

Axel Ebers and Stephan L. Thomsen*

Benefit–Cost Analysis of Social Media Facilitated Bystander Programs¹

Abstract: Bystander programs contribute to crime prevention by motivating people to intervene in violent situations. Social media allow addressing very specific target groups, and provide valuable information for program evaluation. This paper provides a conceptual framework for conducting benefit–cost analysis of bystander programs and puts a particular focus on the use of social media for program dissemination and data collection. The benefit–cost model treats publicly funded programs as investment projects and calculates the benefit–cost ratio. Program benefit arises from the damages avoided by preventing violent crime. We provide systematic instructions for estimating this benefit. The explained estimation techniques draw on social media data, machine-learning technology, randomized controlled trials and discrete choice experiments. In addition, we introduce a complementary approach with benefits calculated from the public attention generated by the program. To estimate the value of public attention, the approach uses the bid landscaping method, which originates from display advertising. The presented approaches offer the tools to implement a benefit–costs analysis in practice. The growing importance of social media for the dissemination of policy programs requires new evaluation methods. By providing two such methods, this paper contributes to evidence-based decision-making in a growing policy area.

Keywords: conceptual framework; benefit-cost analysis; social media; bystander programs; machine learning; discrete choice experiments

JEL classifications: C93; D61; H43; K42

Axel Ebers: Institute of Economic Policy, Leibniz University Hannover, Königsworther Platz 1, 30165 Hannover, Germany, e-mail: ebers@wipol.uni-hannover.de

***Corresponding author: Stephan L. Thomsen:** Institute of Economic Policy, Leibniz University Hannover, Königsworther Platz 1, 30165 Hannover, Germany, e-mail: thomsen@wipol.uni-hannover.de

1 We like to thank the editor, Amanda Ross, and two anonymous referees for valuable comments and suggestions. This research was conducted as part of the joint research project *Security Communication via Online Social Networks – An Innovative Approach to Crime Prevention*, which was funded by the German Federal Ministry of Education and Research.

1. Introduction

Every economist would clearly agree that violence causes immense social damage, but agreement on how society can avoid the damage is less clear. Incarceration appears to be less efficient than prevention (Welsh et al., 2015). However, what type of prevention promises the largest benefit to society? Evidence casts doubt on the prevention program addressing the potential offender or victim (DeGue et al., 2014). The program addressing the potential bystander appears more promising. The bystander program aims to contribute to crime prevention by motivating people to intervene when they observe warning signs or incidents of violence. During the program, people train the skills necessary for safe and effective intervention. To the extent that training is successful, the program reduces violent crime. Related evidence supports this hypothesis. For example, universities with a bystander program in place experienced a significant reduction in the crime rate (Coker et al., 2015). These universities found a way to avoid the social damages from violence.

Despite its success in preventing crime, the bystander program can be a challenge for the budget and the potential participant. The traditional bystander program consists of face-to-face training in small groups. Training a large number of people requires a large number of staff. High staff requirements make the program costly. Attending the program can also be a challenge as training takes place at a scheduled date and place. Someone who has another date or lives far from the training place cannot attend.

These challenges sparked the development of bystander programs that provide the training via the internet. The online bystander program requires less staff to train a large number of people. Low staff requirements make the program cost efficient (Cugelman et al., 2011). People can attend the training on their computer at a flexible date and place. Higher flexibility will increase the number of attendants (White et al., 2010). Early studies found that online bystander programs have the same positive effects as traditional programs. The most discussed examples of online bystander programs are *TakeCare* (Kleinsasser et al., 2015) and *RealConsent* (Salazar et al., 2014). Both programs aimed to address the challenges described above and were successful in promoting bystander behavior.

The online bystander program in turn may not find enough suitable participants. Social media may help with that problem. Firms increasingly rely on social media to reach customers with their ads. *Facebook* dominates the digital advertising market (eMarketer, 2019). On *Facebook*, millions of people post, like, and share information about their personal life. *Facebook* uses this information to understand their most private traits including personality, political views, or sexual orientation (Kosinski et al., 2013). The traits of a person have a strong influence on her future behavior. Just as firms pay *Facebook* to find the people who are most likely to buy their products, government agencies use the platform to find the people who are most likely to

benefit from their program interventions. For the case at hand, these are the people who are the most likely to become a bystander of a violent situation.

Besides targeting and large-scale dissemination of content, social media allow conducting social experiments. For example, Bond et al. (2012) conducted a randomized experiment to show that *Facebook* ads affect voter turnout. Similar experiments would allow to show the effect of a bystander program. The estimated effect can form the basis for a benefit–cost analysis. Benefit–cost analysis is the most important instrument for assessing the efficiency of prevention measures. However, despite the possibilities of program dissemination and data collection, no study to date has analyzed the potential of social media for the implementation and benefit–cost analysis of bystander programs.

We aim to fill this gap by providing a conceptual framework for conducting benefit–costs analysis of bystander programs, and put a special emphasis on the potential of social media in this context. The conceptual framework consists of an extended benefit–cost model, and a set of research approaches to estimating the program benefit. Within this framework, benefit arises from the damages avoided by preventing violent crime. Because the practical implementation demands a lot of time and resources, the concept is better suited for large-scale bystander programs. For small-scale programs, we suggest a complementary approach, where benefit arises from the value of public attention generated by the program.

Our conceptual framework requires that the bystander program and the benefit–cost analysis be implemented simultaneously. Our work is therefore of interest to both the practitioner who is commissioned to implement a bystander program, and the researcher who has to conduct the benefit–cost analysis of that program. Furthermore, benefit–cost analysis should ultimately serve the policy maker who has to decide on program funding. Our work will help her to understand and classify information coming from a benefit–cost analysis.

The rest of this paper is organized as follows: Section 2 lays out the theoretical foundations of our conceptual framework including the mechanics of bystander programs, and the potential outcomes of a violent situation. In Section 3, we develop our extended benefit–cost model, which employs the benefit–cost ratio as the model framework. The value of benefits depends on the number of crimes prevented, and the average damage coming from one of these crimes. Section 4 provides systematic instructions for each step to estimating the number of crimes prevented. First, it explains how to use social media and web analytics tools to estimate the number of program participants. Second, it introduces a machine-learning approach for estimating the share of future bystanders among participants. Third, it explains how to conduct a randomized experiment on *Facebook* to estimate the program effect on bystander behavior. Fourth, it provides an estimator for the number of bystanders that are injured during their intervention. Section 5 describes how to use *discrete choice*

experiments to estimate the average damage from violent crime. Section 6 introduces our complementary approach that uses the bid landscaping method to estimate the value of the public attention generated. The last section provides some conclusions.

2. Theoretical considerations

2.1 Bystander programs and the prevention of violence

Imagine you observe a man about to slap a woman in the face. Bystander behavior includes different reactions to observing physical violence. Positive reactions include the four Ds: direct, distract, delegate, and delay. Direct tactics involve direct intervention aiming to prevent or stop the violence. You may try to come between the two or even hold the man down. Direct tactics are quite risky. Distraction tactics distract the attention of the offender to rescue the victim. You can try to engage the man in a conversation and bring the woman away. Delegation tactics involve other people and a plan for cooperation. You may talk to the man, while another person brings the woman away. Delay tactics apply after the violent situation. You may give first aid or consolation. You may also help the woman to find a good hospital or counseling center. (Banyard et al., 2005). The appropriate tactic depends on the situation at hand. For example, you should save direct tactics for dangerous emergencies.²

Bystander behavior cannot be taken for granted though. Before you intervene, you have to overcome a series of mental obstacles (Latané & Darley, 1970). You have to notice the event in the first place. Assuming you do, you have to understand that the event marks a case of violence, feel responsible for helping, and know you have the skills for intervention. Finally, you have to think that helping is better than not. Only if you overcome all the mental obstacles, you will take action. Different factors decide if you manage to overcome a particular obstacle. Interdependencies and feedback loops connect the different obstacles with each other. Figure 1 illustrates the mental process for better memorization.

When you make a decision, you will not weigh the pros and cons of helping in a purely rational way. The burden to decide comes suddenly and unexpected. Options for influencing the situation are limited. Fear and anxiety hold you back, while anger and indignation push you forward (Halmburger et al., 2017). You do not know all the potential outcomes of helping. Even less do you have measurable criteria to calculate expected utility (Gigerenzer & Selten, 2002; Simon, 1959). According to the concept

² An alternative framework differentiates between *reactive* and *proactive* bystander opportunities as well *primary* (before the assault), *secondary* (during the assault), and *tertiary* (after the assault) prevention in the context of sexual violence (McMahon & Banyard, 2012).

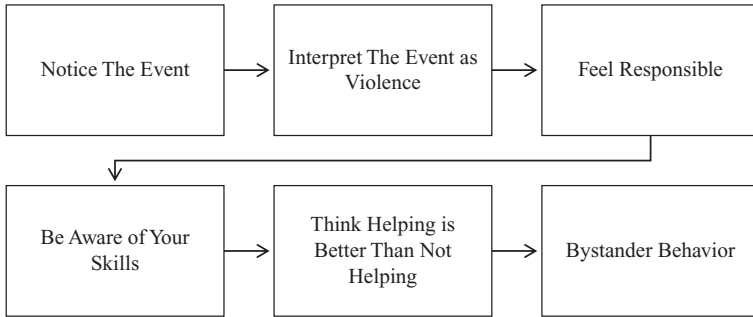


Figure 1 The mental obstacles to bystander behavior.

Notes: The model illustrates the mental obstacles to bystander behavior.

Source: Own representation based on Latané and Darley (1970).

of *bounded rationality*, each human decision results from this interplay of several factors: the time available, the tractability of the situation, and the cognitive resources of the decision-maker.

Given the limited time and resources, you take mental shortcuts to find a practical solution. The final solution may not be optimal, but satisfactory (Simon, 1955).³ The mental shortcuts rely on cues such as emotions, well-known examples, or the behavior of others (Tversky & Kahneman, 1974). If other bystanders are present during the attack, you may think that they should help. If they do not, you may think that they disapprove helping. Remarkably, the other bystanders might think the same. Both thoughts may stop all of you from intervening (Darley & Latané, 1968; Latané & Darley, 1969). In very dangerous emergencies, however, the described *bystander effect* is less pronounced (Fischer et al., 2011).⁴

The typical bystander program tackles the mental obstacles to bystander behavior (Kettrey & Marx, 2020). The trainer will explain that you might fail to notice the event if you are in a hurry or look at your smartphone all the time. She will also describe the warning signs of violence. Knowing the warning signs allows you to assess any situation correctly. Since the trainer must motivate you to act despite personal risk, she will emphasize personal responsibility and the potentially severe health consequences of physical violence. You will get to reflect your own knowledge and attitudes. In role-plays, you will practice the skills for safe and effective intervention. The practice will help you to internalize the skills and strengthen your confidence. The acquired knowledge and hands-on practice will help you overcoming the obstacles to intervention (Jonas et al., 2007).

³ The described concept of satisficing also applies to moral behavior (Gigerenzer, 2010).

⁴ The described factors leading to the *bystander effect* are called *diffusion of responsibility* and *pluralistic ignorance*.

The typical online bystander program will provide the same theoretical knowledge and practice, but rely on different communication channels. Instead of a human trainer, you may receive the information via image, text, or video content. A chat bot or forum allow interaction by posting and discussing questions. Instead of role-plays, you may practice your intervention skills in videogames. Despite the differences in communication channels, the online program will still aim to tackle the mental obstacles to bystander behavior. If you apply the practiced skills to a violent situation, it will likely have a good outcome.

2.2 The potential outcomes of a violent situation

Building our benefit–cost model requires analyzing the potential outcomes of a violent situation. For the analysis, we differentiate between four scenarios. In the first scenario, the bystander fails the mental obstacles described in [Figure 1](#), and does not intervene. Remember that violence causes immense social damage. The victim suffers the pain, cannot go to work, and has to pay the physician. The state has to pay the investigator, prosecutor, and prison guard. The offender may suffer the consequences of incarceration. But since the offender lacks *standing*, we do not count his costs in our benefit–cost analysis (Zerbe & Bellas, 2006). The working hours missed by the victim reduce the total economic production. Public spending on the justice system strains the public budget. The reduced production and additional public spending represent immense social damage (Becker, 1968).⁵

In the second scenario, the bystander overcomes the obstacles and attempts to intervene. During the attempt, the offender injures the bystander instead of the initial victim. Now the bystander suffers the described consequences of violence. The public still has to carry the spending on the justice system. The overall social damage remains the same.

In the third scenario, the bystander is able to help the victim without suffering an injury herself. Successful bystander behavior prevents the social damage of violence. The victim can carry on working and spending her money on food and clothes. The state can build more schools, roads, or fiber optic cables. The benefit of bystander behavior comes from the additional productivity, consumption spending, and infrastructure investment. Each time a participant of the bystander program prevents a violent crime, we can assign the social benefit to the program (Welsh & Farrington, 2000).

⁵ We refer the interested reader to the standard literature on the Costs of Crime (e.g. Cohen, 2000, 2005; Cohen & Bowles, 2010).

In the fourth scenario, the bystander misunderstands the situation. Because of high arousal and strong emotions, she intervenes without danger ahead. Unjustified intervention might spark costly conflicts. Furthermore, the bystander might injure the person she falsely regards as the offender. Since the “false offender” has done nothing wrong, we would have to count his damages. The fourth scenario is unlikely though. Our model only considers the behavior of people who completed the training. We assume that well-trained people will not misunderstand the situation. Based on the assumption, we exclude the fourth scenario from the following consideration.

3. Building the benefit–cost model

We build our model within the framework of the benefit–cost-ratio (BCR). As we have described above, the benefit arises from preventing violent crime and the resulting social damage. The costs include all expenses on implementing the bystander program. The benefit–cost-ratio is the ratio of benefits over costs:

$$BCR = \frac{BENEFITS}{COSTS}. \quad (1)$$

Policy makers can use the benefit–cost-ratio for their funding decisions. Suppose they consider funding a single program. If the benefit–cost-ratio is above one, the program benefits society and should obtain funding. If the benefit–cost-ratio is below one, the program costs society and should obtain no funding.

The benefit–cost-ratio has the advantage that it allows comparing programs, even if they come from different policy areas. In case the policy maker has to choose between several programs, she should choose the program with the highest benefit–cost-ratio. We build our model on the assumption that the benefits and costs of the bystander program accrue in the short term. Because of the short-term view, we refrain from discounting the benefits and costs.

3.1 Estimating program costs

Remember that program costs include the expenses for implementing the program. We can simply read the expenses from the account statement. The bulk of expenses will arise from social network advertising. We must rely on ads for reaching enough people to make an impact on society. Here, we face the problem that *organic reach* on social media is low. On *Facebook*, for example, *organic reach* ranges between 2% and 6%. This means if we have 100 followers and refrain from advertising, only 2 to 6 persons will have our content displayed in their feed (Mochon et al., 2017).

The second largest bulk of expenses arises from labor. We need staff to develop the strategy, set up the online presence, and create appealing content. We also need staff to manage and analyze our media, community, and promotions. Operating costs include license fees for software and design elements. Pro rata costs include real estate, public utilities, maintenance of equipment, office supplies, insurance premiums, depreciation and replacement, and taxes (Milbrath, 2016).

3.2 Estimating program benefits

While we can find the costs on our account statement, we must put more effort into finding the benefits of the bystander program. Suppose benefit arises from crime prevention as described above. In order to estimate the monetary value of the benefit, we must answer two fundamental questions: First, how many crimes are prevented because of the bystander program? Second, what is average monetary damage coming from one of these crimes? If we multiply the number of crimes prevented by the damage per crime, we get the monetary value of program benefits.

In turn, the number of crimes prevented is the product of several factors. The program will only reach a certain number of participants. Only some of the participants will encounter a violent situation and get the chance to intervene. Of those who get the chance, only some will actually intervene and prevent a violent crime. Remember that some bystanders replace the initial victim. In case the bystander is injured instead of the initial victim, the total number of crimes remains unchanged.

If we multiply the total participants by the share of participants who become a bystander of a violent situation, the share of bystanders who intervene in that situation, and the rate of non-replacement, we get the number of crimes prevented. If we then multiply the number of crimes prevented by the average damage per crime, we get the monetary value of benefits:

$$BENEFIT = PARTICIPANTS \times BYSTANDER \times INTERVENTION \\ \times (1 - REPLACE) \times DAMAGE \quad (2)$$

Where *PARTICIPANTS* is the total number of program participants, *BYSTANDER* is the share of participants who become a Bystander, *INTERVENTION* is share of bystanders that intervene because of the program, and *REPLACE* is the share of bystanders who are injured during intervention. Logically, $(1 - REPLACE)$ is the share of safe interventions. *DAMAGE* represents the average monetary damage per crime. In order to calculate the value of program benefits, we must separately estimate each factor in equation (2), and then multiply the factors as described above.

4. How many crimes are prevented?

4.1 Estimating the number of program participants

The number of program participants has a direct influence on the number of crimes prevented. The more people participate, the more people will be motivated and able to intervene when they observe a risky situation. The appropriate technique for estimating the number of participants depends on distribution channel. If the program distributes its content directly to the social media feed, the corresponding social media analytics tools will display the relevant outcome metrics including *reach* and *videos*. While the *reach* metric gives us the number of people who had the opportunity to see a post, the *videos* metric counts how many people watched a video for three seconds or longer. Depending on the form of content, the corresponding metric will provide a good estimate of the number of participants.

Direct distribution to the social media feed may be suited for a brief program intervention, which consists of a series of posts or videos. A program that is more complex may redirect the participant from the social networking site to an external website, which contains the program content. A website specially built for the purpose offers a higher degree of design freedom. For example, it allows using interactive content like videogames. Now the estimation of the number of participants requires installing *Google Analytics* and setting up the website in a particular way. Each time a person completes the bystander program, the website must display a specified page. We must define the page as a destination goal in *Google Analytics*. Now *Google Analytics* will count the number of times that a person completed the program. Online marketers call a completed activity a *conversion*. In our context, the *reach*, *videos*, or *conversions* provide good estimates of the number of participants.

Even though the presented estimation techniques rely on the best technology available to date, interpretation of the metrics must take place with some caution. Instead of perfectly accurate information, the analytics tools provide estimates with an underlying probability distribution. Imperfection arises from technical problems including *cookie deletion*, *multiple-machine browsing*, or *non-human traffic* (Sterne, 2010). To account for the imperfection, sensitivity analyses is mandatory when estimating the number of program participants. Sensitivity analysis involves assuming a margin of error for the outcome metrics, and calculating confidence intervals on that basis. The minimum, average, and maximum values from the confidence interval enter equation (2) to examine how variation in the outcome metric affects the program benefits.

The estimated number of participants forms the upper bound for the number of crimes prevented. In order to provide conservative estimates, we assume that participants will intervene not more than once, and exclude possible spillover effects on their family and friends from further examination.

4.2 Estimating the share of participants that become a bystander

For effective crime prevention, a large number of program participants is necessary but not sufficient. Only the persons that become a bystander of a violent situation get the chance to prevent a crime. A higher share of such persons among program participants will *ceteris-paribus* increase the number of crimes prevented. Every participant that does not become a bystander costs the program without returning any benefit. The program should therefore aim to target future bystanders. Estimating the share of future bystanders relies on the following expression:

$$BYSTANDER = \frac{Risk\ population}{1 - PRECISION(1 - Risk\ population)}, \quad (3)$$

where *Risk population* is the share of the total population that is at risk of becoming a bystander. *PRECISION* We suggest using the rate of violent crime to approximate the share of the total population at risk. A higher general crime rate increases the likelihood of becoming a bystander. Police statistics and victimization surveys provide information on the crime rate. The *PRECISION* coefficient measures the precision in targeting the persons at risk. Appendix A shows the derivation of equation (3).

According to equation (3), the share of the population at risk forms the lower bound for the share of future bystanders among program participants. If targeting is completely imprecise (*PRECISION* = 0), the share of future bystanders among participants is equal to the share of the population at risk (*BYSTANDER* = *Risk population*). In contrast, if targeting is completely precise (*PRECISION* = 1), every program participant will become a bystander in the future (*BYSTANDER* = 1).

In order to target specific consumers for their advertising costumers, *Facebook* and *Google* rely on machine learning technology. Researchers increasingly discover the potential of the technology for prevention policy. In a feasibility study, Hassanpour et al. (2019) use *Instagram* data and machine learning to identify persons at risk of substance abuse. Since we also use social media to target program participants, we suggest applying their approach to identify persons at risk of becoming a bystander. Although the approach uses *Instagram* data, the procedure should also be suited for other social networking sites.

The approach bases on the assumption that *Instagram* profile data contain information, which indicates risk. People share information about their personal life by posting image and text content. Lifestyle choices and social environments influence the level of risk (Calvó-Armengol & Zenou, 2004; Walters, 2017). For example, a lot of violence happens in nightlife. People who routinely engage in nightlife are more likely to encounter a violent situation. Many party pictures in the profile may therefore indicate risk.

Researchers can train a machine-learning architecture to extract the information from *Instagram* profile data that indicates risk, and classify people into risk categories accordingly. Then, they can test if the trained model makes correct predictions about the risk categories. Finally, they can use the extracted information to target the suitable program participants. The test results allow estimating the share of future bystanders among participants. For training the machine-learning architecture, one must first draw a random sample of persons who have an *Instagram* profile. The *application-programing interface* allows downloading the profile data, which consist of pictures, captions, and comments on pictures made by other people. To predict the risk category of a person based on her profile data, the algorithms need information about the true risk level. To gather the information, we use an online survey that asks the person about experiences with violence (Mynard & Joseph, 2000). More experiences with violence indicate a higher risk level. We match the survey data with the profile data using a unique identifier, in order to prepare the next step.

The machine-learning architecture consists of several layers of neural networks. The first layers extract the relevant information from the profile data, while the last layer generates the risk estimation model. Because images and text have different data structures, different neural networks are required to extract information from each data type. Roughly speaking, the neural networks take the relevant information from the image or text, and map it to a vector space. For example, each word in the English language has an index number based on its position in the dictionary. The first word in the dictionary has number 1, the second number 2, and so on. The neural network maps words to a vector space based on their co-occurrences in the *Instagram* captions or comments. The vector space enters the risk estimation model in the next layer.

The risk estimation model is based on a neural network layer with *softmax normalization* and a *cross-entropy loss* function. The *softmax function* takes the vector space of real numbers generated in the previous layer, and normalizes it to a probability distribution. After applying *softmax*, each component of the vector space can be interpreted as a probability. Larger input components will lead to larger probabilities. The probability distribution is the predicted value of risk. The *cross-entropy loss* function provides a measure of dissimilarity between the predicted risk and the true risk. Remember we measured the true risk using the online-survey. When

true risk and predicted risk are the same, the loss function takes on the minimum value. At this point, the model makes a correct prediction of risk categories.

The risk estimation model is trained using a particular subsample – the *training set*. Training involves finding the vector weights that minimize the *cross-entropy loss* function. For this purpose, the authors suggest to use *stochastic gradient descent*, which is an iterative algorithm to find the minimum value. Imagine the loss function as range of mountains: From a starting point, the algorithm goes along the descent until the numerical value does not improve any more. The model may then be trained in several iterations. After each iteration, one should feed back the *cross-entropy loss* to the neural network in order to improve the predictions as explained above.

Another subsample – the *test set* – is used to evaluate if the trained model makes correct predictions about the risk categories. *Precision* is a standard machine learning evaluation metric that measures the share of correct predictions. We can formally express the precision metric using the following equation:

$$PRECISION = \frac{\sum \text{True risk category}}{\sum \text{Predicted risk category}}. \quad (4)$$

The numerator is the number of persons who truly belong to a particular risk category according to their survey data. The denominator is the number of persons classified as belonging to this risk category by the trained model.

Persons in the high-risk category are likely to become future bystanders. With our bystander program, we should therefore target persons with similar Instagram profiles and sociodemographic characteristics. *Instagram* allows targeting such persons based on their location, sociodemographic characteristics, interests, and online behavior. We only have to enter the respective information into the advertising platform. If we use the presented approach for targeting, we can enter the *PRECISION* metric into equation (3) to estimate the share of future bystanders among participants.

Potential shortcomings of the machine learning approach arise from training the models on an imbalanced distribution of risk-classes among people. In the population, we observe a low prevalence of high-risk individuals. A random sample, therefore, is probably skewed towards the lower risk population. The imbalance can negatively affect the performance of the models, which rely on information of each risk group. Nevertheless, we can overcome this shortcoming by oversampling high-risk individuals. For example, the study may include a high proportion of juvenile offenders. Further shortcomings relate to potential biases and privacy issues. To address these issues, we must implement legal regulations and boundaries before real-life implementation. Finally, implementing the approach would require a lot of time and resources. The research team that wants to implement the approach requires at least

one machine-learning expert, but these experts are scarce. We therefore suggest conducting a pilot study to extract the relevant information from social media data. Ideally, this information will be made available to other researchers who want to conduct a benefit-cost analysis of a social media facilitated bystander program or a similar policy measure.

Our approach may remind the informed reader of *predictive policing*. The two approaches are similar, yet distinct. While *predictive policing* includes methods to predict likely offenders or victims (Perry, 2013), our approach aims to predict likely bystanders of a violent crime. Remember that prevention programs that address potential offenders or victims are less effective than programs that address potential bystanders. Furthermore, arresting a potential offender based on a prediction is incompatible with the principles of the rule of law. Our approach therefore presents a useful extension of the idea of *predictive policing*.

4.3 Estimating program effectiveness

Program effectiveness is crucial for the number of crimes prevented. A bystander program is only effective if it changes the behavior of its participants. But only some of the participants who encounter a violent situation will intervene because they have finished the program. Other participants will not intervene at all, while yet others would have intervened in any way. Only for the first group, we can say the program had a causal effect. In other words, the causal effect is the change in behavior we can clearly attribute to the program. The more people intervene because they have finished the program, the larger is the causal effect (Thomsen, 2016).

Actual bystander behavior, however, is hard to observe under experimental conditions. We therefore suggest measuring the causal effect on the willingness to intervene (WTI). The willingness to intervene is an example of behavioral intention, which is the single best predictor of actual behavior (Fishbein & Ajzen, 2011). In order to measure the willingness to intervene, we suggest using the scale developed by Banyard et al. (2005). The scale asks respondents to express the likelihood that they would perform each of 51 bystander behaviors. Possible answers range from 1 (*very unlikely*) to 7 (*very likely*). An exemplary item reads as follows: “*How likely are you to investigate if you are awakened at night by someone calling for help?*” (Banyard et al., 2007, p. 8). The mean score over the 51 items gives a measure of the willingness to intervene.⁶

⁶ Few studies have attempted to observe bystander behavior under experimental conditions. For one of these studies, researchers staged a violent situation, and assessed whether participants actually tried to help the false victim. They also measured the time it took them to decide before helping (Fischer et al., 2006).

In order to estimate the causal effect, we suggest conducting a randomized experiment, which involves randomly assigning participants to either the treatment group (TG) or the control group (CG), and measuring the difference in outcomes between the two groups. The treatment group receives the bystander program, while the control group does not. The suggested experimental design uses *timing* as the randomization method.

We explain the basic procedure of conducting the randomized experiment on *Instagram*. Nevertheless, researchers can apply the procedure to other social networking sites as well. The experiment requires setting up an *Instagram* page and at least two ad campaigns. Remember that advertising is necessary to reach a sufficient sample size. With both ad campaigns, we target persons who are likely to become a bystander. These are persons who have the same characteristics as the high-risk group identified by the machine-learning approach. For targeting these persons, we have to enter their location, sociodemographic characteristics, and interests into *Instagram's* advertising platform.

As explained above, we achieve randomization by the timing of the ad campaigns. The ads of the first campaign take place before the start of the bystander program. The ads contain a link redirecting people to an online-survey that measures their willingness to intervene. Since they have no access to the bystander program, the respondents of the first survey form the control group.

The ads of the second campaign take place after the start of the bystander program. They contain a link that redirects people to the website holding the program content. At the end of the program, another link redirects people to the online-survey measuring the willingness to intervene. Since they received the bystander program, the respondents of the second survey form the treatment group. To ensure that the second ad campaign does not reach any members of the control group, we must set an exclusion parameter in the advertising platform of *Instagram*.

Random assignment prevents people from choosing one of the two groups based on their personal characteristics. So-called *self-selection* would yield biased estimates. If, for example, all people who already have a high willingness to intervene would choose the treatment group, while all people with a low willingness choose the control group, the mean difference between the groups would overestimate the effect of the bystander program. In this context, we speak of *selection bias*. But if random assignment works and the sample is large enough, the two groups are similar in all observable and unobservable characteristics. Now the difference in outcomes does not result from difference in individual characteristics but only from program

Such experiments are, however, subject to serious ethical concerns, and hard to implement within the online environment.

participation. Under randomization, the mean difference in outcomes thus yields a *consistent* estimate of the causal effect:

$$INTERVENTION = \overline{WTI}_{TG} - \overline{WTI}_{CG}. \quad (5)$$

\overline{WTI}_{TG} is the mean willingness to intervene in the treatment group, and \overline{WTI}_{CG} is the mean willingness to intervene in the control group. To be very specific, *INTERVENTION* shows the average causal effect of the program on its participants, the *average treatment effect on the treated* (ATT). For the case at hand, the ATT measures the share of bystanders that intervene because they have finished the program.⁷

Although social sciences widely accept randomized experiments as the gold standard of policy evaluation, the method has particular limitations. The randomized experiment is the best method for estimating the treatment effect only if randomization does not affect the decision to participate in the program. We speak of *randomization bias*. Furthermore, for a multistage program, the mean-difference estimator is only valid conditionally on the stage where randomization takes place (Heckman, 2020). Since we assume a one-stage program where randomization does not affect the decision to participate, the limitations do not apply.

Another limitation arises from the fact that people tend to express strong behavioral intentions but fail to act accordingly. The gap between intention and actual behavior results from differences between perceiving hypothetical situations and perceiving real situations. Because of this so-called *hypothetical bias*, the answers given in a survey might overstate the true underlying willingness to intervene (Ajzen et al., 2004). We can draw on two potential solutions to the problem. The first solution involves explaining *hypothetical bias* before asking the questions on the willingness to intervene. The second involves asking respondents how certain they are about the answers they have just given (Blumenschein et al., 2008).

4.4 Estimating the rate of replacement

The last factor considered in this section is the rate of replacement, which we have denoted as *REPLACE*. As described above, the rate of replacement measures the share of bystanders who are injured during their attempt to intervene. If the bystander replaces the initial victim, the total number of crimes remains unchanged. In reverse, $(1 - REPLACE)$ measures the share of interventions without

⁷ In order to see if randomization was successful, it is mandatory to be *checking for balance*. This means that we test the hypothesis, that the distribution of covariates is similar in both of the groups (TG and CG).

replacement. The extent to which replacement occurs is an open question we have to answer empirically.

For our conceptual framework, we assume that the rate of replacement consists of an *idiosyncratic* component and a systematic component. The *idiosyncratic* component depends on the capacity of the bystander to intervene safely. The capacity in turn depends on the skills acquired during the bystander program. If the bystander program has been able to increase the capacity by teaching the necessary skills, the rate of replacement will decrease accordingly.

To estimate the effect of the bystander program on skills, we include the *bystander efficacy scale* in the experiment described in the previous section. For each of the 51 presented bystander behaviors, the scale asks respondents for their self-confidence in performing the behavior. Possible answers range from 0% (“cannot do”) to 100% (“very certain can do”) in 10% increments. The mean value over the 51 items yields the bystander efficacy score (Banyard et al., 2005).

The mean difference in the bystander efficacy score between the treatment group and the control group provide a consistent estimate of the program’s causal effect on the intervention skills:

$$SKILLS = \overline{EFFICACY}_{TG} - \overline{EFFICACY}_{CG}. \quad (6)$$

$\overline{EFFICACY}_{TG}$ is the mean bystander efficacy score in the treatment group, and $\overline{EFFICACY}_{CG}$ is the mean bystander efficacy score in the control group. The *SKILLS* term is the ATT on the capacity for safe intervention.

The systematic component of the risk of replacement includes all factors that are beyond the control of the bystander. The systematic risk factors include the characteristics of the victim, offender, and situation. Systematic risk increases the risk of replacement. To capture systematic risk, we add a *risk premium* to our consideration, which is denoted as *RISK*. We can express the rate of replacement as a function of skills and the systematic *risk premium*:

$$REPLACE = f(SKILLS, RISK), \quad (7)$$

$$\frac{\partial REPLACE}{\partial SKILLS} < 0, \frac{\partial REPLACE}{\partial RISK} > 0. \quad (8)$$

The first partial derivative shows that a higher level of skills decreases the rate of replacement. The second partial derivative shows that systematic risk increases the rate of replacement. If we assume a linear relation between the rate of replacement, skills, and systematic risk, we can use the following expression:

$$REPLACE = 1 - SKILLS + RISK. \quad (9)$$

On that basis, we can calculate the rate of replacement as follows:

$$(1 - \text{REPLACE}) = 1 - (1 - \text{SKILLS} + \text{RISK}) \quad (10)$$

$$\Leftrightarrow (1 - \text{REPLACE}) = \text{SKILLS} - \text{RISK} \quad (11)$$

In turn, we can insert the rate of replacement into equation (2) in order to calculate the benefits of the bystander program.

Potential limitations of the approach arise from the assumption of linearity and the proper estimation of the risk premium. To account for the limitations, we recommend conducting sensitivity analysis by systematically varying the risk premium in equation (11) in order to see how variation affects the program benefits given by equation (2). We may also consider different functional forms for equation (7).

5. What is the damage per crime?

To calculate the value of program benefits, we have to multiply the number of crimes prevented by the average damage per crime. We denote the average damage as *DAMAGE*. To obtain an estimate for the average damage, we have to make four preliminary considerations. First, we have to decide on the scope of analysis. The scope of analysis determines the perspective from which we consider the benefits and costs of the program. Ultimately, we aim to serve the policy maker who decides on the funding of the program. The policy maker should consider the effect of the program on overall welfare. We therefore consider the damages incurred by the taxpayer and the victim.

Second, we have to decide if we consider social costs or external costs. Social costs reduce the aggregate well-being of society, while external costs refer to negative consequences imposed by one person to another – who does not voluntarily accept the negative consequences. Because we lack a general agreement on which approach to prefer, we suggest providing estimates based on both approaches. On that basis, the policy maker can choose the approach she prefers.

Third, we must decide if our estimates should base on incidence or prevalence. *Incidence-based* estimates count the present and future costs in the year in which the crime takes place. *Prevalence-based* estimates count the costs in the year in which they are realized, regardless of when the crime has taken place. Because *incidence-based* estimates show how much we could save by preventing future incidents, they are better suited for our framework.

Fourth, we must choose the appropriate costing methodology. For our conceptual framework, we suggest the use of *discrete choice experiments* (DCEs). Appendix

B provides an overview of alternative costing methodologies including the *jury awards method*, the *life satisfaction method*, and the *quality-adjusted life years method*.

Discrete choice experiments are a top-down, stated preferences approach to estimating the damages of crime. Top-down approaches attempt to combine all cost components into a single measure by estimating the public willingness-to-pay for crime reduction. Stated preferences methods use surveys to ask respondents for their subjective willingness-to-pay. Picasso & Grand (2019) use *discrete choice experiments* to estimate the willingness-to-pay for a reduction in homicide risk. Picasso & Cohen (2019) extend the approach by including the risk of violent crime. Based on their examples, we explain the basic procedure for estimating the damages of crime. The procedure consists of three general steps. In the first step, we use *discrete choice experiments* to measure the individual preferences for crime prevention. In the second step, we estimate the *discrete choice model*, which consists of several utility functions representing the preferences. Based on the estimated utility functions, we calculate the average value of damages in the third step.

As described above, we use *discrete choice experiments* to measure the preferences for crime prevention. Each experiment consists of several choice tasks. Each task involves the choice between three alternatives: the status quo and two security programs. We describe the security programs with the associated crime risk, policy, and program costs. We illustrate the crime risk using the number of victims among family and friends during one year. The policies differ in the intensity of police presence and the strictness of sentencing. Finally, we exemplify the costs using a hypothetical tax for implementing the security program. Crime risk, policy, and program costs of the security programs are expressed in relation to the status quo.⁸

The status quo and the alternative scenarios are fundamentally different. The status quo is well known, while the alternative scenarios are hypothetical. To consider the difference, we present each choice task in two stages. First, the respondent has to choose whether to stay in the status quo or to invest in a security program. Second, she has to choose between the two security programs. If we assume rationality, people will select the alternative that generates the highest utility. The choices made thus reflect their individual preferences.⁹

The *discrete choice model* represents the decision-making patterns during the choice experiments as a set of utility functions:

⁸ We can implement the choice experiments using *oTree*, which is a software platform for economics experiments.

⁹ To reduce the number of possible choice tasks, an *optimal experimental plan* can be followed. Accordingly, a reduced number of tasks is created using Fedorov's (2013) algorithm. These tasks are arranged into blocks that are manageable for respondents via the algorithm by R. K. Meyer & Nachtsheim (1995). Finally, the different blocks are randomly assigned to the survey participants.

$$\tilde{U}_{ait} = U_{ait} + \varepsilon_{ait}. \quad (12)$$

The term \tilde{U}_{ait} is the unobservable true utility generated by alternative a for individual i in task t . Utility depends on the deterministic component U_{ait} , and the random component ε_{ait} . The random component accounts for issues such as omitted variables, measurement errors, or individual deviations from rationality. With no loss of generality, we can express the deterministic component as a function that is linear in parameters:

$$U_{ait} = \beta' x_{ait}. \quad (13)$$

β is a vector of coefficients associated with the vector x of the explanatory variables. The explanatory variables are the attributes of the alternatives. They include the crime risk V (% of households victimized/year), the strictness of sentencing S (0 for current, 1 for strict), police presence P (0 for current, 1 for extended), and the costs of the security program C (monetary units/month). Note that S , P , and C take on a value of zero for the status quo. The β coefficients capture the sensitivity of the utility to changes in the explanatory variables.¹⁰

Model calibration consists of finding the values of the β coefficients that best reproduce the choices people made during the choice experiments. To account for the two-stage structure of the choice tasks, we employ a *nested logit discrete choice model*. This means we estimate a *multinomial logit model* at each stage of the choice task. We conduct the calibration of the model via *maximum likelihood estimation* using any common statistical packages, e.g. like *R* or *STATA*. In order to explore the functional form, we estimate different model specifications.¹¹ We can compare the different model specifications using *Akaike's Information Criteria* (AIC) and *likelihood-ratio* (LR).¹²

Finally, we use the estimated utility function to calculate the average damage per crime. We can derive the average damage from the *marginal rate of substitution* (MRS) of money for risk:

¹⁰ Choices are likely to differ between individuals. To account for individual differences, the set of explanatory variables can be expanded to demographic and psychographic measures z , with the associated vector of coefficients γ . The individual differences may be hypothesized to influence utility via the vector of β coefficients.

¹¹ For example, Picasso & Cohen (2019) perform quadratic Box & Cox (1964) estimation. This implies including level, square and cross terms for each of the explanatory variables.

¹² AIC provides a means for model selection. Suppose a statistical model aims to represent a data-generating process. The representation will never be exact. Some information will be lost instead. AIC estimates the relative amount of lost information. When comparing several models, researchers should choose the one with the lowest AIC score. In comparison, when conducting a LR test, they should choose the model with the highest log-likelihood.

$$MRS = -\frac{\frac{\partial U}{\partial V}}{\frac{\partial U}{\partial C}}. \quad (14)$$

The numerator is the marginal utility of crime risk, and the denominator is the marginal utility of program costs. Put simply, the *marginal rate of substitution* (approximately) indicates how many monetary units a household is willing to give up in exchange for a reduction in crime risk. The *marginal rate of substitution* thus captures the average willingness to pay for crime reduction per household. Multiplied by the number of households in the population, the *marginal rate of substitution* yields a good estimate of *DAMAGE*. In order to calculate the benefit of our bystander program, we insert the estimate of *DAMAGE* into equation (2).

As a potential limitation, the DCE method is only valid to the extent to which surveys are able to elicit the true preferences. In response, *revealed preferences* methods could be considered as reasonable alternatives. This is because they elicit preferences from actual market transactions.¹³ *Revealed preferences* methods are, however, only valid if we can assume market efficiency and full information on risks and prices. They are also limited to actually observed crime rates and spending patterns. This makes it impossible to value hypothetical risks or policies yet to be implemented.

Furthermore, different crimes are highly collinear. For this reason, revealed preferences cannot isolate the value of specific crimes. They also tend to underestimate the true costs of crime. Finally, the required market data are not always available. This is especially true in developing countries. Because of the stated reasons, revealed preferences methods are difficult to implement in practice. This makes stated preference approaches the usual choice (Cohen & Bowles, 2010; Picasso & Cohen, 2019; Picasso & Grand, 2019).

6. Estimating the value of public attention

In the previous three sections, we have described our conceptual framework for estimating the benefits of a social media facilitated bystander program. Within the framework, program benefit arises from crime prevention. In this section, we

¹³ For example, Thaler (1978) derived property owners' WTP for a safe neighborhood from differences in property values. This approach is based on the assumption that neighborhoods with higher crime rates *ceteris paribus* have lower housing prices (see also Hellman & Naroff, 1979; Rizzo, 1979). A similar revealed preferences approach derives the *value of a statistical life* from compensating wage differentials for on-the-job exposure to the risk of death (Viscusi & Aldy, 2003). This approach evaluates wage-risk tradeoffs in labor markets assuming that workers are willing to accept a marginally higher fatality risk in exchange for appropriate compensation.

introduce a complementary approach where benefit arises from the public attention on the issue of violence. The basic rationale behind the complementary approach is as follows: in the *information age*, attention has a value of its own. In order to determine the value of attention, we suggest using an opportunity cost approach, which draws on methodology from online display advertising.

With the proliferation of digital media, information is abundant, almost free of charge, and available in real time. Because of non-rivalry and non-excludability in consumption, some economists even consider information a public good. Under these circumstances, attention is the limiting factor in the consumption of information and thus considered a scarce resource (Davenport & Beck, 2001). Because we trade consumer attention on the digital advertising market, the market price will reflect the value of attention. We use the market price to estimate the opportunity costs of paying attention to a particular piece of information – the bystander program.

The most common form of trading online ad space is programmatic advertising with *real-time bidding* (RTB). Programmatic advertising accounted for 73% of the U.S. digital advertising market in 2017. In 2023, the market share is expected to be 92% (Statista, 2018). *Real-time bidding* is a complex process, which involves several parties that interact automatically in real time. In essence, each time a person visits a website, the system auctions the available ad space on an ad exchange. See Appendix C for a detailed description of the *real-time bidding* process.

Under *real-time bidding*, the hammer price depends on both the characteristics of the website and the person paying attention to it. We can exploit the fact that the market price partly depends on personal characteristics, to estimate the value of attention of a specific group of people – the participants of the bystander program. For estimating the value, we employ the *bid simulator* in *Google Analytics*. The *bid simulator* simulates a process called *bid landscaping*. In *bid landscaping*, advertisers place bids on several websites, in order to see how varying the bid price influences the probability of winning the bid. Cost curves illustrate the relation between the bid price and the probability of winning the bid as depicted by Figure 2 (Paulson et al., 2018).

The bid price is expressed in *cost per mille* (CPM), which is the price of generating a thousand *impressions*, i.e. a thousand opportunities to see the ad. The probability of winning the bid increases with the bid price. The cost curve is concave meaning the bid price has decreasing marginal effects on the probability of winning.

Notably, the *bid simulator* allows simulating cost curves for specific target groups based on their demographic, psychographic, and behavioral data. As implied above, we can exploit this feature to estimate the value of attention generated by the bystander program. For estimating the value of attention generated, we collect demographic (and psychographic) information from a representative sample of

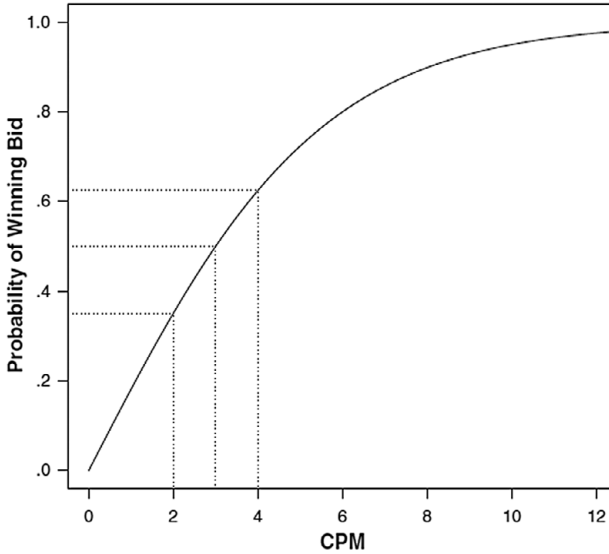


Figure 2 Example cost curve.

Notes: The cost curve illustrates the mathematical relation between the bid price and the probability of winning the bid. This relation is revealed through the process of bid landscaping. The bid price is expressed in cost per mille. This is the price of generating thousand impressions. The probability of winning the bid increases with the bid price. The marginal effect of the bid price on the probability of winning the bid is decreasing. The cost curve, therefore, has a concave shape. *Source:* Paulson et al. (2018), p. 491.

program participants. We enter the collected information in the *bid simulator*, which generates cost curves based on the information. For different bid prices, the generated cost curve will show the probability of winning the bid for people who have the same characteristics as our program participants. The bid prices thus reflect the value of having their attention with a certain probability.

The described approach uses the bid price as an estimate of the value of attention *per capita*. We multiply the bid price by the total number of program participants to calculate the total value of attention generated by the program. For the bystander program to be profitable, the value of attention generated should exceed the expenses for program implementation. Again, we suggest conducting sensitivity analysis. For this purpose, we calculate the bid prices that have 0.5 and 0.95 probability of winning the bid respectively. The bid price with a 0.5 probability of winning is the *average CPM*. We can use the *average CPM* for calculating the lower bound of program benefits. The bid price with a 0.95 probability of winning provides the basis for calculating the upper bound.

Of notice, the approach described above represents a useful extension to the more comprehensive conceptual framework developed in the previous sections.

Generating attention is the first step in the process leading to behavior change. The approach is therefore best suited to assess the benefit–cost ratio in the early stages of program implementation. Furthermore, the approach is useful to evaluate small-scale bystander programs. Small-scale programs may lack the resources necessary to follow the more comprehensive framework. Finally, the direct value of public attention is likely to be well below the costs of crime prevented. If the presented approaches yield different results, we suggest interpreting the value of attention as the lower bound of program benefits.

7. Conclusion

Bystander behavior contributes to preventing violent crime and the associated social damages. It can be a useful complement to traditional police work if done properly. Police campaigns in various countries therefore promote engaging in bystander behavior. Before people decide to intervene in a violent situation, however, they have to overcome a number of mental obstacles. The bystander program tackles the mental obstacles in order to make people intervene despite personal risk. It also teaches the skills necessary for safe and effective intervention. The online bystander program addresses the challenges of traditional programs by reducing staff costs and facilitating attendance.

In this paper, we have shown how social media allow targeting the right people and collecting information for benefit–cost analysis. On that basis, we have developed a coherent framework for conducting a benefit–cost analysis of social media facilitated bystander programs. As with any research approach, there are some limitations. First, the breadth and depth of data collected is constrained by privacy regulation and user consent. Each study that applies our framework has to reconcile data requirements and privacy rights in the studied country. Second, if the framework is implemented in practice, the external validity of results will be limited by political, legal, and cultural differences between countries. Researchers should bear this in mind.

Despite the limitations, social media data, analytics tools, and machine learning technology have a large potential for the advancement of benefit–cost analysis. We believe that our framework has laid the groundwork for further development. The importance of social media for the dissemination and communication of policy measures is constantly increasing. New evaluation methods are required to facilitate evidence-based policymaking. Our intention is to make such a method accessible to a broad readership.

Supplementary Materials

To view supplementary material for this article, please visit <http://dx.doi.org/10.1017/bca.2020.34>.

References

- Ajzen, I., T.C. Brown, and F. Carvajal. 2004. "Explaining the Discrepancy Between Intentions and Actions: The Case of Hypothetical Bias in Contingent Valuation." *Personality and Social Psychology Bulletin*, 30(9): 1108–21.
- Banyard, V. L., M.M. Moynihan, and E. G. Plante. 2007. "Sexual Violence Prevention Through Bystander Education: An Experimental Evaluation." *Journal of Community Psychology*, 35(4): 463–81.
- Banyard, V. L., E.G. Plante, and M.M. Moynihan. 2005. *Rape Prevention through Bystander Education: Bringing a Broader Community Perspective to Sexual Violence Prevention*, 1–206. Washington, DC: US Department of Justice.
- Becker, G.S. 1968. "Crime and Punishment: An Economic Approach." In *The Economic Dimensions of Crime*, 13–68. London: Palgrave Macmillan.
- Blumenschein, K., G.C. Blomquist, M. Johannesson, N. Horn, and P. Freeman. 2008. "Eliciting Willingness to Pay Without Bias: Evidence from a Field Experiment." *The Economic Journal*, 118(525): 114–37.
- Bond, R.M., C.J. Fariss, J. J. Jones, A. D. I. Kramer, C. Marlow, J. E. Settle, and J. H. Fowler. 2012. "A 61-Million-Person Experiment in Social Influence and Political Mobilization." *Nature*, 489(7415): 295.
- Box, G.E.P., and D.R. Cox. 1964. "An analysis of transformations." *Journal of the Royal Statistical Society: Series B (Methodological)*, 26(2): 211–43.
- Calvó-Armengol, A., and Y. Zenou. 2004. "Social Networks and Crime Decisions: The Role of Social Structure in Facilitating Delinquent Behavior." *International Economic Review*, 45(3): 939–58.
- Cohen, M.A. 2000. "Measuring the Costs and Benefits of Crime and Justice." In *Measuring the costs and benefits of crime and justice*, edited by D. Duffee, Vol. 4, 263–316. Washington, DC: National Institute of Justice.
- Cohen, M.A., and R. Bowles. 2010. "Estimating Costs of Crime." In *Handbook of Quantitative Criminology*, 143–162. New York, NY: Springer.
- Coker, A.L., B.S. Fisher, H.M. Bush, S.C. Swan, C.M. Williams, E.R. Clear, and S. DeGue. 2015. "Evaluation of the Green Dot Bystander Intervention to Reduce Interpersonal Violence Among College Students Across Three Campuses." *Violence Against Women*, 21(12), 1507–27.
- Cugelman, B., M. Thelwall, and P. Dawes. 2011. "Online interventions for social marketing health behavior change campaigns: a meta-analysis of psychological architectures and adherence factors." *Journal of Medical Internet Research*, 13(1): e17.
- Darley, J. M., and B. Latané. 1968. "Bystander Intervention in Emergencies: Diffusion of Responsibility." *Journal of Personality and Social Psychology*, 8(4p1): 377.
- Davenport, T. H., and J. C. Beck. 2001. *The Attention Economy: Understanding the New Currency of Business*. Brighton, MA: Harvard Business Press.

- DeGue, S., L.A. Valle, M. K. Holt, G.M. Masetti, J.L. Matjasko, and A.T. Tharp. 2014. “A Systematic Review of Primary Prevention Strategies for Sexual Violence Perpetration.” *Aggression and Violent Behavior*, 19(4): 346–62.
- eMarketer. 2019. Facebook-Google Duopoly Won’t Crack This Year. Available at <https://www.emarketer.com/content/facebook-google-duopoly-won-t-crack-this-year>. (accessed Jul 6, 2020).
- Fedorov, V.V. 2013. *Theory of Optimal Experiments*. Amsterdam, NL: Elsevier.
- Fischer, P., T. Greitemeyer, F. Pollozek, and D. Frey. 2006. “The Unresponsive Bystander: Are Bystanders More Responsive in Dangerous Emergencies?” *European Journal of Social Psychology*, 36(2): 267–78.
- Fischer, P., J.I. Krueger, T. Greitemeyer, C. Vogrincic, A. Kastenmüller, D. Frey, M. Heene, M. Wicher, M. Kainbacher. 2011. “The Bystander-Effect: A Meta-Analytic Review on Bystander Intervention in Dangerous and Non-Dangerous Emergencies.” *Psychological Bulletin*, 137(4): 517.
- Fishbein, M., and I. Ajzen. 2011. *Predicting and Changing Behavior: The Reasoned Action Approach*. New York, NY: Psychology Press; Taylor & Francis Group.
- Gigerenzer, G., and R. Selten. 2002. *Bounded rationality: The adaptive toolbox*. Boston, MA: MIT Press.
- Gigerenzer, G. 2010. “Moral satisficing: Rethinking moral behavior as bounded rationality.” *Topics in Cognitive Science*, 2(3): 528–554.
- Gigerenzer, G., and D. G. Goldstein. 1996. “Reasoning the Fast and Frugal Way: Models of Bounded Rationality.” *Psychological Review*, 103(4): 650.
- Halmburger, A., A. Baumert, and M. Schmitt 2017. “Everyday Heroes: Determinants of Moral Courage.” In *Handbook of Heroism and Heroic Leadership*, edited by S. T. Allison, G.R. Goethals, and R. M. Kramer, 165–84. New York, NY: Routledge; Taylor & Francis Group.
- Hassanpour, S., N. Tomita, T. DeLise, B. Crosier, and L. A. Marsch. 2019. “Identifying Substance Use Risk Based on Deep Neural Networks and Instagram Social Media Data.” *Neuropsychopharmacology*, 44(3): 487.
- Heckman, J.J. 2020. *Randomization and Social Policy Evaluation Revisited*. IZA Discussion Paper No. 12882. <https://ssrn.com/abstract=3521700>
- Hellman, D.A., and J.L. Naroff. 1979. “The Impact of Crime on Urban Residential Property Values.” *Urban Studies*, 16(1): 105–12.
- Jonas, K.J., M. Boos, and B. Brandstätter. 2007. *Training Moral Courage: Theory and Practice*. Göttingen, DE: Hogrefe.
- Kettrey, H. H., and R. A. Marx. 2020. “Effects of Bystander Sexual Assault Prevention Programs on Promoting Intervention Skills and Combatting the Bystander Effect: A Systematic Review and Meta-Analysis.” *Journal of Experimental Criminology*, 1–25.
- Kleinsasser, A., E. N. Jouriles, R. McDonald, and D. Rosenfield. 2015. “An Online Bystander Intervention Program for the Prevention of Sexual Violence.” *Psychology of Violence*, 5 (3): 227–35.
- Kosinski, M., D. Stillwell, and T. Graepel. 2013. “Private Traits and Attributes are Predictable from Digital Records of Human Behavior.” *Proceedings of the National Academy of Sciences*, 110(15): 5802–5.
- Latané, B., and J.M. Darley. 1969. “Bystander “Apathy.”” *American Scientist*, 57(2): 244–68.
- Latané, B., and J.M. Darley. 1970. *The Unresponsive Bystander: Why Doesn’t He Help?* Upper Saddle River, NJ: Appleton-Century-Crofts.

- McMahon, S., and V.L. Banyard. 2012. "When Can I Help? A Conceptual Framework for the Prevention of Sexual Violence Through Bystander Intervention." *Trauma, Violence, and Abuse*, 13(1): 3–14.
- Meyer, R.K., and Nachtsheim, C. J. (1995). The coordinate-exchange algorithm for constructing exact optimal experimental designs. *Technometrics*, 37(1), 60–69.
- Milbrath, S. 2016. The 7 Components of Every Social Media Budget. Available at <https://blog.hootsuite.com/the-7-components-of-every-social-media-budget/>. (accessed April 16, 2019)
- Mochon, D., K. Johnson, J. Schwartz, and D. Ariely. 2017. "What are likes worth? A Facebook page field experiment." *Journal of Marketing Research*, 54(2): 306–17.
- Mynard, H., and S. Joseph. 2000. "Development of the Multidimensional Peer-Victimization Scale." *Aggressive Behavior: Official Journal of the International Society for Research on Aggression*, 26(2): 169–78.
- Paulson, C., L. Luo, and G.M. James. 2018. "Efficient Large-Scale Internet Media Selection Optimization for Online Display Advertising." *Journal of Marketing Research*, 55(4): 489–506.
- Perry, W.L. 2013. *Predictive Policing: The Role of Crime Forecasting in Law Enforcement Operations*. Santa Monica, CA: Rand Corporation.
- Picasso, E., and M. A. Cohen. 2019. "Valuing the Public's Demand for Crime Prevention Programs: A Discrete Choice Experiment." *Journal of Experimental Criminology*, 15(4): 529–50.
- Picasso, E., and M.C. Grand. 2019. "The Value of the Risk to Life in the Context of Crime." *Journal of Benefit-Cost Analysis*, 10(2): 178–205.
- Rizzo, M.J. 1979. "The Cost of Crime to Victims: An Empirical Analysis." *The Journal of Legal Studies*, 8(1): 177–205.
- Salazar, L.F., A. Vivolo-Kantor, J. Hardin, and A. Berkowitz. 2014. "A Web-Based Sexual Violence Bystander Intervention for Male College Students: Randomized Controlled Trial." *Journal of Medical Internet Research*, 16(9): 1–16.
- Simon, H. A. 1955. "A behavioral model of rational choice." *The Quarterly Journal of Economics*, 69(1): 99–118.
- Simon, H. A. 1959. "Theories of decision-making in economics and behavioral science." *The American Economic Review*, 49(3): 253–283.
- Simon, H. A. 1979. "Rational Decision Making in Business Organizations." *The American Economic Review*, 69(4): 493–513.
- Statista. 2018. Ad Spending Split – Programmatic / Non-Programmatic. Available at <https://www.statista.com/outlook/216/109/digital-advertising/united-states#market-revenue-programmatic>. (accessed December 20, 2018)
- Sterne, J. 2010. *Social Media Metrics: How to Measure and Optimize Your Marketing Investment. Media*. Hoboken, New Jersey: John Wiley.
- Thaler, R. 1978. "A Note on the Value of Crime Control: Evidence from the Property Market." *Journal of Urban Economics*, 5(1): 137–45.
- Thomsen, S.L. 2016. On the Economic Analysis of Costs and Benefits of Prevention, In: C. Heinzelmann und E. Marks (Ed.): *International Perspectives of Crime Prevention*, 8, 9–20. Forum Verlag Bad Godesberg. ISBN: 978-3-942865-55-5.
- Tversky, A., and D. Kahneman. 1974. "Judgment Under Uncertainty: Heuristics and Biases." *Science*, 185(4157): 1124–31.
- Viscusi, W.K., and J.E. Aldy. 2003. "The Value of A Statistical Life: A Critical Review of Market Estimates Throughout the World." *Journal of Risk and Uncertainty*, 27(1): 5–76.

- Walters, G.D. 2017. “Choice in a Criminal Lifestyle.” In *Modelling the Criminal Lifestyle*, 153–82. London, UK: Palgrave Macmillan.
- Welsh, B.C., and D.P. Farrington. 2000. “Monetary Costs and Benefits of Crime Prevention Programs.” *Crime and Justice*, 27: 305–61.
- Welsh, B.C., D.P. Farrington, and B. R. Gouar. 2015. “Benefit-Cost Analysis of Crime Prevention Programs.” *Crime and Justice*, 44(1): 447–516.
- White, A., D. Kavanagh, H. Stallman, B. Klein, F. Kay-Lambkin, J. Proudfoot, J. Drennan, et al. 2010. “Online Alcohol Interventions: A Systematic Review.” *Journal of Medical Internet Research*, 12(5): e62.
- Zerbe, R.O., and A.S. Bellas. 2006. *A Primer for Benefit-Cost Analysis*. Cheltenham, UK: Edward Elgar Publishing Limited.