

Religious congregations, charitable giving and welfare

Lessons from an event study of the Catholic-clergy sexual abuse scandals
in the United States*

Nicolas Bottan

Department of Economics, University of Illinois at Urbana-Champaign

Ricardo Perez-Truglia[†]

Department of Economics, Harvard University

This Draft: October 2011. First Draft: May 2011.

Abstract

Many studies have documented a strong positive correlation between religious participation and pro-social behavior. However, no conclusive evidence exists supporting the direction of causality. We present novel evidence by exploiting an event study of the Catholic-clergy sexual abuse scandals in the United States from 1980 to 2010. We created a unique dataset with the exact geo-location of each parish involved in a scandal and the exact date when each accusation became public. First, we show that Catholic religious participation in the community declines permanently in the aftermath of a scandal. Second, we demonstrate that the community affected by a scandal suffers a permanent decline in total charitable contributions and a decline in the private provision of welfare. In addition, we show that abuse scandals in lay organizations do not have a similar effect on pro-social behavior. This pattern is consistent with the hypothesis that religious congregations foster pro-social behavior.

JEL Classification: D64, H41, L31, Z1, Z12.

Keywords: religion, welfare, charitable giving, social capital, media.

*We thank Robert Barro for his advice and encouragement during every stage of the project. We want to acknowledge funding from the Lab for Economic Applications and Policy (Harvard University) and the Warburg Funds (Harvard University). Ricardo Perez-Truglia wants to acknowledge support from the Institute of Human Studies and the Bradley Fellowship. Raj Chetty gave us very valuable feedback in the early stage of the project. We thank Alberto Alesina, Joseph Altonji, Joshua Angrist, Julian Cristia, Guillermo Cruces, Rafael Di Tella, Roland Fryer, Osea Giuntella, Edward Glaeser, Daniel Hungerman, Lawrence Katz, Rachel McCleary, Michael Norton, Nathan Nunn, László Sándor, Ugo Troiano and Rodrigo Wagner for their valuable comments, as well as seminar participants at the the Political Economy of Religion Seminar Series (Harvard), the Labor/Public Lunch (Harvard), the Graduate Student Political Economy Workshop (Harvard), the Inter-American Developing Bank and the Boston University Political Economy Research Group. Fiorella Benedetti, Alejandra Baigun and Ovul Sezer provided excellent research assistance. The usual disclaimer applies.

[†]Corresponding author: rtruglia@fas.harvard.edu. Harvard University, Department of Economics: Littauer Center, 1805 Cambridge Street, Cambridge, MA (02138).

1 Introduction

People in the United States donate more time and money relative to those in other developed nations (e.g. Ruiter and De Graaf, 2006) and, in addition, are more likely to attend religious services (e.g. Iannaccone, 2003). Americans who belong to religious congregations are more likely to act pro-socially, as measured by charitable contributions (e.g. Brooks, 2003) and volunteering (e.g. Becker and Dhingra, 2001). The correlation between religious participation and pro-social behavior has led social observers as far back as Alexis de Tocqueville to conjecture that religious organizations foster pro-social behavior (Polson, 2009). Religious beliefs may foster pro-social behavior, as in the case of people being good to others in order to be saved (e.g. Azzi and Ehrenberg, 1975). However, religious participation can have an effect on pro-social behavior beyond the enactment of religious beliefs themselves. For example, religious congregations may foster social norms that encourage helping others (e.g. Wuthnow, 1991). Furthermore, congregations can increase pro-social behavior through increased socialization (e.g., Putnam, 2000; Putnam and Campbell, 2010), through the diffusion of information about opportunities to volunteer and donate money (e.g. Park and Smith, 2002), through solicitation (e.g. Hodgkinson, 1995), and by helping individuals develop skills and resources useful for volunteering (e.g. Peterson, 1992).

Although several studies indicate a strong positive correlation between religious outcomes and pro-social behavior, no conclusive evidence exists pointing to the direction of causality. For example, if a larger number of altruistic people self-select into religious congregations, this would generate a positive correlation between religious participation and charitable giving—even in the absence of a causal effect of religious participation on charitable giving. This paper contributes novel evidence to this long-standing question by examining how a shock to a religious congregation (i.e. the Catholic-clergy sexual abuse scandals) affects religious participation and pro-social behavior. We created a unique dataset containing the exact geo-location of each Catholic institution involved in a Catholic-clergy sexual abuse scandal and the exact date when each scandal first became public. We identify thousands of unique events—distributed throughout all fifty states—and many characteristics on the abuses, abusers, and victims. We exploit the fine distribution of the scandals over space and time by means of an event-study analysis and perform a number of additional falsification tests.

The first part of the paper demonstrates that a community suffers a sharp permanent decline in Catholic affiliation after exposure to a scandal. We take advantage of the fact that most students enrolled in Catholic schools have Catholic families, and then use enrollment in Catholic schools as a behavioral proxy for the size of the Catholic community in a given area. Our estimates suggest a permanent decline in enrollment in Catholic schools in the aftermath of a scandal. Indeed, almost half of the sharp decline in the number and size of Catholic schools experienced in the United States during the 2000s can be directly associated to the scandals occurring during that decade. Additionally, we use survey measures of religious affiliation. The estimates suggest that, if the 1,125 Type-I scandals documented in the database had never happened, the number of adult adherents to the Catholic Church would currently measure approximately 5 million people higher than the current count of adherents—about 10% higher. The data indicates that the scandals did not affect just nominal Catholics, but Catholics who attended church, reported a strong religious affiliation, prayed, believed in life after death, and believed in God. Moreover, the evidence suggests that those individuals who abandoned

their Catholic affiliation because of the scandals did not convert to any other religious denomination.

The second part of the paper shows that a community suffers a sharp permanent decline in pro-social behavior in the aftermath of a scandal. The mean charitable contribution declines permanently in a zip code after a scandal becomes public. The effect on charitable contributions in the zip code stabilizes after the fifth year at around 4%, which one can translate to a drop of around 12% in the charitable contributions of the subpopulation of Catholic adherents. The scandals affected contributions to charities affiliated to the Catholic Church, but had no effect whatsoever on contributions to charities affiliated to non-Catholic religious denominations. Most important, our estimates suggest that the areas affected by a scandal also suffer a permanent decline of 10% in the presence of charitable establishments that provide social services (e.g. soup kitchens, homeless shelters), which means that the decline in charitable contributions is translated into a lower provision of welfare.

In summary, the evidence suggests that a negative shock to a religious institution has a significant effect on pro-social behavior. The third part of the paper discusses the potential mechanisms that can explain this pattern. Our favorite explanation is that religious congregations foster pro-social behavior. Intuitively, we should not compare pro-social behavior across individuals that participate in religious congregations and individuals that do not, because of the typical selection bias. Instead, we should ask whether the individuals that chose to participate in religious congregations would be less benevolent if we were to reduce their religious participation. In an ideal experimental setting, we would like to expose a random group of communities to a “treatment” that makes religious participation more costly (e.g. banning parking lots in parishes), and then test whether in those communities the decline in religious participation is accompanied by a decline in pro-social behavior. In terms of this ideal experimental framework, our study uses a “treatment” in which a priest from the community is publicly accused of sexual abuse. Nevertheless, this is not the only possible mechanisms that can explain the findings. We present evidence against an alternative mechanism, which asserts that the scandals might have a direct effect on pro-social behavior.

Organizations in the non-profit sector rely heavily on individual donations, which account for three quarters of the total contributions of money (List, 2011). Furthermore, individuals make substantial contributions of time: e.g. in 2009, in addition to \$212 billion in money, individuals donated \$169 billion in volunteer service.¹ Religious organizations comprise a substantial share of the American non-profit landscape, as they receive over one third of total donations of money (Giving USA, 2010) and volunteer time (Bureau of Labor Statistics, 2010). In particular, religious organizations are deeply involved in providing social services (Chaves, 2004): e.g. approximately 90% of churches are actively engaged in providing social services (Cnaan et al., 2002), and it is estimated that religious organizations supply social services to over 70 million Americans each year (Johnson, Tompkins and Webb, 2002).² Due to their importance, the U.S. government recently started to collaborate with religious organizations in the provision of social services (Hungerman, 2004; Hungerman and Gruber, 2005), which generated an important deal of controversy.³ Despite the growing interest among policy-makers, economists have

¹The sources are GivingUSA and Corporation for National and Community Service, respectively.

²Non-profit organizations are an important source of income redistribution in the US, insofar they provide a substantial amount of welfare and social services (e.g. Alesina and Glaeser, 2004).

³In January of 2001 President George W. Bush created the White House Office of Faith-Based and Community Initiatives. This initiative was taken to trial in the Supreme Court, which in 2007 ruled that the new Office is not unconstitutional (Hein v. Freedom From Religion Foundation, 551 U.S. 587).

undertaken little research on this topic. Our paper contributes by looking at the interplay between religious congregations, charitable giving, and the private provision of welfare.

This paper relates to a multi-disciplinary literature that seeks to understand the cultural factors that mediate the formation of social capital (e.g. Banfield, 1958; Putnam et al., 1994; Putnam, 2000). In particular, economists have grown increasingly interested in the importance of culture for understanding economic outcomes, such as long-term growth (Nunn and Wantchekon, 2009), corruption (Fisman and Miguel, 2007), and international trade (Guiso et al., 2009). Although religious congregations is believed to play a major role in the determination of those cultural norms (e.g. Barro and McCleary, 2003; Guiso, Sapienza, and Zingales, 2003; for more references see McCleary and Barro, 2006), the existing evidence is not conclusive about the direction of causality. This paper contributes by showing how historical events, like the scandals in the Catholic Church, can be used to measure how religious congregations affect individual and societal outcomes.⁴

In addition, this paper relates to a multi-disciplinary literature that studies the causes and consequences of the clergy sexual abuses in the Catholic Church. Although a number of studies have addressed the causes and circumstances of the abuses (e.g. JJCCJ, 2011) as well as the victims' psychological effects (e.g. McMackin et al., 2009), there are almost no studies looking at broader consequences of the scandals. Some exceptions are Hungerman (2011), who examines the correlation between the number of allegations in a state and religious adherence in that state; and Dills and Hernández-Julian (2011), who examine the correlation between the number of allegations in a diocese and the enrollment in Catholic schools at that diocese. Our paper contributes by looking at broader consequences of the scandals and by creating a unique dataset that allows for precise identification of these effects.

The paper proceeds as follows. Section 2 introduces the data on Catholic-clergy sexual abuse scandals. Section 3 shows the effect of the scandals on alternative measures of religious participation. Section 4 shows the effect of the scandals on pro-social behavior. Section 5 discusses the potential mechanisms that can explain the findings. The final section concludes.

2 Creation of the database on the Catholic-clergy sexual abuse scandals

In this section, we will introduce the data on Catholic-clergy sexual abuse scandals, and provide some basic descriptive statistics on the distribution of scandals over space and time.

Since the mid-1980s, the Catholic Church has repeatedly experienced revelations of sexual abuse committed by its clergy. The number of accusations grew dramatically after January of 2002, when the topic became a national media phenomenon. On January of 2002 the *Boston Globe* published a story about the defrocked priest John Geoghan and his long record of child sexual abuse, arguing that church leaders put Geoghan in positions where he had access to children even though they were aware of his record of child abuse. Hundreds of media articles mentioning allegations of sexual abuse by Catholic-clergy followed the *Globe* article across the country.⁵ As a response to the events of 2002,

⁴See also Nunn (2010) on the religious conversion in colonial africa.

⁵For further details about the chain of events, see Hungerman (2011) and the references therein.

the full body of Catholic bishops of the United States in the General Meeting in Dallas that same year approved the Charter for the Protection of Children and Young People. One of its main resolutions ordered a study about the crisis, later conducted by the John Jay College of Criminal Justice. All the Roman Catholic dioceses in the United States provided information on each priest accused of sexual abuse and on each of the priest's victims. Investigators collected this information in such a way that would not disclose the names of the accused priests or the dioceses in which they worked. The product of this study (JJCCJ, 2004, 2006, 2011) provides a complete picture of the severity of the problem in the United States.

According to these reports, 5,768 priests had received at least one allegation of abuse in the period 1950–2009, or 5.3% of the 109,694 priests active in the United States since 1950. Since a majority of the priests involved in abuses had multiple accusations on record, the number of victims involved in the accusations amounted to 15,235, or 2.6 victims per priest. The majority of the accusations seem to be serious. For example, out of the 10,667 allegations in the period 1950–2002, 5,681 have ended up in diocesan investigations, 80% of which were found substantiated, 18.5% were found unsubstantiated, and only 1.5% were deemed false (JJCCJ, 2004). Doyle and Rubino (2004) estimate that plaintiffs have filed more than 3,000 civil lawsuits related to the clergy scandals in the United States. According to the NGO Bishop Accountability, the Catholic Church in the United States has made around \$3 billion in awards and settlements. Last, a preliminary list by Bishop Accountability suggests that the church has laicized at least 325 accused U.S. priests for sexual abuse.

In order to create the database containing information on the clergy abuse scandals, we used the database collected by Bishop Accountability (bishopaccountability.org) as our main source of information. The database, originally created to provide a public list of U.S. Catholic clergy accused of abusing children and vulnerable adults, supports each case with media reports, legal documents, photos, and assignment records. The database consists of a list of diocesan and religious order priests, brothers, seminarians, deacons and nuns affiliated to the Catholic Church and working in the United States who have faced an allegation (with or without legal action) relating to the sexual abuse of a child, vulnerable adult, or possession of child pornography.⁶ Bishop Accountability only includes an individual in the database if the organization has obtained appropriate documentation. Their documentation usually consists of a copy of a newspaper article from a reputable newspaper or a copy of legal documents filed in court and maintained in a public file (Bishop Accountability claims to double-check every allegation with the cited source document). The total number of Catholic clergy with sexual abuse scandals in the Bishop Accountability database is 3,397 (as of December of 2010)—to be more precise: 3,100 priests, 182 bishops, 42 deacons, and 73 nuns. Note, however, that we cannot compare the counts from Bishop Accountability with those of the Nature and Scope study in a direct fashion, since their

⁶A child is defined as a person under 18 at the time the alleged offense occurred, although they give place to a state-specific definition of sexual abuse when questions arise. Incidents of alleged sexual abuse of adults, murder, theft, drug use, or other crimes are not included. Alleged acts of sexual abuse or possession of child pornography by lay teachers, church volunteers, church administrators, or other diocesan or religious order employees are excluded. If an individual is "cleared" or "exonerated" by an internal church investigation, or if it is returned to ministry, or if a criminal investigation dismissed the case because the alleged offense is beyond the statute of limitations, the individual remains in the database as long as the victim has not withdrawn the allegation. If an individual is found not guilty or not liable after a trial, but other victims have come forward with allegations, the individual is listed in the database. If the individual faces an allegation for an act which occurred after the individual has left the church, the individual is listed in the database. If the individual was visiting from another country and faces an allegation, the individual is listed in the database.

definitions differ in many dimensions.⁷

We complemented the information provided by Bishop Accountability with several other data sources. For instance, we compared the list of newspaper articles with databases of historical newspapers (e.g. LexisNexis Academic). We used the *Official Catholic Directory*, the most authoritative historical resource available today on the Catholic Church, to obtain information on the appointment of the accused at the time of the scandal and at the time when the abuse took place. We crosschecked the addresses of the organizations involved in each scandal (e.g. parishes) using the *Official Catholic Directory* and a variety of online sources: Google Maps, the official websites of the Catholic institutions and dioceses, and several public-listing websites (e.g. parishesonline.com, yelp.com). We also collected data on many characteristics of the scandals: e.g. characteristics of perpetrator, the victims, the abuse, and the newspapers.

We are interested in the date of the scandals, meaning the date when the allegation becomes public in the news as opposed to the date when the alleged abuse occurred. For the purposes of our paper, we do not need to assume that the allegations were true. The only important fact for the identification strategy is that these scandals impose a discontinuous negative shock to the religiosity in the community of the scandal; indeed, it is far more important that people believe that the accusations are true than having substantiated the accusations. In order to be considered a scandal, we require each event to satisfy several basic conditions. The most important is that the allegation (or group of allegations) must be public. In almost all cases the publicity of the allegation responds to one or (most often) multiple newspaper articles. Other, alternative forms of publicity, such as appearances in TV news, may also be included. We define the date of a scandal as the first newspaper article (or other public event) covering an allegation of sexual abuse by a given priest in a given Catholic institution (e.g. parish, school, hospital). We only consider articles from newspapers whose circulation area specifically includes the location of the scandal (i.e., local, statewide, or nation-wide newspapers). The Bishop Accountability list includes many priests who do not appear in newspapers articles but only in long lists of clergyman with allegations provided by the diocese, which does not meet our definition of a scandal event.

Despite our having matched the data provided by Bishop Accountability with the databases of historical newspaper articles, the possibility remains that, in a few cases, we do not observe one or more newspaper articles released some months earlier than the ones that we do observe. In addition, some scandals may have become public a little earlier than the date of the first news article. For example, we identified some cases in which a “quasi-public” event occurred prior to the publishing of the first newspaper article, as in the case of the police interrogating the priest one month before the news article appeared. We do not use the dates of the “quasi-public” events because of the difficulty in determining the level of publicity of each of these earlier events. In any case, we conduct our analysis using data aggregated by year, so the measurement error of the order of few months should not make a significant difference.

⁷If anything, the count of Bishop Accountability seems to fall short from the 5,700 priests identified by the Nature and Scope study during the period 1950-2009. The priests included in the Nature and Scope study but not in the Bishop Accountability list are probably the ones where the accusations had no legal or mediatic repercussions. Since we only care about the abuses that had repercussion in the media, and since media documents were considered “sufficient” information to be included in the Bishop Accountability database, we are confident that our data gives a fairly complete account of the scandals across the US. Furthermore, all the econometric exercises rely on data post-1990, a period where we are most confident about the completeness of the bishop-accountability list.

A few scandals include multiple accusations of a given priest in the same parish, but separated by some years. We record those accusations as a new scandal event if more than 5 years separate the two events. In a similar fashion, some scandal events relate to a given individual with accusations in two locations geographically close to each other. These cases typically include a priest who accused of abusing children both in a parish and in the Catholic school next to that parish. We record those events as a single scandal, and use the address of the place where the first accusation took place as the location (since the empirical analysis uses zip code-level data, it will not make a difference which address we use). In addition, note that if there are accusations to multiple clergymen in a given location, then each of them will count as a separate scandal event, even if the accusations are all from the same victim. We decided to exclude some cases from the definition of a scandal. First, we do not include cases in which the priest is only accused of having pornography (39 clergymen). We also exclude those events wherein the accused priest works for an organization affiliated with a non-Catholic religious denomination or in the Catholic judiciary system (10 clergymen). Since these are very few observations, this paper's results remain virtually identical if we do not make these exclusions.

An allegation can affect a Catholic institution in two ways. A priest who is currently working in the institution may be accused, even if he committed the abuse in some other institution. On the other hand, a priest may be accused of having abused while working in the institution in the past, even if he is not working at that same institution at the time the scandal breaks. Therefore, we consider two types of scandals:

Type-I scandal: The location is given by the address of the institution where the priest is working when he is first accused of committing sexual abuse (if working). The date of this scandal is given by the date of the first article mentioning an abuse committed by this priest from a newspaper with a circulation that reaches the place of the scandal.

Type-II scandal: The location is given by the address of the institution where a priest is accused of having committed the abuse. The assigned date uses the date of the first article mentioning the abuse in that location from a newspaper with a circulation that reaches the place of the abuse.

For instance, consider a priest abused in 1975 during his appointment in a parish in town A, and abused again in 1982 during an appointment in town B. This priest is publicly accused for the first time in March of 1997, during his appointment in town C, for having abused in town A. Later in 1999, once defrocked, he is accused for his abuse in town B. The priest will have one Type-I scandal in 1997 in town C, and two Type-II scandals: one in 1997 in town A and one in 1999 in town B. Many individuals have no Type-I scandals, often because they were either retired or deceased at the time of the scandal (a likely situation, given the average time gap between the abuses and the accusations).⁸

Since they are different in nature, the two types of scandals could have effects of a different magnitude. Therefore, we include the two types of scandals as separate variables. In addition, we believe that the data on Type-I scandals is of a higher quality. In any case, the effects of Type-I and Type-II scandals have consistently remained qualitatively similar, which provides reassurance about the robustness of the results. Some individuals have no Type-II scandals, usually because the accusations

⁸In some cases, the priest was forced to retire a couple of months before the scandal, most likely because the Church had private information on an accusation and was reacting to that. Whenever this is the case, we consider the priest as working at that location, so it counts as a Type-I scandal. Note that in some cases the institution of the Type-II scandal is closed at the time the scandal becomes public, but it will still count as a Type-II scandal.

do not have media repercussion in the place where the abuse took place.⁹ Out of the 3,407 clergy in the Bishop Accountability database, 1,285 (38%) have neither Type-I nor Type-II scandals. Of the remaining 2,122 individuals, 458 (22%) have only Type-I scandals, 995 (47%) have only Type-II scandals, and 669 (31%) have both Type-I and Type-II scandals.¹⁰ These amount to 1,125 Type-I events and 1,899 Type-II events, totaling 3,024 events during the 1980–2010 period.¹¹

Our identification strategy relies on an event-study analysis, by testing whether the evolution of the dependent variable before the date of the scandal is any different between those communities that suffer a scandal and those communities that do not. We show this to be the case. Intuitively, it means that the timing of the scandals appears random. If that were not the case, then it would be difficult (or even impossible) to ascertain causal effects from the event-study analysis. We believe at least two institutional factors contribute to this finding. First, the 2002 *Boston Globe* article imposed an exogenous force that triggered a substantial number of the scandals in the database. Figure 1.a shows the distribution of scandals over the years, where the outbreak of scandals following the 2002 *Boston Globe* article stands out. Figure 1.b shows the monthly distribution of new scandals during the 2001–2004 period, giving a closer look at the events of 2002.

The second institutional factor is the large time lapse between the perpetration of the abuses and the surfacing of the accusations. The number of abuse incidents per year increased steadily from the mid-1960s through the late 1970s, then declined in the 1980s and has remained low since then. Prior to 1985, usually the parents of the abused youths made the allegations soon after the incident took place. However, after 1985, mostly the victims themselves made the accusations of abuse, often decades after the date of the incident (JJCCR, 2011). While most of the scandals took place in the 2000s, most of the abuses took place before 1985. Thus, the incidents often became public one or more decades after they happened, with some reports describing incidents that happened thirty to forty years earlier.

Figure 2 shows the geographic distribution of the scandals across the territory of the United States at four different points in time (for a video go to people.fas.harvard.edu/~rtruglia/), where a darker color represents a higher density of Catholic adherents per square mile in the state (in log scale). Although the map does not show Alaska and Hawaii, these states do appear in the database. We can analyze the characteristics of a geographic area that predicts the number of scandals that will happen in that area. Appendix A presents this analysis. The results suggest that, once we control for the size of the Catholic congregation in the county, additional characteristics (e.g. income, racial composition) have no explanatory power to predict the number of scandals. This finding is not relevant for the internal validity of the results, since we assess such validity more directly by means of an event study. Nonetheless, this provides reassurance on the external validity of our results, suggesting that, if we take the congregations that by chance did not have a scandal and “treat” them with one, we should

⁹By definition, a clergyman can have no more than one Type-I scandal. In virtually all cases the priest does not continue working after the Type-I scandal. If we generalize the definition of Type-I scandal to comprise more than the first accusation, only a handful of priests have a second Type-I scandal. We do not consider those cases because those priests started or continued working in those locations in spite of the pre-existing allegations, so the effect of the marginal allegation is debatable.

¹⁰Note that often a priest is accused of having abused in parish A while he is working in that same parish, so that will generate both a Type-I scandal and Type-II scandal in the same location. In the regressions we include Type-I and Type-II scandals in a separate basis. Notwithstanding, the results are robust if we use alternative criteria to bundle the scandals: e.g. if we include Type-I scandals along with scandals that are “Type-II but not Type-I.”

¹¹Not all of these events will appear in all the applications below, as some datasets may not cover some geographic areas and/or time periods.

find an effect that is qualitatively similar.

Appendix A also discusses many of the details associated to the zip code-level data. For example, it details the creation of the database that identifies which zip codes are adjacent to a given zip code. We also present a comparison of characteristics at the zip code level by classifying them into three groups: those with a scandal, those without a scandal but are located adjacent to a zip code with a scandal, and those that do not have nor are located adjacent to one with a scandal. Scandals tend to occur mechanically in more populated areas, because of the larger pool of people who can be abusers or victims. The differences in other characteristics are small in magnitude. All the regressions in the paper include zip code fixed effects, so we control for any observable and unobservable differences between zip codes with scandals (i.e. treatment group) and without scandals (i.e. control group). Furthermore, we will always control for the interaction between the time effects and the logarithm of population, logarithm of land area, and share of urban population, all taken from the 1990 U.S. Population Census. This will account for any difference in the evolution over time of a given dependent variable associated with differences for zip codes with these characteristics. We find the results to be the same in practice when we also include the interaction between the time effects and other zip code characteristics (e.g. racial composition).

3 Effect of the scandals on religious congregations

This section demonstrates how the local communities affected by the Catholic-clergy sexual abuse scandals suffered a sharp permanent decline in religious adherence, religious participation, and religious beliefs. We are not going to disentangle the specific channels through which the negative shock to Catholicism operated, but we can briefly enumerate what these may be.

The scandals can change participation in religious congregations by making it less attractive. Ever since 2002, many nationally representative polls have asked Catholics about their reaction to the scandals. The results indicate that not only are Catholics well aware of the scandals¹² but also that Catholics report that the scandals affected their relationship with the congregation.¹³ Members of the congregations may stop attending services because they perceive that the priests are a danger to their children.¹⁴ A scandal can generate feelings of betrayal and spite among members of the congregation, which will drive followers away. Much like in non-religious congregations, one of the reasons why people value their membership in religious congregations is because of the signal that it sends the rest of society. By damaging the image of the Catholic community, the scandals make it less attractive to stay or become part of that community. In addition, the religious services and other religious activities in the congregation shape deep religious beliefs. Thus, the scandals can have an indirect effect on religious beliefs through the effect on religious participation, in particular when those who have stopped participating in the Catholic congregation do not join other religious congregations.

¹²For instance, 84% (49%) of Catholics paid at least “some attention” (“a great deal of attention”) to the issue of sexual abuse in the Catholic Church; and 82% (35%) of Catholics would be at least “somewhat interested” (“strongly interested”) in a story about clergy sexual abuse in the news on television, in print, or on the Internet (Gray et al., 2006).

¹³For example, when asked how much has the issue of sexual abuse by priests hurt the credibility of Church leaders, 40% of Catholics responded “A great deal,” 40% responded “Somewhat,” 15% responded “Only a little” and only 5% responded “Not at all” (FADICA, 2002).

¹⁴This mechanism may entail some kind of irrationality by the agents (e.g. over-reaction to the scandals), since there is evidence that the abuse incidence rates have been very low since the late 1990s (JJCCJ, 2011).

The scandals could also affect the supply of religious activities: e.g. by affecting the number of priests available, or by increasing financial stress through lawsuits and other abuse-related costs. However, because of the centralized structure of the Catholic Church, there is a high likelihood that most of these supply-side shocks are smoothed at the diocese level. For example, according to the *Official Catholic Directory* there were 171 Catholic Dioceses and Archdioceses in the year 2000, which had an average of 108 parishes each.¹⁵ Since plaintiffs address the lawsuits to the diocese, the negative income shock from a marginal lawsuit would be shared across the other 100 parishes in the diocese, and not by the specific parish mentioned in the lawsuit.¹⁶ In a similar manner, a defrocked priest can be replaced by another priest from the diocese’s pool of priests or even by a priest from a different diocese. Since we measure the effect of the scandals at a much finer geographical level—the zip code level, our estimates would not capture these supply-side effects.

In subsection 3.1, we use data on the number of Catholic schools and enrollment in Catholic schools as proxies for the size of the Catholic community in a given area, and examine how the scandals affected those measures. In subsection 3.2, we use look at alternative measures of religious belonging, using survey data from the General Social Survey on religious affiliation, frequency of church attendance, frequency of prayer, and so forth. When compared to each other, the datasets from subsections 3.1 and 3.2 have advantages and disadvantages, so the evidence from one subsection complements the evidence from the other. Moreover, the main findings remain consistent with the evidence presented in Appendix B, which explores the effect of the scandals on the number of religious establishments at the zip code level.

3.1 The effect of the scandals on enrollment in Catholic schools

Catholic schools comprise the largest non-public school system in the United States, and serve as one of the landmarks of the Catholic community in the United States. The share of non-Catholic enrollment in Catholic schools is relatively small: e.g. for the school year 2010-2011 the National Catholic Educational Association (NCEA) reports a 15% share of non-Catholic students in Catholic schools.¹⁷ Thus, the number and size of Catholic schools can serve as a proxy for the size of the Catholic community in a given area. In this subsection, we will look at the effects of the scandals on the number of Catholic schools, and on the number of students enrolled in Catholic schools. We recognize that this is not the best measure of Catholic adherence. For instance, measuring the number of people who go to Catholic mass would seem more appropriate. Indeed, in the following subsection we analyze a different data source that contains measure of identification with religious denominations, religious participation, and beliefs. Appendix B presents the analysis of zip code-level data on the number of religious establishments. However, the data on Catholic schools has at least two important advantages with respect to the other data sources. First, enrollment in Catholic schools serves as a

¹⁵And a median of 90, a minimum of 24 and a maximum of 412.

¹⁶In the case of non-diocesan clergy, the lawsuits seem to be against the order and not against a particular parish.

¹⁷In spite of the big changes in the number of schools and enrollment, this share has changed very little over the last couple of decades. For example, in the 2000-2001 school year, before the big outburst in scandals, the share of non-Catholic enrollment was 13.6% (NCEA, 2001). The religious composition of students may have changed as a response to the scandals. However, since the estimates from this section are very large in magnitude, this is not a major concern. In any case, if there was an increase in the share of non-Catholic students, it would mean that we are under-estimating the effect of the scandals on the number of Catholic students.

behavioral measure of affiliation to Catholicism, so it is not subject to some usual caveats common to self-reported data. Second, such administrative data for the universe of all public and private schools in the United States covering almost two decades translates into more precise estimates—particularly valuable in an event study.

For data on private schools, of which religious and Catholic schools are subsets, we use the Private School Survey (PSS). The U.S. Department of Education’s National Center for Education Statistics (NCES) has biannually conducted this census of private elementary and secondary schools in the United States since the 1989–1990 school year, and it is available until the 2007–2008 school year. The target population for the survey consists of all private schools in the United States meeting the NCES definition. The PSS consists of a single survey completed by administrative personnel in private schools; it includes such information as name, address, religious orientation, enrollment, number of teachers, level of school, and so forth. In order to obtain data on public schools, we also employ the Common Core of Data (CCD). Also administered by the NCES since the school year 1986–1987, it collects annual data about virtually all public schools in the United States. State-level education-agency officials supply the data, which includes information on school characteristics almost identical to those in the PSS (indeed, the design of the PSS was based on the CCD). Dills and Hernández-Julian (2011) use this same PSS dataset to examine the correlation between a measure of the number of allegations of sexual abuse by Catholic-clergy in a diocese and the enrollment in Catholic schools at that diocese. We are able to achieve a much more precise identification of the effect of the scandals by introducing the rich dataset on the geographic and temporal distribution of the scandals. Also, while their study examines the correlation between allegations and enrollment, we provide strong causal evidence by means of the event-study analysis and the additional falsification tests.

In virtually all cases, parents make enrollment decisions prior to the start of the school year, and, in most cases, they make these decisions more than six months before classes start. For example, we would expect a scandal that took place in November of 1994 to affect the number of schools and enrollment no earlier than the 1995–1996 school year. For the sake of simplicity, we are going to define the number of schools and students during the *calendar year* t as the number of schools and students as of the *school year* from September 1 of $t + 1$ to August 31 of $t + 2$. In this way, we can interpret the coefficients on Post-Scandal exactly as we do in the rest of the paper. Another important detail about the data is that we observe the zip code of the schools where the students attend, and not the zip code in which those students live. Because of the geographic distribution of Catholic schools, a majority of students in Catholic schools live in a different zip code than that of the school they attend. Since many Catholics go to Catholic schools in neighboring zip codes, we should expect that even if a scandal affected Catholicism in that zip code only, it would still significantly affect the enrollment in Catholic schools in neighboring zip codes. In order to consider this fact, instead of the stock of scandals in the same zip code we are going to look at the stock of scandals in the same zip code and adjacent zip codes.

The variable $\text{Post-Scandal}_{z,t}$ is the “stock” of Type-I scandals in zip code z and its adjacent zip codes at time t . Thus, if a zip code had its first and only scandal in year t_0 , Post-Scandal for that zip code will take the value 0 for $t < t_0$ and the value 1 for $t \geq t_0$. If the effects of the scandals are permanent and stable over time, then the coefficient on Post-Scandal will identify such effects, just

like in the typical difference-in-difference framework., The problem with regressing a given dependent variable on Post-Scandal is that the coefficient on the latter may simply reflect that the zip codes with scandals had different pre-trends than those without scandals. We define the variable Pre-Scandal as the number of scandals in zip code z during years $t + 1$ and $t + 2$. If the coefficient on Post-Scandal simply captured differential pre-trends between the control and treatment groups, then the coefficient on Pre-Scandal should also pick up those differential pre-trends. Thus, if the coefficient on Pre-Scandal is not significant, that would be evidence that the coefficient on Post-Scandal is not driven by differences in pre-trends.

Table 2 shows descriptive statistics about the data. The sample includes all zip codes that had at least one school (public or private) thorough the years in the sample. Those 22,727 zip codes include an average of 0.34 Catholic schools and 4.43 non-Catholic schools. Table 2 shows the difference-in-difference estimates, where the dependent variables are the number of Catholic schools and the enrollment in Catholic schools, and the independent variables are those that describe the timing of the scandals. The regressions are linear regressions with heteroskedastic-robust standard errors clustered at the zip code level, and they include control variables such as zip code fixed effects, time effects, and the interaction between the time effects and the logarithm of population in the zip code, the logarithm of land area and the share of urban population. The coefficient on Post-Scandal from column (1) suggests that, on expectation, a scandal permanently decreases the number of Catholic schools in the same and adjacent zip codes by 0.044. Column (2) adds the variable Pre-Scandal, whose coefficient is not statistically significant. This suggests that there are no differential pre-trends between zip codes affected by a scandal and unaffected zip codes.

In addition, we perform the typical analysis shown in event studies. We want to see how the dependent variable changes one year after the scandals become public, two years after they become public, ..., and one year before they become public, two years before they become public, and so forth. For that, we created a set of variables $d_{z,t}^s$ that, for each integer s , zip code z and time t , takes the value of the number of scandals in zip code z and adjacent zip codes at time $t + s$. As the omitted category, we choose $d_{z,t}^{1,2}$: i.e. the variable that captures the effect of a scandal two years before it becomes public. First, we expect the coefficients on $d_{z,t}^s$ with $s < 0$ to have a negative coefficient, meaning that Catholic adherence decreases after a scandal in the same or adjacent zip code. Second, we expect the coefficients on $d_{z,t}^s$ with $s > 0$ not to be statistically different from zero, meaning that, before the date when the scandal becomes public, Catholic adherence evolves similarly in zip codes with and without scandals. As is typical in event studies, we use the point estimates and the 95% confidence intervals from the coefficients $d_{z,t}^s$'s to present the results in graphical form. Figure 3.a (b) corresponds to a regression of the number of Catholic schools (students enrolled in Catholic Schools) a zip code on the set of $d_{z,t}^s$'s and a set of control variables: zip code fixed effects, time effects, and so forth. Figure 3.a (b) confirms that the effect on the number of Catholic schools (enrollment in Catholic schools) starts exactly when the scandal becomes public.

In addition, Figure 3 suggests that the effects of the scandals intensify over time. Indeed, we will always observe a pattern like this. There are multiple explanations for this finding. Most important, the date of the scandal is the date of the first newspaper article, but that does not mean that the scandal is a "binary" event (i.e. there is a scandal or there is not). Although the first news article represents the

transition from “no scandal” to “some scandal,” the years following the scandal can include events that usually intensify the scandal’s severity. Some examples of the events following to the first news article include further allegations, a trial, further proof about the accusation, a conviction, defrocking of the priest, and so forth. Because of this, one may naturally expect that the effect of the scandal intensifies during the first 1 to 5 years. Nevertheless, other, more mechanical, reasons explain this pattern of intensification: e.g. Catholic organizations, much like all organizations, will strive to survive for as long as possible, so they may successfully delay an imminent closure at least for a couple of years.

Column (3) of Table 2 adds a variable that counts the stock of scandals that took place in zip codes separated by two degrees—zip codes adjacent to those adjacent to a zip code with a scandal. The coefficient on this variable should be smaller than the former coefficient on Post-Scandal, or even zero, since we expect the effect of the scandals to fall per the distance to the epicenter. This is indeed the case: the coefficient is statistically significant at the 5% level, but small in magnitude.¹⁸ Column (4) simultaneously introduces the Type-I and Type-II scandals. Both types of scandals have a significant negative effect on the number and size of Catholic schools. Note that since 60% of the Type-I scandals are also Type-II scandals, we cannot compare the coefficient on “Post-Scandal (Type-I)” in column (4) to those in columns (1) through (3). If anything, the coefficients on “Post-Scandal (Type-I)” in columns (1) to (3) should be compared to the coefficient on “Post-Scandal (Type-I)” from column (4), plus 0.6 times the coefficient on “Post-Scandal (Type-II)” from column (4). Using this formula, the results in column (4) fall perfectly in line with those in columns (1) to (3). Given this, when a priest now working in town A is publicly accused of having abused 20 years ago when working in town B (often in a different state than A), the number of schools goes down not only in town A but also in town B. According to column (4), the magnitude of the effect of the Type-II scandals seems to be roughly 80% of the effect of Type-I scandals.

In order to explore whether there were any (positive or negative) spillovers to other religious denominations, we use the number of non-Catholic religious schools as dependent variable. If former Catholics participate in other religious denominations, then we should see an increase in enrollment in non-Catholic religious schools, insofar it serves as a proxy for the size of non-Catholic religious congregations in the area. The scandals have no effect on schools affiliated with non-Catholic religious denominations. This remains consistent with the findings in other subsections of the paper, where we find the scandals did not affect the adherence to non-Catholic denominations. We also look at the effect on the number of non-Catholic schools, religious or not. A positive coefficient on Post-Scandal would mean that at least some of the students who leave Catholic schools due to the scandals attend non-Catholic schools from the same zip code; and a null coefficient on Post-Scandal would mean that all those students are moving to schools in different zip codes. The coefficient on Post-Scandal is positive, but not statistically significant.

Apart from the effect on the number of catholic schools, we can look at the effect on enrollment in Catholic schools. On the one hand, some of the students leaving Catholic schools that are closing could attend other Catholic schools in the area, so the effect of the scandals on the number of Catholic schools

¹⁸The results are similar if instead of this variable we introduce a variable that counts the number of scandals in the county. Additionally, notice that we could split the Post-Scandal variable in two variables: one for scandals in the same zip code, and other for scandals in adjacent zip codes. Even though the point estimate for the scandals in the same zip code is a slightly larger, the difference between the two is not statistically significant, which explains our decision of combining those two variables into a single variable.

would over-estimate the true impact of the scandals. On the other hand, there could be a net outflow of students from the Catholic schools that remained open, so the effect of the scandals on the number of schools would underestimate the true impact of the scandals. Column (7) of Table 2 reproduces the regression from column (2) but uses enrollment instead of number of schools as the dependent variable. In a similar manner, Figure 3.b reproduces Figure 3.a but uses enrollment instead of number of schools. In both cases, the results are qualitatively similar: e.g. the coefficient on Post-Scandal from column (2) is equivalent to 12% of the mean of the respective dependent variable, while the coefficient on Post-Scandal from column (7) suggests a decrease of 15% of the mean of the respective dependent variable.

The post-2002 scandal crisis coincided with a period of sharp decline of the Catholic school system in the U.S.: according to the Annual Statistical Report by the NCEA, between the 2000 and the 2011 school years, 1,755 Catholic schools were reported closed or consolidated (21.5%), and the number of students in Catholic schools declined by 587,166 (22.1%). We can assess to what degree the scandals can explain the sharp decline in the Catholic school system. According to our data, the number of Catholic schools decreased from 7,688 in the 1999–2000 school year to 6,695 schools in the 2007–2008 school year (the most recent year available in the PSS), so there was a net decrease of 993 Catholic schools. We can use the coefficients on Post-Scandal from column (4) of Table 2 to calculate how different the total number of Catholic schools would be if all of the Type-I and Type-II scandals that happened in that period never would have occurred. Our estimate is 456 Catholic schools, or 46% of the net decline of 993 schools during that period. The picture is very similar if we instead look at the number of students enrolled in Catholic schools. The data indicates that the number of students in Catholic schools fell from 2,340,847 in the 1999–2000 school year to 1,930,693 in the 2007–2008 school year. We can attribute the decline of 169,150 students in Catholic schools, or equivalently 41% of net decline of 410,154 students, to the Type-I and Type-II scandals that occurred during that period. Thus, this strongly supports the conclusion that the Catholic-clergy sexual abuse scandals played a key role during the sharp decline of the U.S. Catholic school system.

3.2 The effect of the scandals on self-reported religious affiliation

In the previous subsection, we presented strong evidence about how the scandals affected enrollment in Catholic schools. This measure of Catholic affiliation has some advantages, like the fact that it is a behavioral measure. However, we would like to check the robustness of the results using a more conventional measure of religious affiliation: self-reported survey data. In addition, we want to test whether the people that abandoned their Catholic affiliation because of the scandals were not just nominal Catholics (e.g. Catholics that did not go to church). However, the data on religious affiliation, participation, and beliefs is of a lower quality than the data on school enrollment. Although the estimates are not as precise, when taken together with the findings from the previous subsection, the evidence from this subsection confirms that the scandals led to a permanent decline in the size of the Catholic congregation.

We use data from the General Social Survey (GSS). Since 1974, this survey has interviewed about 1,500 individuals each year from a nationally representative sample, and it is widely regarded as the single best source of data on societal trends. The county identifiers for the respondents are available

since 1994, when the survey became biannual, until the last available round in 2010. This sample covers 25,604 individual observations and 333 counties. Between 1994 and 2010, 173 out of the 333 counties covered by the sample had at least one Type-I scandal, totaling 442 scandals. The county-level data does presents some challenges, however. As will be discussed below, the large size of some counties makes it a less straightforward process to measure the kind of “local” effects that we want to measure. In addition, only 333 counties make up the entire sample, and the smaller half of those counties (which happen to measure the local effects in a more direct fashion) have included a few interviews and often a few observations within each of those years. Although the variation in the sample is enough to run many sharp falsification tests, we can run some falsification tests only with the zip code-level data and not the county-level data; given this, we cannot explore the differential effects of Type-I and Type-II scandals.

Counties in the United States significantly differ in terms of population. For example, among the 333 counties in this sample, the 1990 population figures range from a minimum of 4,802 inhabitants to a maximum of 8,863,164 inhabitants, with a median of 187,768 inhabitants. The smallest counties have populations so small that the scandals can affect the entire population of the county. However, the size of the largest counties can be several times the size of those “local” communities, so one scandal in a big county will only affect a small share of that county’s entire population. This issue does not arise in the rest of the econometric exercises in this paper, since the data is always at a more disaggregated geographical level (zip code level). In order to account for the local nature of the scandals with the county-level data, we need to weight the data on scandals properly. Dividing the number of scandals by the number of inhabitants in the county serves as the most straightforward method. To understand this weighting, imagine that we have a “small” county and that a scandal hits that county, affecting its entire population. Next, we create a second county by pasting together two communities, each of them equal in size to the first county. When a scandal hits this second county, it is only going to affect half of the population in the county (i.e. one of the two communities). Thus, the effect of one scandal in the second county should be half as large as the effect of one scandal in the first county. The division by the county population implies exactly this effect. Although the qualitative results are robust if we use different ways to weight the scandals—or even no weighting at all, dividing by the population size seems to yield estimates that are more precise.

The definition of the variables Post-Scandal and Pre-Scandal remains the same as before, except that we use counties instead of zip codes, and we divide those variables by the county population (in 100,000s) as of 1990. For example, if a county with 100,000 inhabitants had its first and only scandal in year t_0 , Post-Scandal in that county will take the value 0 for $t < t_0$ and the value 1 for $t \geq t_0$. The interpretation of Post-Scandal and Pre-scandal remains the same as those explained earlier. We always use linear regressions with heteroskedastic-robust standard errors clustered at the county level, and include county-specific fixed effects, time effects, and a group of additional control variables: gender, age, age squared, dummies for black and white, three dummies for marital status, household income, and a set of four dummies for employment status, education, and number of children. The results are robust if we use nonlinear models instead of linear models and are robust to the inclusion of additional control variables.

We will employ some measures of religious affiliation as dependent variables. We created a dummy

variable that indicates whether the respondent reported to be Catholic, and another dummy variable for whether the respondent reported adhering to a non-Catholic religious denomination. The omitted category (i.e. the dummy variable that takes the value 1 if and only if the two former dummies are zero) would indicate whether the respondent is atheist or agnostic. By construction, the effect of a regressor on the latter dependent variable is equal to (minus) the sum of the effects on the first two dependent variables. For the sake of presentation, we will omit these (redundant) coefficients. Table 3 displays the descriptive statistics for the main variables employed in this subsection. Around 24% of the respondents are Catholics, 61% adhere to non-Catholic religious denominations, and the remaining 15% are atheists or agnostic.

Table 4 shows the difference-in-difference estimates of the effect of scandals on religious adherence. The coefficient on Post-Scandal in column (1) suggests that if a county with 100,000 inhabitants is affected by a scandal, it would suffer a permanent decrease of 6 percentage points in the share of Catholic adherents (note that the GSS data includes adults only). This coefficient is both statistically and economically very significant. Hungerman (2011) also uses GSS data to look at how the scandals affected religious adherence, and our results are consistent with his. However, there are many important differences between the two approaches. One difference is that he uses a measure of the state-level number of allegations, while we focus on a much geographically localized definition of scandals. This is very important for the purposes of this paper, insofar in the following section we will also be looking at localized effects. Also, Hungerman (2011) looks at the correlation between the current number of allegations and current religious adherence, while we try to identify the permanent effect of the scandals. Last, the richer nature of our dataset allows us to make many falsifications tests. Notwithstanding, our evidence is by no means a substitute to the evidence presented in Hungerman (2011). On the contrary, we believe the evidence in the two papers are strongly complementary to each other, where one focuses on more geographically localized effects and the other on less localized effects.

One interesting question is whether the individuals who left Catholicism joined a non-Catholic religious congregation afterwards. According to the coefficient on Post-Scandal in column (2), the scandals did not affect significantly the share of adherents to non-Catholic denominations.^{19,20} By construction, that means that the 6 percentage point decrease in the share of Catholic adherents translates into a 6 percentage point increase in the share of non-adherents (i.e., atheists and agnostics). Indeed, the result that the scandals did not affect non-Catholic denominations is consistent with the findings in other subsections of the paper. In subsection 3.1, we reveal that the scandals did not affect enrollment in non-Catholic religious schools, and, in subsection 4.2, we demonstrate that the scandals did not affect charitable contributions to non-Catholic religious-related charities.

The rest of the specifications in Table 4 offer some sharp falsification tests. One may speculate that affiliation to the Catholic church may have been in decline in the counties affected by the scandals

¹⁹Since the coefficient on Post-Scandal is not extremely precise, we cannot reject the hypothesis that the scandals had a small (positive or negative) effect on non-Catholic religious adherence. Also, note that we are identifying a very localized effect of the scandals. For example, the national trend in the number of scandals may make some people less attracted to Catholicism, independently of whether their own community was affected by a scandal, which is not captured by our estimates. Indeed, Hungerman (2011) presents evidence that the number of allegations at the state-level is positively associated to membership and contributions to other religious denominations.

²⁰According to the evidence on religious conversion (e.g. Barro, Hwang and McCleary, 2010), Catholics are relatively more likely to convert to Protestant denominations than to non-Protestant denominations. In light of this evidence, we replaced the dummy for non-Catholic religion by a dummy for Protestant religion, and the results were practically the same.

even before they became public, so the coefficient on Post-Scandal could just be capturing differential pre-trends between treatment and control groups. If this is true, then the number of scandals in the following two years, Pre-Scandal, should be able to predict changes in religious adherence. Column (3) includes Pre-Scandal in the regression. The coefficient on Post-Scandal is virtually the same than that of column (1), and coefficient on Pre-Scandal is zero and not statistically significant. As an additional falsification test, we exploit a question in the GSS that asked respondents about the religion in which they were raised. We should expect the Post-Scandal variable not to have any effect on religious adherence while growing up. We created a dummy for whether the respondent was raised Catholic and a second dummy for whether she was raised in a non-Catholic denomination. As expected, column (5) shows that the Post-Scandal variable does not have a significant effect on religious adherence while growing up.

We also perform the typical analysis performed in event-studies. We created a set of variables $d_{c,t}^s$ that, for each integer s , county c and time t , takes the value of the number of scandals in county c at time $t+s$ (divided by 100,000 inhabitants in the county as of 1990). We run a regression of the dummy for Catholic affiliation on the set of $d_{c,t}^s$'s and a set of control variables similar to the one presented before: county fixed effects, time effects, and so forth. Figure 4 shows the coefficients on the $d_{c,t}^s$'s in graphical form. As expected, none of the coefficients with $s > 0$ is statistically significant. Right after a scandal becomes public, there is a permanent decrease in Catholic adherence of around 5 percentage points. The first three coefficients with $s < 0$ are negative and statistically significant. Although the last two coefficients with $s < 0$ are not statistically significant, the magnitude of the point estimates is about the same as the magnitude of the first three coefficients. The loose confidence intervals in those two coefficients is probably due to the sample frame. Since most of the scandals covered in the GSS sample happened after 2001 and the sample ends in 2010, there is relatively less variation to identify those parameters. This figure suggests that the effects of the scandals on Catholic adherence were permanent and relatively stable, which validates the use of our difference-in-difference specifications in Table 4.

We can provide a general picture of the importance of the Catholic-clergy sexual abuse scandals in shaping the religious landscape in the United States. According to the estimates presented before, a scandal decreases the share of Catholics by around 5 percentage points in a county that had 100,000 inhabitants as of 1990 (the size of the county does not matter for this calculation). Given the population growth in 1990–2010, a county of 100,000 inhabitants in 1990 has approximately 124,000 inhabitants in 2010; and 72% of those inhabitants are over 18 years old (2010 U.S. Population Census). Therefore, the scandal would decrease the number of Catholics in the county as of 2010 by $0.72 \cdot 0.05 \cdot 124,000 = 4,464$. According to our database, there were 1,125 Type-I scandals over the last 30 years (only 442 of those happen in the 333 counties during the sample period 1994–2010, so this exercise relies heavily on the extrapolation of the results). The scandals are responsible for a permanent decline of $1,125 \cdot 4,464 = 5,022,000$ Catholic adherents. According to the GSS data, 24% of adults were Catholic in 2010. Since the current U.S. population is around 300 million, and the share of people over 18 years is around 72%, the current count of Catholic adherents is $0.24 \cdot 0.72 \cdot 300,000,000 = 51,840,000$. Thus, if the 1,125 Type-I scandals never occurred, the number of Catholic adherents in the United States today would have totaled about 9.7% higher. Moreover, the results suggest that this is a

permanent decline in Catholic affiliation.

We can also compare this calculation with the effect of the scandals on the enrollment in Catholic schools estimated in the previous subsection. The number of students enrolled in Catholic schools in the 2007–2008 school years was 1,930,693. We estimate that the scandals in the database that occurred prior to that school year resulted in decreasing the number of students in Catholic schools by 208,948. Therefore, if not for the scandals, enrollment in Catholic schools would have been 10.8% higher. The magnitude of this effect is perfectly consistent with the 9.7% effect found on Catholic adherence estimated using the GSS data.

Nonetheless, religious adherence tells only one part of the story. The Catholic adherents who abandoned Catholicism could have largely been “nominal” adherents who did not go to church, believe in God, or pray. In order to test this hypothesis, we construct a measure of affiliation to the Catholic Church weighted by the intensity of the affiliation. Let g_i be the variable measuring the intensity of individual’s i religious life, like the frequency of church attendance, and c_i a dummy variable indicating whether the individual is Catholic. Instead of the dichotomous dependent variable for Catholicism, c_i , we will use $g_i^c = g_i \cdot c_i$ as dependent variable.²¹ For instance, in the case of religious attendance, this variable would take the value 0 if the individual is not Catholic, and, if the individual were Catholic, it would equal his/her frequency of church attendance. Then, if an individual is Catholic but does not go to church at all, he would get the same score as if he were not Catholic. The effects of the scandals on this variable are twofold: part of the effect comes from Catholics who remain affiliated but reduce the intensity of their affiliation, and part of the effect comes from Catholics who drop their affiliation entirely. Just for comparison purposes, we also report the results using the dependent variables $g_i^o = g_i \cdot o_i$, where o_i denotes adherence to non-Catholic religious denominations.

Of great importance to our research, the GSS has dozens of questions about strength of religious affiliation. However, it only asks most of those questions for one year or just a couple of years. There are only two questions about religious life that were asked in all of the waves from our 1994–2010 sample: attendance to religious services and the strength of religious affiliation. It measures religious attendance with the question: “How often do you attend religious services?” Since the responses are recorded in categories, we created a variable that proxies the number of times the respondent goes to religious services during the course of one year: Never (0); Less than once a year (0.5); About once or twice a year (1.5); Several times a year (6); About once a month (12); 2–3 times a month (30); Nearly every week (41); Every week (52); Several times a week (60). The results are robust if we code this variable in a different fashion. We construct the strength of affiliation using the answer to the question “Would you call yourself a strong (*R’s religious preference*) or a not very strong (*R’s religious preference*)?” This variable can take the following values: no religion (0); somewhat strong (1); not very strong (2); and strong (3).²²

The survey did ask a few questions about religion many times during our sample period although not

²¹Note that if we regressed a variable measuring the intensity of religious life on Post-Scandal for the sub-sample of Catholics, the coefficient would suffer from a severe selection bias. Intuitively, those individuals that dropped from Catholicism were likely among the least religious Catholics. Thus, the scandals would have a (selection) effect that increases the average level of religiosity among those that remain Catholic, which can even override the negative causal effect of the scandals. Fortunately, because of the way in which this variable is constructed, if we regress d_i^c on Post-Scandal the corresponding coefficient will not suffer from the selection bias discussed above.

²²Note that even though a Catholic can have zero frequency of church attendance, it cannot score a 0 in strength of affiliation. As a consequence, this variable does not provide the same test than the rest of the variables considered.

every year: frequency of prayer, belief in life after death, and belief in God. While church attendance and strength of affiliation have over 23,000 individual responses each, frequency of prayer has only 15,000 observations, belief in life after death has 18,000 observations, and belief in God has just 10,000 observations. Hence, the estimates for these variables will not be precise, and we must interpret the results with more care. For example, we define frequency of prayer is defined as the answer to the question “About how often do you pray?”: never (0); less than once a week (1); once a week (2); several times a week (3); once a day (4); several times a day (5). We construct the variable belief in life after death using the answer to the question “Do you believe there is a life after death?”: no (0); undecided (1); yes (2). Last, the variable “belief in God” can take one of the following values: I don’t believe in God (0); I don’t know whether there is a God and I don’t believe there is any way to find out (1); I don’t believe in a personal God, but I do believe in a Higher Power of some kind (2); I find myself believing in God some of the time but not at others (3); While I have doubts, I feel that I do believe in God (4); I know God really exists and I have no doubts about it (5).

Table 3 provides some summary statistics about the g_i^c 's (and g_i^o 's). Table 5 shows the difference-in-difference estimates. If the people that abandoned their Catholic affiliation because of the scandals were nominal Catholics (i.e. people that did not go to church, or pray, and so forth), then we should find no effect of Post-Scandal on the g_i^c 's. As before, we include the variable Pre-Scandal along with Post-Scandal in order to show that the results are not the product of differential pre-trends between control and treatment groups. Column (1) shows that there is a significant permanent decrease in local Catholic church attendance after the scandals become public, which amounts to 30% of the mean of the corresponding dependent variable. Column (3) suggests that the scandals imposed a permanent decline in the subjective evaluation of the strength of Catholic affiliation, whose magnitude is around 10% of the mean of the corresponding dependent variable. Column (5) suggests that the scandals imposed a significant permanent decrease in the frequency of prayers by Catholics, which amounts to 19% of the average of the corresponding dependent variable. Column (7) suggests that the scandals imposed a decline in Catholic belief in life after death, whose magnitude is around 10% of the mean of the corresponding dependent variable (although the coefficient on Post-Scandal is not statistically significant at the 10% level). Last, column (9) suggests a decrease in Catholic belief in God of around 18% of the mean of the corresponding dependent variable. Columns (2), (4), (6), (8) and (10) show that, as expected, the scandals did not have a significant effect on the religious life of non-Catholic adherents.

The evidence in this section suggests that the scandals significantly reduced the share of the population affiliated to the Catholic church located in the community, as well as measures of Catholic participation and beliefs. The effects of the scandals are not only significant from a statistical point of view, but also substantial in magnitude. Appendix B offers further evidence on this respect, using administrative data on the number and employee size of religious organizations at the zip code level.

4 The effect of the scandals on pro-social behavior

The singular goal of this section is to demonstrate the permanent decline in pro-social behavior in the aftermath of the scandals. We will leave for the next section the discussion of the precise mechanisms

that can jointly explain the evidence in the two sections (i.e. that both religious participation and pro-social behavior declined in the aftermath of the scandals).

We are going to use three different sources of administrative data to disentangle the effect of the scandals on pro-social behavior. In subsection 4.1, we show that the scandals generate a permanent decline in the mean charitable contributions in the zip code. In subsection 4.2, we show that charitable organizations affiliated to the Catholic Church suffered a permanent drop in contributions, but charities affiliated to other religious denominations did not. Last, we want to examine whether the scandals affected the output of the charitable organizations. In subsection 4.3, we show that the scandals generated a permanent decline in the private provision of welfare, as measured by the number and employee size of the charitable establishments that provide social services in a zip code (e.g. soup kitchens, homeless shelters).

4.1 The effect of the scandals on individual charitable contributions

We use zip code-level data on individual charitable giving prepared by the Statistics of Income Division (SOI) of the Internal Revenue Service (IRS). We base the data on administrative records of individual income tax returns from the IRS's Individual Master File (IMF) system, which includes a record for every Form 1040, 1040A, and 1040EZ filed with the IRS. The data for a given tax year includes returns filed between January 1 and December 31 of the following calendar year. Almost all of these are returns for the corresponding tax year, although they also include a limited number of late-filed returns for previous tax years.²³ Data is available for tax years 1997, 1998, 2001, 2002, 2004, 2005, 2006, and 2007. The data also include variables on adjusted gross income, charitable contributions, and number of dependents, among others. However, the SOI did not collect all the variables during all the tax years.²⁴ Data on charitable contributions was only collected for the tax years 1997, 2002, 2004, 2005, 2006, and 2007. Since that is the dependent variable, we can only use data for those years in the regressions. Because a great deal of the scandals happened during that sample period, we do have substantial variation in the variables of interest for the purpose of estimation.

We define charitable contributions as the amount taxpayers reported as charitable contributions on line 18 of Schedule A from IRS Form 1040 (as of tax year 2001). Taxpayers use Schedule A to report their total itemized deductions after being limited on Form 1040 for high-income taxpayers and, for most taxpayers, when this amount is larger than their standard deduction. These deductions include deductions for medical expenses, state and local taxes, deductible interest expenses, charitable contributions, and other miscellaneous deductions. Taxpayers whose standard deduction exceeds these amounts generally do not file Schedule A. Thus, charitable contributions do not measure the totality of charitable contributions, but only charitable contributions by itemizers. According to data from the Panel Study of Income Dynamics for 2000–2008, around 54% of the households that give to charity are itemizers, and the total contributions by non-itemizers constitute just 33% of the contributions by itemizers. This issue does not threaten the internal validity of the results. Nevertheless, the individuals who itemize are different in some respects from those who do not itemize (e.g. on average they are

²³If a taxpayer filed returns for multiple tax years during a given calendar year, only the most recent return is included.

²⁴When a control variable (e.g. number of dependents) is not available for a year, we assign it the value 0 and we add a dummy variable that takes the value 1 for the missing cases.

richer), so the effect of the scandals on itemizers could be a little different in magnitude than the effect on non-itemizers (e.g. Wilhelm, 2005).

We focus on the subset of zip codes with positive mean charitable contributions in all the years in the sample, which includes almost all zip codes in the database. We detail the descriptive statistics for the key variables in the analysis in Table 6. The mean adjusted gross income is \$44,500, and the mean charitable contribution is \$850 (both in 2008 U.S. dollars). Note that both mean income and mean charitable contributions are higher in zip codes with scandals or those adjacent to scandals, since those zip codes are, on average, more urban and, therefore, richer. Also, note that the data on charitable giving is an annual total, so we do not know precisely when during each the year the contributions occurred. Thus, if a zip code is affected by a scandal in May of 1999, we do not know which charitable contributions for that year were made prior to news of the scandal. In order to ensure that this uncertainty does not contaminate the coefficient on Post-Scandal, all the regressions include an extra variable that takes the value of the number of scandals occurring in that zip code during that year. Thus, the coefficient on Post-Scandal only relies on the comparison between the years before and after the year of the scandal, but not the same year of the scandal. We proceed in the same manner in the following subsections. In practice, it makes virtually no difference whether we include this additional variable or not.

Table 7 shows the difference-in-difference estimates. The dependent variable is the logarithm of mean charitable contributions in the zip code, and the main independent variables are the usual variables that describe the timing of the scandals. As usual, all regressions include zip code fixed effects, time effects and the interaction between the time effects and zip code characteristics (e.g. log of population, share of urban population). The regressions also include state-specific time trends, and some other variables taken from the SOI data (e.g. the log of mean adjusted gross income). The coefficient on Post-Scandal from column (1) suggests a significant permanent drop of roughly 2% in charitable contributions after a scandal becomes public. Column (2) shows that the coefficient on Pre-Scandal is insignificant, which suggests that there were no differences in pre-trends between zip codes with a scandal and zip codes without a scandal.

Figure 5.a illustrates the event-study analysis. As expected, the coefficients preceding the date when the scandal becomes public are all statistically insignificant, while the coefficients after the scandal becomes public are negative and (mostly) significant. In order to improve the statistical power, Figure 5.b presents the same event study, but pools the Type-I and Type-II scandals. The qualitative results are the same. The coefficient on Post-Scandal in columns (1) to (3) from Table 7 is an average of all the coefficients to the right of the date of the scandal shown in Figure 5. Because of the sample frame, where most of the observations are from 2002 to 2007, this unweighted average tends to underestimate the actual magnitude of the permanent effect of the scandals. Figure 5 suggests that the long run effects of the scandals on charitable giving lie somewhere between 3% and 4%. It is important to note that this 4% comprises the effect on charitable contributions for the whole population, not just the subpopulation of Catholics. If the effect of the scandals on charitable giving is limited to Catholic adherents, then the effect of the scandals on the subpopulation of Catholics will be a multiple of 4%. Using data from the Panel Study of Income dynamics, we estimated that, in the zip codes affected by at least one scandal, about one third of adults are Catholic adherents. Thus, if the effect of the scandals

is driven entirely by the behavior of Catholics (as suggested by subsection 3.2), a rough estimate would say that the permanent effect of the scandals on the charitable contributions by Catholics is in the order of 12% (i.e. 4% divided by 1/3).

Column (3) introduces variants to the variables Post-Scandal and Pre-Scandal: instead of counting the scandals in the same zip code, they count the scandals that take place in adjacent zip codes. Thus, the coefficient on “Post-Scandal, Adjacent” measures the permanent change in charitable contributions in a zip code after an adjacent zip code has been affected by a scandal. In an equivalent manner, the coefficient on “Pre-Scandal, Adjacent” measures whether zip codes adjacent to scandals have different pre-trends than zip codes not adjacent to scandals. The results suggest that the effects of the scandals on charitable contributions are concentrated mostly on the same zip code where the scandal takes place. In column (4), we introduce Type-I and Type-II scandals simultaneously. Both types of scandals have a significant effect on charitable donations [recall that we cannot compare the coefficient on “Post-Scandal (Type-I)” in column (4) to that of columns (1) through (3)], and both Pre-Scandal variables are close to zero and not statistically significant. Intuitively, when a priest working in town A is publicly accused of having abused while working in town B (often in a different state than A) 20 years ago, charitable giving goes down both in town A and in town B. The difference-in-difference estimate for the Type-II scandals is slightly over half of the estimate for the Type-I scandals. As a final falsification test, we reproduced the regression in column (2) but used the log of mean income instead of the log of mean charitable contributions as the dependent variable. As expected, the scandals do not have any effect on the average income in the zip code. The results are the same if we use other dependent variables (e.g. number of claimed dependents).

Last, we can give a rough calculation of the magnitude of the total effect of the scandals on charitable contributions in the United States. Pooling together the Type-I and Type-II scandals, the 2,857 events in our database took place in the zip codes covered by the IRS SOI data. The total itemized charitable contributions, as of 2002, in those zip codes totaled \$22.2 billion (all figures are in 2008 dollars). According to Figure 5.b, each of those events permanently decrease charitable contributions by about 3.7% per year, so the total decline in itemized charitable contributions due to the scandals equals about \$800 million per year. If the effect of the scandals on non-itemizers is similar, and since itemized charitable contributions comprise only 75% of the total charitable contributions (according to PSID data for 2002), that would mean that the permanent effect of the scandals on the totality of charitable contributions would be slightly over \$1 billion per year.²⁵

4.2 The effect of the scandals on charitable contributions by religious affiliation

In the previous subsection, we showed that there was a sharp and permanent drop in charitable contributions in the aftermath of the scandals. However, we did not pin down the religious affiliation of the individuals responsible for the decrease in charitable contributions. If the decline of participation in Catholic congregations were responsible for the drop in charitable contributions, we would expect the decline in contributions to be concentrated almost exclusively among Catholic donors. This subsection

²⁵The effect would probably be higher if we could also include the effects of the scandals on the value of volunteered time.

provides evidence on this respect.

We will use the NCCS Core Files (Core), available every year for the period 1989–2009. The National Center for Charitable Statistics (NCCS) produces these using administrative data provided by the IRS.²⁶ The data covers all 501(c)(3) organizations that complied with the requirement to file a Form 990 or Form 990-EZ. The IRS does not keypunch financial data for those organizations that filed Form 990 but were not required to do so because they had less than \$25,000 in gross receipts or are religious congregations. NCCS also excludes a small number of other organizations, such as foreign organizations or those that generally considered part of the government. 501(c)(3) organizations represent the largest part of the nonprofit sector. A minority of these organizations are mutual-benefit organizations. Since those are essentially providers of private services to paying customers, we exclude them from the analysis.

We need to identify a group of charities oriented towards social services to which Catholics would be likely to make cash donations and an equivalent group of charities for other religious denominations. The non-profit organizations in the data are divided into categories according to the National Taxonomy of Exempt Entities (NTEE) system, created and used by the IRS and NCCS to classify nonprofit organizations. We are going to focus on the set of organizations under code X, which agglomerates religion-related organizations. The subcategories within the X code denote the particular denomination or groups of denominations with which the organization identifies: X22 denotes Roman Catholic, X21 denotes Protestant denominations, X20 denotes other Christian denominations, X30 denotes Jewish, X40 denotes Islamic, X50 denotes Buddhist, and X70 denotes Hindu. The rest of the subcategories within the X code agglomerate other minor religion-related organizations (e.g. X82 denotes “religious television”).

Since we are going to include organization-specific fixed effects, an organization will be useful for estimation if and only if it reports charitable contributions at least two years in a row. Out of the 664 organizations in the data under the Roman Catholic NTEE code (X22), 478 (72%) satisfy this condition. We define charitable contributions as the total public support, as defined in line 1d of Form 990.²⁷ For a given return, we define the year to be the start year of a reporting organization’s fiscal year.²⁸ Some organizations report zero contributions during one or more years. Upon inspection of the data and the documentation compiled by the NCCS, we concluded that those zeros are likely due to misreporting rather than actually reflecting no contributions.²⁹ Out of the 478 catholic organizations with at least two years of contributions, we present results for the 388 organizations that always report positive contributions (the qualitative results do not depend on this criterion).

Some of the NTEE codes agglomerate relatively similar organizations: e.g. the NTEE code B25

²⁶It combines descriptive information from the IRS’s Business Master File and financial variables from the IRS’s Return Transaction Files after they have been cleaned by NCCS.

²⁷This is the sum of direct public support (line 1a), indirect public support (line 1b), and government contributions and grants (line 1c). Unfortunately, the data for these separate components is only available for a limited number of organizations and years.

²⁸In more than half of the returns the fiscal year coincides perfectly with the calendar year (i.e. January through December). Of the remainder, the vast majority have fiscal years ending in June.

²⁹For example, an organization reports around \$100,000 in contributions every year for several years, except two random years when it reports exactly \$0. The documentation compiled by NCCS alerts about this source of error: e.g. “a cursory review of the records shows that a prominent university which accounted for a substantial percentage of all assets for reporting public charities in one state in 1993 did not file in 1995, yet the university is alive and well” (Guide to Using NCCS Data, 2006, page 14).

agglomerates only secondary and high schools. The subcategories under code X are much more heterogeneous. Since religious congregations are exempt from filling the IRS form 990, no churches are listed among the organizations under NTEE code X. Whenever an organization fulfills a specific role, like a university or an orphanage, it is classified in the NTEE code of its main role, not in the X NTEE code; therefore, a Catholic High School would be classified as B25, with all the other secondary schools. Among the 388 Catholic organizations in the sample, most focus mainly on providing social services, although a few of the organizations may have more focus on the diffusion of a religious message.³⁰ This focus of faith-based organizations on delivering services to disadvantaged populations rather than facilitating worship and ritual practices is in line with the findings of comprehensive studies of the religious charitable sector (Chaves, 2004; Cnaan et al., 2002).

Unlike the previous subsections, where the structure of the data was a balanced panel of zip codes, in this subsection, we work with an unbalanced panel of individual charities. Thus, we are going to measure the effect of the scandals on charitable contributions, conditional on the charity continuing to exist after the scandal becomes public.³¹ These organizations are relatively large (e.g. the Catholic charities have on average contributions of half a million dollars), so they may receive contributions from donors living in different zip codes. Because of these issues, we define the Post-Scandal variable as the stock of scandals in the same zip code where the charitable organization is located and all of its adjacent zip codes (just as in our analysis of the effect of the scandals on the Catholic schools).³² While the databases used before consisted of thousands of zip codes observed over many years, in this database there are fewer than 400 organizations affiliated with the Catholic Church; each of them appears in the sample for an average of only 9 years. Thus, the number of Type-I scandals that occur near these organizations remains relatively low, which introduces a concern for statistical power. As a solution, we pool the Type-I and Type-II scandals in the creation of the variables Post-Scandal and Pre-Scandal.³³

Table 8 details the descriptive statistics about the data. Table 9 shows the difference-in-difference estimates. We regress the logarithm of charitable contributions on Post-Scandal, including the typical set of control variables: organization fixed effects, time effects, and the interaction between the time effects and zip code characteristics (e.g. population size). Column (1) suggests that whenever a scandal affects a Catholic-affiliated charity, a permanent decrease in charitable contributions in the order of 4% results. Column (2) adds the Pre-Scandal variable. The coefficient on Pre-Scandal is zero and statistically insignificant, which means that there were no differences in pre-trends among charities affected by a scandal and unaffected charities. Columns (3) and (4) provide some robustness tests. Column (3) excludes 95 organizations that change their city address during at least one of the years that the organization appears in the sample. The results are even more pronounced. Column (4)

³⁰We came to this conclusion after the inspection of the websites of a random sample of these organizations.

³¹The scandals will also increase the likelihood that a given charity disappears, which has a trivial negative effect on the total charitable contributions in the community. Thus, the effects measured here can be thought as a lower bound to the overall effect of the scandals on charitable giving.

³²The observed address of an organization corresponds to its legal address. An organization with a single location may have a legal address different than the physical address, most likely in an adjacent zipcode. Also, an organization working in several locations may report only one form under the legal address of a “central agency.” This issue is going to introduce some measurement error in the geographical dimension of the data.

³³The results are robust if instead we create a variable that adds the number of Type-I scandals and the number of scandals that are Type-II but not Type-I (so we do not double-count when a single priest is simultaneously involved in Type-I and Type-II scandals in the same zipcode).

excludes 31 charities whose mean contributions are above \$1 million (in 2009 dollars). The results are practically the same.

Most important, columns (5) and (6) correspond to the same regression in column (2), but instead of Catholic-affiliated charities they focus on charities affiliated to other religious denominations. Column (5) looks at organizations related to non-Catholic denominations: i.e., organizations that belong to NTEE codes X20 (Christian), X21 (Protestant), X30 (Jewish), X40 (Islamic), X50 (Buddhist) or X70 (Hindu). The coefficient on Post-Scandal is zero and precisely estimated, which means that the Catholic scandals do not affect charitable donations to non-Catholic religious-related organizations. Column (6) looks at the subset of organizations under NTEE code X21 (Protestant) with the exact same results. In sum, the results suggest that the drop in charitable contributions in the aftermath of the scandals is concentrated entirely among donors affiliated with the Catholic Church.

4.3 The effect of the scandals on the private provision of welfare

In the previous subsections we demonstrated that a permanent drop in charitable contributions follow the scandals. In this subsection, we present evidence that the drop in charitable contributions affected the output of the charitable organizations: i.e. the provision of welfare in the community.

We use administrative data from the Zipcode Business Patterns (ZBP), an annual series prepared by the U.S. Census Bureau using several sources of administrative data.³⁴ The data consists of the number and size of establishments in each zip code for each industry. We define an establishment as a single physical location where business transactions take place and which keeps payroll and employment records. Note that, if a firm or organization operates in many locations, it will have as many entries as establishments. Although the zip code data is available annually for the period 1994–2008, definitional changes in establishments, activity status, and industry classifications can affect the comparability of data over time. The sample from 1998–2008 employed the North American Industry Classification System (NAICS) to classify the businesses. Prior to 1998, it classified the businesses according to the Standard Industrial Classification (SIC) system. NAICS identifies new industries, redefines concepts, and develops classifications to reflect changes in the economy. However, the correspondence between NAICS and SIC for the categories used in this subsection is imperfect, so we focus on the 1998–2008 sample only. Although the NAICS system did introduce some changes in 2002 and 2007, they did not affect any of the NAICS codes used in this subsection.

We focus on establishments oriented towards the provision of social services, comprised of the following categories: community food services (NAICS code 624210), temporary shelters (624221), other community housing services (624229), emergency and other relief services (624230), vocational rehabilitation services (624310), child and youth services (624110), services for elderly and disabled (624120), other individual and family services (624190), grantmaking foundations (813211 and 813219),

³⁴The data is extracted from the Business Register, the Census Bureau's file of all known single- and multi-establishment companies. The Annual Company Organization Survey and quinquennial Economic Censuses provide individual establishment data for multi-location firms. Data for single-location firms are obtained from various programs conducted by the Census Bureau, such as the Economic Censuses, the Annual Survey of Manufactures, and Current Business Surveys, as well as from administrative records of the Internal Revenue Service (IRS), the Social Security Administration (SSA), and the Bureau of Labor Statistics (BLS). The data covers all NAICS industries except some industries that are not relevant for the purposes of this subsection (e.g. crop and animal production).

human rights (813311) and other social advocacy (813319).³⁵ Most of the organizations that fall in these categories are primarily concerned with providing services to low-income individuals and other disadvantaged groups. The following are just some examples of the many types of establishments that fall into the above categories: soup kitchens, homeless shelters, housing assistance agencies, sheltered workshops, child welfare services, youth centers, teen outreach services, disability support groups, senior citizens activity centers, family social service agencies, family welfare services, alcoholism counseling, immigrant resettlement services, and so forth.

The data not only specifies the number of establishments in each industry for a given zip code, but also indicates the approximate employment size of each establishment in categorical form. We construct a proxy variable for the number of employees, defined as the sum of the number of establishments in each employment-size-group weighted by the average number of employees in the corresponding category. For instance, if a zip code has one establishment with 1 to 4 employees and five establishments with 5 to 9 employees, the proxy for the number of employees takes the value $2.5+7=9.5$. We use the total number of employees in a zip code as the dependent variable, which allows for the measurement of effects on the extensive margin (i.e. change in the number of establishments) as well as on the intensive margin (i.e. change in the average size of the establishments). Table 10 presents descriptive statistics about the data. We focus on the subset of 10,480 zip codes that consistently have some charitable establishments throughout the sample period.³⁶ In those zip codes, 8.5 establishments, on average, provide social services with the total number of employees being 164, on average.

Table 11 shows the difference-in-difference estimates. We regress the logarithm of the number of employees working in charitable organizations on the variables describing the timing of the scandals. As usual, all regressions include zip code fixed effects, time effects, and the interaction between the time effects and zip code characteristics (e.g. log of population, share of urban population). The coefficient on Post-Scandal in column (1) suggests that a scandal reduces the number of charitable employees in the zip code by around 8.5%. Column (2) adds the variable Pre-Scandal. Its coefficient is not significantly different from zero, which suggests that no differences exist in pre-trends between zip codes with scandals and zip codes without scandals. Figure 6 shows the event-study analysis. As expected, no differences appear in the evolution of welfare services before the scandals become public (note that the standard errors get very large when we look at seven or more years before the scandals, because of the sample period). The effect of a scandal intensifies over time, although it seems to stabilize after the fifth year.

Column (3) adds the variable that measures the stock of scandals in adjacent zip codes. The results suggest that the effects are concentrated on the same zip code of the scandal, as in the case of charitable contributions. Column (4) introduces the Type-I and Type-II scandals simultaneously. Both types of scandals have a statistically and economically very significant effect on the number of employees in charitable establishments [recall that we cannot compare the coefficient on “Post-Scandal (Type-I)” in column (4) to that of columns (1) through (3)], and both Pre-Scandal variables are close to zero and not statistically significant. The effect of the Type-II scandals on the number of employees constitutes slightly more than half of the effect of the Type-I scandals, in line with the findings on charitable

³⁵The results are robust to the use of alternative groupings.

³⁶We focus on this subset mainly because we want to take the log of the dependent variable, so we can estimate the effect of the scandals as semi-elasticities.

contributions from the previous section. As such, when a priest working in town A is publicly accused about having abused 20 years ago while working in town B, the size of the charitable sector goes down not only in town A but it also goes down (roughly half as much) in town B. We also created a category that agglomerates an arbitrarily chosen group of retail establishments: supermarkets (NAICS code 445110), new car dealers (441110) and used car dealers (441120). The average zip code has seven of these establishments and 259 employees. As a falsification test, column (5) uses the number of employees in the retail sector as dependent variable. The coefficient on Post-Scandal is very close to zero and not statistically significant, which confirms that the scandals did not have any significant effect on the size of the retail sector.

Last, column (6) reproduces column (2) but uses the number of establishments instead of the number of employees as the dependent variable. Compared to the coefficient in column (2), the coefficient on Post-Scandal from column (6) is smaller, and, furthermore, its p-value is slightly over 0.1. One could interpret this as evidence that the effects of the scandals are felt relatively more in the intensive margin (i.e. shrinking of establishments) than in the extensive margin (i.e. closing of establishments). However, the number of establishments in a zip code is not the best measure, since it gives the same weight to types of establishments with different sizes (e.g. the average child and youth services organization has three times the size of the average community food services organization). Instead, we can use a measure of the number of establishments that weights the number of establishments within each NAICS code by the average establishment size in that code as of 1998. If we use that measure as dependent variable (not reported), we see a statistically significant drop in the number of establishments of about 4% after the scandals.

5 Discussion

The evidence from the last two sections strongly suggests that a negative shock to a religious institution has a significant effect on pro-social behavior. In this section we will provide our interpretation of the evidence. In subsection 5.1, we enumerate and discuss the potential mechanisms that can explain the patterns observed in the data. In subsection 5.2, we discuss a particular mechanism, which asserts that the scandals might have a direct effect on pro-social behavior. Furthermore, we perform two different tests of that conjecture.

5.1 Interpretation of the findings

Several studies have documented a strong positive correlation between religion and pro-social behavior; individuals that participate actively in religious congregations donate more time and money (e.g. Brooks, 2003; Becker and Dhingra, 2001), and the share of religious adherents in a county is positively correlated to the number of nonprofit organizations (Polson, 2009). However, the correlation between religious participation and pro-social behavior could be spurious. For example, if more altruistic people self-select into religious congregations, this will generate a trivial correlation between religious participation and charitable giving. This identification challenge is commonplace in studies regarding religion, as many measures of religious life have been found to be strongly correlated to a wide variety of social and economic indicators (e.g. Putnam and Campbell, 2010), but the evidence on the direction

of causality has remained quite limited (e.g. Gruber and Hungerman, 2007; Hungerman, 2011).

In an ideal experimental setting, we would measure what happens to pro-social behavior in a randomly chosen group of communities exposed to a “treatment” that makes religious participation more costly. For example, we could ban parking lots in places of worship in a randomly selected set of communities, hoping that it will make people go to church less frequently. We can interpret the evidence presented in this paper as a quasi-experiment wherein the “treatment” consists of whether a priest is publicly accused of sexual abuse. Since the scandals reduced both religious participation and pro-social behavior, the evidence favors the hypothesis that religious congregations foster pro-social behavior. However, that is our favorite interpretation, but not the only possible interpretation of the findings. For example, it might be possible that the scandals affected pro-social behavior directly, and not only indirectly through the decline in religious congregations. We will first enumerate many mechanisms that are consistent with our interpretation of the findings, and in the following section we will test the power of an alternative mechanism.

First, higher religious participation can affect pro-social behavior through an effect on religious beliefs (e.g. belief in God). This mechanism, widely discussed in the sociology literature, has been formalized in economic models (e.g. the “salvation motive” in Azzi and Ehrenberg, 1975). As a rough metaphor, consider a model in which people have uncertainty about the existence of God. People are told that, if God exists, they will be rewarded (or equivalently, not punished) in the afterlife only if they help others in need (e.g. Matthew 25:40; Proverbs 11:24–25; Matthew 6:2–4; Mark 12:41–44). If God does not exist, then such rewards and punishments do not apply. Therefore, pro-social behavior should be an increasing function of the probability belief in God (or as an equivalent, belief in life after death, when the punishments and rewards are realized).

However, religious participation can also affect pro-social behavior through channels that are independent from religious beliefs. To begin with, higher participation in a social group can make people incorporate the social norms of the group (Akerlof and Kranton, 2000), which, in the case of religious congregations, can explain the increased proclivity towards pro-social behavior (e.g. Wuthnow, 1991). Higher religious participation can increase socialization, which strengthens the links in the community and may foster volunteering and charitable giving (e.g. Putnam, 2000; Putnam and Campbell, 2010). Higher participation in the congregation can increase the demand for pro-social behavior as a signal of unobservable personal traits, such as benevolence (e.g. Perez-Truglia, 2010). In addition, pro-social behavior can play an instrumental role in the congregation, helping to alleviate a free-riding problem (Iannaccone, 1992).

By participating in the congregation, people can learn about more opportunities to volunteer and donate money (e.g. Park and Smith, 2002). Some authors have suggested that religious congregations function as civic training grounds, where individuals can develop skills and resources useful for such other social activities as volunteering (e.g. Peterson, 1992). In addition, individuals who actively participate in the religious communities may be more likely to be solicited for donations of time and money. This channel can be very important, since solicitation is a powerful factor driving pro-social behavior (e.g. Hodgkinson, 1995; Landry et al., 2006). Religious congregations, since they are deeply involved in the provision of social services in the United States, are naturally better at disseminating information and training and soliciting volunteers (e.g. Chaves, 2004; Cnaan et al., 2002). Indeed,

many contemporary nonprofit service organizations seeking to combat such numerous social problems as poverty and homelessness trace their origins back to reform movements that emerged from religious organizations (Day, 2000). In particular, the Catholic Church has, throughout its history, established such institutions to provide a wide array of social services as orphanages, hospitals, schools, and homes for the sick and elderly (Bane, 2005).

The observed effect of religious participation on pro-social behavior probably responds to a combination of many of the above mechanisms. Some of them are complementary: religious organizations may foster benevolent social norms and provide opportunities to fulfill the increased interest in benevolent behavior. Disentangling the precise mechanisms in play is not only interesting from a sociological point of view, but can also be important for policy analysis. For example, the recent interest of policy-makers in the religious organizations that provide social services could likely be due to a belief that such organizations are better than the government in providing and/or financing such services (Gruber and Hungerman, 2005). By identifying which mechanisms contribute to the success of religious organizations, the policy-makers will be in a better position when collaborating with religious organizations, and they may even incorporate some of those advantages to the non-religious organizations.

Many of the mechanisms discussed above predict that the decline in religious participation made people less attracted to helping others (e.g. through an effect on religious beliefs or social norms) or just less likely to be solicited and learn about volunteer opportunities. According to this view, as long as the drop in religious participation is permanent, the effects on charitable contributions and social welfare will also be permanent. However, it is also possible that part of the decline in charitable contributions is due to the difficulties former members of the Catholic congregation faced in finding similar non-Catholic charities to which to donate their money and time. According to this view, even if the drop in religious participation were permanent, eventually there would be a partial recovery in pro-social behavior, as the former Catholics identify non-religious organizations to which they can contribute.³⁷ The estimates suggest that the decline in pro-social behavior lasts for as many years as we can look into the future, which is over a decade. If anything, the data suggests that the decline in pro-social behavior intensifies more and more over time. However, it is possible that, in the very-long-run, new charitable organizations will find a way to reach former Catholics and recover part of the decline in pro-social behavior.³⁸

5.2 Did the scandals have a direct effect on pro-social behavior?

The news about the sexual abuse of minors may change people's perception about social justice or their trust in others, which can directly affect pro-social behavior, either positively or negatively. If the decline in the perception of social justice increases the desire to contribute to society, in the hope that they can make a difference, then the estimates from this paper would under-estimate the effect of religious participation on pro-social behavior. If, on the contrary, the changed perception of social justice makes people less likely to behave pro-socially, then this paper's estimates would over-estimate

³⁷Intuitively, something like this would probably happen if the scandals were against a non-religious charitable organization (e.g. United Way).

³⁸It is possible to perform some empirical analysis on this respect. For example, we could test whether the effects of the scandals are stronger or weaker in places with more quantity and variety of non-profit organizations. If this channel was important, then we should find the drop in charitable giving to be deeper and more lasting in areas with fewer alternatives to the Catholic-related charitable organizations.

the true effect of religious participation on pro-social behavior.

To us it appears unlikely that the clergy abuse scandals can have such a significant “direct” effect on pro-social behavior. First, the news articles about the clergy sexual abuse scandals are just one of the many news articles that depict situations of bad (and also good) character, from which people may form beliefs about things like justice and trust in others. Moreover, all the news articles taken together are just a small fraction of all the situations that can influence such beliefs, including what happens to the individuals in their everyday lives. Second, the news about the scandals reached both Catholics and non-Catholics. If these had a direct effect on beliefs about justice, we would expect them to affect the belief of non-Catholics, maybe not by the same magnitude than for Catholics, but still in the same direction. However, this prediction is in sharp contrast with the above evidence indicating that the scandals affected only the charitable contributions made by Catholics. Furthermore, we provide two tests of this particular mechanism.

5.2.1 The effect of the scandals on self-reported trust in others

If the scandals had a direct effect on pro-social behavior, we would expect them to affect measures of trust in others. We use standard survey questions about trust in others obtained from the GSS (we described the dataset already in subsection 3.2). We use all the questions employed in the literature on social capital (Glaeser et al., 2000). We define the variable Trust as the answer to the question “Generally speaking, would you say that most people can be trusted or that you can’t be too careful in dealing with people? Can’t be too careful (0); Depends (1); Most people can be trusted (2).” We define the variable Helpful as the answer to the question: “Would you say that most of the time people try to be helpful, or that they are mostly just looking out for themselves? Just look out for themselves (0); Depends (1); Try to be helpful (2).” We define the variable Fair as the answer to the question: “Do you think most people would try to take advantage of you if they got a chance, or would they try to be fair? Would take advantage of you (0); Depends (1); Would try to be fair (2).” We also constructed a variable that is the average of the first three variables (the results are the same if we instead use the principal component). Note that all variables are constructed in such a manner so that a higher value indicates more trust in others. However, the GSS did not include these questions in every questionnaire of the GSS, but only in random subsamples. Therefore, the sample size is lower than the one used when looking at the effect of the scandals on religious adherence.

We want to examine whether the scandals affected trust in others and, in particular, among Catholics. Looking at the effect of the scandals on the subset of respondents self-identified as Catholic would introduce a selection bias, since we know that the scandals affect Catholic identification. In order to avoid this problem, we instead examine the effect of the scandals on the subset of the respondents who were raised as Catholic, a grouping that precedes the scandals. For the sake of comparison, we also report the effect on the entire population. Table 12 presents the difference-in-difference estimates, as well as the mean and standard deviation of each dependent variable. For the subpopulation of Catholics, the coefficients on Post-Scandal are never statistically significant, and, in addition, the point estimates are always small and, in all but one case, positive. Note that the results for the entire population are very similar to the ones for the subpopulation of Catholics. In conclusion, we find no evidence that those raised Catholic reacted to the scandals by lowering their trust in others.

5.2.2 The effect of abuse scandals in lay organizations

If part of the effect from the religious scandals is not due to the decline in religious participation, then scandals in non-religious organizations should have a qualitatively-similar effect. Thus, in order to assess the direct effect of scandals on pro-social behavior, we examine the effect of abuse scandals on pro-social behavior when the abuses take place in lay organizations. We created a dataset similar to that of the Catholic-clergy sexual abuse scandals, but about child abuse scandals in non-religious schools. We used a variety of data sources in order to create a list of teachers publicly accused of having abused students in non-religious schools.³⁹ We followed the same criteria that we used for the database on Catholic-clergy abuses.⁴⁰ The final database consists of 607 scandals over the period 1990-2010. The number of events is over half of the number of Type-I scandals in the database on Catholic abuse scandals. Although the statistical power will be lower relative to the analysis of religious scandals, it should be more than enough for the difference-in-difference estimates.

We will examine whether the scandals had an effect on the mean charitable contributions and the provision of welfare (i.e. the dependent variables from subsections 4.1 and 4.3). It would be reasonable to find a negative effect on charitable contributions, since schools receive donations from parents and those can be affected by the scandals. However, if scandals have no direct effect on pro-social behavior, we should find that the scandals in lay organizations did not decrease the provision of welfare. Table 13 presents the difference-in-difference estimates. The first two columns show the effect of the lay abuse scandals on the log of mean charitable contributions, using the IRS SOI database. The last two columns show the effect of the lay abuse scandals on the log of employees in charitable establishments, using the ZBP database. The specifications are almost identical to those presented in Table 7 and Table 11.⁴¹ Column (1) illustrates a negative but statistically insignificant effect of the lay scandals on mean charitable contributions. Column (2) breaks down the effect of the scandals between own and adjacent zip codes. The coefficient on “Post-Scandal, Same” is still statistically insignificant, and the coefficient on “Post-Scandal, Adjacent” seems to be generated by differential pre-trends between treatment and control groups. Most importantly, we want to assess the effect of the lay scandals on the provision of welfare in the zip code. The coefficients on “Post-Scandal, Same” in columns (3) and (4) are statistically insignificant. The point estimate in column (3) is positive, and the point estimate in column (4) is practically zero. The qualitative results are robust if we use alternative specifications. In summary, the evidence is consistent with the hypothesis that abuse scandals did not have a direct effect on pro-social behavior.

³⁹For example, we used databases of historical newspapers to look for articles related to teacher sexual abuse: e.g. LexisNexis Academic, Google News, NewsLibrary.

⁴⁰One difference between the two databases is that lay teacher abuse scandals become public soon after the abuses take place, while the abusers are still working at the schools of the abuse. As a consequence, all the events considered here are simultaneously Type-I and Type-II scandals.

⁴¹We found statistically significant pre-trend differences in charitable contributions between zip codes with lay scandals and zip codes without lay scandals: i.e., charitable giving in a zip code with a scandal was going down even before the scandal erupted, which means that the difference-in-difference estimates would be invalid. That difference in pre-trends disappears by simply including a time trend that is specific to those zip codes with scandals. Because of the need for this correction, the results should be interpreted with more care.

6 Conclusions

Although several studies have documented a strong positive correlation between participation in religious organizations and pro-social behavior, no conclusive evidence supports the direction of causality. In order to contribute new evidence to this long-standing question, we performed an event-study analysis of the Catholic-clergy sexual abuse scandals in the United States. The evidence strongly suggests that a negative shock to a religious institution has a significant negative effect on pro-social behavior, since the scandal-affected communities suffered a permanent decline in the Catholic congregation accompanied by permanent declines in charitable contributions and the provision of welfare. This pattern is consistent with the widespread prior belief that religious congregations foster pro-social behavior. We do not attempt to identify the precise mechanisms responsible for this relationship, leaving that for future research. However, we do present evidence that the scandals do not have a direct influence on pro-social behavior: people's trust in others is not affected by the scandals, and abuse scandals in lay organizations do not have an effect on pro-social behavior.

One can identify many possible extensions to this paper. For example, we can exploit heterogeneity in the characteristics of the scandals to explore which characteristics are associated with the scandals' more severe consequences. We can potentially exploit data about how each Catholic institution dealt with the scandals in order to measure the consequences of following adequate crisis-management strategies. In addition, we can examine whether some characteristics of the newspapers that investigated and reported the abuses are associated with more severe consequences of the scandals. Furthermore, one can use this same Natural Experiment to explore the relationship between religious congregations and other outcomes. For instance, in ongoing research, we are exploring whether the scandals had an effect on political outcomes (e.g. contributions to political campaigns); we are also exploiting the distribution of scandals over space and time as an instrumental variable to assess whether Catholic schools attain better student outcomes.

Last, our results suggest that a promising line of research could examine how scandals in general, not only sexual scandals in religious organizations, can have sizeable consequences for the society as a whole. The abuses directly affected several individuals and institutions. Some estimate the lawsuits and other abuse-related costs for the past 40 years to total around \$3 billion (bishopaccountability.org) and the number of victims to number more than 15,000 (JJCCJ, 2011). On top of that, the abuses introduced indirect effects when they became public, often decades after the perpetration of the abuse. Our paper demonstrates that those indirect effects can be sizable. Indeed, the scandals permanently changed the religious affiliation of around 5 million Americans, contributed significantly to the sharp decline of the Catholic school system, and caused permanent billion-dollar losses in charitable contributions. We hope that future researchers will make use of an identification strategy similar to the one proposed in this paper to explore the indirect effects of non-sexual scandals in non-religious organizations, such as corruption scandals in governments and corporations.

References

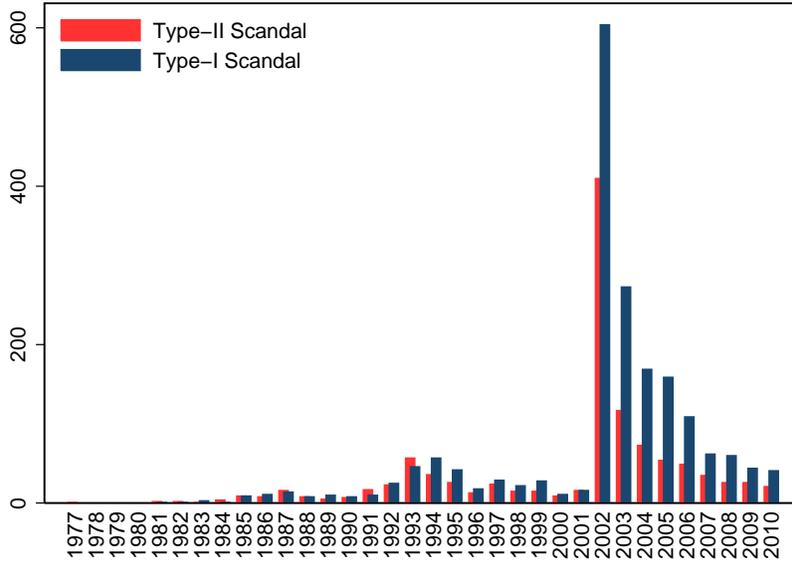
- [1] Akerlof, George A. and Kranton, Rachel E. (2000), "Economics and Identity," *Quarterly Journal of Economics*, Vol. 115 (3), pp. 715-733.

- [2] Alesina, Alberto and Glaeser, Edward L. (2004), "Fighting Poverty in the US and Europe: A World of Difference." Oxford: Oxford University Press.
- [3] Azzi, C. and Ehrenberg, R. (1975), "Household Allocation of Time and Church Attendance," *Journal of Political Economy*, Vol. 83 (1), pp. 27-56.
- [4] Bane, Mary Jo (2005), "The Catholic puzzle: Parishes and civic life." In Mary Jo Bane, Brent Coffin, & Richard Higgins, Eds., "Taking faith seriously." Cambridge: Harvard University Press.
- [5] Banfield, E. (1958), "The Moral Basis of a Backward Society." New York: Free Press.
- [6] Barro, R.J. and McCleary, R.M. (2003), "Religion and Economic Growth across Countries," *American Sociological Review*, Vol. 68 (5), pp. 760-781.
- [7] Barro, R.J.; Hwang, J. and McCleary, R.M. (2010), "Religious Conversion in 40 Countries," *Journal for the Scientific Study of Religion*, Vol. 49 (1), pp. 15-36.
- [8] Becker, Penny Edgell, and Dhingra, Pawan H. (2001), "Religious involvement and volunteering: Implications for civil society," *Sociology of Religion*, Vol. 62 (3), pp. 315-335.
- [9] Beyerlein, Kraig and Hipp, John R. (2006), "From pews to participation: The effect of congregation activity and context on bridging civic engagement," *Social Problems*, Vol. 53 (1), pp. 97-117.
- [10] Brooks, Arthur C. (2003), "Religious faith and charitable giving," *Policy Review*, Vol. 121 (5), pp. 39-50.
- [11] Chaves, Mark (2004), "Congregations in America." Cambridge: Harvard University Press.
- [12] Cnaan, Ram A.; Boddie, Stephanie C.; Handy, Femida; Yancey, Gaynor and Schneider, Richard (2002), "The invisible caring hand: American congregations and the provision of welfare." New York: New York University Press.
- [13] Day, Phyllis (2000), "A new history of social welfare." Boston: Allyn and Bacon.
- [14] Dills, Angela and Hernández-Julian, Rey (2011), "Negative Publicity and Catholic Schools," *Economic Inquiry*, forthcoming.
- [15] Doyle, Thomas and Rubino, Steven (2004), "Catholic Clergy Sexual Abuse Meets the Civil Law," *Fordham Urban Law Review*, January 1.
- [16] Finke, Roger and Scheitle, Christopher P. (2005), "Accounting for the Uncounted: Computing Correctives for the 2000 RCMS Data," *Review of Religious Research*, Vol. 47, pp. 5-22.
- [17] Fisman, Ray and Miguel, Edward (2007), "Corruption, Norms and Legal Enforcement: Evidence from Diplomatic Parking Tickets," *Journal of Political Economy*, Vol. 115(6), pp. 1020-1048.
- [18] Glaeser, Edward; Laibson, David; Scheinkman, Jose and Soutter, Christine (2000), "Measuring trust," *Quarterly Journal of Economics*, Vol. 115 (3), pp. 811-846.
- [19] Gruber, Jonathan (2004), "Pay or Pray? The Impact of Charitable Subsidies on Religious Attendance," *Journal of Public Economics*, Vol. 88 (12), pp. 2635-2655.
- [20] Gruber, Jonathan and Hungerman, Daniel M. (2007), "Faith-based charity and crowd-out during the great depression," *Journal of Public Economics*, Vol. 91, pp. 1043-1069.

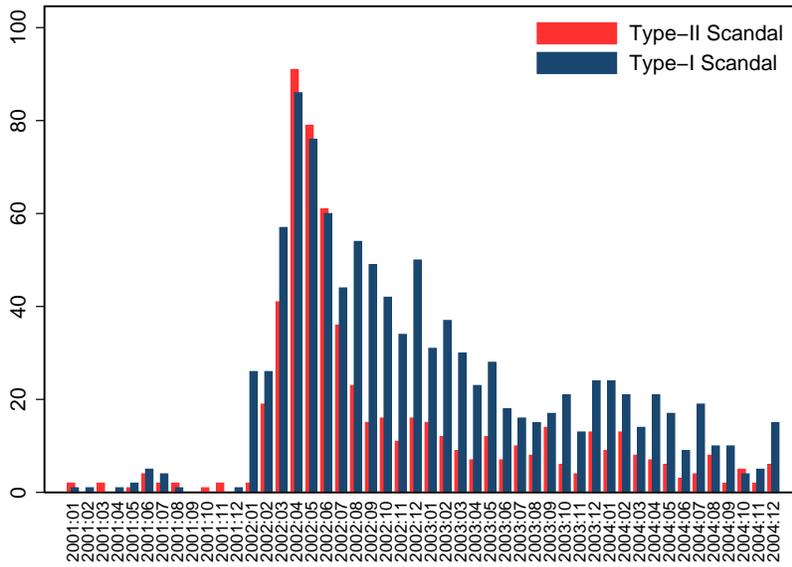
- [21] Gruber, Jonathan and Hungerman, Daniel (2008), "The Church vs. The Mall: What Happens When Religion Faces Increased Secular Competition?" *Quarterly Journal of Economics*, Vol. 123, pp. 831-862.
- [22] Guiso, L.; Sapienza, P. and Zingales, L. (2003), "People's Opium? Religion and Economic Attitudes," *Journal of Monetary Economics*, Vol. 50 (1), pp. 225-82.
- [23] Guiso, Luigi; Sapienza, Paola and Zingales, Luigi (2009), "Cultural Biases in Economic Exchange?" *The Quarterly Journal of Economics*, Vol. 124 (3), pp. 1095-1131.
- [24] Hodgkinson, V. (1995), "Key factors influencing caring, involvement, and community." In P. Schervish, V. Hodgkinson and M. Gates, Eds., "Care and community in modern society." San Francisco: Jossey-Bass.
- [25] Hungerman, Daniel (2011), "Does Religious Proscription Cause People to Act Differently? Evidence from a Theory-Based Test," NBER Working Paper No. 17375.
- [26] Hungerman, Daniel (2011), "Substitution and Stigma: Evidence on Religious Competition from the Catholic Sex-Abuse Scandal," Mimeo.
- [27] Hungerman, Daniel (2005), "Are church and state substitutes? Evidence from the 1996 welfare reform," *Journal of Public Economics*, Vol. 89, pp. 2245-2267.
- [28] Iannaccone, Laurence (1992), "Sacrifice and Stigma: Reducing Free-riding in Cults, Communes, and Other Collectives," *Journal of Political Economy*, Vol. 100 (2), pp. 271-291.
- [29] Iannaccone, Laurence R. (2003), "Looking backward: A cross-national study of religious trends," Working paper, Center for Study of Public Choice.
- [30] John Jay College of Criminal Justice (2004), "The Nature and Scope of the Problem of Sexual Abuse of Minors by Catholic Priests and Deacons in the United States." Washington DC: United States Conference of Catholic Bishops.
- [31] John Jay College of Criminal Justice (2006), "The Nature and Scope of the Problem of Sexual Abuse of Minors by Catholic Priests and Deacons in the United States: Supplementary Report." Washington DC: United States Conference of Catholic Bishops.
- [32] John Jay College of Criminal Justice (2011), "The Causes and Context of Sexual Abuse of Minors by Catholic Priests in the United States, 1950-2010," Washington DC: United States Conference of Catholic Bishops.
- [33] Johnson, B.; Tompkins, R. and Webb, D. (2002), "Objective Hope. Assessing the Effectiveness of Faith-Based Organizations: A Review of the Literature," Center for Research on Religion and Urban Civil Society.
- [34] Landry, Craig; Lange, Andreas; List, John A.; Price, Michael K. and Rupp, Nicholas G. (2006), "Toward an Understanding of the Economics of Charity: Evidence from a Field Experiment," *Quarterly Journal of Economics*, Vol. 121 (2), pp. 747-782.
- [35] List, John A. (2011), "The Market for Charitable Giving," *Journal of Economic Perspectives*, Vol. 25 (2), pp. 157-80.

- [36] McCleary, R.M. and Barro, R.J. (2006), "Religion and Economy," *Journal of Economic Perspectives*, Vol. 20 (2), pp. 49-72.
- [37] McMackin, R.A. , Keane, T.M., Kline P.M. (2009), "Understanding the Impact of Clergy Sexual Abuse: Betrayal and Recovery." London: Routledge.
- [38] Nunn, Nathan (2010), "Religious Conversion in Colonial Africa," *American Economic Review Papers and Proceedings*, Vol. 100 (2), pp. 147-152.
- [39] Nunn, Nathan and Wantchekon, Leonard (2009), "The Slave Trade and the Origins of Mistrust in Africa" *American Economic Review*, forthcoming.
- [40] Park, Jerry Z. and Smith, Christian (2000), "To whom much has been given...: Religious capital and community voluntarism among churchgoing Protestants," *Journal for the Scientific Study of Religion*, Vol. 39 (3), pp. 272-86.
- [41] Perez Truglia, R.N. (2010), "Conspicuous Consumption in the Land of Prince Charming." Available at SSRN: <http://ssrn.com/abstract=1078523>.
- [42] Peterson, Steven (1992), "Church participation and political participation: The spillover effect," *American Politics Quarterly*, Vol. 20 (1), pp. 123-139.
- [43] Polson, Edward C. (2009), "Cultivating the Common Good: Civic Life and Religious Contexts in American Society," Mimeo, Baylor University.
- [44] Putnam, R.; Leonardi, R. and Nanetti, R. (1994), "Making Democracy Work: Civic Traditions in Modern Italy." New Jersey: Princeton University Press.
- [45] Putnam, Robert. (2000), "Bowling alone: The collapse and revival of American community." New York: Simon and Schuster.
- [46] Putnam, Robert and Campbell, David E. (2010), "American Grace: How Religion Divides and Unites Us." New York: Simon and Schuster.
- [47] Ruiter, Stijn and De Graaf, Nan Dirk (2006), "National Context, Religiosity and Volunteering: Results from 53 Countries," *American Sociological Review*, Vol. 71 (2), pp. 191-210.
- [48] Wilhelm, Mark (2005), "Basic Facts about Charitable Giving from the Center on Philanthropy Panel Study." Mimeo.
- [49] Wuthnow, Robert (1991), "Acts of compassion: Caring for others and helping ourselves." Princeton: Princeton University Press.

Figure 1: The distribution of Catholic-clergy sexual abuse scandals over time



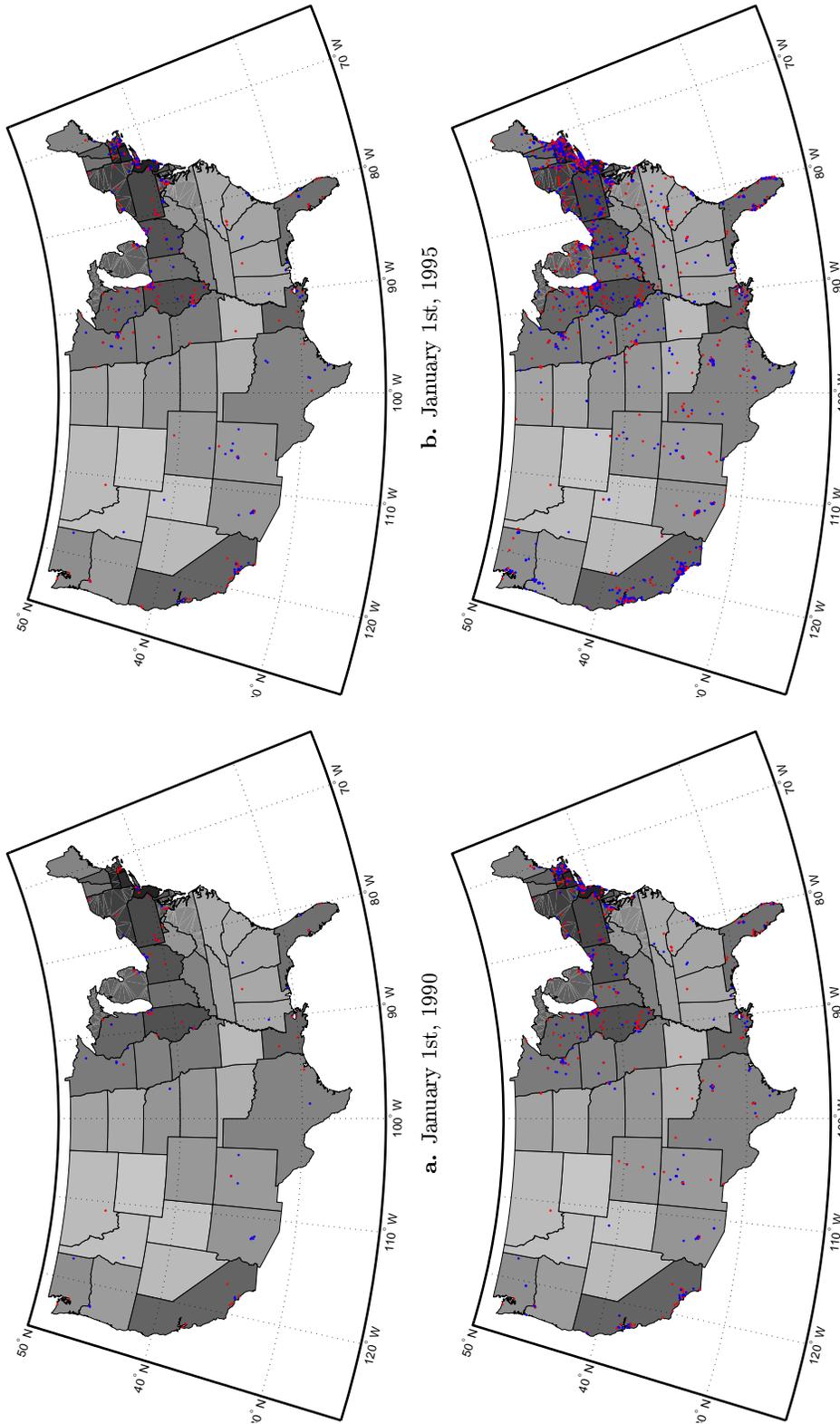
a. By year: 1977-2010



b. By month: Jan-01 to Dec-04

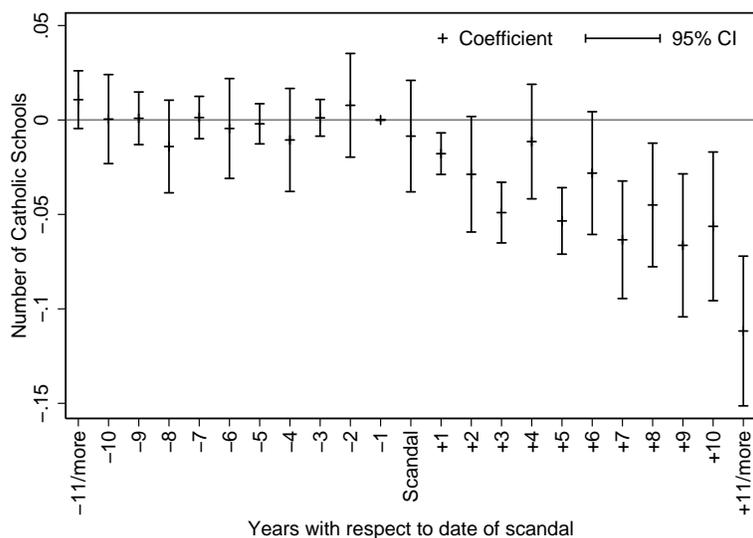
Notes: Data on scandals was compiled by the authors. See Section 2 for a description of the data, and footnote to Figure 2 for a brief definition of Type-I and Type-II scandals.

Figure 2: The distribution of scandals over space and time

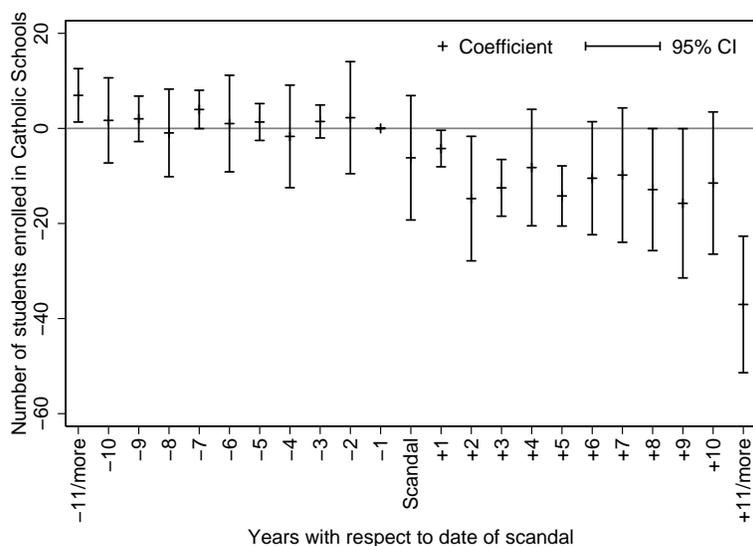


Notes: Data on scandals was compiled by the authors. See Section 2 for a description of the data. Type-I scandals denote the location (e.g. parish) where a priest is working when he is first accused (if working). Type-II scandals denote the location (e.g. parish) where a priest is accused of having committed the abuse. The date is when the accusation first became public in the community surrounding the corresponding location (e.g. an article in a local newspaper). Although Alaska and Hawaii do not appear in the maps, the database does include several events in those states.

Figure 3: Effect of scandals on number and size of Catholic schools



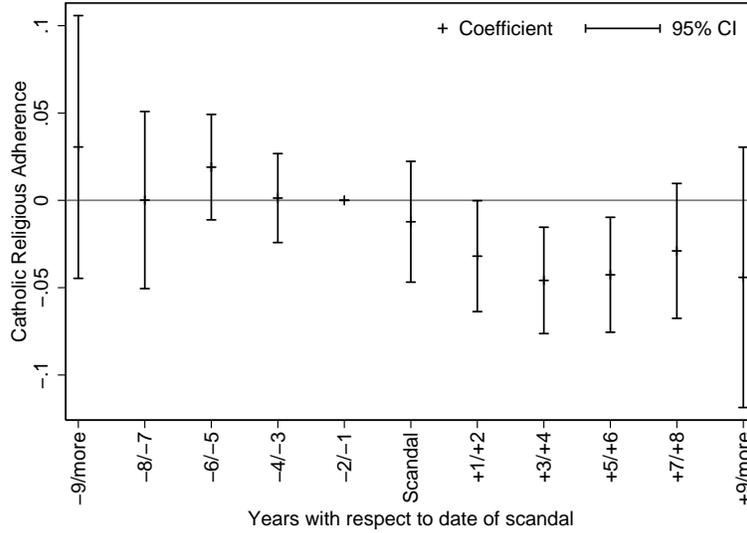
a. Number of Catholic schools



b. Number of students enrolled in Catholic schools

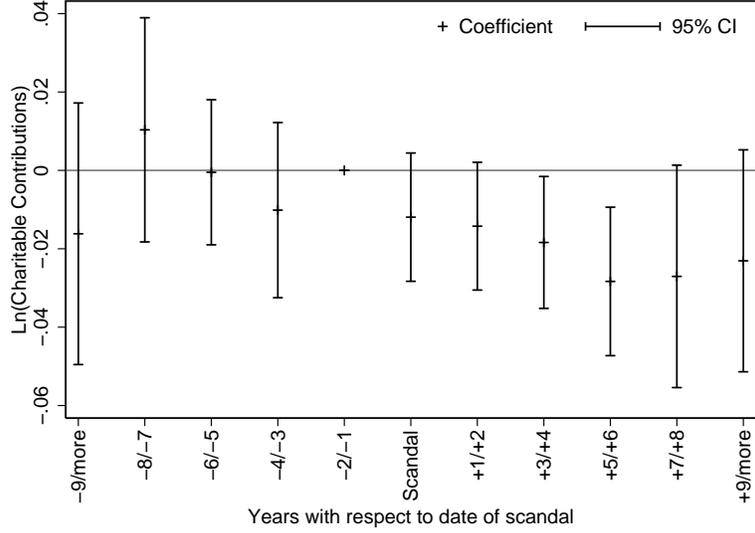
Notes: This is a graphical representation of an OLS regression of the number of Catholic schools (students enrolled in Catholic schools) in the zip code on a set of variables describing the timing of the scandals (i.e. the $d_{z,t}^s$'s). The "-1" is the omitted category (i.e. its coefficient is normalized to zero). The regression also includes zip code fixed effects, time effects and the interaction between the time effects and the logarithm of population in the zip code, the logarithm of land area, and the share of urban population, all taken from the 1990 US Population Census. Each bar represents the 95% confidence interval, and the center of the bar represents the corresponding point estimate. Confidence intervals are constructed with heteroskedasticity-robust standard errors, clustered at the zip code level. Data on private schools obtained from the Private School Survey, taken every two years since the academic year 1989-90 until the academic year 2007-2008. See Table 1 for descriptive statistics, and its footnote for data definitions. The sample excludes zip codes that are non-standard, and zip codes that changed substantially over time.

Figure 4: Effect of scandals on religious adherence (GSS)

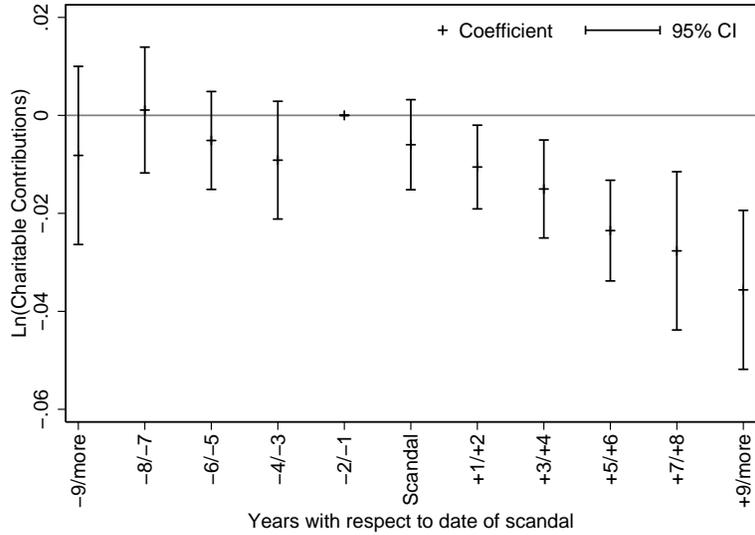


Notes: This is a graphical representation of an OLS regression of the dummy on Catholic adherence on a set of variables describing the timing of the scandals (i.e. the $d_{c,t}^s$'s). The “-2/-1” is the omitted category (i.e. its coefficient is normalized to zero). Each bar represents the 95% confidence interval, and the center of the bar represents the corresponding point estimate. Confidence intervals are constructed with heteroskedasticity-robust standard errors, clustered at the county level. Data from the General Social Survey, 1994-2010. See Table 3 for descriptive statistics, and its footnote for more details about the data. Regressions include county-specific fixed effects, time effects and control variables: gender, age, age squared, dummies for black and white, three dummies about marital status, household income, a set of four dummies about employment, education and number of children.

Figure 5: Effect of scandals on charitable contributions (IRS SOI)



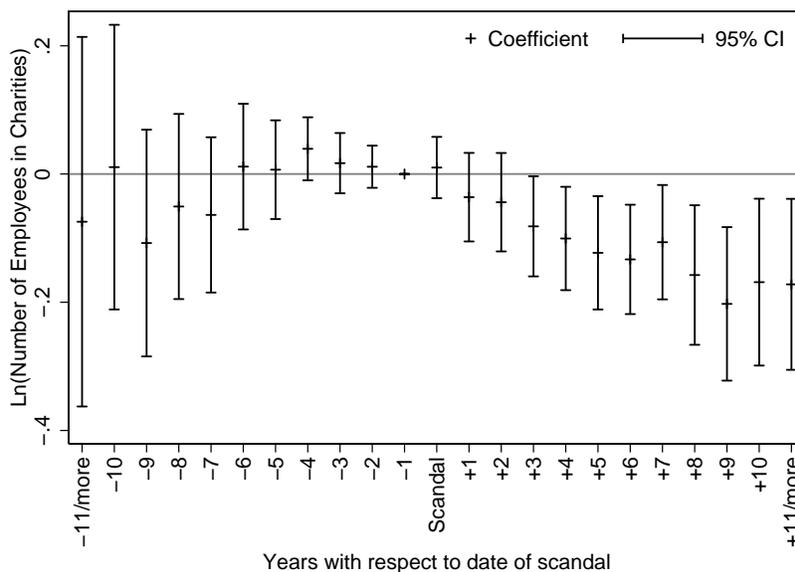
a. Effect of Type-I scandals



b. Effect of Type-I and Type-II scandals

Notes: This is a graphical representation of an OLS regression of the log of mean itemized charitable contributions in a zip code on a set of variables describing the timing of the scandals (i.e. the $d_{z,t}^s$'s). The “-2/-1” is the omitted category (i.e. its coefficient is normalized to zero). The regression includes zip code fixed effects, time effects, state-specific time trends, the logarithm of the number of returns, the logarithm of mean income (if positive), the logarithm of the number of returns with EITC, the stock of scandals in adjacent zip codes, and the interaction between the time effects and the logarithm of population in the zip code, the logarithm of land area, and the share of urban population, all taken from the 1990 US Population Census. Each bar represents the 95% confidence interval, and its center represents the corresponding point estimate. Confidence intervals are constructed with heteroskedasticity-robust standard errors, clustered at the zip code level. Data on charitable contributions obtained from the Statistics of Income Division of the IRS. See Table 6 for descriptive statistics, and its footnote for data definitions. The sample excludes zip codes that are non-standard, and zip codes that changed substantially over time.

Figure 6: Effect of scandals on charitable establishments (ZBP)



Notes: This is a graphical representation of an OLS regression of the log of the number of employees in charitable establishments that provide social services in the zip code on a set of variables describing the timing of the scandals (i.e. the $d_{z,t}^s$'s). The “-1” is the omitted category (i.e. its coefficient is normalized to zero). The regression also includes zip code fixed effects, time effects, state-specific time trends, and the interaction between the time effects and the logarithm of population in the zip code, the logarithm of land area, the share of urban population, and the shares of white, black, Asian and Hispanic population, all taken from the 1990 US Population Census. Each bar represents the 95% confidence interval, and the center of the bar represents the corresponding point estimate. Confidence intervals are constructed with heteroskedasticity-robust standard errors, clustered at the zip code level. Data on number of establishments and number of employees is from the Zipcode Business Patterns. See Table 10 for descriptive statistics, and its footnote for data definitions. The sample excludes zip codes that are non-standard, and zip codes that changed substantially over time.

Table 1: Descriptive statistics, Catholic schools

	Obs	Mean	Sd	Min	Max
Number of Catholic Schools	227270	0.34	0.80	0.00	14.00
Scandal in same zipcode	9360	1.36	1.53	0.00	12.00
Scandal in adjacent zipcode	42230	0.60	1.04	0.00	14.00
Neither	175680	0.22	0.60	0.00	9.00
Enrollment in Catholic Schools	227270	100.47	296.84	0.00	6900.00
Scandal in same zipcode	9360	463.72	649.74	0.00	6900.00
Scandal in adjacent zipcode	42230	190.20	409.51	0.00	6060.00
Neither	175680	59.55	201.37	0.00	4531.00
Number of Non-Cath. Rel. Schools	227270	0.53	1.17	0.00	37.00
Number of Non-Cath. Schools	227270	4.43	4.62	0.00	69.00
Post-Scandal (Type I), same zip and adj.	227270	0.14	0.51	0.00	14.00
Pre-Scandal (Type I), same zip and adj.	227270	0.04	0.24	0.00	10.00
Post-Scandal (Type I), zip adj to adj	227270	0.33	0.81	0.00	12.00
Pre-Scandal (Type I), zip adj to adj	227270	0.09	0.37	0.00	10.00
Post-Scandal (Type II), same zip and adj.	227270	0.21	0.78	0.00	23.00
Pre-Scandal (Type II), same zip and adj.	227270	0.06	0.38	0.00	15.00

Notes: Data on private schools from the Private School Survey, and data on public schools from the Common Core of Data. The dependent variable is the number of schools (enrollment) in the zip code in a given year. The data is biannual, starting in the academic year 1989-90 until the academic year 2007-2008. The sample includes all zip codes that had at least one school (public or private) thorough the years in the sample. Post-Scandal is the stock of past scandals, and Pre-Scandal is the number of scandals during the following two years. The sample excludes zip codes that are non-standard, and zip codes that changed substantially over time.

Table 2: Effect of scandals on number of Catholic schools

	Number of Schools						Enrollment	
	(1) Cath.	(2) Cath.	(3) Cath.	(4) Cath.	(5) Oth. Rel.	(6) Non-Cath.	(7) Cath.	(8) Cath.
Post-Scandal (Type I), same zip and adj.	-0.044*** (0.006)	-0.044*** (0.006)	-0.042*** (0.006)	-0.026*** (0.007)	0.000 (0.007)	0.010 (0.023)	-15.636*** (2.471)	
Pre-Scandal (Type I), same zip and adj.		0.001 (0.004)	0.001 (0.004)	0.003 (0.005)	-0.001 (0.007)	0.012 (0.019)	-2.224 (1.758)	
Post-Scandal (Type I), zip adj to adj			-0.006** (0.003)					
Pre-Scandal (Type I), zip adj to adj			0.001 (0.002)					
Post-Scandal (Type II), same zip and adj.				-0.021*** (0.004)				
Pre-Scandal (Type II), same zip and adj.				-0.006* (0.003)				
Observations	22720	22720	22720	22720	22720	22720	22720	22720
R-Squared	0.04	0.04	0.05	0.05	0.12	0.11	0.04	
No. of zipcodes	22727	22727	22727	22727	22727	22727	22727	22727

Notes: Heteroskedasticity-Robust standard errors in parentheses, clustered at the zip code level. Stars indicate significance level: * p<0.1, ** p<0.05, *** p<0.01. All regressions include zip code fixed effects, time effects, and the interaction between the time effects and the logarithm of population in the zip code, the logarithm of land area and the share of urban population. Data on private schools from the Private School Survey, and data on public schools from the Common Core of Data. The data is biannual, starting in the academic year 1989-90 until the academic year 2007-2008. See Table 1 for descriptive statistics, and its footnote for data definitions. The sample includes all zip codes that had at least one school (public or private) through the years in the sample. The sample excludes zip codes that are non-standard, and zip codes that changed substantially over time.

Table 3: Summary statistics (GSS)

	Obs	Mean	Sd	Min	Max
Catholic Adherent	25565	0.24	0.43	0.00	1.00
Other Rel. Adherent	25565	0.61	0.49	0.00	1.00
Raised Catholic	24062	0.30	0.46	0.00	1.00
Raised Other Rel.	24062	0.63	0.48	0.00	1.00
Religious Attendance, Catholic	25297	5.46	14.63	0.00	60.00
Religious Attendance, Other Rel.	25297	14.84	21.78	0.00	60.00
Strength of Affiliation, Catholic	24764	0.55	1.01	0.00	3.00
Strength of Affiliation, Other Rel.	24764	1.43	1.27	0.00	3.00
Frequency of Prayer, Catholic	15393	0.82	1.60	0.00	5.00
Frequency of Prayer, Other Rel.	15393	2.20	2.10	0.00	5.00
Belief in Afterlife, Catholic	18204	0.39	0.78	0.00	2.00
Belief in Afterlife, Other Rel.	18204	1.00	0.97	0.00	2.00
Belief in God, Catholic	10689	1.07	1.95	0.00	5.00
Believes in God, Other Rel.	10689	2.68	2.33	0.00	5.00
Post-Scandal	25604	0.22	0.61	0.00	7.22
Pre-Scandal	25604	0.03	0.16	0.00	2.50

Notes: Data from the General Social Survey, 1994-2010. Data on scandals compiled by the authors. Post-Scandal is the stock of scandals in the county of the respondent at the moment of the interview. Pre-Scandal is the number of scandals during the two years following the interview. Post- and Pre-Scandal are normalized by dividing by 100,000 inhabitants in the county, using data from the 1990 US Population Census. Religious attendance is a proxy measure for how many weeks the respondent goes to church per year (constructed based on a categorical question). The scale for strength of religious affiliation is: no religion (0); somewhat strong (1); not very strong (2); and strong (3). The scale for frequency of prayer is: never (0); less than once a week (1); once a week (2); several times a week (3); once a day (4); several times a day (5). The scale for belief in Afterlife is: No (0); Undecided (1); and Yes (2). And the scale for belief in God goes from “I don’t believe in God” (0) to “I know God really exists and I have no doubts about it” (5). For the measures of religious intensity (e.g. “Religious Attendance”), we present descriptive statistics for the raw variables and the interaction between those variables and dummy variables for whether the respondent is Catholic (e.g. “Religious Attendance, Catholic”) and whether the respondent is of a non-Catholic religion (e.g. “Religious Attendance, Other Religions”).

Table 4: Effect of scandals on religious adherence (GSS)

	Current adherence				Adherence growing up			
	(1) Catholic	(2) Other	(3) Catholic	(4) Other	(5) Catholic	(6) Other		
Post-Scandal	-0.062*** (0.023)	0.007 (0.021)	-0.063*** (0.024)	0.020 (0.023)	-0.021 (0.029)	0.014 (0.031)		
Pre-Scandal			-0.004 (0.020)	0.032 (0.020)	0.003 (0.025)	0.007 (0.025)		
Observations	25387	25387	25387	25387	23903	23903		
R-Squared	0.05	0.06	0.05	0.06	0.05	0.06		
No. of counties	333	333	333	333	333	333		

Notes: Heteroskedasticity-Robust standard errors in parentheses, clustered at the county level. Stars indicate significance level: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Data from the General Social Survey, 1994-2010. See Table 3 for descriptive statistics, and its footnote for more details about the data. Each column is a separate OLS regression. All regressions include county-specific fixed effects, time effects and control variables: gender, age, age squared, dummies for black and white, three dummies about marital status, household income, a set of four dummies about employment, education and number of children. Dependent variables are dummies for current adherence (e.g. in column (1) the dummy equals 1 if respondent declares being Catholic, 0 if not). Columns (5) and (6) are similar dummies, but constructed using a question on what religion the respondent was raised in.

Table 5: Effect of scandals on religious intensity (GSS)

	Church Attendance		Strength of Affiliation		Frequency of Prayer		Belief in Afterlife		Belief in God	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Catholic	Other	Catholic	Other	Catholic	Other	Catholic	Other	Catholic	Other
Post-Scandal	-1.600** (0.739)	1.054 (0.901)	-0.103** (0.042)	0.070 (0.048)	-0.152* (0.091)	-0.133 (0.092)	-0.039 (0.031)	0.022 (0.051)	-0.196* (0.109)	-0.015 (0.115)
Pre-Scandal	0.651 (0.738)	0.325 (0.641)	-0.026 (0.034)	0.099** (0.045)	-0.006 (0.068)	-0.059 (0.096)	-0.012 (0.031)	0.063 (0.080)	0.120 (0.155)	-0.032 (0.081)
Observations	23602	23602	23085	23085	15244	15244	18028	18028	10572	10572
R-Squared	0.22	0.14	0.46	0.33	0.40	0.30	0.38	0.26	0.45	0.33
No. of counties	333	333	333	333	333	333	333	333	322	322

Notes: Heteroskedasticity-robust standard errors in parentheses, clustered at the county level. Stars indicate significance level: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Data from the General Social Survey, 1994-2010. See Table 3 for descriptive statistics, and its footnote for more details about the data. Each column is a separate OLS regression. All regressions include county-specific fixed effects, time effects and control variables: gender, age, age squared, dummies for black and white, three dummies about marital status, household income, a set of four dummies about employment, education and number of children. In the even (odd) columns the dependent variables are the interaction between the measure of religiosity and the dummy for Catholic (other religions).

Table 6: Descriptive Statistics, charitable contributions (IRS SOI)

	Obs	Mean	Sd	Min	Max
Mean Charitable Cont. (/1000)	149178	0.85	1.51	0.00	89.34
Scandal in same zipcode	5670	1.04	1.34	0.02	28.16
Scandal in adjacent zipcode	27210	1.02	1.90	0.01	83.33
Neither	116298	0.80	1.40	0.00	89.34
Mean Income (/1000)	149178	44.62	33.67	-107.51	2258.84
Scandal in same zipcode	5670	51.31	36.34	13.96	597.50
Scandal in adjacent zipcode	27210	50.81	42.97	-50.83	1227.79
Neither	116298	42.85	30.71	-107.51	2258.84
Post-Scandal (Type I)	149178	0.03	0.19	0.00	6.00
Pre-Scandal (Type I)	149178	0.00	0.06	0.00	2.00
Post-Scandal (Type I), Adjacent	149178	0.23	0.60	0.00	8.00
Pre-Scandal (Type I), Adjacent	149178	0.02	0.16	0.00	5.00
Post-Scandal (Type II)	149178	0.05	0.28	0.00	11.00
Pre-Scandal (Type II)	149178	0.01	0.10	0.00	7.00

Notes: Data on charitable giving, income and dependents at the zip code level obtained from the Statistics of Income Division of the IRS. Charitable giving is the amount taxpayers reported as charitable contributions on line 18 of Schedule A from IRS Form 1040 (as of tax year 2001). Income denotes gross income, equivalent of line 33 of IRS Form 1040 (as of tax year 2001). The statistics are for the subset of zip codes that always have positive charitable contributions. The zip code means for contributions, income and dependents are computed by dividing the zip code totals over the total number of returns in the zip code. Charitable contributions and income are converted to 2008 dollars using the CPI-U from the Bureau of Labor Statistics. Post-Scandal is the stock of past scandals, and Pre-Scandal is the number of scandals during the following two years. The sample excludes zip codes that are non-standard, and zip codes that changed substantially over time.

Table 7: Effect of scandals on charitable contributions (IRS SOI)

	Ln(Charitable Contributions)				Ln(Income)
	(1)	(2)	(3)	(4)	(5)
Post-Scandal (Type I)	-0.019*** (0.006)	-0.023*** (0.007)	-0.017** (0.008)	-0.020*** (0.007)	0.007 (0.004)
Pre-Scandal (Type I)		-0.005 (0.008)	0.003 (0.009)	-0.006 (0.008)	0.001 (0.004)
Post-Scandal (Type I), Adjacent			-0.006* (0.003)		
Pre-Scandal (Type I), Adjacent			-0.007** (0.004)		
Post-Scandal (Type II)				-0.011** (0.005)	
Pre-Scandal (Type II)				0.007 (0.005)	
Observations	149178	149178	149178	149178	149159
R-Squared	0.49	0.49	0.49	0.49	0.68
No. of zipcodes	24863	24863	24863	24863	24863

Notes: Heteroskedasticity-Robust standard errors in parentheses, clustered at the zip code level. Stars indicate significance level: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. All the above are OLS regressions for the subset of zip codes that always have positive charitable contributions (income, in the case of column (5)). All regressions include zip code fixed effects, time effects, state-specific time trends, the logarithm of the number of returns, and the interaction between the time effects and the logarithm of population in the zip code, the logarithm of land area, and the share of urban population, all taken from the 1990 US Population Census. Columns (1) through (4) also include the logarithm of mean income (if positive), and the logarithm of the number of returns with Earned Income Tax Credit. Data on mean charitable giving and income by zip code is from the Statistics of Income Division of the IRS. See Table 6 for descriptive statistics, and its footnote for data definitions. Income denotes gross income. Charitable contributions and income are converted to 2008 dollars using the CPI-U from the Bureau of Labor Statistics. The sample excludes zip codes that are non-standard, and zip codes that changed substantially over time.

Table 8: Descriptive statistics, charitable contributions by religious affiliation (IRS Core)

	Obs	Mean	Sd	Min	Max
Charitable Contributions, Catholic	3379	0.52	2.06	0.00	48.11
Charitable Contributions, Non-Catholic	105289	0.39	3.03	0.00	292.72
Charitable Contributions, Protestant	30039	0.29	0.92	0.00	51.87
Post-Scandal	146599	0.99	2.00	0.00	26.00
Pre-Scandal	146599	0.20	0.74	0.00	21.00

Notes: Charitable contributions are expressed in million dollars, and are converted to 2009 dollars using the CPI-U from the Bureau of Labor Statistics. Data on charitable contributions for the period 1989-2009 is from the NCCS Core Files. Catholic denotes organizations with NTEE code X22 (Roman Catholic), Protestant denotes organizations with NTEE code X21 (Protestant), and non-Catholic denotes organizations within NTEE code X20 (Christian), X21 (Protestant), X30 (Jewish), X40 (Islamic), X50 (Buddhist) or X70 (Hindu). Post-Scandal is defined as the sum of Type-I and Type-II scandals that happened in the same zip code or adjacent zip codes in the past. Equivalently, Pre-Scandal is the sum of Type-I and Type-II scandals that will happen in the same zip code or adjacent zip codes during the following two years. The data is for the subsample of charities that report some contributions for all the years that they appear in the data. The sample excludes zip codes that are non-standard, and zip codes that changed substantially over time.

Table 9: Effect of scandals on charitable contributions by religious affiliation (IRS Core)

	Dep. var.: Ln(Charitable Contributions)					
	(1) Cath.	(2) Cath.	(3) Cath.	(4) Cath.	(5) Non-Cath.	(6) Prot.
Post-Scandal	-0.038** (0.018)	-0.038** (0.019)	-0.048** (0.021)	-0.036** (0.018)	-0.001 (0.005)	-0.008 (0.008)
Pre-Scandal		-0.001 (0.019)	0.004 (0.025)	0.002 (0.019)	0.000 (0.005)	0.009 (0.009)
Observations	3379	3379	2405	3028	105289	30039
R-Squared	0.07	0.07	0.11	0.09	0.01	0.02
No. of charities	389	389	290	358	15259	3966

Notes: Heteroskedasticity-Robust standard errors in parentheses, clustered at the zip code level. Stars indicate significance level: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. All the regressions above are OLS regressions for the subset of zip codes that have positive contributions in all of the years that they appear in the sample. All regressions include organization fixed effects, time effects, and the interaction between the time effects and the logarithm of population in the zip code, the logarithm of land area, the share of urban population, and the shares of whites, blacks, Asians and Hispanic, all taken from the 1990 US Population Census. Column (3) excludes charities that changed the city of address at least once. Column (4) excludes charities whose mean contributions are above one million (in 2009 dollars). Data on charitable contributions for the period 1989-2009 is from the NCCS Core Files. Post-Scandal is the stock of Type-I and Type-II scandals that took place in the same zip code of the charity or in an adjacent zip code. See Table 8 for descriptive statistics, and its footnote for data definitions. The data is for the subsample of organizations that report some contributions for all the years that they appear in the data. The sample excludes zip codes that are non-standard, and zip codes that changed substantially over time.

Table 10: Descriptive statistics, number and size of charities (ZBP)

	Obs	Mean	Sd	Min	Max
No. of establishments, Charitable	125760	8.53	11.42	1.00	385.00
Scandal in same zipcode	9432	13.03	13.35	1.00	183.00
Scandal in adjacent zipcode	29712	10.33	14.11	1.00	385.00
Neither	86616	7.43	9.86	1.00	186.00
No. of employees, Charitable (/100)	125760	1.64	2.94	0.03	97.15
Scandal in same zipcode	9432	2.62	3.28	0.03	47.31
Scandal in adjacent zipcode	29712	2.15	3.54	0.03	97.15
Neither	86616	1.36	2.61	0.03	62.40
No. of establishments, Retail	180852	7.08	7.51	1.00	152.00
No. of employees, Retail (/100)	180852	2.59	3.31	0.03	37.44
Post-Scandal (Type I)	312348	0.03	0.18	0.00	6.00
Pre-Scandal (Type I)	312348	0.01	0.07	0.00	4.00
Post-Scandal (Type I), Adjacent	312348	0.20	0.56	0.00	8.00
Pre-Scandal (Type I), Adjacent	312348	0.04	0.23	0.00	6.00
Post-Scandal (Type II)	312348	0.04	0.26	0.00	11.00
Pre-Scandal (Type II)	312348	0.01	0.11	0.00	7.00

Notes: Zip code-level data on number of establishments and number of employees is from the Zipcode Business Patterns. The statistics for the number of establishments and employees is for the subset of zip codes that always have positive values for the respective variable. The number of employees is a proxy variable constructed as the sum of the number of establishments in each employment-size-group, weighted by the average number of employees in the corresponding category. For instance, if a zip code has 1 establishment with 1 to 4 employees and 5 establishments with 5 to 9 employees, the proxy for number of employees takes the value $2.5+7=9.5$. The charitable establishments include organizations for human rights, other social advocacy organizations, child and youth services, among many others (see the document for a full list and NAICS codes). Retail establishments are: supermarkets, new car dealers and used car dealers. The data on charitable (retail) organizations is for the subset of zip codes that always have at least one charitable (retail) establishment. Post-Scandal is the stock of past scandals, and Pre-Scandal is the number of scandals during the following two years. The sample excludes zip codes that are non-standard, and zip codes that changed substantially over time.

Table 11: Effect of scandals on number and size of charities (ZBP)

	Number of Employees				Establishments	
	(1) Charity	(2) Charity	(3) Charity	(4) Charity	(5) Retail	(6) Charity
Post-Scandal (Type I)	-0.085*** (0.028)	-0.103*** (0.034)	-0.099*** (0.038)	-0.084** (0.035)	-0.007 (0.020)	-0.024 (0.016)
Pre-Scandal (Type I)		-0.013 (0.022)	-0.008 (0.024)	-0.006 (0.022)	-0.000 (0.013)	-0.009 (0.011)
Post-Scandal (Type I), Adjacent			-0.003 (0.014)			
Pre-Scandal (Type I), Adjacent			-0.005 (0.010)			
Post-Scandal (Type II)				-0.047*** (0.018)		
Pre-Scandal (Type II)				-0.017 (0.014)		
Observations	125760	125760	125760	125760	180852	125760
R-Squared	0.07	0.07	0.07	0.07	0.01	0.11
No. of zipcodes	10480	10480	10480	10480	15071	10480

Notes: Heteroskedasticity-Robust standard errors in parentheses, clustered at the zip code level. Stars indicate significance level: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. All the regressions above are OLS regressions. The data on charitable (retail) organizations is for the subset of zip codes that always have at least one charitable (retail) establishment. All regressions include zip code fixed effects, time effects, and the interaction between the time effects and the logarithm of population in the zip code, the logarithm of land area, the share of urban population, and the shares of whites, blacks, Asians and Hispanic, all taken from the 1990 US Population Census. Data on number of establishments and number of employees is from the Zipcode Business Patterns. See Table 10 for descriptive statistics, and its footnote for data definitions. The sample excludes zip codes that are non-standard, and zip codes that changed substantially over time.

Table 12: Effect of scandals on measures of trust

	Trust		Helpful		Fair		Average	
	(1) All	(2) Cath.	(3) All	(4) Cath.	(5) All	(6) Cath.	(7) All	(8) Cath.
Post-Scandal	-0.084 (0.056)	-0.018 (0.098)	0.075 (0.064)	0.093 (0.084)	0.051 (0.064)	0.038 (0.082)	0.023 (0.045)	0.040 (0.061)
Pre-Scandal	-0.043 (0.048)	-0.162* (0.084)	0.131** (0.052)	0.203** (0.096)	0.121** (0.053)	0.073 (0.105)	0.073** (0.035)	0.041 (0.066)
Observations	16459	4557	14079	4268	14002	4248	13918	4225
R-Squared	0.09	0.09	0.06	0.05	0.08	0.06	0.12	0.11
No. of counties	332	313	332	313	332	313	332	313
Mean of Dep. Var.	0.74	0.76	1.01	1.02	1.12	1.14	0.96	0.97
Sd of Dep. Var.	0.94	0.94	0.96	0.95	0.95	0.94	0.73	0.72

Notes: Heteroskedasticity-Robust standard errors in parentheses, clustered at the zip code level. Stars indicate significance level: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Each column is a separate OLS regression. All regressions include county-specific fixed effects, time effects and control variables: gender, age, age squared, dummies for black and white, three dummies about marital status, household income, a set of four dummies about employment, education and number of children. Data from the General Social Survey, 1994-2010. The dependent variable Trust is the answer to the question “Generally speaking, would you say that most people can be trusted or that you can’t be too careful in dealing with people? Can’t be too careful (0); Depends (1); Most people can be trusted (2).” The variable Helpful is the answer to the question: “Would you say that most of the time people try to be helpful, or that they are mostly just looking out for themselves? Just look out for themselves (0); Depends (1); Try to be helpful (2).” The variable Fair is the answer to the question: “Do you think most people would try to take advantage of you if they got a chance, or would they try to be fair? Would take advantage of you (0); Depends (1); Would try to be fair (2).” The variable Average is just the average responses to the previous three variables. The header “All” means that the regression includes all respondents. The header “Cath.” means that the regression only includes individuals that declared to be raised as Catholic. Post-Scandal is the stock of scandals in the county of the respondent at the moment of the interview. Pre-Scandal is the number of scandals during the two years following the interview. Post- and Pre-Scandal are normalized by dividing by 100,000 inhabitants in the county, using data from the 1990 US Population Census.

Table 13: Effect of abuse scandals in lay organizations on pro-social behavior

	Ln(Char. Cont.)		Ln(Char. Employ.)	
	(1)	(2)	(3)	(4)
Post-Scandal, Same	-0.025 (0.016)	-0.013 (0.016)	0.043 (0.056)	-0.003 (0.061)
Pre-Scandal, Same	-0.005 (0.009)	0.002 (0.010)	-0.029 (0.033)	-0.072* (0.038)
Post-Scandal, Adjacent		-0.011** (0.005)		0.042* (0.022)
Pre-Scandal, Adjacent		-0.007* (0.004)		0.041** (0.018)
Observations	149178	149178	129624	129624
R-Squared	0.50	0.50	0.07	0.07
No. of zipcodes	24863	24863	10802	10802

Notes: Heteroskedasticity-Robust standard errors in parentheses, clustered at the zip code level. Stars indicate significance level: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Post-Scandal is the stock of past scandals, and Pre-Scandal is the number of scandals during the following two years. The dependent variable in the first two columns is the log of mean charitable contributions in the zip code, using the IRS SOI dataset. The dependent variable in the last two columns is the number of employees in charitable establishments in the zip code, using the ZBP dataset. The specifications are the same than those explained in the footnote of Tables 7 and 11, except that we include a time-trend specific to the zip codes with scandals.

A Explaining the geographic distribution of the scandals

The following subsections compare the characteristics of counties (zip codes) according to the number of scandals that they experience, which is useful to understand the geographic distribution of scandals.

A.1 Across counties

As a thought experiment, imagine that we go back in time to the year 2000 and we are asked to predict the number of scandals that will take place in the period 2001-2010 in each of the counties in the US. In order to provide an answer, we can estimate a Negative Binomial Regression Model where the dependent variable is the number of scandals that actually happened during that period and the independent variables are characteristics of those counties (e.g. population size) as of the year 2000. The most obvious predictor of the number of scandals is the size of the Catholic congregation: i.e., if more Catholics live there, then there will be more priests there, and thus the number of potential victims and perpetrators is higher. But we will also see whether the size of other religious congregations has any predictive power.

In order to obtain measures of the size of the religious congregations in the US counties, we use the Religious Congregations and Membership Study (RCMS), plausibly the most complete census of religious congregations and their members in the United States. The data is collected by the Association of Statisticians of American Religious bodies and distributed by the Association for Religion Data Archives. The data is decennial and offers county-level aggregates. In the year 2000, participants included 149 Christian denominations, associations, or communions (including Latter-day Saints and Unitarian/Universalist groups); two specially defined groups of independent Christian churches; Jewish and Islamic totals; and counts of temples for six Eastern religions.⁴² Apart from the number of congregations, they measure the number of adherents, which include all full members, their children, and others who regularly attend services or participate in the congregation.⁴³ We dropped from the sample 30 counties that had either zero or missing value in the number of Catholic congregations or Catholic adherents, leaving us with a sample of 2956 counties.

We report the results for the subset of scandals that happened in 2001-2010, but the results are qualitatively the same if we use a different time period instead. Figure 7.a shows the distribution of scandals across counties: while 89% (87.5%) of counties have no Type-I (Type-II) scandals, 7% (6.5%) of counties have exactly one Type-I (Type-II) scandal, and the remaining 4% (6%) counties have 2 or more Type-I (Type-II) scandals. There are a handful of counties that have several scandals. For instance, the counties with most scandals are Cook county in Illinois (15 Type-I scandals and 46 Type-II scandals), Jefferson County in Kentucky (12 and 44), Queens County in New York (10 and 22), Philadelphia County in Pennsylvania (7 and 25), and many counties in Massachusetts (e.g. Middlesex, Suffolk, Essex, Norfolk, Worcester). The large number of scandals, however, simply reflects

⁴²While these data contain membership data for many religious groups in the United States, including most of the larger groups, they do not include every group. See for example the discussion in Finke and Scheitle (2005).

⁴³When religious groups reported only adult membership, the RCMS used the following formula to derive the number of adherents: the total county population was divided by the total county population less children 13 years and under (from US Census data), and the resulting figure was multiplied by the confirmed members. Using total adherents allows for more meaningful comparisons between groups that count children as members (e.g. Catholics) and those that don't (e.g. Baptists).

the fact that those are counties with a large Catholic population, and thus the pool of clergymen and churchgoers that can potentially become involved in a scandal is larger. This can be clearly seen from Figure 7.b, which shows that the number of scandals is strongly correlated to the number of Catholic adherents in the county.

The results from the regression analysis are shown in Table 15. Columns (1) to (3) show the results using the number of Type-I scandals as dependent variable, while columns (4) to (6) reproduce the results for the Type-II scandals. The coefficients reported are incidence-rate ratios. For example, the coefficient on the log of number of Catholics in column (1) is 2.15, which means that an increase of 0.1 logs in the number of Catholic adherents (around 10%) would increase the expected number of scandals by about 115%. Holding constant the number of Catholics and the number of non-adherents, the expected number of scandals does not seem to be influenced by the number of non-Catholic adherents. The expected number of scandals does seem to be increasing in the number of non-adherents. There are many potential explanations for that correlation. We believe that a great deal of the correlation responds to the fact that the number of Catholic adherents is imperfectly measured by the survey data, so the number of non-adherents captures the residual effect of such measurement error.⁴⁴ Indeed, in columns (2) and (4) we introduce the number of Catholic congregations as an additional proxy for the size of the Catholic congregation, and their coefficients are statistically and economically significant. This suggests that the variable on the number of Catholic adherents was not capturing the totality of the effect of the size of the Catholic congregation on the number of scandal events.

In columns (3) and (6) we included some county characteristics from the 2000 census, like the racial composition and the median income. All these variables are standardized by dividing them by their corresponding standard deviation, so the coefficients in the table can be directly interpreted as the incidence-rate ratios for a one-standard deviation increase in the corresponding variable. Column (3) shows that, other than the size of the Catholic population, Type-I scandals are not significantly associated with any of the basic county characteristics introduced. In terms of the previous thought experiment, if we were asked to predict the number of Type-I scandals in a particular county, we would only focus on total population and the size of its Catholic congregation. Those same basic county characteristics do have predictive power in explaining the Type-II scandals (e.g. they are more likely to happen in counties with higher population density).

A.2 Across zip codes

Zip codes were created by the U.S. Postal Service as a tool to help deliver the mail more efficiently. Unsurprisingly, they are not ideal from the perspective of research in social science. Fortunately, if we observe thousands of zip codes over several years, the statistical power is so high that the measurement error introduced by the imperfections of the zip code geography is not a major concern. We took some measures in order to ameliorate these sources of measurement error. For obvious reasons, we always drop all the non-standard zip codes.⁴⁵ Unfortunately, unlike census tracts, when a zip code changes its

⁴⁴ Additionally, note that the scandals are probably not a function of the size of the Catholic congregation as of 2000, but some average between the size in year 2000 and the size over the past decades, when the abuses took place (this is specially important for the Type-II scandals).

⁴⁵ According to our data on zip code characteristics, of the 42,127 zip codes in the year 2000, 81% were standard zip codes, 12% were P.O. Box only, 5% were unique (assigned to a single high-volume address, like some universities), and the rest were military.

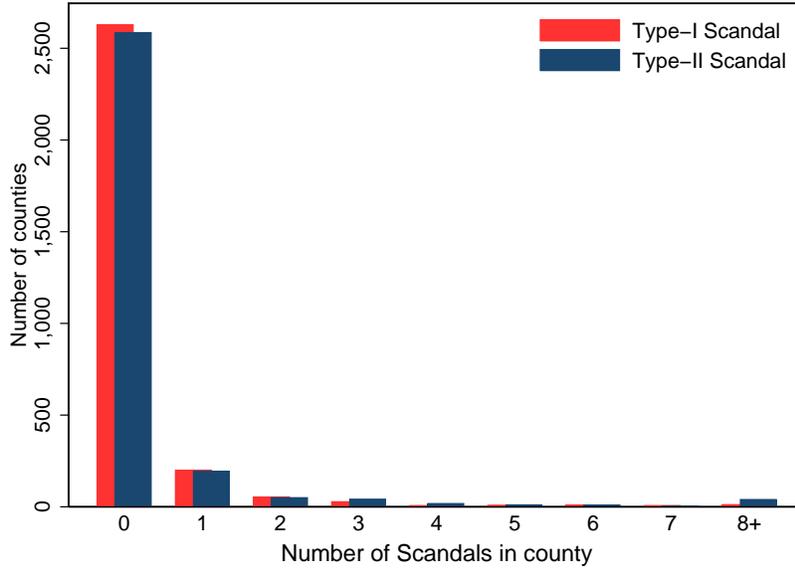
definition it does not change its name. In a few cases these changes are dramatic, but more commonly they are small and subtle. To the best of our knowledge, there is no existing geographic equivalency file relating zip codes at different points in time. In order to minimize this source of measurement error, we dropped zip codes from the data whenever we find evidence that they changed their definition over time in a substantial manner.⁴⁶

We are also interested in measuring the spillovers from one zip code to its neighboring zip codes. To create a database of adjacent zip codes, we used the cartographic boundaries of the 2000 ZIP Census Tabulation Areas (ZCTA) created by the US Census Bureau, where a 5-digit ZCTA is typically nearly identical to a 5-digit USPS zip code. Two zip codes are adjacent to each other if their boundaries touch. For example, Figure 8 shows zip code 02121 and those zip codes that are adjacent to it. All zip codes have at least one adjacent zip code. Among all the standard zip codes in year 2000, the average (standard deviation) of the number of adjacent zip codes is 5.39 (2.85). For those zip codes with at least one Type-I scandal, the average (standard deviation) of the number of adjacent zip codes is 6.78 (3.45).

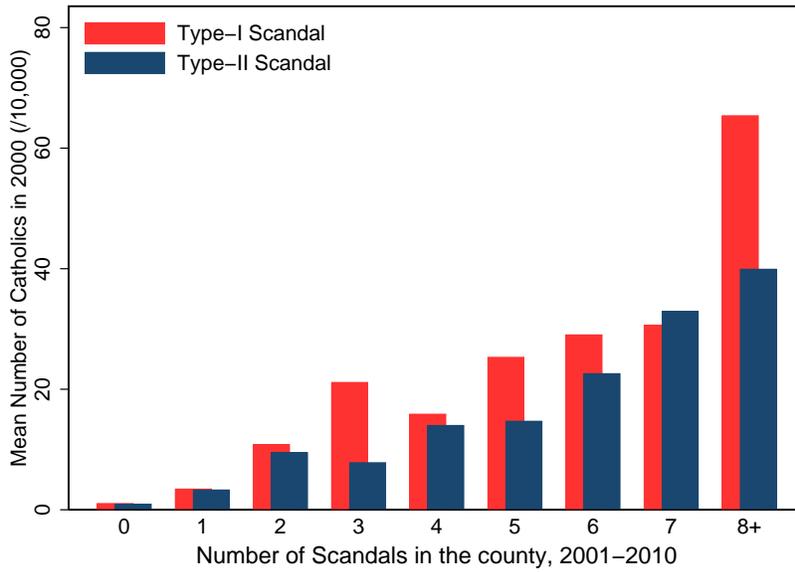
We can examine whether the zip codes that experienced scandals are different in some dimension from the zip codes that did not suffer any scandals, and from their adjacent zip codes. Unfortunately, we do not have zip code-level data on the size of religious congregations, so we cannot perform the same analysis that we did for the county-level data in the previous subsection. Nevertheless, we can use Census data on several interesting socio-economic indicators. Table 16 provides a comparison of a selection of zip code characteristics (e.g. population, land size) among those where there was a scandal, zip codes that did not suffer any scandals but are adjacent to a zip code that suffered a scandal, and zip codes that did not suffer scandals and are not adjacent to a zip code with scandals either. The statistics are shown separately for Type-I and Type-II scandals, but the findings are qualitatively similar across the two. The first row provides the most important fact: scandals tend to happen in zip codes that are more populous. The difference is not only statistically significant, but also economically significant. The reason is straightforward: more people means more Catholic churchgoers and priests, so the potential number of abusers and victims is higher. The second and third rows are an immediate consequence of the first row: since more populous zip codes are on average smaller in area and more urban, scandals tend to happen in zip codes that are smaller and more urban. The differences in other dimensions are sometimes statistically significant, but economically not very significant. Furthermore, the differences in those other variables are most likely explained by the differences in population size: e.g. median income is increasing in population density, and since zip codes with scandals have higher density it is not surprising that median income is a little higher in the zip codes that had scandals. Finally, note that zip codes with scandals are more similar to those adjacent to the scandals than to the rest of the zip codes.

⁴⁶We employ two databases from the Census Bureau, one about zip codes in 1990 and the other about zip codes as of 2000. Around 1,400 zip codes appear in one database but not in the other, which probably means that the zip codes were created in the interim, so we drop them from the data. Those two databases include a proxy for the centroid of each zip code. The centroids may differ because one zip code was terminated and then that same zip code number was used somewhere else in the US territory. But they may also differ because of measurement error, or because of meaningless changes in geography: e.g. a zip code may “incorporate” a great portion of water area, or a great portion of land area that is not inhabited (e.g. desert). We drop 70 zip codes whose centroids differ considerably (50 miles or more). The results that follow are practically the same if we use very different thresholds to drop zip codes, which suggests that the results in this paper are not sensitive to this source of measurement error.

Figure 7: Distribution of scandals across counties, 2001-2010



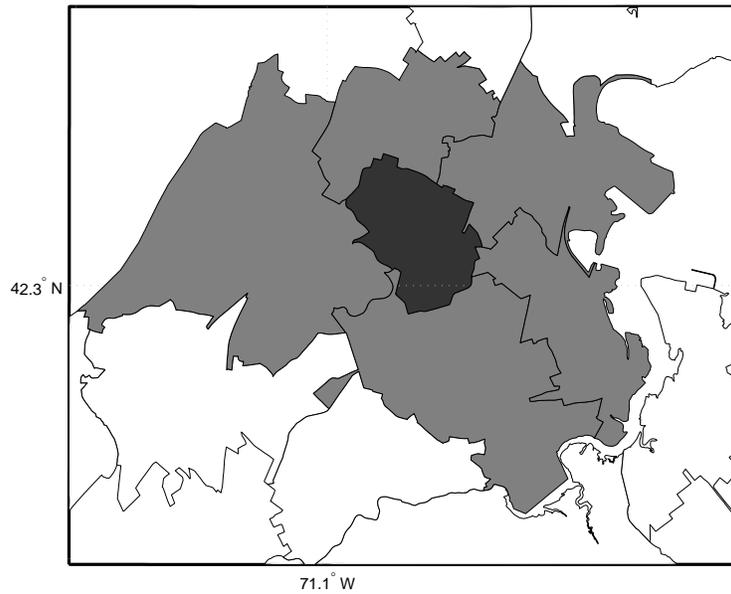
a. Histogram of scandals



b. Number of scandals and Catholic population

Notes: Data on mean number of Catholics by county in year 2000 obtained from the Religious Congregations and Membership Study (RCMS). The count of scandals correspond to the totality of scandals in the period 2001-2010. The data on scandals was compiled by the authors. See Section 2 for a description of the data, and footnote to Figure 2 for a brief definition of Type-I and Type-II scandals.

Figure 8: Example of adjacent zip codes



Notes: The shapes denote a group of zip codes in the Boston metropolitan area. The darkest shape is zip code 02121, and the shapes in grey denote its adjacent zip codes. The geographic boundaries for ZIP Code Tabulation Areas were obtained from the 2010 TIGER/Line® Shapefiles, prepared by the Census Bureau.

Table 14: Summary statistics (RCMS)

	Obs	Mean	Sd	Min	Max
No. of Type-I Scandals	2986	0.24	1.02	0.00	15.00
No. of Type-II Scandals	2986	0.43	2.21	0.00	46.00
No. of Catholic Adherents (/10000)	2986	2.08	10.38	0.00	380.64
No. of Non-Catholic Adherents (/10000)	2986	2.63	6.61	0.00	172.24
No. of Non-Adherents (/10000)	2986	4.66	14.31	0.00	399.05
No. of Catholic Congregations	2986	7.30	16.53	0.00	370.00
No. of Non-Catholic Congregations	2986	81.00	133.50	0.00	3766.00
Land Area	2986	2.97	9.55	0.01	377.88
Share Urban Pop. (%)	2986	41.60	30.63	0.00	100.00
Share White Pop. (%)	2986	84.79	16.24	4.51	99.74
Share Hispanic Pop. (%)	2986	6.37	12.25	0.08	97.54
Share College Graduates (%)	2986	16.74	7.81	4.92	63.75
Share Unemployed (%)	2986	5.77	2.69	0.21	33.03
Mean HH Income (/1000)	2986	36.76	9.23	13.12	85.72

Notes: Data on county-level religious adherence obtained from the 2000 Religious Congregations and Membership Study (RCMS). The number of scandals is the count of scandals that became public after the year 2000. Non-Catholic adherents (congregations) are the sum of adherents (congregations) over the 148 non-Catholic denominations in the data. The number of non-adherents is obtained by subtracting the number of Catholic and non-Catholic adherents from the total population in the county. Other county characteristics obtained from the 2000 Census. Share college is the share of the population 25 years old or older that completed college education, and share unemployed is the unemployment rate.

Table 15: County characteristics associated with the scandals

	Type-I Scandals			Type-II Scandals		
	(1)	(2)	(3)	(4)	(5)	(6)
Ln(Catholic Adherents)	2.158*** (0.114)	1.545*** (0.148)	1.657*** (0.202)	2.332*** (0.126)	1.476*** (0.136)	1.423*** (0.167)
Ln(Non-Catholic Adherents)	0.897 (0.082)	0.842 (0.149)	0.679* (0.156)	0.868 (0.087)	0.828 (0.141)	0.672* (0.153)
Ln(Non-Adherents)	1.199** (0.111)	1.218* (0.133)	1.246* (0.166)	1.414*** (0.133)	1.493*** (0.160)	1.464*** (0.178)
Ln(Catholic Congregations)		1.662*** (0.200)	1.586*** (0.219)		1.970*** (0.247)	1.987*** (0.277)
Ln(Non-Catholic Congregations)		1.120 (0.286)	1.434 (0.464)		1.068 (0.272)	1.353 (0.437)
Land Area			1.013 (0.090)			0.626*** (0.083)
Share Urban Pop.			0.980 (0.124)			1.507*** (0.218)
Share White Pop.			1.150 (0.131)			0.999 (0.143)
Share Black Pop.			1.110 (0.140)			0.699** (0.117)
Share Hispanic Pop.			0.869* (0.064)			0.776*** (0.056)
Share College Graduates			0.973 (0.063)			1.047 (0.069)
Share Unemployed			0.903 (0.099)			1.196* (0.117)
Ln(Mean HH Income)			1.068 (0.102)			0.938 (0.091)
Observations	2956	2956	2956	2956	2956	2956
Pseudo-R-Squared	0.27	0.28	0.29	0.29	0.30	0.32

Notes: The dependent variable is the number of scandals in the county that became public after the year 2000. Coefficients are incidence-rate ratios from a Negative Binomial Regression. Heteroskedasticity-robust standard errors in parentheses. Stars indicate significance level: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Data at the county level presented for the year 2000 on religious adherence obtained from the Religious Congregations and Membership Study (RCMS). We excluded counties that had missing values in one or more of the adherence variables. Data on county characteristics obtained from the 2000 US Population Census. The county characteristics that appear in columns (3) and (6) have been standardized by dividing them by their corresponding standard deviation. See Table 14 for descriptive statistics about the data, and its footnote for data definitions.

Table 16: Zip code characteristics associated to the scandals

	Type-I Scandal			Type-II Scandal		
	(1) Same	(2) Adjacent	(3) Neither	(4) Same	(5) Adjacent	(6) Neither
Population (/1000)	23.049 (16.107)	12.954 (14.174)	8.257 (11.293)	26.381 (17.547)	12.335 (13.215)	7.283 (10.068)
Area (miles)	57.762 (129.759)	66.829 (133.443)	95.198 (184.029)	43.895 (114.392)	67.761 (133.902)	100.663 (193.264)
Share Urban (%)	77.209 (34.002)	49.500 (45.021)	31.427 (41.198)	84.416 (28.789)	50.022 (44.573)	27.943 (39.420)
Share White (%)	83.657 (21.599)	86.984 (20.282)	88.428 (18.279)	81.299 (22.811)	86.839 (20.174)	89.147 (17.528)
Share Black (%)	9.730 (18.477)	7.963 (17.186)	7.636 (15.991)	10.564 (19.571)	7.604 (16.313)	7.364 (15.583)
Share Indian/Eskimo (%)	0.597 (2.152)	0.889 (4.402)	1.107 (5.187)	0.590 (1.388)	0.801 (3.215)	1.132 (5.379)
Share Asian (%)	2.512 (5.787)	1.775 (5.366)	1.008 (3.667)	3.101 (7.658)	1.899 (5.690)	0.815 (3.074)
Share Hispanic (%)	7.982 (15.908)	5.458 (12.560)	3.985 (10.424)	10.135 (18.073)	5.978 (13.505)	3.447 (9.278)
Average HH Size	2.625 (0.327)	2.673 (0.345)	2.663 (0.325)	2.627 (0.363)	2.674 (0.361)	2.663 (0.313)
Share college (%)	7.463 (5.770)	6.199 (5.418)	5.061 (4.567)	8.050 (6.510)	6.097 (5.131)	4.877 (4.488)
Share school (%)	17.319 (4.159)	18.221 (4.552)	18.927 (4.554)	17.003 (4.340)	18.277 (4.690)	19.033 (4.481)
Median Income (/1000)	32.337 (12.796)	31.209 (13.390)	27.093 (10.704)	32.784 (12.515)	31.451 (13.705)	26.636 (10.371)
Observations	851	3288	18462	1072	3219	17160

Notes: Data on zip code characteristics is taken from the 1990 US Population Census. Data on scandals was compiled by the authors. See Section 2 for a description of the data, and footnote to Figure 2 for a brief definition of Type-I and Type-II scandals. The zip codes under the title “Same” are those that were affected by at least one scandal. The zip codes under “Adjacent” are those that were never affected by a scandal, but are adjacent to at least one zip code in the group “Same.” Zip codes under “Neither” are those zip codes that do not belong to “Same” nor “Adjacent.” The sample excludes zip codes that are non-standard, and zip codes that changed substantially over time.

B Effect of the scandals on the number and size of religious establishments in the zip code

In this appendix we look at the effects of the scandals on the number of religious establishments in a zip code. The data comes from the Zip Business Patterns (ZBP), which is described in subsection 4.3. We focus on establishments that belong to the category “religious organizations” (SIC code 8660 and NAICS code 813110): bible societies, churches, convents, missions, monasteries, mosques, places of worship, religious organizations, religious retreat houses, religious shrines, synagogues and religious temples. Other establishments maintained by religious organizations, such as educational institutions, hospitals, publishing houses, reading rooms, social services, and secondhand stores, are classified according to their primary activity, and therefore do not belong to this category. The zip code data is available annually since 1994. The sample from 1998-2008 employed the North American Industry Classification System (NAICS) to classify the businesses, while the 1994-1997 sample employed the Standard Industrial Classification (SIC) system. Fortunately, there is a perfect one-to-one correspondence between the SIC and NAICS codes for religious establishments. Also, the category of religious establishments was not affected by the changes in the NAICS system introduced in 2002 and 2007. As a result, we can use the full sample 1994-2009.

There are some caveats with the data. First, it corresponds to the totality of religious establishments, and not only to the Catholic establishments. However, the evidence in the rest of the paper suggested that the scandals did not affect non-Catholic religious organizations, and thus the effect on the total number of religious establishments should serve as an approximation for the effect of the scandals on the total number of Catholic establishments. A second issue with the data is that it aggregates over establishments that, even if they were all Catholic, would still be quite diverse: from churches to retreat houses to missions. This source of measurement error will certainly weaken the precision of the estimates, and it will also make the interpretation of the results less straightforward, but we have no reasons to believe that it will introduce systematic biases.

Just like in subsection 4.3, we created a proxy variable for the number of employees working in the religious sector in a zip code, defined as the sum of the number of establishments in each employment-size-group, weighted by the average number of employees in the corresponding category. As a robustness test, apart from the number of religious establishments in a zip code, we will also look at the total number of employees working in religious establishments in that zip code. This can ameliorate some forms of measurement error: e.g. when two or more parishes are merged into a single parish, the number of religious establishments goes down, but it is not necessarily because there was a decrease in the number of adherents. Also, note that if a zip code is affected by a scandal in may of 1999, we do not know if the number of establishments that we observe in 1999 in that zip code corresponds to the number of establishments before or after that scandal took place. In order to ensure that the coefficient on Post-Scandal is not contaminated by this uncertainty, in all the regressions in this paper we introduce an extra variable which takes the value of the number of scandals that happened in that zip code during that year. Thus, the coefficient on Post-Scandal will only rely on the comparison between the years before and after the year of the scandal, but not the same year of the scandal. In practice, it makes virtually no difference whether we include this additional variable or not.

Since the dependent variable is the number of all religious organizations, and not just the number of Catholic parishes, it is less clear how we expect the scandals to affect this variable. We suspect that zip codes with more religious establishments will experience a larger drop in the number of establishments, since the scandal may lead to the closure of the institution involved in the scandal (e.g. the parish) and also the closure of other religious establishments frequented by the same Catholics that attend the institution of the scandal (e.g. a bible society close to the parish of the scandal). In order to let the regression account for this patterns, we use as dependent variable the square root of the number of religious establishments (and employees). Intuitively, the marginal effect of a scandal will be proportional to the square root of the number of establishments in the zip code, so scandals in zip codes with more religious establishments experience a larger marginal effect from a scandal.⁴⁷

Table 17 provides basic descriptive statistics. Just to mention some, the average zip code has 6.59 religious establishments and 64.11 employees. The regression results are shown in Table 18. According to the coefficient on Post-Scandal from column (1), following a scandal there is a permanent decrease of 0.04 in the square root of the number of religious establishments in the zip code of the scandal. This coefficient implies that if we take a zip code with just 1 religious establishment, then the expected effect of one scandal would be a decrease of $2 \cdot 0.04 \cdot \sqrt{1} = 0.08$ in the number of religious establishments (i.e. one every twelve scandals are followed by the closure of a religious establishment). We will see in Figure 9 that the effect intensifies over time, so the permanent effect is greater than the one reported by the coefficient on Post-Scandal. Column (2) adds the variable Pre-Scandal. The coefficient on Post-Scandal is very similar to that in column (1), and the coefficient on Pre-Scandal is not statistically significant, which suggests that there are no differential pre-trends between zip codes that were affected by scandals and zip codes that were not. Figure 9 shows the event-study analysis. Figure 9.a confirms that the effect on the number of religious establishments starts exactly when the scandal becomes public. As seen repeated times in the paper, the effects of the scandals on the number of religious establishments intensify over time.

Next, we want to explore whether the effects of the scandals are confined to the very same zip code where the scandal took place, or whether they extend to neighboring zip codes. Given the definition of the dependent variable, it is reasonable to expect some spillovers. Just to give an example, people attending parishes in the zip code of a scandal are likely to participate in religious organizations located in neighboring zip codes, and thus a negative shock to the parishioners in a given zip code will spillover to the religious establishments in adjacent zip codes. However, we should find that the effects of the scandals on the adjacent zip codes are not larger than the effects on the very same zip code where the scandal takes place. Indeed, the coefficients in column (3) suggest that the effect of a scandal on the number of religious establishments in the adjacent zip code is statistically significant, and in magnitude it is roughly half as large as the effect of a scandal on the number of religious establishments in the very same zip code of the scandal.

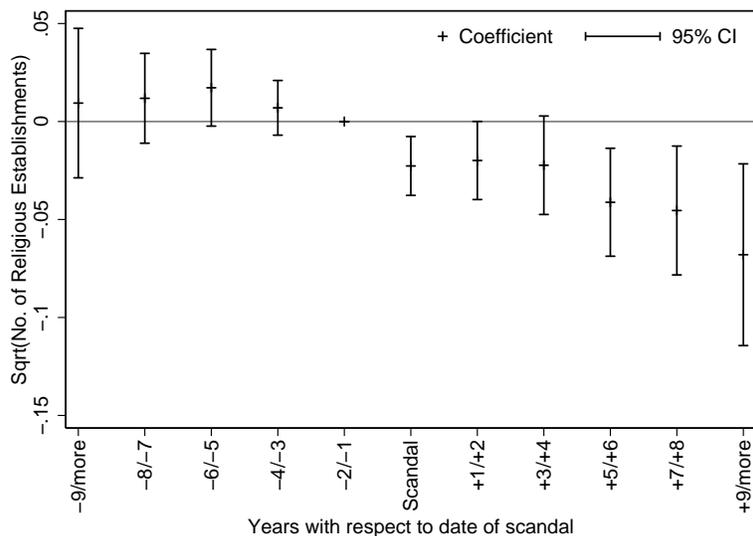
Column (4) introduces simultaneously the Type-I and Type-II scandals, and both types seem to have an effect on the number of religious establishments (recall that we cannot compare directly the coefficient on “Post-Scandal (Type-I)” in column (4) to those in columns (1) through (3)). Intuitively,

⁴⁷The basic results are qualitatively the same if we do not use the square root transformation, or if we use a Poisson Model. We presented this specification because it is by far the one most preferred by the data, as manifested by more precise estimates.

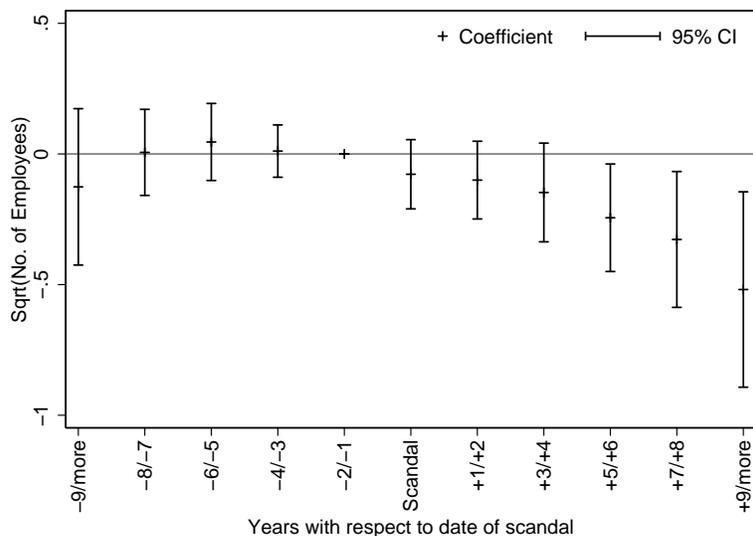
when a priest working in town A is publicly accused about having abused 20 years ago while working in town B (often from a different state than A), the number of religious establishments goes down not only in town A but it also goes down in town B. Even though the coefficient on “Post-Scandal (Type-II)” is larger than that of “Post-Scandal (Type-I)”, the coefficient on “Pre-Scandal (Type-I)” is statistically significant, so we should not jump to any conclusions about the relative importance between the Type-I and Type-II scandals.

Column (5) of Table 18 and Figure 9.b show that the results are practically the same when we use the number of employees as dependent variable instead of the number of establishments. The coefficient on Post-Scandal in column (2) amounts to 2% of the average of the corresponding dependent variable. The coefficients in column (5), which uses the number of employees instead of the number of establishments, suggests a similar magnitude: the coefficient is around 3% of the mean of the corresponding dependent variable. It is important to note that the effect is not very large relative to the mean of the dependent variable because the dependent variable is the number of *all* religious establishments, not only the *Catholic* ones. The magnitude of the effect relative to the mean number of *Catholic* establishments is probably many times higher. In summary, the results from this appendix are consistent with the findings presented in Section 3.

Figure 9: Effect of Scandals on number and size of religious establishments



a. Number of religious establishments



b. Number of employees in religious establishments

Notes: This is a graphical representation of an OLS regression of the number of establishments (employees) in the religious industry on a set of variables describing the timing of abuse scandals, $d_{c,t}^s$'s. The “-2/-1” is the omitted category (i.e. its coefficient is normalized to zero). The regression also includes zip code fixed effects, time effects and the interaction between the time effects and the logarithm of population in the zip code, the logarithm of land area, and the share of urban population, all taken from the 1990 US Population Census. Each bar represents the 95% confidence interval, and the center of the bar represents the corresponding point estimate. Confidence intervals are constructed with heteroskedasticity-robust standard errors, clustered at the zip code level. Data is from the Zipcode Business Patterns for the period 1994-2009. See Table 17 for descriptive statistics, and its footnote for data definitions. The sample excludes zip codes that are non-standard, and zip codes that changed substantially over time.

Table 17: Descriptive statistics, religious establishments (ZBP)

	Obs	Mean	Sd	Min	Max
No. of Rel. Establishments	382000	6.32	7.24	0.00	71.00
Scandal in same zipcode	15088	12.43	8.61	0.00	54.00
Scandal in adjacent zipcode	70192	7.91	7.67	0.00	71.00
Neither	296720	5.64	6.84	0.00	68.00
No. of Rel. Employees	382000	64.11	117.83	0.00	2761.50
Scandal in same zipcode	15088	155.04	175.76	0.00	2017.50
Scandal in adjacent zipcode	70192	90.72	136.32	0.00	2117.50
Neither	296720	53.19	105.82	0.00	2761.50
Sqrt(No. of Establishments)	382000	2.12	1.35	0.00	8.43
Sqrt(No. of Rel. Employees)	382000	5.90	5.41	0.00	52.55
Post-Scandal (Type I)	382000	0.02	0.17	0.00	6.00
Pre-Scandal (Type I)	382000	0.00	0.07	0.00	4.00
Post-Scandal (Type I), Adjacent	382000	0.17	0.52	0.00	8.00
Pre-Scandal (Type I), Adjacent	382000	0.03	0.21	0.00	6.00
Post-Scandal (Type II)	382000	0.04	0.24	0.00	11.00
Pre-Scandal (Type II)	382000	0.01	0.10	0.00	7.00

Notes: Data from the Zipcode Business Patterns for the period 1994-2009. The data is zip code-level data for the subset of religious establishments (SIC code 8660 and NAICS code 813110): bible societies, churches, convents, missions, etc. Other establishments maintained by religious organizations (e.g. schools) do not belong here. The number of employees is a proxy variable constructed as the sum of the number of establishments in each employment-size-group, weighted by the average number of employees in the corresponding category. For instance, if a zip code has 1 establishment with 1 to 4 employees and 5 establishments with 5 to 9 employees, the proxy for number of employees takes the value $2.5+7=9.5$. The sample excludes zip codes that are non-standard, and zip codes that changed substantially over time.

Table 18: Effect of scandals on number of religious establishments

	Number of establishments				Employees
	(1)	(2)	(3)	(4)	(5)
Post-Scandal (Type I)	-0.039*** (0.012)	-0.042*** (0.013)	-0.026* (0.014)	-0.026* (0.014)	-0.191** (0.097)
Pre-Scandal (Type I)		-0.011 (0.009)	-0.012 (0.010)	-0.007 (0.010)	0.012 (0.068)
Post-Scandal (Type I), Adjacent			-0.014*** (0.005)		
Pre-Scandal (Type I), Adjacent			0.001 (0.004)		
Post-Scandal (Type II)				-0.043*** (0.009)	
Pre-Scandal (Type II)				-0.017** (0.007)	
Observations	382000	382000	382000	382000	382000
R-Squared	0.08	0.08	0.08	0.08	0.08
No. of zipcodes	23875	23875	23875	23875	23875

Notes: Heteroskedasticity-Robust standard errors in parentheses, clustered at the zip code level. Stars indicate significance level: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. The regression also includes zip code fixed effects, time effects and the interaction between the time effects and the logarithm of population in the zip code, the logarithm of land area, and the share of urban population, all taken from the 1990 US Population Census. Data from the Zipcode Business Patterns for the period 1994-2009. See Table 17 for descriptive statistics, and its footnote for data definitions. The sample excludes zip codes that are non-standard, and zip codes that changed substantially over time.