IZA Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 10654

Learning about the Enforcement of Conditional Welfare Programs: Evidence from Brazil

Fernanda Brollo
Katja Maria Kaufmann
Eliana La Ferrara

MARCH 2017

I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 10654

Learning about the Enforcement of Conditional Welfare Programs: Evidence from Brazil

Fernanda Brollo
University of Warwick, CAGE and CEPR

Katja Maria Kaufmann
Mannheim University, CESifo and IZA

Eliana La Ferrara
Bocconi University, IGIER and LEAP

MARCH 2017

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Learning about the Enforcement of Conditional Welfare Programs: Evidence from Brazil*

We study the implementation of Bolsa Familia, a program that conditions cash transfers to poor families on children's school attendance. Using unique administrative data, we analyze how beneficiaries respond to the enforcement of conditionality. Making use of random variation in the day on which punishments are received, we find that school attendance increases after families are punished for past noncompliance. Families also respond to penalties experienced by peers: a child's attendance increases if her own classmates, but also her siblings' classmates (in other grades or schools), experience enforcement. As the severity of penalties increases with repeated noncompliance, households' response is larger when peers receive a penalty that the family has not (yet) received. We thus find evidence of spillover effects and learning about enforcement.

JEL Classification: I25, I38, O15

Keywords: enforcement, conditional welfare programs, learning, Brazil

Corresponding author:
Katja Maria Kaufmann
Mannheim University
Department of Economics
L 7, 3-5
68131 Mannheim
Germany

E-mail: kaufmann@vwl.uni-mannheim.de

* We thank Manuela Angelucci, Josh Angrist, Felipe Barrera-Osorio, Esther Duflo, Caroline Hoxby, Robert Jensen, Leigh Linden, Marco Manacorda, Karthik Muralidharan, Johannes Rincke, Je Smith, Michele Tertilt, Eric Weese, conference and seminar participants at Dartmouth, EUDN, Harvard, Michigan, MIT, NBER, Namur, Paris School of Economics, Toulouse School of Economics, Uppsala, Universitat Pompeu Fabra and Yale for helpful comments. Giulia Zane, Simone Lenzu and Emanuele Colonnelli provided excellent research assistance.

1 Introduction

Governments around the world increasingly rely on 'conditional' welfare programs in many different areas, including unemployment and social assistance benefits, maternity grants, child support and support for asylum seekers. Existing evaluations of conditional programs provide evidence on the combined effect of formal rules and the enforcement of these rules, but not on the role of enforcement itself. Yet understanding such role is indispensable for offering policy advice on program implementation, given that monitoring and enforcement are costly. Furthermore, political or administrative constraints may limit the extent to which governments are able (or willing) to enforce conditionalities. This paper investigates how the enforcement of program conditions affects beneficiaries' behavior in the context of conditional cash transfer (CCT) programs. CCT's provide a stipend to poor families as long as they meet certain conditions and have become a widely used tool to fight poverty in low and middle-income countries.

We study the implementation of the "Bolsa Familia Program" (BFP) in Brazil, a large-scale CCT that provides monthly subsidies to poor families conditional on school attendance of all school-aged children being at least 85 percent of school days every month. Failure to comply with this condition implies the receipt of up to five warnings and penalties which increase in severity with the number of past violations. The goal of our paper is to analyze whether people respond to the enforcement of BFP conditions by adjusting school attendance and whether they learn about the quality of enforcement from experiences of their peers.

From a theoretical point of view, it is not obvious that one should observe an increase in school attendance in response to the receipt of a warning. First, the family may not be able to (or want to) adjust on this margin, e.g., if the fall in attendance was due to a particularly severe shock which persists. Second, even if there is room for adjustment, the timing of it may not necessarily coincide with the receipt of a warning. If families expect that the government will detect and punish noncompliance with probability one, they should react already at the time when they fail to meet the program conditions. If, on the other hand, families believed that this probability were smaller than one, then they should respond after they receive the warning. Apart from responding to their own warnings, families might also respond to their peers' experiences: learning that someone else was punished for not attending school should lead to updating one's beliefs about the strictness of enforcement and to increasing school attendance in response.

To test these hypotheses we make use of a unique dataset that we compiled from administrative sources. It covers the universe of BFP beneficiaries in the Northeast of Brazil and includes information on child and family characteristics, monthly records on school attendance of each child, warnings received and benefits disbursed for the years 2008-2009. By merging these data with the Brazilian school census, we are also able to identify children of other BFP beneficiaries who are in the same class, grade or school as the family's own children. This allows us to construct different sets of peers.

We present our results in two parts. In both, our outcome of interest is the likelihood that individuals fail to comply with the school attendance condition in a given month (in short, "noncompliance"). In the first part we analyze the extent to which people respond to their own experiences of enforcement. We find that the likelihood of noncompliance decreases after receiving a warning or penalty. The estimated size of the effect is a decrease of 4 percentage points, which represents a 30 percent decrease over the average monthly noncompliance rate.

To provide evidence that we identify a causal effect we exploit random variation in the timing of arrival of warnings within each month. The date at which households can withdraw their monthly transfers –and receive the warning message, if applicable– depends on the last digit of the head's social security number, which is random. We can thus compare households who receive warnings earlier in the month with households who receive them later: the former should display a stronger contemporaneous reaction because they have more time to adjust their children's school attendance in that month, which is what we find.

In the second and main part of the paper we study whether beneficiaries react to warnings received by *other* households. The goal is to shed light on the importance of information transmission and learning in the context of program enforcement and to quantify potential spillovers that matter for estimating the benefits of stricter enforcement. We start by estimating if the likelihood that a child does not comply responds to the share of "peers" who received warnings in that month or in the month before, controlling for the household's own warnings. We use different definitions of peers, including BFP beneficiaries who are (i) a child's own classmates; (ii) classmates of the child's siblings (typically of different age); and (iii) classmates of the child's siblings who attend a different school. In all cases we find that warnings received by peers induce a decrease in noncompliance in the order of 1.1 (2.3) percentage points in the same (subsequent) month when all of one's peers receive a warning. This effect does not seem to be driven by correlated shocks: effects are very similar whether we focus on a child's own classmates (who are likely to experience similar shocks) or on the classmates of a child's sibling who attends a different school. Also, as a falsification test we show that there is no evidence of an anticipated response.

To investigate the learning process more in depth, we make use of the fact that families progress through different "warning stages" and receive increasingly severe penalties if they repeatedly fail to comply. When we distinguish between peers who get penalized at a lower warning stage than the family's own, and peers who receive warnings for higher stages, the latter induce a larger decrease in individual noncompliance. This is consistent with our learning interpretation because those warnings convey *new* information on the likelihood that the government implements higher order punishments.

Finally, to further corroborate our learning interpretation, we analyze a different set of peers, that is, BFP beneficiaries who live in the same zip code as the family. People living in the same area are likely to pick up the transfer from the same local service point, so we exploit the quasi-random variation in the day of the month on which beneficiaries can withdraw the money to construct an exogenous set of peers with whom recipients may interact. We then test whether warnings to beneficiaries who live in the same zip code and can cash in the benefit *on the same day* as family *i* affect the likelihood of noncompliance of *i*'s children, after controlling for warnings to BFP families in *i*'s zip code.[1] We find significant effects in the same and in the following month, consistent with the existence of information spillovers among families living in the same area. Also in this case, warnings received by

[1]The latter is included to control for possible geographically correlated shocks.

others for higher stages than one's own have stronger effects. The fact that we find evidence of information transmission using two sets of peers that are based on two entirely different identification strategies increases our confidence in the validity of our results.

Our paper contributes to several strands of literature. First, there is a relatively small literature on estimating behavioral responses to imperfect program enforcement. Banerjee, Glennerster and Duflo (2008) discuss the implications of imperfect enforcement in the Indian health sector; Rincke and Traxler (2011) and Drago, Mengel and Traxler (2015) analyze spillovers from enforcement of compliance with TV license fees; Lochner (2007) investigates the effect of learning about arrest probabilities on criminal behavior. Earlier work by Black, Smith, Berger and Noel (2003), Van den Berg, van der Klaauw and van Ours (2004) and Lalive, van Ours and Zweimüller (2005) studies the effects of unemployment insurance programs that monitored job search effort. Aside from the difference in context, our paper contributes to the above literature by focusing on the process through which beneficiaries learn about the quality of enforcement, such as the importance of information transmission through different types of signals and via different types of peers.[2]

Our paper is also related to a large literature on CCT's. Among others, Schultz (2004), Todd and Wolpin (2006) and Attanasio, Meghir and Santiago (2012) estimate the impact of Progresa/Oportunidades in Mexico; Bourguignon, Ferreira and Leite (2003), De Janvry, Finan and Sadoulet (2011) and Bursztyn and Coffman (2012) study *Bolsa Escola*, the predecessor of *Bolsa Familia* in Brazil, while Bastagli (2008), De Brauw et al (2015a,b) and Chioda, de Mello and Soarez (2016) focus on Bolsa Familia. In contrast to the above literature, we focus on the enforcement aspect of CCT's. There are a few papers which analyze the enforcement of eligibility criteria, i.e. errors of type one and two in terms of families included in or excluded from the program (e.g., Cameron and Shah, 2014), but there is very little evidence on the enforcement aspect of *conditionalities*. One notable exception is DeBrauw and Hoddinott (2010) who test the direct effect of program conditions in

[2]Also, since we do not have direct measures of *beliefs* about enforcement, we infer them from individuals' behavior, akin to the approach in Chetty, Friedman and Saez (2013).

the context of Progresa by exploiting the fact that some beneficiaries who received transfers did not receive the forms needed to monitor school attendance. They use matching and fixed effect methods to show that the absence of these forms significantly reduced compliance.

Within the CCT literature Schady and Araujo (2006), Baird, McIntosh and Özler (2011) and Benhassine et al. (2015) compare the effectiveness of conditional versus unconditional cash transfers (UCT). The first two papers find that conditionality (or the belief of it) improves school attendance, while the latter paper concludes that an unconditional transfer which is "labeled" as education support is equally effective. Our contribution differs from the CCT versus UCT comparison. In a world where conditionality nominally exists but program recipients anticipate that it will not be enforced, one may fail to find a difference in the impact of the two types of programs.[3] But this does not imply that conditionality, if enforced, would be uneffective. Our paper is precisely an attempt to understand how people update their beliefs about enforcement after experiencing it (directly or through peers) and change their behavior in response.

Finally, several papers analyze peer effects and spillovers in the context of conditional cash transfer programs, e.g., Barrera-Osorio, Bertrand, Linden and Perez-Calle (2011) for Colombia; Angelucci et al (2010), Angelucci and De Giorgi (2009) and Bobonis and Finan (2009) for the Mexican program Progresa/Oportunidades. Compared to this literature, our focus is not to identify conventional or "direct" peer effects (such as the effect of peers' school attendance on own attendance), but on the effect of peer warnings (a signal about the quality of enforcement) on individual attendance. This has the advantage of mitigating some of the identification problems that the conventional peer effect literature has to face.[4]

The remainder of the paper is organized as follows. In section 2 we provide background information on Bolsa Familia, while in section 3 we present the data and descriptive statistics. In sections 4 and 5 we discuss our empirical strategy and

[3]For example, Schady and Araujo (2006) conclude that cash transfers can be effective without monitoring but hypothesize that the effect of unenforced conditions might dissipate once households realize that they will not be punished.

[4]Our interest in how people learn about program features from their peers is shared, for example, by Duflo and Saez (2003).

results on whether individuals respond to experiencing enforcement and whether they learn from peers, respectively. Section 6 concludes.

2 Background Information on Bolsa Familia

The Bolsa Familia Program (BFP) reaches around 14 million Brazilian families, that is 60 million poor people (equivalent to about 30 percent of the Brazilian population) with an annual budget of over 24 billion reais (USD 11 billion, about 0.5 percent of GDP). Thus BFP reaches nearly three times as many people and is about three times as large in terms of budget as the well-known conditional cash transfer program Progresa/Oportunidades.

BFP was launched by the Brazilian president Inácio Lula da Silva in 2003 to consolidate four different programs (Federal Bolsa Escola Program, Auxilio Gas, Bolsa Alimentação, Fome Zero) into a single program.[5] The implementation of the program has seen a gradual evolution over the years. The election-free year 2005 was used to strengthen the core architecture of the program and to improve the registry of families. In 2006, the Ministry of Social Development (MDS) embarked on initiatives to promote further vertical integration with sub-national CCTs and integrated the conditional transfers paid under the Child Labor Eradication Program (PETI). The agenda for 2007 and beyond was to reinforce the monitoring and verification of conditionalities, to strengthen oversight and control mechanisms and to continue improving the program's targeting system. Notably, monitoring and enforcement of conditionalities had been relatively weak in the first years of the program, and were strengthened after 2006.

The targeting of the program was conducted in two steps. First, there was geographic targeting at the municipal level: the federal government allocated BFP quotas to municipalities according to estimates of poverty. Within municipalities, spatial maps of poverty were used to identify and target geographic concentrations of the poor. The second step was to determine eligibility at the household level. Eligibility was determined centrally by MDS based on household registry data that was collected locally and transmitted into a central database known as the Cadastro

[5] For a detailed description of the features of BFP, see Lindert et al. (2007).

Unico.

Benefits. BFP provides two types of benefits, a "base" and a "variable" transfer that depend on family composition and income. Families with a monthly per capita family income of up to R$60 (US$30) are classified as "extremely poor", while families with between R$60 and R$120 are classified as "moderately poor". The base benefit is provided only to families in extreme poverty, regardless of their demographic composition. Both extremely poor and moderately poor families receive a "variable" benefit which depends on the number of children in the family (capped at three to avoid promoting fertility) and on whether the mother is pregnant or breast-feeding. Benefit amounts are as follows. The base benefit amounts to R$ 60 (approximately US$ 30, which is also the per capita income threshold for the extremely poor), and the variable benefit to R$ 20. Benefit amounts and eligibility thresholds are periodically adjusted for inflation.

To illustrate the magnitude of the program transfers, think of a family with three children that is right at the threshold to be classified as extremely poor, i.e., they have a monthly per capita income of R$60 and a family income of R$ 300. This family would receive monthly transfers of around R$120, which amounts to 40 percent of their total family income.

Conditionality. BFP cash transfers are conditional on all age-relevant family members complying with requirements in terms of school attendance. Each school-aged child between 6 and 15 has to attend at least 85 percent of school days each month (absence due to health reasons is justified and does not count towards the number of days absent). If a single child fails to meet this requirement in a given month, the family is affected for the whole amount of the transfer, i.e. also for quotas that pertain to other children. This element of "joint responsibility" for children in the same family is a unique feature of BFP, e.g., compared to other well-known CCTs such as Progresa/Oportunidades.[6]

BFP also has conditions related to health behavior, such as health check-ups for pregnant women and vaccinations for children below age 5. Based on our data,

[6] In 2008 a new subprogram, Benefício Variável Jovem (BVJ), was added to BFP to increase school attendance of teenagers aged 16 and 17. BVJ has its own set of conditions and warning system and does not affect the family's warning stage in the main program (i.e. for the children aged 6 to 15). We only have data for a few months on BVJ, hence we focus on children aged 6 to 15.

those conditions are rarely enforced and we do not observe individual responses to enforcement of those conditions (in contrast to having monthly data on children's school attendance). Therefore we focus on children aged 6 to 15 and analyze responses to enforcement of school attendance conditions.

Penalties. The consequences of noncompliance vary depending on the historical record of compliance of each family. In the first case of noncompliance the family receives a warning without any financial repercussion. With the second instance of noncompliance, the family receives a second warning and benefits are blocked for 30 days, after which the family receives the accumulated benefit of the previous and the current month. The third and fourth warnings lead to a loss of benefits for 60 days each time and these benefits are never recovered. Finally, after the fifth warning the benefit is canceled and the family loses eligibility. According to the general rules, the family can re-apply to the program 18 months later. The 18 months rule is also used to "reset" a family warning history, e.g., if a family is in warning stage 2 but then complies with the rules for 18 consecutive months, their warning history goes back to stage zero.

Timing of penalties. Each month families can withdraw their transfer money at a local service point with a Bolsa Familia "electronic benefit card".[7] In case of noncompliance, at the time of withdrawal the family receives a warning message which specifies the penalty that the family incurs. To avoid bottlenecks at the local service point, when families pick up their BFP transfers, each family is assigned a pre-specified date each month starting from which the family can withdraw the transfer money (and contextually receive the warning). The exact date in each month is determined by the last digit of the social security number of the legally responsible adult of the family, which is basically random, as we will show. In particular, the ten digits (0,1,...,9) correspond to the ten weekdays in the last two weeks of the month.[8] We will exploit this randomness to identify the causal impact

[7]Local service points are either Caixa branches or Caixa "correspondents". Caixa Econômica Federal, a savings and credit union, is the government agency responsible for transferring BFP benefits to the beneficiary families. Caixa has a large network of banking correspondents, which are commercial establishments with a different business focus, used to expand access to remote and particularly poor areas (Kumar et al., 2006).

[8]Every family receives a benefit calendar with the information on when they can pick up the

of warnings received by a family on school attendance of its children.

Quality of information about the BFP program.

Families are well informed about transfer amounts, conditionalities and penalties: these aspects are widely and regularly publicized on TV, radio and newspapers in Brazil, and are spelled out in a booklet issued to each beneficiary family (*Agenda de Compromissos*). In case of noncompliance, at the time of withdrawal of the transfer money, the family receives a message that reports the family's warning stage, the month of noncompliance to which the warning refers, the names of the child(ren) who failed and which type of warning they might receive in the next instances of noncompliance.

An example of a warning message for a family receiving their first warning is: "The family has not complied with the conditionalities of Bolsa Familia for the first time. At this moment the payment will not be blocked. But if you fail to comply again your benefit payment may be blocked, suspended or even cancelled."[9] Note that the text does not say that the payment will be suspended with probability one, hence there is a degree of uncertainty about the actual enforcement of conditionality which is embedded in the system. This uncertainty is further reinforced by the experience of lax enforcement in the early years of the program.

The implementation of this conditionality scheme involves different actors. First of all, children's attendance is recorded by school teachers. The school sends the attendance lists of students to the municipality, reporting the exact fraction of school days attended in case attendance was below 85 percent, otherwise only reporting that the student complied. Each municipality collects the lists and sends them to the Ministry of Education (MEC), which determines whether the family as a whole complied or not in a given month, i.e. whether all children between 6 and 15 attended at least 85 percent of school days. MEC sends a detailed report to the Ministry of Social Development (MDS), which establishes which warning the family should receive in case of noncompliance and whether the family is entitled to the

---

transfer in the different months.

[9]In addition to this short paragraph, the warning message repeats the general rules of BFP in a clear and salient manner. It briefly mentions again the program conditions (i.e. of at least 85% of school attendance of all school aged children) and which instance of noncompliance may lead to which type of penalty.

transfer for that month (based on the warning stage reached).

The MDS sends this information to the Caixa Econômica Federal, a savings and credit union that transfers the benefit amount to the bank account of the family if the family is entitled to receiving the transfer for that month.

[Insert Figure 1]

These different steps of the process involve time and lead to a delay in terms of the month in which the warning is received compared to the month of noncompliance. Figure 1 illustrates an example where the family failed to comply in the month of February, right after the beginning of the school year, and received a warning in May. We denote the difference between the month of warning and the month of noncompliance as “delay”. There is substantial variation in the extent of delay, which we make use of in our graphical motivation for analyzing the effect of warnings on families' behavior and in some robustnes checks which we conduct to support our main findings. During our sample period this variable ranges from 2 to 6, with a mean of 3.8 months and a median of 3 months. Appendix Figure A.1 shows the frequency of different delays. As we will show, the variation in delay is orthogonal to household and child characteristics and is entirely driven by time and area effects.

3 Data

We make use of a unique dataset that we assembled combining several sources of administrative data. The first source is the household registry (*Cadastro Unico*) held by the Brazilian Ministry of Social Development (MDS), which contains socio-economic characteristics of all BFP beneficiary households.[10] We also know the eight digit zip code (*Código de Endereçamento Postal - CEP*) of the neighborhood where the household lives, which we will exploit in our analysis of learning. We have information on the universe of households enrolled in the program during

[10]For each household member the *Cadastro Unico* comprises information on age, gender, race, marital status, education and employment status and for the household as a whole it reports expenditures, house property, garbage collection, connection to running water, electricity etc.

2008-2009 in the Northeast of Brazil, one of the poorest and largest regions in the country, comprising 30 percent of the Brazilian population, and with more than half of its inhabitants living in poverty. To conduct our analysis we extracted a 10 percent random sample, yielding a total of 478,511 households in the program.[11]

A second dataset from MDS contains monthly records of school attendance during 2008-09 for each child monitored for the attendance conditionality, i.e., each child aged 6 to 15. We know whether the child complied with the attendance condition in a given month, and in case of noncompliance we know the exact fraction of days attended below the threshold of 85 percent. This dataset also includes two categories that qualify the absence: one is justifiable absences such as sickness, for which the household does not incur any penalty; the other category is for unjustifiable absences (e.g., child labor, teen pregnancy, etc.) which instead count towards the warnings.

A third dataset from MDS contains monthly information on warnings, which allows us to create the complete warning history of each family. In particular, we know in which month they received a warning, the month in which attendance failed to meet the threshold and that gave rise to that particular warning, and the warning stage in which the family is at any given point in time.

The fourth administrative dataset we use is payroll data, which contains monthly information on the benefits that the family received and whether the benefit was blocked or suspended in a given month. All four of these datasets can be linked through the social security number *(Número de Identificação Social - NIS)* of the household member legally responsible for the child or of the child him/herself.

We complement the above administrative records on BFP with two datasets on schools. The first is the School Census compiled by the Ministry of Education, which contains information on all children who are enrolled in a given school, grade and class. We merge the School Census with the BFP administrative data using the NIS code of the child when available in the School Census and otherwise based on

[11] To improve our precision in estimating the effects of different warning stages, we over-sampled households that were warned at least once during our sample period, and use regression weights to correct for this. Since we include household or individual fixed effects in all our regressions, households that are never warned do not contribute to estimating the effect of the main variable of interest, namely the receipt of a warning.

area code, school code, grade, full name and date of birth of the child.[12] This allows us to identify the peers who are in the same school, grade and class for each child and who are also recipients of BFP. This is important for our analysis to identify the role of learning from classmates.

The final dataset we employ is Prova Brasil, collected by the *Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira* (INEP) to evaluate the quality of education across Brazilian schools. These data include a questionnaire administered to principals of schools, containing, among other things, information on the policies that the school adopts to inform parents about children's unexcused absences. We exploit this information in one of the robustness checks for our analysis.

[Insert Table 1]

Table 1 presents summary statistics of the main variables of interest. A full set of summary statistics for all variables is provided in Appendix Table A.1. The outcome variable we use in our analysis is whether a family fails to comply with BFP rules, i.e., at least one child attends less than 85 percent of school days in a given month. We refer to this variable as "noncompliance".

Panel A of Table 1 shows that on average the likelihood of noncompliance for the household is 8.6 percent in any given month, and the likelihood of receiving a warning is 3.5 percent. In the average month 29 percent of our households are in warning stage 1, which means they have received a message about noncompliance but not lost money; 8.5 percent are in warning stage 2, where the money is blocked and then returned; 3 and 1 percent are in warning stages 3 and 4, respectively, where the transfer is lost for two months, and 0.2 percent reach warning stage 5, when they are expelled from the program.

Panel B of Table 1 shows that the average child who is a BFP recipient in our sample has 16 other BFP recipients in her class, 52 in the same grade and 279 in the same school. About 1 percent of one's classmates who are BFP recipients ('Peers') fail to meet the attendance conditionality in the average month, and 1.4 percent receive a warning message (for any of the five warning stages).[13]

[12] We are able to match 68% of individuals, 41% based on their NIS.

[13] Since warnings are received at the household level, while the fraction of peers not complying

4 Own Experience of Enforcement

Before analyzing whether individuals learn about enforcement from peers, we investigate how individuals respond to experiencing enforcement themselves. In particular, we test whether households react to a warning by increasing the school attendance of their children. We conduct this analysis at the household level because warnings are applied to the entire family and not to individual children. In Appendix Table A.2 we show that results are very similar if we use the child as the unit of observation.

From a theoretical point of view, it is not obvious that one should observe an increase in school attendance in response to the receipt of a warning. The first reason is that the family may not be able to (or not want to) adjust on this margin, e.g., if the fall in attendance was due to a particularly severe and persistent shock.

The second reason has to do with the priors held by the household and its updating process. When deciding whether to fulfill program conditions, families compare the costs and benefits of compliance. An essential part of this calculation is the belief they hold about the probability that the government will detect noncompliance and punish them for it. If the family believed this probability to be one, then it should already anticipate the consequences of noncompliance at the time when the child fails to attend school, and should respond accordingly. The increase in attendance –if any– would not necessarily coincide with the receipt of a warning several months later, but plausibly come earlier. If, on the other hand, the household believed that the probability of punishment were smaller than one, then on top of any earlier response it should adjust its behavior upon receipt of a warning, in order to avoid further penalties.

Identifying the causal effect of warnings on school attendance is empirically challenging because the receipt of a warning is a function of past school attendance. We therefore resort to a number of alternative strategies to establish a causal link between receipt of a warning and attendance response.

are calculated at the individual child level, the fraction of peers in noncompliance is lower than the fraction of peers warned, since not all children in a household fail to comply at the same time.

4.1 Graphical analysis

We start by visually exploring the pattern of noncompliance in the months before and after the warning.

[Insert Figures 2 and 3]

Figure 2 pools data from the different types of warnings received by different households and plots the probability that at least one child in the household fails to meet the 85 percent attendance threshold against the number of months since the warning was received. The value 0 on the horizontal axis indicates the month of the warning, -1 is the month before, and $1, 2, ...5$ indicate the number of months after the warning was received. As can be seen from the figure, there is a clear discrete downward jump in noncompliance immediately after the receipt of a warning: households respond by increasing attendance of their children, hence the likelihood of noncompliance goes down. The probability of noncompliance remains relatively constant in the months after the warning.

Appendix figure A.2 presents the same data disaggregated by warning stage, i.e. separately for households that received their first warning (Warning Stage 1), their second warning (Warning Stage 2), etc. until the fifth and last warning. The pattern displayed in figure 2 is found across all warning stages except for the last one, where the decrease in noncompliance is less discontinuous (note, however, that very few people reach Warning Stage 5).

In figure 3 we repeat the exercise conditioning on the time elapsed between noncompliance and the receipt of the warning, i.e. on "delay". In this figure, we indicate with 0 the month when the family failed to comply with the attendance requirement and with a vertical line the month in which the warning is received. Moving from left to right and from top to bottom, figure 3 shows what happens to rates of noncompliance when the delay is $2, 3, ...$ up to 6 months.[14] In $t = 0$ the probability of noncompliance is one by construction; from $t = 1$ onwards this probability decreases to a 'normal' level, but once the warning is received, it jumps down and stays at this lower level given the new (higher) warning stage.

[14]See Appendix Figure A.1 for the frequency of different delays.

In all cases except for the delay of 6 months, we see a discrete jump in the rate of noncompliance immediately after the receipt of a warning: if the delay is 2 months, then immediately after the arrival of the warning in month 2 the likelihood of noncompliance jumps down; if the delay is 3 months, it jumps down in month 3, etc. This is important because it suggests that the adjustment was *caused* by the receipt of the warning. Other possible reasons (e.g., mean reversion of the shock) would generate a pattern where the adjustment would depend on the distance from the month of noncompliance (month 0) and not necessarily correspond to the month of the warning, i.e., the (shifting) vertical line. In other words, there is no reason why mean reversion should cause a discrete jump downwards which happens with the exact same delay as the warning.[15] Also, the fact that we see a response of the household upon arrival of the warning (i.e. a further reduction in the likelihood of noncompliance) suggests that the warning was not fully anticipated.

4.2 Multivariate analysis

We next estimate the effects of warnings on noncompliance using multivariate regressions. As a first step, we want to regress noncompliance on the warning stage that a household is in, since the warning stage determines the cost of not complying with conditionality. One challenge in identifying this effect is that some households have a much higher propensity to miss attendance than others, and these are the households who reach higher warning stages. If we simply used cross-sectional variation in the data, this would bias the (negative) coefficient on warnings towards zero or even induce a positive correlation between receiving warnings and failure to comply. We therefore use household fixed effects throughout our analysis to control for time-invariant differences in the propensity to attend school. This implies that we exploit variation within family over time in the warning stage reached, i.e. in the receipt of new warnings.

[15] Moreover, as we show in the next section, delay is orthogonal to household and child characteristics (see Appendix Table A.3).

Our baseline specification for estimating the effects of own warnings is:

$$Y_{ht} = \sum_{k=1}^{5} \alpha_k WS_{ht}^{k} + \gamma X_{ht} + D_t + D_h + \epsilon_{ht} \quad (1)$$

where h denotes the household, t the month, Y is a dummy equal to 1 when at least one child in the household attends less than 85 percent of the school days in a month ("noncompliance"); WS^k denotes a dummy equal to one if the household is in warning stage k (with $k = 1, ..., 5$); X is a vector of household level controls including fraction of male children in the household, number of boys and girls in different age brackets (6-10, 11-15, 16-18); D_t denotes month and year dummies to control for seasonality and time effects; D_h denotes household fixed effects and ϵ is the error term. Given the panel nature of our data, we estimate equation (1) and all other regressions in this section using a linear probability model and clustering the standard errors at the household level.

[Insert Table 2]

Table 2 shows our results. The probability of noncompliance decreases after the household receives a warning and advances to a higher warning stage, confirming the findings of our graphical analysis. The negative coefficients are of increasing magnitude the higher the warning stage, but one should be careful when comparing the effects of different warning stages because those coefficients are estimated by averaging over different subsets of families. For example, the coefficient on warning stage 5 is only estimated for those families who ever get to warning stage 5 in our two-year period of observation, while the coefficient on warning stage 1 is estimated over the set of families who receive at least the first warning during our sample period. These households are likely to differ in unobservable ways, hence we cannot attribute differences in the magnitude of the coefficients solely to the incentive effects of larger penalties. Since our focus here is on whether and how people respond to and learn about enforcement and not on the effects of different warning stages, we leave a more detailed analysis of this issue for future analysis.[16]

[16]To explore the effect of larger penalties, in Appendix Table A.4 we estimate the same regression

Appendix Table A.2 shows that the results are qualitatively similar if we conduct the analysis at the individual (child) level, including child level fixed effects and time varying controls.[17]

Causality

To show that we identify a causal effect of warnings on attendance behavior, we exploit random variation in the timing of warning arrival. In particular, we make use of the fact that the exact day of the receipt of warnings depends on the last digit of the social security number of the legally responsible adult in the family, as described in section 2. We conjecture that families who receive warnings earlier should have more time to react to the new information, hence we expect a larger behavioral response for these families compared to those who receive the warning later in the month. In the limit, someone who receives a warning the last day of the month could only adjust if her child's attendance was just one day below the threshold. To test this conjecture, we split the sample into "early" and "late" warning receivers, where "early" are the households that fall in the first half of the period during which money can be withdrawn within a month and "late" are those that fall in the second half. Appendix Table A.5 shows that observable characteristics are almost perfectly balanced across the two groups (in the few cases where there are significant differences, these are extremely small, at the third decimal digit). We estimate the following modified version of Equation (1):

$$Y_{ht} = \sum_{k=1}^{5} \alpha_k^E WS_{ht}^{E,k} + \sum_{k=1}^{5} \alpha_k^L WS_{ht}^{L,k} + \sum_{k=1}^{5} \beta_k^E WS_{h,t-1}^{E,k} + \sum_{k=1}^{5} \beta_k^L WS_{h,t-1}^{L,k} + \gamma X_{ht} + D_t + D_h + \epsilon_{ht} \tag{2}$$

holding the set of households constant by conditioning on the initial and last warning stage that the households reach during our sample period. For example, for households that reach warning stage 3, we can compare the magnitude of the effects of warning stages 1 and 2 (the effect of warning stage 3 cannot be interpreted for this subset of households, since by construction warning stage 3 is the highest warning stage they will reach, which means that their noncompliance rate mechanically decreases thereafter). Appendix Table A.4 shows that higher warning stages have stronger effects in terms of reducing noncompliance for all subgroups.

[17]The magnitude of the coefficients is smaller when estimated with individual level data because not all children in a given household fail to comply at the same time.

where the superscripts E and L stand for "early" and "late", respectively. The variable $WS_{ht}^{E,k}$ ($WS_{ht}^{L,k}$) is the interaction between the warning stage and a dummy taking value one if household h is an "early" ("late") warning receiver. We also include first lags of these variables to estimate delayed household responses. Our prior is that $\left|\alpha_k^E\right| > \left|\alpha_k^L\right|$ because during the month when the warning is received households who are warned earlier have more time to adjust the attendance of their children. On the other hand, since "early" and "late" households on average fail to comply at the same rate, we expect "late" households to catch up during the following month, hence we expect $\left|\beta_k^E\right| < \left|\beta_k^L\right|$.

[Insert Table 3]

Table 3 reports our estimates.[18] Coefficients $\widehat{\alpha_k^E}, \widehat{\beta_k^E}$ for "early" households are displayed in column 1; coefficients $\widehat{\alpha_k^L}, \widehat{\beta_k^L}$ for "late" ones in column 2, while they are both estimated in the same equation. Column 3 reports the differences $\widehat{\alpha_k^E} - \widehat{\alpha_k^L}$ and $\widehat{\beta_k^E} - \widehat{\beta_k^L}$, with p-values in square brackets. The pattern of the estimated coefficients is clear: households that receive warnings early respond more strongly in the month of warning receipt, while households that receive warnings late catch up in the following month and respond more strongly then. Given the randomness of the "early" vs. "late" classification, it is difficult to rationalize this pattern with explanations other than the fact that the household is responding to the actual receipt of the warning and to its informational content.

To provide further support for this interpretation we show in Online Appendix A that the same conclusion holds when using variation in the arrival of warnings across months. In particular, we make use of the fact that the delay of warnings varies over time in a non-monotonic way due to bureaucratic complexities and is orthogonal to household characteristics.[19] In Appendix Table A.6 we thus conduct a difference-in-difference analysis comparing households who received warnings

[18]The smaller number of observations in Table 3 compared to Table 2 is due to the fact that, by introducing a lag in warning stage, we lose the first month of our sample period.

[19]In Appendix Table A.3 we regress delay on household and child level controls, month and year dummies and municipality dummies. We find that time and area fixed effects explain more than 98 percent of the variation and only one out of 23 household and child controls is significant at the 10 percent level. The p-value for the joint test that the coefficients on all household and child characteristics are zero is 0.91.

with different lags and show that there is a significant decrease in the probability of noncompliance precisely *after* the receipt of the warning, in that the receipt of a warning decreases noncompliance by 3.8 percentage points (equivalent to a 30 percent decrease).

4.2.1 Interpretation

A possible interpretation of our results so far is that when households receive a warning, they learn not so much that the program is being enforced, but that their children have missed school – something they may not have been aware of. While this is certainly not a concern in our main analysis about learning from peers' warnings (as discussed in the next section), we test the importance of this concern in Appendix Table A.7 by exploiting variation in school policies regarding communications with families. Specifically, we merge our administrative data with a dataset on schools called PROVA Brasil. PROVA contains some school-level variables that capture whether parents receive information about non-attendance of their children from the school, independently of BFP warnings. We compare schools that inform parents with schools that do not. Under the hypothesis that warnings have an effect because parents learn about non-attendance of their children, the effect of warnings should be smaller in schools where parents are already informed about non-attendance. The results show, however, that there is basically no difference in the effect of warnings between schools which inform parents and schools that do not. This suggests that families' response to warnings is unlikely to be entirely due to learning about children's non-attendance.[20]

Behavioral response vs teachers' reporting One remaining interpretation issue is the extent to which the behavioral responses we identify correspond to actual

[20]Bursztyn and Coffman (2012) present evidence that the conditionality embedded in Brazil's CCT is a valuable source of information for parents and empowers them in the negotiation with their children. The effect of warnings that we estimate could then be interpreted as resulting from the improved ability of parents to commit to punishing their children when the parents themselves are penalized by the government. While this interpretation would not be inconsistent with the results in Appendix Table A.7, it cannot explain the learning from peers results that we present in the next section.

increases in attendance, and not simply to more lenient reporting by teachers, e.g., because parents convince teachers to be less strict in registering non-attendance after they receive a warning.[21] A first reaction to this is that even in this case our hypothesis that families learn about the strictness of enforcement would be valid: it would be the action taken by the family that would differ (e.g., persuading teachers instead of sending children to school). Secondly, our results on warnings received by peers –as discussed in the next section– help address this point to some extent. In particular, we find that child *i*'s noncompliance goes down when the classmates of *i*'s siblings get warned or when *i*'s neighbors receive warnings. Since the teacher of child *i* would not know about those children's warnings (in particular if the siblings and neighbors' children attend different schools, for example because they are of a different age), this result is more likely explained in terms of actual child's attendance as opposed to misreporting. In fact, it may be difficult for families to ask teachers to falsify attendance records simply because someone else in another school has been warned. Third, children have several different teachers who have to register attendance: bribing or convincing all of them may not be easy in the presence of generally high stakes.[22] In fact, municipalities face important incentives for proper monitoring and reporting of school attendance: since 2005 the Ministry of Social Development (MDS) has exerted massive efforts to strengthen the monitoring of conditionalities, introducing mechanisms to promote incentives for quality management and rewarding innovations in the decentralized management of BFP (Lindert et al., 2007).

Taken all the evidence together, our results suggest that families respond to the enforcement of program conditions and that at least an important part of their behavioral response corresponds to an actual increase in terms of school attendance.

To summarize our results so far, we have shown that the receipt of warnings reduces the likelihood that households fail to comply with the 85 percent school attendance requirement. We provided evidence that this effect can be interpreted as

[21]Unfortunately, we cannot directly test for misreporting, differently from Linden and Shastri (2012) who rely on external monitors' verification in a sample of Indian schools.

[22]Teachers earn a relatively high salary and it is unclear that parents from poor households could offer a sufficiently large amount of money to induce teachers to falsify records and risk complaints by other students or teachers, and/or risk punishment by the school principal.

causal, i.e. that the increase in school attendance is a response to the warning. We now move to the main focus of this paper, namely the question of whether people not only respond to experiencing enforcement themselves, but also learn about the strictness of enforcement from warnings of their peers.

5 Peers' Experience of Enforcement

In this section we analyze whether BFP beneficiaries learn about enforcement from the experiences of peers who receive warnings. In particular, children who are in the same class as one's own children and their parents are likely to be key sources of information regarding the implementation of the program. To explore these effects we rely on individual children as the unit of observation and construct different sets of peers, comprising the child's own classmates and the classmates of a child's siblings.

5.1 Identification

Identifying the impact of the warnings received by classmates is a non-trivial issue. To discuss our empirical strategy it is useful to start from a simple specification, which is *not* the one we estimate, but which helps highlight identification challenges. Consider the following model:

$$Y_{iht} = \sum_{k=1}^{5} \alpha_k WS_{ht}^{k} + \beta\, PEERWARN_{iht} + \gamma X_{iht} + D_t + D_i + \epsilon_{iht} \tag{3}$$

where i denotes the child, h the household, t the month; Y is an indicator for whether child i failed to comply in month t, WS^k is a set of dummies to denote if the household is in warning stage k; X is a vector of child level controls including age and number of brothers and sisters in different age brackets (6-10, 11-15, 16-18); D_t denotes month and year fixed effects; D_i denotes child fixed effects and ϵ is the error term. The key regressor of interest is $PEERWARN_{iht}$, which is the fraction of i's classmates who receive a warning in month t.

Suppose we found –as we do– a negative correlation between a child's failure to attend in a given month and the fraction of peers who are warned ($\widehat{\beta} < 0$). Before we can interpret this correlation as learning we need to address several identification challenges.

Correlated shocks. The first threat to identification are correlated shocks that may directly affect a student and her peers, thus inducing a correlation between $PEERWARN_{iht}$ and ϵ_{iht} in equation (3). Consider for example an economic shock leading to an increase in the opportunity cost of schooling in the area where individual i lives. In response to such a shock, both individual i and her peers would be more likely not to comply and thus more likely to receive a warning.[23]

If the shock was persistent, it would induce individual i and her peers to still not comply several months later, when the warning is received. However, this type of mechanism would generate a *positive*, not negative, correlation between a child's noncompliance and her peers' warnings ($\widehat{\beta} > 0$), while what we find in the data is a negative correlation ($\widehat{\beta} < 0$).

If, on the other hand, the shock was mean-reverting, this would lead to a negative correlation between a child's noncompliance and her peers' warnings and could thus be confounded with the learning interpretation we offer. To address this concern, we include the lead of peers' warnings as a falsification test, i.e., we add $PEERWARN_{ih,t+1}$ among the controls in equation (3). Our reasoning exploits the time lag from the moment in which noncompliance occurs and the moment in which the warning arrives (as discussed in section 2, this delay has a median value of 3 months and a mean of 3.8 months). If a child's attendance increases because of mean-reversion after the initial instance of noncompliance, there is no reason why the attendance would start reverting to the mean with the exact same delay as the warnings of her peers, which vary between 2 and 6 months. In the case of mean-reversion, one would typically expect changes in attendance to start occurring *before* the arrival of the warnings, which implies that we should find a negative and significant coefficient on $PEERWARN_{ih,t+1}$. Failing to find an effect of the lead variable would be hard to reconcile with the interpretation of mean-reverted shocks: it

[23]To address the concern that the arrival of an individual's own warning might be correlated with the arrival of her peers' warnings, we always control for own warnings and analyze if peers' warnings have an additional effect.

would mean that the initial shock, which led to noncompliance, starts reverting exactly, say, four months afterwards (when peers happen to receive the warning) but not three months afterwards.

An additional strategy we employ to deal with grade or school-specific correlated shocks is to exploit warnings received not only by individual *i*'s own classmates, but by the classmates of *i*'s *siblings*. Since siblings typically have different ages, the effect of warnings received by siblings' peers should not reflect class or grade-specific shocks. This should be particularly true for a second specification we use, in which we focus on peers of siblings in *other schools* (being in different schools is mostly due to siblings' age difference). The warnings received by students of those schools should not be correlated with shocks experienced by child *i*'s own school, conditional on the warnings of the child's family and of her own peers. On the other hand, warnings received by siblings' peers do contain information about the quality of BFP enforcement. We include the *maximum* fraction of peers who got warned across child *i*'s siblings, as the strongest signal of enforcement.[24]

Conventional peer effects. A separate concern relates to direct or 'conventional' peer effects. As we have shown in the first part of the paper, once an individual receives a warning, she reacts by reducing the likelihood of noncompliance (i.e. increasing attendance). This means that when a child's classmates receive warnings, they will attend school more in response to their own warnings. The child may then start attending more because she observes her peers doing so. This response would imply a spillover effect of enforcement on peers' school attendance, but cannot necessarily be interpreted as learning about enforcement: it might be due to learning about the benefits of schooling or to a preference to attend school with more peers.

To identify "learning about enforcement" we pursue two approaches. First, we directly control for the fraction of child *i*'s classmates who fail to comply with conditionality, to analyze if *warnings* of *i*'s peers have an independent effect on *i*'s likelihood to fail.[25] Second, we analyze whether the probability that child *i* fails

[24]The more families in a class receive a warning, the more likely it is that this becomes a topic of discussion among children and parents of beneficiary households.

[25]This may lead to an underestimate of the "true" learning effect, if part of the learning about

to comply decreases when her *siblings*' peers (possibly at a different school) get warned. Since those peers are not child i's classmates (or not even in child i's school), the direct effect of their school attendance should not be important. On the other hand, warnings received by siblings' peers contain relevant information about the strictness of enforcement.

To sum up, the two specifications we estimate are:

$$Y_{iht} = \sum_{k=1}^{5} \alpha_k WS_{ht}^{k} + \sum_{n=-1}^{0} \beta_n PEERWARN_{ih,t+n} + \sum_{n=-1}^{0} \zeta_n PEERFAIL_{ih,t+n} + \gamma X_{iht} + D_t + D_i + \epsilon_{iht} \tag{4}$$

$$Y_{iht} = \sum_{k=1}^{5} \alpha_k WS_{ht}^{k} + \sum_{n=-1}^{0} \beta_n PEERWARN_{ih,t+n} + \sum_{n=-1}^{0} \delta_n MaxPEERWARN_{ih,t+n} + \gamma X_{iht} + D_t + D_i + \epsilon_{iht} \tag{5}$$

where $PEERFAIL_{ih,t+n}$ is the fraction of i's classmates who are BFP recipients and fail to comply in month t; $MaxPEERWARN_{ih,t}$ is the maximum fraction of classmates who got warned in month t among i's siblings. This variable is constructed alternatively from all of i's siblings or from siblings who attend a different school than i. For both peers' warnings and peers' noncompliance we augment the specification with a one-month lag, because it is ex ante difficult to say how persistent the effects may be. We estimate equations (4) and (5) using a linear probability model and clustering the standard errors at the household level (or, as a robustness check, at the school level).

5.2 Main results

Table 4 reports our main results on classmates' warnings. The coefficient on the fraction of classmates warned in column 1 is −0.011 for the contemporaneous variable and −0.023 for the lagged one, both significant at the 1 percent level. This indicates that after more classmates receive a warning, the child is less likely to fail

enforcement happens through observing one's peers' higher attendance.

to comply with the attendance requirement. To assess the magnitude of the effect in relation to the mean: if all of *i*'s classmates who are BFP recipients received a warning in a given month, child *i*'s noncompliance would decrease by 13 percent in the same month and 27.5 percent in the following month. [26]

[Insert Table 4]

In column 2 we include the fraction of classmates warned as lead variable to conduct a falsification test and rule out the importance of correlated shocks and mean reversion. As explained above, if mean reversion were driving our results, we would expect a negative coefficient on the lead of classmates' warnings. We find instead a precisely estimated zero, lending support to our interpretation.

In column 3 we control for the fraction of classmates who fail to meet the attendance threshold and find that both the contemporaneous and the lagged variable are positively correlated with own noncompliance. At the same time, we still find a negative effect of classmates' warnings that is equally significant and comparable in size to that of the previous specifications. Again, this is consistent with our learning interpretation and suggests that 'conventional' peer effects operating through peers' attendance are not driving our results.

In the remaining part of table 4 we focus on warnings received by siblings' peers. For this analysis we naturally need to restrict the sample to children who have at least one sibling in the age range of BFP conditionality. It turns out that a large fraction of our original sample satisfies this condition, but in column 4 we report our benchmark estimates for the reduced sample to show that the magnitude of the coefficients is virtually the same as in column 1.

Column 5 shows that not only an individual's own peers but also her siblings' peers matter for her attendance decision: both the contemporaneous and the lagged fraction of siblings' classmates warned significantly reduces a child's likelihood not to comply, while controlling for the warnings of the child's own classmates. The results remain significant and comparable in size for the lag when we exclusively focus on siblings going to a different school (column 6), while the coefficient on

[26] Appendix table A.8 presents the same results as table 4, but displays in addition the coefficients on the family's own warning stage, which we control for in all specifications.

the contemporaneous variable is also negative but insignificant. The findings in columns 5 and 6 lend further support to our learning interpretation, as the peers who receive warnings in columns 5 and 6 are attending different grades or schools and are exposed to different shocks. In Appendix table A.9 we present results clustering standard errors at the school level. Standard errors remain virtually unchanged.

5.2.1 Interpretation: informational content of different warnings

To gain further insights into the role played by the new information contained in warnings, we analyze how the effect of classmates' warnings varies depending on the peers' warning stage relative to the family's own warning stage. We expect that warnings received by peers who are in a lower warning stage than the family's own should carry relatively less information, because the family has already experienced first hand that the government enforces the program up to that level. On the other hand, warnings of a level *higher* than one's own carry new information, because the family could still hold a prior that the government punishes up to some point, but will not go through with more severe and costly punishments. Peers' warnings of the same stage as the family's could still have an effect, for example, if other families receive them earlier in the month or because they improve the precision of the signal.

[Insert Table 5]

In Table 5 we distinguish between classmates who receive warning 1, 2, ...5 (the variables listed by row) and estimate four different regression equations, conditional on whether a family is currently in warning stage 1, 2, 3 or 4 (the different columns). Clearly, all stages higher than one's own should be relevant, but the next stage might be particularly important for at least two reasons. First, in our data, the higher the warning stage, the fewer the households that reach it. This implies that more households receive a warning for the next stage, compared to subsequent warnings. Second, if people are present-oriented, the closer in the future is the punishment, the more they care about it.

The estimates in Table 5 lend support to our hypothesis that warnings received by peers in higher warning stages have a stronger impact on school attendance than warnings for the same or lower warning stage. For example, in column 1 the coefficient on peers' warnings for stage 2 is twice as large as that for stage 1. In column 2 the coefficient on peers' warnings for stage 3 is about one and a half times as large as for stage 2 (similarly for column 3). Only in column 4 we find that while the coefficient of peers' warnings for stage 4 is significant, the one for stage 5 is not, which is likely due to the fact that extremely few families reach warning stage 5 at all. Overall, these results corroborate our interpretation of learning about enforcement, as it would be difficult to find alternative explanations that produce the asymmetric pattern that we uncover.

5.2.2 Interpretation: different peers

While classmates are an important source of information about the strictness of enforcement, another potential source are adults who are BFP beneficiaries and who live in the same area. We hypothesize that families who receive warnings may interact directly, and not only through their children in schools, and focus on a dimension of interaction for which we can find exogenous variation.

Consider the interaction among adults living in the same neighborhood. These adults and their families may experience common shocks and hence their children's attendance behavior may be correlated. However, as we explained in section 2, the day of the month in which households receive warnings is determined by the last digit of the social security number (NIS) of the designated household member. Since this number is as good as random, the set of families living in a given zip code who receive the warning on the same day is also as good as random. Why would this be relevant from an information point of view? Households receive warnings when they pick up the transfer from the local Caixa service point.[27] Households that visit the same Caixa point on the same day are more likely to meet and communicate about warnings on the day when any information about enforcement is most salient. We thus test whether compliance of a household is affected by warnings received by other households that live in the same zip code *and cash the transfer on the same*

[27] Kumar et al (2006) report that in the early 2000's Caixa started expanding their banking correspondent network to reach even remote locations, and that by 2008 (the period studied in this paper) there was roughly one Caixa correspondent per zip code.

day, conditional on warnings received by all households in the neighborhood. We estimate the following specification:

$$Y_{iht} = \sum_{k=1}^{5} \alpha_k WS_{ht}^{k} + \sum_{n=-1}^{0} \beta_n BANKWARN_{ih,t+n} + \sum_{n=-1}^{0} \zeta_n NEIGHBORWARN_{ih,t+n} + \delta NEIGHBORS_{iht} + \gamma X_{iht} + D_t + D_i + \epsilon_{iht} \tag{6}$$

where all variables are defined as in equation (3), except for the following. We denote the number of BFP beneficiaries in i's zip code in month t as $NEIGHBORS_{iht}$, while $NEIGHBORWARN_{iht}$ denotes the fraction of BFP beneficiaries living in i's zip code who receive a warning in month t. The key regressor of interest is $BANKWARN_{iht}$, which is the fraction of BFP beneficiaries living in i's zip code who receive a warning in month t among those whose families can pick up the transfers at the local Caixa point on the same day. By construction, this variable is not based on a peer group which is chosen endogenously.

[Insert Table 6]

Table 6 reports the results. Column 1 shows that there is indeed a strongly significant negative relation between the fraction of BFP beneficiaries from the same zip code who can pick up their transfers on the same day and individuals' decision of noncompliance. Based on our estimates, if 10 percent of the BFP recipients who live in the same zip code and can pick up their transfers on the same day received a warning, a child's probability of noncompliance would decrease by 3.2 percent in the same month and 6.3 percent in the following month.

In column 2 we include *BANKWARN* as lead variable to conduct a falsification test and rule out the importance of correlated shocks and mean reversion. The coefficient on the lead variable is a precisely estimated zero (with a coefficient of −0.0007), lending support to our learning interpretation.

In column 3 we control for the fraction of BFP beneficiaries from the same neighborhood who are warned in a given month (*NEIGHBORWARN*). The fraction of neighbors who are warned *on the same day* (*BANKWARN*) is exogenous when controlling for the overall fraction of neighbors warned within that month, as discussed above. We find that *NEIGHBORWARN* is positively correlated with an individual's noncompliance, while the lag of this variable is negatively correlated, which might be due to correlated shocks that have some persistence but mean-revert at some point. More importantly, controlling for *NEIGHBORWARN* does not change our key result, which is that a larger fraction warned *on the same day* significantly decreases an individual's likelihood of noncompliance. The point estimates imply that if 10 percent of the BFP recipients who live in the same zip code and can pick up their transfers on the same day received a warning, a child's probability of noncompliance would decrease by 4.1 percent in the same month and 3.2 percent in the following month.[28]

We also analyzed how the effect of other households' warnings varies depending on their warning stage relative to the family's own warning stage. Analogous to our analysis in section 5.2.1, we find that warnings received by other households who live in the same zip code, can pick up their transfer on the same day and are in higher warning stages have a stronger impact on compliance than warnings for the same or lower warning stage. This is consistent with our learning interpretation: warnings for lower stages than one's own carry relatively less information, because the family has already experienced them. The results are reported in Appendix Table A.11.[29]

Overall, the fact that we find consistent effects for two different types of peers that are based on two entirely different identification strategies (partially overlapping networks in the case of classmates and quasi-random variation in the case of

[28] In Appendix table A.10 we present results clustering standard errors at the zip-code level. Standard errors are somewhat larger, but all conclusions are unchanged. In particular, the coefficients on the main variable of interest, i.e. on the "fraction of neighbors warned on the same day", and on its lag are all significant at the one percent level.

[29] In Appendix Table A.11, for households that are in stage 1 (column 1) the coefficients on other households' warnings for stages 2 and 3 are three and five times larger, respectively, than that for stage 1 (the difference is significant at the 2 and 1 percent level, respectively). In column 2 the coefficient for stage 3 is similar as for stage 2, but the one for stage 4 is more than four times larger (the difference is significant at the 10 percent level). Only in columns 3 and 4 we do not find significant differences, possibly due to the fact that the punishment in the case of the fourth warning is in fact the same as in the case of the third warning, while extremely few families reach warning stage 5 at all.

households living in the same zip code), strengthens our confidence in the validity of our approach and interpretation.

6 Conclusions

In this paper we study the implementation of the large-scale conditional cash transfer program "Bolsa Familia" in Brazil. This program conditions transfers to poor families on children's school attendance: when families fail to comply with this requirement, they receive a series of warnings and financial penalties. We analyze how people learn about and respond to the enforcement of program conditions. We find that warnings not only have a direct effect on the families warned, but also important spillover effects on other families, who learn from the experiences of their children's classmates and from other households living nearby who are warned on the same day.

Our finding that people adjust their behavior to the strictness of enforcement implies that not only formal rules but actual enforcement of program conditions is crucial for program effectiveness. This aspect seems particularly important for developing countries, as they might lack administrative capacity or political will to strictly enforce the rules.[30] Thus the design of conditional welfare programs should take into account the important dimensions of monitoring and enforcement.

References

[1] Angelucci, M. and G. D. Giorgi (2009), "Indirect effects of an aid program: How do cash transfers affect ineligibles' consumption?", *American Economic Review,* 99(1), 486–508.

[2] Angelucci, M., G. D. Giorgi, M. Rangel, and I. Rasul (2010). "Family networks and school enrollment: Evidence from a randomized social experiment", *Journal of Public Economics,* 94(3-4), 197–221.

[30] Brollo, Kaufmann and La Ferrara (2015) study the electoral costs of enforcement and the political incentives for manipulating the implementation of BFP conditionality in Brazil.

[3] Attanasio, O. P., C. Meghir, and A. Santiago (2012), “Education choices in Mexico: Using a structural model and a randomized experiment to evaluate Progresa”, *Review of Economic Studies*, 79(1), 37-66.

[4] Baird, S., C. McIntosh and B. Özler (2011), “Cash or condition? Evidence from a cash transfer experiment”, *The Quarterly Journal of Economics,* 126(4), 1709–1753.

[5] Banerjee, A. V., R. Glennerster, and E. Duflo (2008), “Putting a band-aid on a corpse: Incentives for nurses in the Indian public health care system”, *Journal of the European Economic Association,* 6(2-3), 487–500.

[6] Barrera-Osorio, F., M. Bertrand, L. L. Linden, and F. Perez-Calle (2011), “Improving the design of conditional transfer programs: Evidence from a randomized education experiment in Colombia”, *American Economic Journal: Applied Economics,* 3(2), 167–195.

[7] Bastagli, F. (2008). “Conditionality in public policy targeted to the poor: Promoting resilience?”, *Social Policy & Society,* 8 (1).

[8] Benhassine, N., F. Devoto, E. Duflo, P. Dupas and V. Pouliquen (2015) “Turning a Shove into a Nudge? A ”Labeled Cash Transfer” for Education”, *AEJ: Economic Policy*.

[9] Black, D., J. Smith, M. Berger and B. Noel (2003) “Is the Threat of Reemployment Services More Effective than the Services Themselves? Evidence from Random Assignment in the UI System.”, *The American Economic Review,* 93(4), 1313–1327.

[10] Bobonis, G. J. and F. Finan (2009). “Neighborhood peer effects in secondary school enrollment decisions”, *The Review of Economics and Statistics,* 91(4), 695–716.

[11] Bourguignon, F., F. Ferreira, and P. G. Leite (2003). “Conditional cash transfers, schooling and child labor: Micro-simulating brazil’s bolsa escola program”, *The World Bank Economic Review,* 17(2).

[12] Brollo, F., K. M. Kaufmann and E. La Ferrara (2015), "The Political Economy of Program Enforcement: Evidence from Brazil", mimeo, Bocconi University, Mannheim University and University of Warwick.

[13] Bursztyn, L. and L. C. Coffman (2012). "The Schooling Decision: Family Preferences, Intergenerational Conflict, and Moral Hazard in the Brazilian Favelas", *Journal of Political Economy,* 120(3), 359-397.

[14] Cameron, L. and M. Shah (2014). "Mistargeting of Cash Transfers, Social Capital Destruction, and Crime in Indonesia", *Economic Development and Cultural Change*, 62(2), 381-415.

[15] Chetty, R., J. Friedman and E. Saez (2013), "Using Differences in Knowledge Across Neighborhoods to Uncover the Impacts of the EITC on Earnings", *American Economic Review*, 103(7), 2683-2721.

[16] Chioda, L., J. M. P. de Mello and R. Soarez (2016) "Spillovers from Conditional Cash Transfer Programs: Bolsa Familia and Crime in Urban Brazil", *Economics of Education Review*.

[17] De Janvry, A., F. Finan, and E. Sadoulet (2011). "Local Electoral Incentives and Decentralized Program Performance", *Review of Economics and Statistics*, 94(3), 672-685.

[18] DeBrauw, A. and J. Hoddinott (2010). "Must conditional cash transfer programs be conditioned to be effective? the impact of conditioning transfers on school enrollment in Mexico", *Journal of Development Economics*.

[19] DeBrauw, A., D. Gilligan, J. Hoddinott and S. Roy (2015a). "The impact of Bolsa Familia on education", *World Development* , 70(6): 303-316.

[20] DeBrauw, A., D. Gilligan, J. Hoddinott and S. Roy (2015b). "Bolsa Familia and household labor supply", *Economic Development and Cultural Change* , 63(3): 423-457.

[21] Drago, F., F. Mengel, and C. Traxler (2015) "Compliance Behavior in Networks: Evidence from a Field Experiment", *IZA DP No. 9443, October 2015*.

[22] Duflo, E. and E. Saez (2003). "The role of information and social interactions in retirement plan decisions: Evidence from a randomized experiment", *Quarterly Journal of Economics*.

[23] Kumar, A., A. Nair, A. Parsons and E. Urdapilleta (2006) "Expanding Bank Outreach through Retail Partnerships Correspondent Banking in Brazil", *World Bank Working Paper No. 85* .

[24] Lalive, R., J. van Ours and J. Zweimüller (2005) "The Effect of Benefit Sanctions on the Duration of Unemployment", *Journal of the European Economic Association*, 2005, 3(6): 1386-1417.

[25] Linden, L. and K. Shastri (2012), "Grain Inflation: Identifying Agent Discretion in Response to a Conditional School Nutrition Program."*Journal of Development Economics*, 99(1), 128-138.

[26] Lindert, K., A. Linder, J. Hobbs, and B. de la Brière (2007). "The nuts and bolts of brazil's bolsa família program: Implementing conditional cash transfers in a decentralized context", *The World Bank, Social Protection Working Paper No. 0709*.

[27] Lochner, L. (2007). "Individual perceptions of the criminal justice system", *American Economic Review,* 97(1), 444–460.

[28] Rincke, J. and C. Traxler (2011). "Enforcement spillovers", *The Review of Economics and Statistics,* 93(4), 1224–1234.

[29] Schady, N. and M. C.Araujo (2006). "Cash transfers, conditions, school enrollment, and child work: Evidence from a randomized experiment in Ecuador", *World Bank Policy Research Working Paper 3930*, The World Bank.

[30] Schultz, P. T. (2004). "School subsidies for the poor: Evaluating the mexican Progresa poverty program", *Journal of Development Economics*.

[31] Todd, P. E. and K. I. Wolpin (2006). “Assessing the impact of a school subsidy program in Mexico: Using a social experiment to validate a dynamic behavioral model of child schooling and fertility”, *American Economic Review*.

[32] Van den Berg, G. J., B. van der Klaauw, and J. C. van Ours (2004). “Punitive sanctions and the transition rate from welfare to work”, *Journal of Labor Economics,* 22, 211–241.

APPENDIX

Figure 1: Timing of Noncompliance and Warning

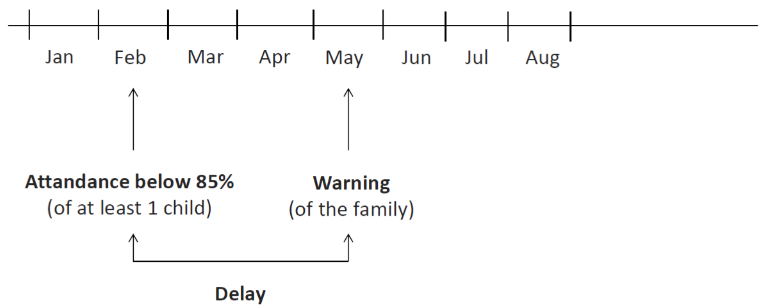


Figure 2: Response to warning

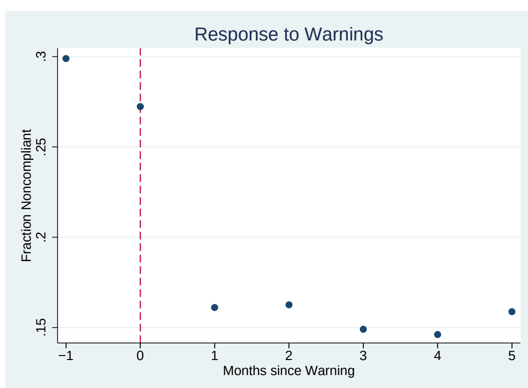


Figure 3: Response to warning, by delay

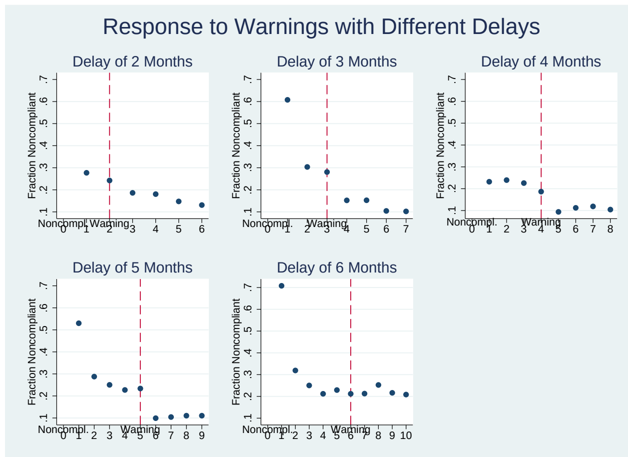


Table 1: Summary statistics

Variable	Observations	Mean	Std. Dev.
A. Own Warnings			
Noncompliance	8090769	0.086	0.280
Warning Stage 1	8090769	0.292	0.455
Warning Stage 2	8090769	0.085	0.279
Warning Stage 3	8090769	0.029	0.169
Warning Stage 4	8090769	0.011	0.103
Warning Stage 5	8090769	0.002	0.041
Warned	8090769	0.035	0.184
Delay of Warnings	267626	3.809	1.398
B. Peer Warnings			
Frac of Classmates Warned	5914381	0.014	0.059
Frac of Classmates Noncompl.	5914381	0.010	0.046
No of Peers in School	5914381	279.407	222.976
No of Peers in Grade	5914381	51.970	50.544
No of Peers in Class	5914381	16.533	6.752

Table 2: Effect of own warnings on attendance

Dependent Variable:	**Noncompliance** (in a Given Month)
Warning Stage 1	-0.0624***
	(0.0000)
Warning Stage 2	-0.1303***
	(0.001)
Warning Stage 3	-0.2050***
	(0.002)
Warning Stage 4	-0.2656***
	(0.003)
Warning Stage 5	-0.3477***
	(0.008)
Controls	Yes
Household FE	Yes
Time FE	Yes
No. Obs.	8,090,769
R-squared	0.23

Notes: Robust standard errors in parentheses (clustered at the household level). *** p<0.01, ** p<0.05, * p<0.1. Included controls are: fraction of male children, number of boys and girls in the household in the age categories 6 to 10, 11 to 15, and 16 to 18, month and year dummies and household fixed effects.

Table 3: Effect of own warnings, random variation in timing

Dependent Variable:	**Noncompliance in a Given Month**		
Timing:	Early	Late	Diff (Early-Late)
	Coeff/(S.E.)	Coeff/(S.E.)	Diff/[p-val]
Warning Stage 1 * Timing	-0.0253***	-0.0112***	-0.0141
	(0.001)	(0.001)	[0.0000]
Warning Stage 2 * Timing	-0.0474***	-0.0052***	-0.0422
	(0.002)	(0.002)	[0.0000]
Warning Stage 3 * Timing	-0.1267***	-0.0633***	-0.0634
	(0.004)	(0.003)	[0.0000]
Warning Stage 4 * Timing	-0.1887***	-0.1168***	-0.0719
	(0.007)	(0.006)	[0.0000]
Warning Stage 5 * Timing	-0.1641***	-0.0955***	-0.0686
	(0.016)	(0.013)	[0.0002]
Lag Warning Stage 1 * Timing	-0.0356***	-0.0484***	0.0128
	(0.001)	(0.001)	[0.0000]
Lag Warning Stage 2 * Timing	-0.0895***	-0.1230***	0.0335
	(0.002)	(0.002)	[0.0000]
Lag Warning Stage 3 * Timing	-0.0930***	-0.1438***	0.0508
	(0.004)	(0.003)	[0.0000]
Lag Warning Stage 4 * Timing	-0.0877***	-0.1513***	0.0636
	(0.007)	(0.006)	[0.0000]
Lag Warning Stage 5 * Timing	-0.2134***	-0.2715***	0.0581
	(0.018)	(0.015)	[0.0086]
Controls	Yes		
Household FE	Yes		
Time FE	Yes		
No. Obs.	7,551,546		
R-squared	0.23		

Notes: ”Early” (”Late”) indicates households that withdraw the transfer -and receive warnings, if applicable- in the first (second) half of the withdrawing period during a month. Robust standard errors in parentheses (clustered at the household level). *** p<0.01, ** p<0.05, * p<0.1. Included controls are: fraction of male children, number of boys and girls in the household in the age categories 6 to 10, 11 to 15, and 16 to 18, month and year dummies. All specifications include household fixed effects.

Table 4: Effect of classmates' warnings

Dependent Variable:	Noncompliance in a Given Month					
Sample	*All Children*			*Children with Siblings*		
	Benchmark	Placebo	Control for Peers' Noncompl.	Benchmark	Siblings' Classmates	Siblings' Classmates Other Schools
	(1)	(2)	(3)	(4)	(5)	(6)
Fraction of Classmates Warned	-0.0112***	-0.0120***	-0.0100***	-0.0106***	-0.0085***	-0.0099***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)
Lag of Frac of Classmates Warned	-0.0231***	-0.0236***	-0.0199***	-0.0227***	-0.0189***	-0.0191***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)
Lead of Frac of Classmates Warned		0.0001				
		(0.002)				
Fraction of Classmates Noncompliant			0.2807***			
			(0.021)			
Lag of Frac of Classmates Noncompliant			0.0952***			
			(0.009)			
Max Frac of Siblings' Classmates Warned					-0.0047**	
					(0.002)	
Lag Max Frac of Siblings' Classmates Warned					-0.0076***	
					(0.002)	
Max Frac of Siblings' Classmates Warned						-0.0013
(Only Sibs at Other Schools)						(0.002)
Lag Max Frac of Siblings' Classmates Warned						-0.0061**
(Only Sibs at Other Schools)						(0.003)
Controls for Own Warning Stage	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Individual FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
No. Obs.	5,914,114	5,628,700	5,914,114	4,902,847	4,902,847	4,902,847
R-squared	0.28	0.27	0.28	0.28	0.28	0.28

Notes: Robust standard errors in parentheses (clustered at the household level). *** p<0.01, ** p<0.05, * p<0.1. Included controls are: age dummies, birth order dummies, number of brothers and sisters in the age categories 6 to 10, 11 to 15, and 16 to 18, month and year dummies. All specifications include individual fixed effects.

Table 5: Information content of classmates' warnings

Dependent Variable:	**Noncompliance in Given Month**			
Own Warning Stage:	1	2	3	4
Frac Classmates Warned (WS 1)	-0.0123***	-0.0139	-0.0118	0.0318
	(0.003)	(0.009)	(0.016)	(0.028)
Frac Classmates Warned (WS 2)	-0.0245***	-0.0203**	-0.0171	-0.0693
	(0.007)	(0.009)	(0.033)	(0.044)
Frac Classmates Warned (WS 3)	-0.0188	-0.0308*	-0.0501**	-0.0392
	(0.012)	(0.017)	(0.021)	(0.051)
Frac Classmates Warned (WS 4)	-0.0197	-0.0386	-0.0787*	-0.0590**
	(0.022)	(0.026)	(0.043)	(0.027)
Frac Classmates Warned (WS 5)	0.0274	0.0840	0.0392	0.1151
	(0.036)	(0.070)	(0.088)	(0.123)
Controls	Yes	Yes	Yes	Yes
Individual FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
No. Obs.	1,993,576	639.362	232.623	88.697
R-squared	0.46	0.52	0.52	0.54

Notes: Robust standard errors in parentheses (clustered at the household level). *** p<0.01, ** p<0.05, * p<0.1. Included controls are: age dummies, birth order dummies, number of brothers and sisters in the age categories 6 to 10, 11 to 15, and 16 to 18, month and year dummies.

Table 6: Effect of warnings of neighbors in the bank

Dependent Variable:	**Noncompliance in Given Month**		
	Benchmark	Placebo	Controls for Neighbors Warned
	(1)	(2)	(3)
Frac in Bank Warned	-0.0275***	-0.0275***	-0.0349***
	(0.004)	(0.004)	(0.006)
Lag of Frac in Bank Warned	-0.0548***	-0.0547***	-0.0284***
	(0.003)	(0.003)	(0.004)
Lead of Frac in Bank Warned		0.0007	
		(0.004)	
Frac Neighbors Warned			0.0164*
			(0.009)
Lag of Frac Neighbors Warned			-0.0609***
			(0.007)
No of BFP Neighbors	-0.0000***	-0.0000***	-0.0000***
	(0.000)	(0.000)	(0.000)
Controls for Own Warning Stage	Yes	Yes	Yes
Controls	Yes	Yes	Yes
Individual FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
Observations	4,813,029	4,813,029	4,813,029
R-squared	0.21	0.21	0.21

Notes: Robust standard errors in parentheses (clustered at the household level). *** p<0.01, ** p<0.05, * p<0.1. Included controls are: age dummies, birth order dummies, number of brothers and sisters in the age categories 6 to 10, 11 to 15, and 16 to 18, month and year dummies and individual fixed effects.

Online Appendix – Not For Publication

A. Difference-in-Difference Analysis

Further insights into the causal nature of the estimates can be gained by making use of variation in the delay of warnings across months. As discussed, the variation in delay is almost entirely explained by time and area fixed effects and there is no correlation with household or child characteristics. We thus estimate a difference-in-differences model that is the regression counterpart of figures 2 and 3. In particular, we compare two groups of households: both groups failed to comply in the same month, but they receive their warnings with different delays. We then compare the behavior of households who have already been warned (the "treated" households) with those who have not yet received their warning (controls). Let "Post" be an indicator taking value 1 in the calendar month(s) after the warning of the households receiving the earlier warnings and 0 otherwise. We estimate:

$$Y_{ht} = \alpha \cdot Treated_{ht} + \beta \cdot Post_{ht} + \gamma(Treated_{ht} \cdot Post_{ht}) + \delta X_{ht} + D_t + D_h + \epsilon_{ht}. \quad (7)$$

Our coefficient of interest is γ, which we expect to be negative if households react to warnings by decreasing noncompliance (i.e. increasing attendance) of their children. The results are reported in Appendix Table A.6 using different time horizons for pre- and post-warning.

[Insert Appendix Table A.6]

In the first column, we consider an interval of three months around a warning and let "*Post*" take value 1 in the month following the warning, and 0 in the month of the warning and the month before. The estimated coefficient suggests that households who receive a warning decrease their noncompliance rate by 3.8 percentage points compared to households that have not yet received their warning. This represents almost a 30 percent decrease over the mean. The estimates are of a similar order of magnitude if we expand the analysis to a period of five months around a warning and define as "*Post*" the two months following the warning: the estimated coefficient in that case is -0.045, significant at the 1 percent level.

To provide further evidence that this is actually identifying a behavioral response to the warning, we show that pre-trends are the same for both groups supportin the common trend assumption underlying our difference-in-difference analysis. In columns 3 and 4 we conduct the following falsification tests. We define as "*Post*" the month of the warning and use as pre-period the month before (column 3) or the two months before (column 4). If our negative estimate were picking up a differential trend for warned households, this would result in a negative coefficient on "*Treated*∗*Post*" in columns 3 and 4. The fact that we instead obtain small and insignificant coefficients on the interaction term lends credibility to the interpretation that our effect is a causal response to the receipt of the warning.

B. Appendix Figures and Tables

Figures

Figure A.1: Frequency of delays

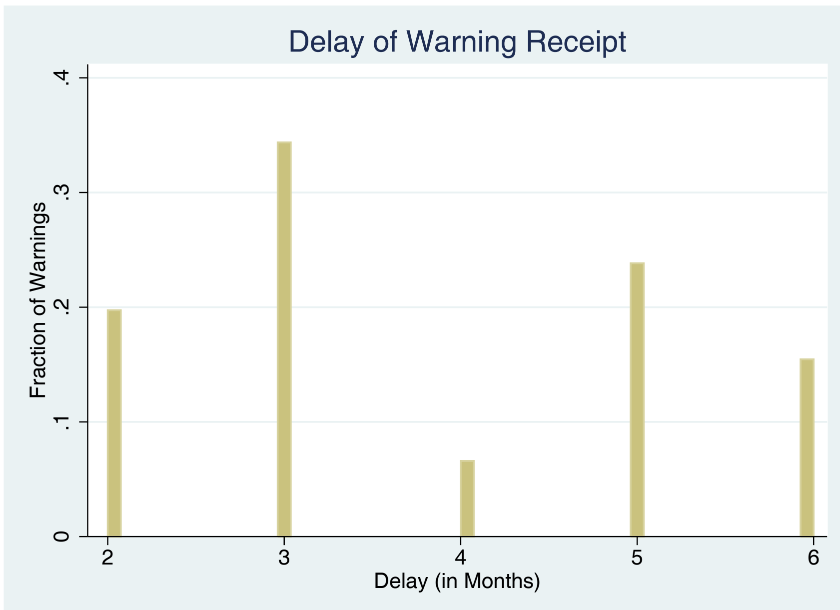
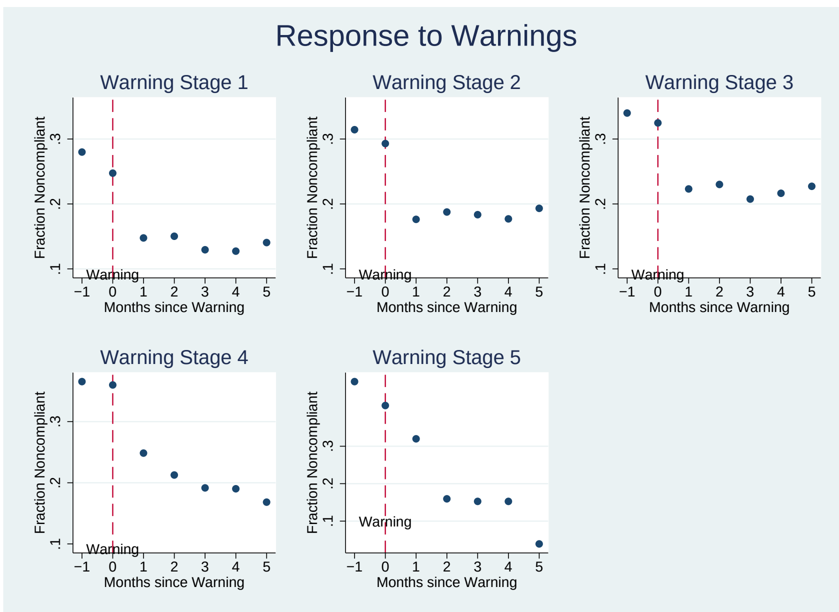


Figure A.2: Response to warning, by warning stage



Tables

Table A.1: Summary statistics

Variable	Observations	Mean	Std. Dev.
Own Warnings			
Noncompliance	8090769	0.086	0.280
Warning Stage 1	8090769	0.292	0.455
Warning Stage 2	8090769	0.085	0.279
Warning Stage 3	8090769	0.029	0.169
Warning Stage 4	8090769	0.011	0.103
Warning Stage 5	8090769	0.002	0.041
Warned	8090769	0.035	0.184
Delay of Warnings	267626	3.809	1.398
Frac Male	8090769	0.530	0.387
N of Sisters 6 to 10	8090769	0.206	0.386
N of Brothers 6 to 10	8090769	0.222	0.400
N of Sisters 11 to 14	8090769	0.250	0.412
N of Brothers 11 to 14	8090769	0.271	0.432
N of Sisters 15 to 17	8090769	0.154	0.340
N of Brothers 15 to 17	8090769	0.172	0.360
Month	8090769	6.538	2.864
Year 2009	8090769	0.492	0.500
Parents informed in Writing	5044299	0.790	0.408
Parents informed in Meetings	5044299	0.940	0.169
Parents informed in Home Visits	5044299	0.703	0.457
Peer Warnings			
Frac of Peers Warned	5914381	0.014	0.059
Frac of Peers Noncompl	5914381	0.010	0.046
Max Fraction of Siblings' Peers Warned			
All Siblings	4903020	0.014	0.058
Siblings in Other Schools	4903020	0.013	0.058
N of Peers in School	5914381	279.407	222.976
N of Peers in Grade	5914381	51.970	50.544
N of Peers in Class	5914381	16.533	6.752
Predetermined Characteristics			
Head Female	478511	0.933	0.249
Head Married	478511	0.337	0.468
Head Single	478511	0.531	0.495
Head Age	478511	38.063	10.028
Head White	478511	0.182	0.384
Head Yrs of Educ	478511	4.517	3.227
Head Working	478511	0.510	0.487
Spouse Age	261318	40.924	10.536
Spouse White	261318	0.167	0.371
Spouse Yrs of Educ	261318	3.519	3.057
Spouse Working	261318	0.738	0.430
Dep Ratio	478511	1.092	1.033
Pct Indio	478511	0.002	0.040
HH Expenditures	478511	2254.6	15356.5
House Property	478511	0.695	0.453
Garbage Collected	478511	0.594	0.487
Running Water	478511	0.610	0.482
Electricity	478511	0.793	0.398

Table A.2: Effect of own warnings, individual level

Dependent Variable:	**Noncompliance** (in a Given Month)
Warning Stage 1	-0.0336***
	(0.0000)
Warning Stage 2	-0.0657***
	(0.001)
Warning Stage 3	-0.0970***
	(0.001)
Warning Stage 4	-0.1147***
	(0.002)
Warning Stage 5	-0.1366***
	(0.005)
Controls	Yes
Individual FE	Yes
Time FE	Yes
No. Obs.	16,056,547
R-squared	0.23

Notes: Robust standard errors in parentheses (clustered at the household level). *** p<0.01, ** p<0.05, * p<0.1. Included controls are: age dummies, birth order dummies, number of brothers and sisters in the age categories 6 to 10, 11 to 15, and 16 to 18, month and year dummies and individual fixed effects.

Table A.3: Correlates of delay

Dependent Variable:	Delay of Warning		
	(1)	(2)	
		Coeff	(Std.Err.)
N of Girls 6 to 10		-0.0009	(0.001)
N of Boys 6 to 10		-0.0005	(0.001)
N of Girls 11 to 14		0.0019	(0.001)
N of Boys 11 to 14		-0.0009	(0.001)
N of Girls 15 to 17		-0.0019	(0.002)
N of Boys 15 to 17		0.0002	(0.001)
Head Female		0.0006	(0.002)
Head Married		-0.0033	(0.004)
Head Single		-0.0026	(0.004)
Head White		0.0009	(0.002)
Head Age		0.0001	(0.000)
Head Yrs of Educ		0.0000	(0.000)
Head Work		0.0001	(0.001)
Spouse White		-0.0003	(0.002)
Spouse Age		0.0001	(0.000)
Spouse Yrs of Educ		0.0000	(0.000)
Spouse Work		-0.0009	(0.001)
Dependency Ratio		0.0001	(0.001)
Expenditures		0.0000	(0.000)
House Property		-0.0023*	(0.001)
Garbage Collection		-0.0004	(0.001)
Water Connection		-0.0012	(0.001)
Electricity		-0.0009	(0.001)
Time FE	Yes	Yes	
Municipality FE	Yes	Yes	
P-val of F-test			
(joint sig of HH charac)		0.910	
Observations	267,626	111,462	
R-squared	0.98	0.98	

Notes: Robust standard errors in parentheses (clustered at the household level). *** p<0.01, ** p<0.05, * p<0.1.

Table A.4: Effect of own warnings, by highest warning stage reached

Dependent Variable	**Noncompliance in a Given Month**				
Highest WS Reached	1	2	3	4	5
Warning Stage 1	-0.1863***	-0.1677***	-0.1716***	-0.1385***	-0.1464***
	(0.001)	(0.002)	(0.004)	(0.009)	(0.027)
Warning Stage 2		-0.5120***	-0.4907***	-0.2874***	-0.2960***
		(0.003)	(0.005)	(0.012)	(0.035)
Warning Stage 3			-0.9837***	-0.7029***	-0.6010***
			(0.006)	(0.017)	(0.048)
Warning Stage 4				-1.1119***	-0.8757***
				(0.020)	(0.059)
Warning Stage 5					-1.2264***
					(0.072)
Controls	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes
Household FE	Yes	Yes	Yes	Yes	Yes
P-Values of Tests:					
Coeff of WS 1 vs 2		0.000	0.000	0.000	0.000
Coeff of WS 2 vs 3			0.000	0.000	0.000
Coeff of WS 3 vs 4				0.000	0.000
Coeff of WS 4 vs 5					0.000
No. Obs.	2,977,415	814,758	221,764	55,233	8,350
R-Squared	0.14	0.21	0.27	0.27	0.27

Notes: Initial warning stage is 0 (i.e. warning stage at the beginning of our period of observation, namely January 2008). Robust standard errors in parentheses (clustered at the household level). *** p<0.01, ** p<0.05, * p<0.1. Included controls are: fraction of male children, number of boys and girls in the household in the age categories 6 to 10, 11 to 15, and 16 to 18, month and year dummies. All specifications include household fixed effects.

Table A.5: Balance test

Timing:	Early Mean/SE	Late Mean/SE	Diff/[p-val]
Time-Varying Characteristics			
N of Sisters 6 to 10	0.1643	0.1644	-0.0001
	(0.3432)	(0.3421)	[0.588]
N of Brothers 6 to 10	0.1738	0.1738	-0.0001
	(0.3529)	(0.3521)	[0.755]
N of Sisters 11 to 14	0.1894	0.1902	-0.0009
	(0.3643)	(0.3654)	[0.002]
N of Brothers 11 to 14	0.1949	0.1948	0.0001
	(0.3708)	(0.3687)	[0.617]
N of Sisters 15 to 17	0.1157	0.1142	0.0014
	(0.2984)	(0.2953)	[0.000]
N of Brothers 15 to 17	0.1186	0.1191	-0.0003
	(0.3017)	(0.3037)	[0.184]
No. Obs.	3,021,032	4,530,514	7,551,546
Pre-Program Characteristics			
Head Married	0.3366	0.3383	-0.0017
	(0.4683)	(0.4689)	[0.221]
Head Single	0.5333	0.5321	0.0012
	(0.4951)	(0.4950)	[0.429]
Head Female	0.9351	0.9336	0.0015
	(0.2457)	(0.2480)	[0.043]
Head Age	38.0615	38.0472	0.0143
	(10.0252)	(10.0127)	[0.630]
Head White	0.1821	0.1819	0.0003
	(0.3839)	(0.3836)	[0.814]
Head Years of Education	4.5259	4.5145	0.0114
	(3.2404)	(3.2181)	[0.251]
Head Work	0.5114	0.5079	0.0036
	(0.4870)	(0.4874)	[0.014]
Spouse Age	40.9540	40.9009	0.0531
	(10.5281)	(10.5078)	[0.209]
Spouse White	0.1653	0.1686	-0.0032
	(0.3697)	(0.3725)	[0.035]
Spouse Years of Education	3.5286	3.5158	0.0127
	(3.0540)	(3.0580)	[0.311]
Spouse Work	0.7389	0.7395	-0.0006
	(0.4299)	(0.4294)	[0.725]
Dependency Ratio	1.0941	1.0903	0.0039
	(1.0307)	(1.0354)	[0.207]
Pct Indio	0.0016	0.0017	-0.0001
	(0.0389)	(0.0402)	[0.452]
Expenditure	2099.51	2227.50	-127.987
	(15129.77)	(15070.91)	[0.775]
House Property	0.6946	0.6953	-0.0007
	(0.4538)	(0.4533)	[0.620]
Garbage Collected	0.5931	0.5946	-0.0016
	(0.4878)	(0.4874)	[0.285]
Water Connection	0.6115	0.6096	0.0019
	(0.4815)	(0.4822)	[0.192]
Electricity	0.7937	0.7925	0.0012
	(0.3976)	(0.3984)	[0.330]
No. Obs.	188,970	284,527	472,029

Notes: ”Early” (”Late”) indicates households that withdraw the transfer -and receive warnings, if applicable- in the first (second) half of the withdrawing period during a month.

Table A.6: Effect of own warnings, difference-in-differences

	Main result		**Placebo**	
	Definition 1 Warning: t=0 Before: t=-1,0 Post: t=1	Definition 2 Warning: t=0 Before: t=-2,-1,0 Post: t=1,2	Definition 1 Warning: t=0 Before: t=-1 Post: t=0	Definition 2 Warning: t=0 Before: t=-2,-1 Post: t=0
Treat * Post	-0.0379***	-0.0447***	0.002	-0.0037
	(0.002)	(0.003)	(0.003)	(0.003)
Treat	0.0014	0.0061***	0.0021	0.0078***
	(0.001)	(0.002)	(0.002)	(0.002)
Post	-0.1853***	-0.0153***	-0.2642***	-0.0269***
	(0.001)	(0.002)	(0.002)	(0.002)
No. Obs.	763,847	458,967	458,224	309,942
R-squared	0.05	0.03	0.06	0.03

Notes: Robust standard errors in parentheses (clustered at the household level). *** p<0.01, ** p<0.05, * p<0.1. Included controls are: fraction of male children, number of boys and girls in the household in the age categories 6 to 10, 11 to 15, and 16 to 18, month and year dummies.

Table A.7: Learning about attendance

Dependent Variable:	Noncompliance in a Given Month			
Parents Informed in:	Writing	Meeting	Home Visit	Benchmark
Warning Stage 1 * Parents Informed	-0.0003	-0.0021	0.0002	
	(0.001)	(0.003)	(0.001)	
Warning Stage 2 * Parents Informed	-0.0008	-0.0039	0.0006	
	(0.003)	(0.007)	(0.002)	
Warning Stage 3 * Parents Informed	-0.0100**	0.0204	-0.0066	
	(0.005)	(0.013)	(0.005)	
Warning Stage 4 * Parents Informed	-0.0096	0.0472*	-0.0009	
	(0.009)	(0.024)	(0.009)	
Warning Stage 5 * Parents Informed	-0.0293	0.1017*	0.0159	
	(0.023)	(0.054)	(0.021)	
Warning Stage 1		-0.0619***		-0.0640***
		(0.003)		(0.0000)
Warning Stage 2		-0.1284***		-0.1324***
		(0.006)		(0.001)
Warning Stage 3		-0.2106***		-0.2034***
		(0.013)		(0.002)
Warning Stage 4		-0.3002***		-0.2626***
		(0.024)		(0.004)
Warning Stage 5		-0.4269***		-0.3401***
		(0.051)		(0.0100)
Controls		Yes		Yes
Household FE		Yes		Yes
Time FE		Yes		Yes
No. Obs.		5,044,299		5,044,299
R-Squared		0.24		0.24

Notes: Robust standard errors in parentheses (clustered at the household level). *** p<0.01, ** p<0.05, * p<0.1. Included controls are: fraction of male children, number of boys and girls in the household in the age categories 6 to 10, 11 to 15, and 16 to 18, month and year dummies. All specifications include household fixed effects.

Table A.8: Effect of classmates' warnings

Dependent Variable:	**Noncompliance in a Given Month**						
Sample		*All Children*				*Children with Siblings*	
	Own Warnings	Benchmark	Placebo	Control for Peers' Noncompl.	Benchmark	Siblings' Peers	Siblings' Peers Other Schools
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Fraction of Classmates Warned		-0.0112***	-0.0120***	-0.0100***	-0.0106***	-0.0085***	-0.0099***
		(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)
Lag of Frac of Classmates Warned		-0.0231***	-0.0236***	-0.0199***	-0.0227***	-0.0189***	-0.0191***
		(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)
Lead of Frac of Classmates Warned			0.0001				
			(0.002)				
Fraction of Classmates Failed				0.2807***			
				(0.021)			
Lag of Frac of Classmates Failed				0.0952***			
				(0.009)			
Max Frac of Siblings' Classmates Warned						-0.0047**	
						(0.002)	
Lag Max Frac of Siblings' Classmates Warned						-0.0076***	
						(0.002)	
Max Frac of Siblings' Classmates Warned							-0.0013
(Only Sibs at Other Schools)							(0.002)
Lag Max Frac of Siblings' Classmates Warned							-0.0061**
(Only Sibs at Other Schools)							(0.003)
Warning Stage 1	-0.0110***	-0.0107***	-0.0110***	-0.0103***	-0.0092***	-0.0092***	-0.0092***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Warning Stage 2	-0.0190***	-0.0188***	-0.0192***	-0.0182***	-0.0175***	-0.0175***	-0.0175***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Warning Stage 3	-0.0229***	-0.0227***	-0.0232***	-0.0219***	-0.0222***	-0.0222***	-0.0222***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Warning Stage 4	-0.0214***	-0.0213***	-0.0213***	-0.0207***	-0.0217***	-0.0217***	-0.0217***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Warning Stage 5	-0.0249***	-0.0248***	-0.0270***	-0.0249***	-0.0242***	-0.0242***	-0.0242***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
No. Obs.	5,914,114	5,914,114	5,628,700	5,914,114	4,902,847	4,902,847	4,902,847
R-squared	0.28	0.28	0.27	0.28	0.28	0.28	0.28

Notes: Robust standard errors in parentheses (clustered at the household level). *** p<0.01, ** p<0.05, * p<0.1. Included controls are: age dummies, birth order dummies, number of brothers and sisters in the age categories 6 to 10, 11 to 15, and 16 to 18, month and year dummies. All specifications include individual fixed effects.

Table A.9: Effect of classmates' warnings (clustering of standard errors at the school level)

Dependent Variable:	Noncompliance in a Given Month						
Sample	*All Children*				*Children with Siblings*		
	Own Warnings	Benchmark	Placebo	Control for Peers' Noncompl.	Benchmark	Siblings' Peers	Siblings' Peers Other Schools
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Fraction of Classmates Warned		-0.0115***	-0.0124***	-0.0103***	-0.0109***	-0.0086***	-0.0101***
		(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)
Lag of Frac of Classmates Warned		-0.0239***	-0.0244***	-0.0207***	-0.0233***	-0.0194***	-0.0197***
		(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)
Lead of Frac of Classmates Warned			-0.0003				
			(0.002)				
Fraction of Classmates Noncompliant				0.2860***			
				(0.022)			
Lag of Frac of Classmates Noncompliant				0.0964***			
				(0.009)			
Max Frac of Siblings' Classmates Warned						-0.0049**	
						(0.002)	
Lag Max Frac of Siblings' Classmates Warned						-0.0081***	
						(0.002)	
Max Frac of Siblings' Classmates Warned							-0.0014
(Only Sibs at Other Schools)							(0.003)
Lag Max Frac of Siblings' Classmates Warned							-0.0063**
(Only Sibs at Other Schools)							(0.003)
Warning Stage 1	-0.0111***	-0.0108***	-0.0112***	-0.0104***	-0.0093***	-0.0093***	-0.0093***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Warning Stage 2	-0.0193***	-0.0190***	-0.0195***	-0.0184***	-0.0178***	-0.0177***	-0.0177***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Warning Stage 3	-0.0231***	-0.0230***	-0.0234***	-0.0221***	-0.0224***	-0.0223***	-0.0223***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Warning Stage 4	-0.0218***	-0.0218***	-0.0218***	-0.0211***	-0.0220***	-0.0220***	-0.0220***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Warning Stage 5	-0.0248***	-0.0246***	-0.0269***	-0.0248***	-0.0241***	-0.0241***	-0.0241***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
No. Obs.	5,833,861	5,554,369	5,833,861	4,837,760	4,837,760	4,837,760	
R-squared	0.28	0.28	0.29	0.29	0.29	0.29	

Notes: Robust standard errors in parentheses (clustered at the school level). *** p<0.01, ** p<0.05, * p<0.1. Included controls are: age dummies, birth order dummies, number of brothers and sisters in the age categories 6 to 10, 11 to 15, and 16 to 18, month and year dummies. All specifications include individual fixed effects.

Table A.10: Effect of warnings of neighbors in the bank (standard errors clustered at zip-code level)

Dependent Variable:	**Noncompliance in Given Month**		
	Benchmark	Placebo	Controls for Neighbors Warned
	(1)	(2)	(3)
Fraction in Bank Warned	-0.0276***	-0.0276***	-0.0349***
	(0.008)	(0.009)	(0.006)
Lag of Frac in Bank Warned	-0.0549***	-0.0548***	-0.0284***
	(0.011)	(0.011)	(0.004)
Lead of Frac in Bank Warned		0.0004	
		(0.008)	
Frac Neighbors Warned			0.0160
			(0.023)
Lag of Frac Neighbors Warned			-0.0612**
			(0.030)
No of BFP Neighbors	-0.0000	-0.0000	-0.0000
	(0.000)	(0.000)	(0.000)
Controls for Own Warning Stage	Yes	Yes	Yes
Controls	Yes	Yes	Yes
Individual FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
No. Obs.	4,797,371	4,797,371	4,797,361
R-squared	0.21	0.21	0.21

Notes: Robust standard errors in parentheses (clustered at the zip-code level). *** p<0.01, ** p<0.05, * p<0.1. Included controls are: age dummies, birth order dummies, number of brothers and sisters in the age categories 6 to 10, 11 to 15, and 16 to 18, month and year dummies and individual fixed effects.

Table A.11: Effect of warnings of neighbors in the bank

Dependent Variable:	Noncompliance in Given Month			
Own Warning Stage	1	2	3	4
Frac in Bank Warned (WS 1)	-0.0190***	-0.0160	-0.0083	0.0916
	(0.004)	(0.016)	(0.033)	(0.071)
Frac in Bank Warned (WS 2)	-0.0640***	-0.0430***	-0.0914**	-0.0598
	(0.014)	(0.008)	(0.038)	(0.056)
Frac in Bank Warned (WS 3)	-0.1121***	-0.0414	-0.0472***	-0.1801
	(0.023)	(0.040)	(0.014)	(0.121)
Frac in Bank Warned (WS 4)	-0.0506	-0.2090**	-0.0476	-0.0397*
	(0.043)	(0.098)	(0.092)	(0.023)
Frac in Bank Warned (WS 5)	-0.0218	-0.0301	0.5936	0.5250*
	(0.096)	(0.192)	(0.496)	(0.300)
Controls for Frac				
Neighbors Warned (By WS)	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Individual and Time FE	Yes	Yes	Yes	Yes
No. Obs.	3,842,873	1,169,159	420,041	155,993
R-Squared	0.27	0.32	0.34	0.36

Notes: Robust standard errors in parentheses (clustered at the household level). *** p<0.01, ** p<0.05, * p<0.1. Included controls are: age dummies, birth order dummies, number of brothers and sisters in the age categories 6 to 10, 11 to 15, and 16 to 18, month and year dummies and individual fixed effects.