AG*	*Listed*	*NIS1*	*DSG*	*Not_V*	*Not_C*
R_{prev}	1	0.0427	0.0182	-0.0007	0.0207	0.0249	0.0148	0.0028	0.0014	-0.0032	0.0142	0.2378
R	0.0427	1	0.0053	0.0006	0.0074	0.0089	0.0047	0.0007	0.0016	0.0006	0.0044	0.0797
N_Emp	0.0182	0.0053	1	-0.0021	-0.0010	-0.0007	0.0032	0.1789	-0.0012	-0.0013	0.0083	0.1016
Covid	-0.0007	0.0006	-0.0021	1	0.0039	0.0088	-0.0217	0.0014	0.6441	0.3806	-0.0233	-0.0022
Online	0.0207	0.0074	-0.0010	0.0039	1	0.3693	-0.0376	-0.0021	0.0031	0.0029	-0.0089	0.0632
Retail	0.0249	0.0089	-0.0007	0.0088	0.3693	1	-0.0811	-0.0006	0.0071	0.0076	0.0096	0.0793
AG	0.0148	0.0047	0.0032	-0.0217	-0.0376	-0.0811	1	0.0246	-0.0179	-0.0146	-0.0100	0.0265
Listed	0.0028	0.0007	0.1789	0.0014	-0.0021	-0.0006	0.0246	1	0.0017	0.0003	0.0038	-0.0012
NIS1	0.0014	0.0016	-0.0012	0.6441	0.0031	0.0071	-0.0179	0.0017	1	0.2451	-0.0194	-0.0015
DSG	-0.0032	0.0006	-0.0013	0.3806	0.0029	0.0076	-0.0146	0.0003	0.2451	1	-0.0168	-0.0020
Not_Virtual	0.0142	0.0044	0.0083	-0.0233	-0.0089	0.0096	-0.0100	0.0038	-0.0194	-0.0168	1	0.0211
Not_Canton	0.2378	0.0797	0.1016	-0.0022	0.0632	0.0793	0.0265	-0.0012	-0.0015	-0.0020	0.0211	1

This table shows the correlation between all used observed variables for cybercrime in Canton Zug.

E Additional Summary Statistics

Table 28: Summary Statistics Simulated Data

	Mean	Median	SD	Min	Max	Skew	Obs
Z (IV affecting Crime)	0.9986	1	0.71	0	2	0	1'000'000
W (IV affecting Enforcement)	1.0009	1	0.71	0	2	0.00	1'000'000
R_{prev} (Recorded earlier period)	0.5005	1	0.50	0	1	0.00	1'000'000
R (Recorded)	0.5976	1	0.49	0	1	-0.40	1'000'000
E (Enforcement)	0.8306	1	0.38	0	1	-1.76	1'000'000
C (Crime)	0.7205	1	0.45	0	1	-0.98	1'000'000

This table provides summary statistics for the simulated variables. Here, both instruments, Z and W, independently take the binomial distributed values for 0, 1 and 2 as set in the simulation. Around 72% of all entities simulated in the population are subject to crime (C). About 83% of the entities experience high detection effort / reporting incentive (E). This gives a rate of recorded crime $R = C \cdot E$ of nearly 60%. Around 50% of all entities have been recorded in crime in a previous time period (R_{prev}).

Table 29: Summary Statistics Sample of Cybercrime Reporting Firms Canton St Gallen, registered in St Gallen (Upper Panel) and registered in other canton (Lower Panel)

	Mean	Median	SD	Min	Max	Skew	Obs
Mean(*N_Emp*)	49.6770	4	358.58	1	5820.8	15.48	270
Mid_Sized	0.1296	0.0	0.34	0	1.0	2.19	270
Large_Sized	0.0074	0.0	0.09	0	1.0	11.43	270
Retail	0.1407	0	0.35	0	1	2.05	270
Online	0.0333	0	0.18	0	1	5.17	270
Not_Virtual	0.0000	0.0	0	0	0.0		270
AG	0.6667	1	0.47	0	1	-0.70	270
Listed	0	0	0	0	0		270
Not_Canton	0.0000	0.0	0	0	0.0		270

	Mean	Median	SD	Min	Max	Skew	Obs
Mean(*N_Emp*)	517.4782	4	6583.73	1	106370.5	15.53	269
Mid_Sized	0.2156	0	0.41	0	1.0	1.38	269
Large_Sized	0.0223	0	0.15	0	1.0	6.43	269
Retail	0.4275	0	0.50	0	1	0.29	269
Online	0.1413	0	0.35	0	1	2.05	269
Not_Virtual	0.0000	0	0	0	0.0		269
AG	0.6989	1	0.46	0	1	-0.86	269
Listed	0.0112	0	0.11	0	1	9.26	269
Not_Canton	1.0000	1	0	1	1.0		269

This table provides summary statistics for the companies reporting cybercrime in Canton St Gallen. The Upper Panel shows the reporting companies registered in Canton St Gallen, the Lower Panel shows reporting companies registered in another canton.

Table 30: Summary Statistics Sample of Cybercrime Reporting Firms Canton Zug, registered in Zug (Upper Panel) and registered in other canton (Lower Panel)

	Mean	Median	SD	Min	Max	Skew	Obs
Mean(*N_Emp*)	27.6342	4	109.10	1	1'340	9.93	185
Mid_Sized	0.1081	0	0.31	0	1	2.50	185
Large_Sized	0.0108	0	0.10	0	1	9.38	185
Retail	0.0595	0	0.24	0	1	3.70	185
Online	0.0162	0	0.13	0	1	7.60	185
AG	0.7676	1	0.42	0	1	-1.26	185
Listed	0.0054	0	0.07	0	1	13.38	185
Not_Virtual	0.0000	0	0	0	0		185
Not_Canton	0.0000	0	0	0	0		185

	Mean	Median	SD	Min	Max	Skew	Obs
Mean(*N_Emp*)	1'411.5600	15	11'174.31	1	96'793	8.29	75
Mid_Sized	0.2933	0	0.46	0	1	0.89	75
Large_Sized	0.0533	0	0.23	0	1	3.90	75
Retail	0.3867	0	0.49	0	1	0.46	75
Online	0.1067	0	0.31	0	1	2.50	75
AG	0.8400	1	0.37	0	1	-1.82	75
Listed	0	0	0	0	0		75
Not_Virtual	0.0000	0	0	0	0		75
Not_Canton	1.0000	1	0	1	1		75

This table provides summary statistics for the companies reporting cybercrime in Canton Zug. The Upper Panel shows the reporting companies registered in Canton Zug, the Lower Panel shows reporting companies registered in another canton.

Table 31: Summary Statistics Cybercrime Population Canton St Gallen and Zug

	Mean	Median	SD	Min	Max	Skew	Obs
Year	2019.6	2020	2.30	2016	2023	-0.07	3'690'564
R	0.0003	0	0.02	0	1	61.41	3'690'564
R_{prev}	0.0023	0	0.05	0	1	20.74	3'690'564
N_Emp	11.7716	1	542.24	1	106'622	172.09	3'690'564
Mid_Sized	0.0192	0.0	0.14	0	1	7.01	3'690'564
Large_Sized	0.0006	0.0	0.02	0	1	40.85	3'690'564
Covid	0.2456	0	0.43	0	1	1.18	3'690'564
Retail	0.0849	0	0.28	0	1	2.98	3'690'564
Online	0.0090	0	0.09	0	1	10.41	3'690'564
AG	0.3368	0	0.47	0	1	0.69	3'690'564
Listed	0.0003	0	0.02	0	1	55.99	3'690'564
DSG	0.0341	0	0.18	0	1	5.14	3'690'564
NIS1	0.7137	1	0.45	0	1	-0.95	3'690'564
Not_Virtual	-0.0069	0	0.08	-1	0	-11.95	3'690'564
Not_Canton	0.0068	0	0.08	0	1	12.05	3'690'564
Both_Cantons	0.0009	0	0.03	0	1	33.40	3'690'564

	Mean	Median	SD	Min	Max	Skew	Obs
Year	2020.2	2020	2.59	2016	2024	-0.10	3'968'712
R	0.0001	0	0.01	0	1	108.03	3'968'712
R_{prev}	0.0009	0	0.03	0	1	33.51	3'968'712
N_Emp	12.8076	4	678.09	1	106'622	127.73	3'968'712
Mid_Sized	0.0150	0	0.12	0	1	7.99	3968712
Large_Sized	0.0006	0	0.02	0	1	40.83	3968712
Covid	0.2158	0	0.41	0	1	1.38	3'968'712
Retail	0.0392	0	0.19	0	1	4.75	3'968'712
Online	0.0055	0	0.07	0	1	13.34	3'968'712
AG	0.5493	1	0.50	0	1	-0.20	3'968'712
Listed	0.0008	0	0.03	0	1	35.83	3'968'712
DSG	0.1574	0	0.36	0	1	1.88	3'968'712
NIS1	0.7565	1	0.43	0	1	-1.20	3'968'712
Not_Virtual	-0.1847	0	0.39	-1	0	-1.62	3'968'712
Not_Canton	0.0020	0	0.04	0	1	22.54	3'968'712
Both_Cantons	0.0009	0	0.03	0	1	32.67	3'968'712

This table provides summary statistics for the population of all firms per month incorporated in Canton St Gallen (Upper Panel) and Zug (Lower Panel). It is a total of 3'690'564 firm months aggregated from 2015 to 2023 for St Gallen and 3'968'712 from 2016 to 2024 for Zug. It describes the data we use to estimate the monthly dark figure of cyberattacks. On average, firms employ 12 workers with with the same largest firm having 106'622 employees in both cantons. However, most companies in St Gallen have only one employee, and four employees in Zug. This makes the distribution of the number of employees highly skewed with a standard deviation of 542 and 678, respectively. Out of all companies incorporated in St Gallen, only 0.03% report per month. The monthly reporting rate in Zug is only half as high, at 0.01%. Reporting (R) is a binary variable being 0 most often. Relatively to reporting in a previous period (R_{prev}), a large share reports at least one more time in a later stage.