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

Fernanda Brollo Katja Maria Kaufmann Eliana La Ferrara

This version: July 2015

Abstract

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 conditionality enforcement. We find that school attendance increases after families are punished for past noncompliance. Families also respond to penalties experienced by peers (spillover effects): 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.

*Fernanda Brollo, Warwick University and CAGE, f.brollo@warwick.ac.uk; Katja Maria Kaufmann, Bocconi University and IGIER, katja.kaufmann@unibocconi.it; Eliana La Ferrara, Bocconi University and IGIER, eliana.laferrara@unibocconi.it.

†We thank Manuela Angelucci, Josh Angrist, Oriana Bandiera, Felipe Barrera-Osorio, Esther Duflo, Eric Edmonds, Erika Field, Maitreesh Ghatak, Caroline Hoxby, Dan Keniston, Asim Khwaja, David Lam, Leigh Linden, Marco Manacorda, Karthik Muralidharan, Rohini Pande, Imran Rasul, Johannes Rincke, Paul Schultz, Jeff Smith, Rodrigo Soarez, Chris Udry, Gerard Van den Berg, 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 and child support. 'Conditional cash transfer' (CCT) programs, which provide a stipend to poor families as long as they meet certain conditions, have become a widely used tool to fight poverty in low and middle-income countries. Existing evaluations of conditional programs provide evidence on the combined effect of formal rules and the enforcement of these rules, but do not provide insights 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 aims at providing evidence on how the enforcement of program conditions affects beneficiaries' behavior.

We study the implementation of the "Bolsa Familia Program" (BFP) in Brazil, a large-scale conditional cash transfer program that provides monthly subsidies to poor families conditional on the fulfillment of certain obligations, the most important being that all school-aged children in the family must attend at least 85 percent of school hours 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, until the household is ultimately removed from the program. The goal of our paper is to analyze whether people respond to the enforcement of BFP conditions 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 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. 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 their noncompliance and punish them for it. If this probability were one, then at the time when the child fails to meet the school attendance requirement, the family would already anticipate that it will be penalized in the future, and the actual receipt of a warning or penalty several months later would contain no new information. If, on the other hand, the household believed that this probability were smaller than one, then it should respond to the receipt of a warning and increase school attendance to avoid further penalties. [1] Apart from responding to their own experiences of enforcement, families might also respond to their peers' experience of enforcement. A family that learns from peers that their noncompliance was penalized may update their beliefs about the strictness of enforcement and also respond in terms of increasing their school attendance.

To test these hypotheses we make use of a unique dataset that we compiled from administrative sources including the Brazilian Ministry of Social Development and the Ministry of Education for the years 2008-2009. These data include information about child and family characteristics for the universe of BFP beneficiaries in the Northeast of Brazil, as well as monthly records on school attendance of each child, warnings received (from which we create the complete warning history of each family), benefits disbursed and whether the benefit was suspended in a given month. By merging these data with the 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 study how people react to warnings and punishments received by other beneficiaries, i.e. to explore the dimension of learning from peers.

We present our results in two parts. In both, our outcome of interest is the likelihood that individuals fail to comply with program conditions in a given month (in short, "failure"). In the first part we analyze the extent to which people respond to own experiences of enforcement. We find that the likelihood of failure decreases after receiving a warning or penalty. The estimated size of the effect is a decrease in

[1] As we discuss in section 2, both the language used in the official warning messages and the experience in the early years of BFP implementation (when conditionality was essentially not enforced) make us hypothesize that during the period of our study there was a substantial degree of uncertainty about how strict enforcement would be.

failure of 4 percentage points, which represents a 30 percent decrease over the average monthly failure rate. We interpret this effect as causal for two reasons. First, we start by observing that the lag between the month in which a child failed to attend school and the month in which the household receives a warning varies over time in a non-monotonic way. This variation is driven by bureaucratic complexities (many different actors are involved in the implementation of the program) and is orthogonal to household characteristics. We can compare households who received warnings with different lags and we show that there is a significant behavioral response (decrease in "failure") precisely after the receipt of the warning. This allows us to rule out that, for example, mean-reverting shocks may be responsible for the adjustment we see.

To provide further evidence that we identify a causal effect we exploit actual random variation in the timing of arrival of warnings within each month. In fact, the date at which the households can withdraw the money –and receive the warning message while withdrawing, if applicable– depends on the last digit of a beneficiary'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. Indeed, we find that "early" warning receivers decrease their failure by more than "late" ones, which suggests that we are identifying a causal response to the warning, given that the two groups are otherwise observationally equivalent.

In the second part of the paper we study whether people learn from the experience of their peers, i.e. whether beneficiaries react to warnings received by *other* households. This is of interest in and of itself, because it allows us to quantify potential spillovers that matter for estimating the benefits of stricter program enforcement. But focusing on punishments received by others also has the following important advantage: to the extent that we are able to rely on shocks orthogonal to those experienced by the household itself, it allows us to analyze the importance of information transmission and learning in the context of program enforcement.

We estimate if a child's propensity to fail (i.e., not meet the 85 percent attendance threshold) 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 failure 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 whether we focus 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 in more depth, we make use of the fact that families progress through different "warning stages" if they repeatedly fail to comply with program conditions, i.e. they receive warnings/punishments that increase in severity with every additional instance of noncompliance. In particular, first the household receives a warning without penalty, then benefits are blocked for one month and recovered the following month, then twice suspended for two months and ultimately the household is removed from the program. When we condition on a family's own warning stage and distinguish between peers who get warned but are at the same or a lower warning stage, and peers who receive warnings for higher stages, the latter induce a larger decrease in failure. This is consistent with our learning interpretation, because warnings for stages higher than one's own convey new information on the likelihood that the government implements higher order punishments.

Finally, we discuss two alternative interpretations of our findings. The first is that warnings may convey information not about enforcement but about the fact that one's child has missed school. While this cannot explain reactions to warnings received by other children (i.e., our results on peers), we also provide direct evidence that this channel does not seem to drive the results on own warnings. Using information on the policies that school principals adopt to communicate with families about attendance, we show that the effect of warnings is not smaller for households who are directly informed by the school about missed attendance.

The second alternative interpretation is that people respond to warnings simply because they learn about the existence of the rules, and not because they update their priors on enforcement. This interpretation cannot explain some of our empirical results. First, we find that people still respond (and in fact, more strongly) to subsequent warnings after having experienced warnings in the past. Since every warning message lists program conditions and penalties in a clear and salient way, it seems unlikely that BFP beneficiaries would repeatedly ignore this information. Second, to the extent that knowledge of program rules is (inversely) related to education, we should find that less educated parents respond more to warnings, while we find very similar responses across the education spectrum (if anything, the results go in the opposite direction).

To summarize, our findings suggest that the quality of enforcement of government policies is important. People respond to and learn about enforcement, not only with respect to own experience of warnings but also with respect to the experiences of their peers. The evidence of important spillover effects further underlines the importance of monitoring and enforcement of program conditions for the effectiveness of conditional programs.

Our paper contributes to several strands of the 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 and show that health workers respond to incentives, but that the local administration ultimately undermined the incentive program by the government. Rincke and Traxler (2011) analyze spillovers from enforcement of compliance with TV license fees, using weather conditions as an instrument for local inspections. Black, Smith, Berger and Noel (2003) and Van den Berg, van der Klaauw and van Ours (2004) study the effects of unemployment insurance programs that monitored the job search effort of unemployed workers and threatened to impose the uptake of mandatory reemployment services or punished financially if they did not meet the effort requirements. Aside from the difference in context, our paper contributes to the above literature by focusing on the process through which program beneficiaries learn about the quality of enforcement and adjust their behavior. In this respect, our work is similar in spirit to Lochner (2007), who investigates the effect of learning about arrest probabilities on criminal behavior. Also, since we do not have direct measures of knowledge of program rules or beliefs about enforcement, we infer them from individuals' behavior. This approach is akin to Chetty, Friedman and Saez (2013), who infer individuals' knowledge of the Earned Income Tax Credit from their endogenous reporting of self-employment income.

Since we analyze the implementation of a conditional cash transfer program, our paper is also related to a large literature on CCT's. Among others, Attanasio, Meghir and Santiago (2012), DeBrauw and Hoddinott (2010), Schultz (2004), and Todd and Wolpin (2006) estimate the impact of Progresa/Oportunidades in Mexico; Bourguignon, Ferreira and Leite (2003), Bursztyn and Coffman (2012), De Janvry, Finan and Sadoulet (2011) study *Bolsa Escola*, the predecessor of *Bolsa Familia* in Brazil, while Bastagli (2008) focuses on Bolsa Familia.[2] 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 –to the best of our knowledge– not the enforcement of conditionalities.

Two other related papers on CCT's by Baird, McIntosh and Özler (2011) and by Benhassine et al (2015) compare the effectiveness of conditional versus unconditional cash transfers. While the former paper finds that the conditional arm of their intervention was more effective in terms of increasing children's school attendance, the latter paper concludes that a "labeled cash transfer", which is unconditional but labeled as an education support program, can be as effective. While those papers investigate whether transfers should be made conditional on certain behaviors or should be provided unconditionally, they do not provide evidence on the importance of enforcement when program conditions are in place.

Several papers analyze peer effects and spillovers in the context of conditional cash transfer programs, e.g., Barrera-Osorio, Bertrand, Linden and Perez-Calle

[2] Notice that there is relatively less evidence on Bolsa Familia compared to Progresa/Oportunidades, most likely due to the fact that BFP was not implemented in a randomized fashion and is therefore more difficult to evaluate.

(2011) for Colombia; Angelucci, De Giorgi, Rangel and Rasul (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, for example, the effect of peers' school attendance on own attendance). Instead we are interested in identifying the effect of peers receiving warnings (a signal about the quality of enforcement) on individual attendance decisions. This has the advantage of mitigating some of the identification problems that the conventional peer effect literature has to face. Our interest in how people learn about program features from their peers is shared, for example, by Duflo and Saez (2003) who analyze retirement savings decisions in the United States. Differently from these authors, we study learning about enforcement of program conditions in a very different context.

The outline of the paper is as follows: In Section 2 we provide background information on the Bolsa Familia program, in particular on program conditions and the mechanisms of enforcement. In Section 3 we present the data and descriptive statistics. In Sections 4 and 5 we discuss our empirical strategy and results on whether individuals respond to experiencing enforcement and whether they learn from peers, respectively. Section 6 concludes.

2 Background Information on the Bolsa Familia Program

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, Lula, in 2003 to consolidate four different programs (Federal Bolsa Escola Program, Bolsa Alimentação, Auxilio

Gas, Fome Zero) into a single program.[3] 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 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.

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

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 has to attend at least 85 percent of school hours each month (absence due to health reasons is justified and does not count towards the number of hours). 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.

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, those conditions are rarely enforced and we do not observe individuals' responses to enforcement of those conditions, while we have monthly data on all children's school attendance. Therefore we focus on analyzing responses to enforcement of school attendance conditions, which comprise the vast majority of the conditionalities enforced.

Penalties. The consequences of non-compliance vary depending on the historical record of compliance of each family. In the first case of non-compliance the family receives a warning without any financial repercussion. With the second instance of non-compliance, 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 reapply 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.

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). Families can withdraw their transfer money starting at a pre-specified date each month with a Bolsa Familia "electronic benefit card". In case of non-compliance, at the time of withdrawal the family receives a message that reports the families' warning stage, the month of failure 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 non-compliance.

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."[4] 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 hourly attendance is recorded by school teachers. The school sends the attendance lists of students to the municipality, reporting the exact fraction of school hours 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

[4]The original text in Portuguese reads as follows: "A Familía descumpriu pela primeira vez as condicionalidades do Bolsa Familía. Neste momento o pagamento não será interrompido. Caso continue em falta com os compromissos, o benefício poderá ser bloqueado suspenso ou até cancelado." 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 (i.e. the first instance of noncompliance may lead to a warning, the second instance to the transfers being blocked, the third and fourth instance to suspension of the payment and the fifth instance to the cancellation of benefits).

15 attended at least 85 percent of school hours. 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 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 can lead to a significant delay in terms of the month in which the warning is received compared to the month of failure. 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. Throughout the paper, we denote the difference between the month of warning and the month of failure as "delay". There is substantial variation in the extent of delay, which we make use of in our analysis: 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.

Importantly, the variation in delay is orthogonal to household and child characteristics and is entirely driven by time and area effects. In Appendix Table A.1 we regress delay on household and child level controls, month and year dummies and municipality dummies. We find that time and area 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. The variation in delay with which the warning is received will allow us to address concerns related to mean reversion of the original shocks and to show that household response occurs precisely after the receipt of the warning, independently of the month when the child failed to comply with the attendance requirement.

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 information on all BFP beneficiary households. In particular, 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. We have information on the universe of households enrolled in the program during 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.[5]

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 85 percent attendance condition in a given month, and in case of non-compliance we know the exact fraction of hours attended below the threshold of 0.85. 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.

[5]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.

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 of the household member legally responsible for the child or of the child 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 administrative data described above using the social security number of the child when available in the School Census and otherwise based on area code, school code, grade, full name and date of birth of the child.[6] 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 because we will rely not only on the warnings received by a given child, but also on those received by her peers.

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 attendance. 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.2. 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 hours in a given month. For ease of exposition, we refer to this variable as "failure" and we study the probability to "fail", but it should be clear that this term is not used in the traditional meaning of grade repetition, but rather it refers to failure to comply with the conditionality embedded in BFP.

[6] We are able to match 68% of individuals, 41% based on their social security number.

Panel A of Table 1 shows that on average the likelihood of failure 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 non-compliance 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).[7]

4 Own Experience of Enforcement

In this section we study how households react to the receipt of a warning; in particular, we test whether they react by increasing school attendance of their children.[8] From a theoretical point of view, it is not obvious that one should see an increase in attendance following a warning. The first reason is that households may not have room to adjust, e.g., if they have been hit by a particularly severe and persistent shock.

The second reason has to do with the priors held by the household and its updating process. Suppose the household assigned probability one to the fact that the government detects and punishes deviations from BFP conditionality. In this case, the household should already anticipate the consequences of noncompliance at the time when the child fails to attend school, and should respond accordingly. The

[7]Since warnings are received at the household level, while the fraction of peers not complying are calculated at the individual child level, the fraction of peers who fail is lower than the fraction of peers warned, since not all children in a household fail to comply at the same time.

[8]The analysis is conducted at the household level because warnings and penalties are applied to the entire family and not to individual children. Nonetheless, in Appendix Table A.3 we show that results are very similar if we use the child as the unit of observation.

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.

4.1 Graphical analysis

We start by visually exploring the pattern of "failure" rates (i.e., rates of non-compliance) 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 $1, 2, ...5$ months after the warning was received. As can be seen from the figure, there is a clear discrete downward jump in failure immediately after the receipt of a warning: households respond by increasing attendance of their children, hence the likelihood of failure goes down. The probability of failure 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 failure 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 failure and the receipt of the warning, i.e. on "delay". In this figure, we indicate with 0 the month of failure (when the family failed to comply with the school 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 failure rates when the delay is $2, 3, ...$ up to 6 months.[9] In $t = 0$ the probability of failure is one by construction; from $t = 1$ onwards the likelihood of failing decreases to a 'normal' level, but once the warning is received, the likelihood of failure jumps down and stays at this lower level given the new higher warning stage.

In all cases except for the delay of 6 months, we see a discrete jump in the failure rate 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 failure jumps down; if the delay is 3 months, the likelihood to fail 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 failure (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. 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 failure) suggests that the warning was not fully anticipated.

4.2 Multivariate analysis

We next estimate the effects of warnings on non-compliance using multivariate regressions. As a first step, we want to regress "failure " on the warning stage that a household is in, since the warning stage determines the cost of noncompliance. One challenge in identifying this effect is that some households have a much higher propensity to fail 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

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

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 fail. This implies that we exploit variation within family over time in the warning stage reached, i.e. in the receipt of new warnings.

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

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

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 hours in a month ("failure"); 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. In column 1 we include household level controls and time dummies, but not household fixed effects: as discussed above, this produces a spurious positive correlation between warnings and failure rate. In column 2 we add household fixed effects and now the probability of failure decreases after the household receives a warning and advances to a higher warning stage. This confirms the findings of our graphical analysis. The negative effect is of increasing magnitude the higher the warning stage, but one should be careful when comparing the effects of different warning stages for the following reason. When estimating the coefficients on different warning stages α_k ($k = 1, ..., 5$) in the full sample, 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.[10]

Appendix Table A.3 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.

Causality

To show that we really identify a causal effect of warnings on attendance behavior, we use two strategies. First, we exploit actual 'random' variation in the timing of warning arrival within a month. Second, we make use of variation in the timing of warning arrival (more precisely, the delay thereof) across months and conduct a difference-in-differences analysis.

To obtain random variation in the timing of the warning, 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. In fact, BFP benefits are disbursed in the second half of each month but are not transferred to all families on the same day, to avoid bottlenecks at the bank branches. The last digit of the social security number of the person who holds the BFP benefit card determines on what day within a month the family can withdraw the transfer. Because warning messages -if any- are received contextually to the money withdrawal, the day of

[10]To explore the effect of larger penalties, in Appendix Table A.4 we estimate the same regression 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 failure rate mechanically decreases thereafter). Appendix Table A.4 shows that higher warning stages have stronger effects in terms of reducing non-compliance for all subgroups.

the month in which warnings may be received depends on the last digit of a social security number, which is essentially random, as we show below.

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 (larger decrease in failure rate) 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 children'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):

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

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, if "early" and "late" households on average fail 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 4]

Table 4 reports our estimates.[11] 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

[11]The smaller number of observations in Table 4 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.

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.

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:

$$Y_{ht} = \alpha \cdot Treated_{ht} + \beta \cdot Post_{ht} + \gamma(Treated_{ht} \cdot Post_{ht}) + \delta X_{ht} + D_t + D_h + \epsilon_{ht}. \quad (3)$$

Our coefficient of interest is γ, which we expect to be negative if households react to warnings by increasing attendance (decreasing failure) of their children. The results are reported in Table 3 using different time horizons for pre- and post-warning.

[Insert Table 3]

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 failure 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 show that this is actually identifying a behavioral response to the warning, 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.

Interpretation

A possible interpretation of our results 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 (see, e.g., Bursztyn and Coffman, 2012). To assess this conjecture we exploit variation in school policies regarding communications with families. Specifically, we merge our administrative data with a dataset on schools called PROVA Brasil. Since PROVA was conducted only in grades five to nine of public schools, we lose about 40 percent of our observations.[12] 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. These variables describe whether schools inform the parents in writing, in individual meetings, by visiting parents at home or not at all.[13]

[12]Note that the schools we can merge are not very different from the schools in our BFP data that we cannot merge. For example, overall in our BFP data the fraction of Bolsa Familia children in schools is 47 percent, while the fraction of BFP children in the schools we can merge is 43 percent. Moreover, we show that results of the benchmark specification are extremely similar for this subsample (compare the last column of Table 5 with the second column of Table 2).

[13]According to the information that principals provided in the PROVA questionnaire, 79 percent of the schools inform parents in writing, 94 percent inform parents in meetings, and 70 percent do

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. If, on the other hand, the effect of warnings is due to families learning about enforcement, there should not be a significant difference in the effect of warnings between the two types of schools. We therefore estimate the following specification:

$$Y_{hst} = \sum_{k=1}^{5} \alpha_k WS_{ht}^{k} + \sum_{k=1}^{5} \sum_{j=1}^{3} \beta_k^j (WS_{ht}^{k} \cdot Inform_{st}^{j}) + \delta X_{ht} + D_t + D_h + \epsilon_{hst} \tag{4}$$

where $Inform_{st}^{j}$ is a set of three dummies for whether the school that a child is attending in month t has a policy of informing the parents (i) in writing, (ii) through individual meetings, or (iii) through home visits. If the effect we uncovered was driven by revelation of information about non-attendance instead of enforcement, we should find $\widehat{\beta_k^j} > 0$; otherwise we would expect $\widehat{\beta_k^j} = 0$.

[Insert Table 5]

The results are reported in Table 5.[14] The last column in the table replicates our benchmark specification (column 2 of Table 2) using the reduced sample for which we have information from PROVA. The magnitude and significance of the estimated coefficients of the five Warning Stage dummies are virtually unchanged. The remaining columns in the table report the results obtained from estimating Equation (4), i.e. they refer to a single regression, although for readability purposes the coefficients on the interaction terms are displayed across three columns. The results show that there is basically no difference in the effect of warnings between schools which inform parents about non-attendance and schools that do not. Out of 25 coefficients, only three are significant at the 5 or 10 percent level, and with opposite signs. There are only two instances in which “already informed” parents

home visits.

[14]In Table 5 we omit the coefficients on the main effects of families’ warning stages due to space constraints. Appendix Table A.6 presents the full set of coefficients.

respond less strongly, that is, the response to warning stages 4 and 5 in the case of schools informing parents in meetings. Nevertheless, even in these cases the overall effect of warnings is negative and the magnitude of the counteracting effect of meetings is quite small.[15] These results suggest that warnings contain information above and beyond children's school attendance, and that the "enforcement" aspect of the information appears to be quantitatively more relevant.

Another possible interpretation of our results is that our estimates may pick up learning about the *existence* of the rules. For example if parents did not read the guidelines that explain the program conditions and potential penalties and did not pay attention to information on the radio and on TV, then the receipt of a warning may be the first instance in which they get to realize that there are sanctions associated with non-attendance. We believe this interpretation is unlikely to explain our findings for a number of reasons.

First, in all our estimates we find that families still respond to warnings and punishments, even after having received (multiple) warnings before. The magnitude of the reaction is actually *increasing* with higher warning stages. Notice that with each additional warning message, the program conditions and the penalties resulting from every instance of noncompliance are explained again in a brief and salient way (see our discussion in section 2). Thus, given that BFP transfers make up an important fraction of these households' income, it is unlikely that recipients would repeatedly ignore the information in warning messages including the description of the rules.

Second, if the effects we find were the result of families being previously unaware of the existence of conditionality, they should be more pronounced for less educated households, where the adult recipients may have more difficulty reading the rules in the booklet or the text of the warning messages. To test this conjecture, we interact the warning stage dummies with an indicator for whether the education of the household head is above the median. If the above conjecture were correct, the coefficients on the interaction terms should be positive. The results, reported in

[15]The counteracting effect of meetings is less than 15 (25) percent of the magnitude of the main negative effect of the warning for warning stage 4 (5) (see Table A.6), and relatively few households receive a fourth or fifth warning.

Appendix Table A.7, show that the coefficients on the interaction terms are actually negative, i.e., if anything more educated parents respond more strongly to the warnings, although the magnitude of the difference is very small compared to the main effect.

To summarize our results so far, we showed 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 causal, i.e. that the increase in school attendance is a response to the information contained in the warning. We also argued that this information is not purely a revelation of children's past non-attendance or of the existence of rules, but rather a signal that the government is enforcing program rules. In the remainder of the paper, we provide evidence that people not only respond to experiencing enforcement themselves, but that they also learn about the strictness of enforcement from warnings of their peers, i.e. that enforcement can have important spillover effects.

5 Peers' Experience of Enforcement

In this section we test if BFP recipients learn from the experience of others who receive warnings and possibly lose their transfers. 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 will 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, etc. For this reason the analysis will be conducted at the child and not at the household level.

5.1 Identification

Identifying the impact of the warnings received by other children 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:

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

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, birth order, 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 these correlations as learning we need to address several important 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 (5). 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 to fail and thus more likely to receive a warning.[16]

If the shock was persistent, it would induce individual i and her peers to still fail more several months later, when the warning is received. However, this type of mechanism would generate a *positive*, not negative, correlation between a child's failure 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 failure and her peers' warnings and could thus be confounded with the interpretation of learning about enforcement from one's peers. To address this concern, we include the lead of peers' warnings as a falsifi-

[16]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.

cation test, i.e., we add $PEERWARN_{ih,t+1}$ among the controls in equation (5). Our reasoning exploits the time lag from the moment in which failure occurs and the moment in which the warning arrives (as discussed, this delay has a median value of 3 months and a mean value of 3.8 months). If a child's attendance increases because of mean-reversion after the initial failure, 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 would mean that the initial shock, which led to a failure, starts reverting exactly, say, four months afterwards (when peers 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* (which mostly depends on the 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.[17]

Conventional peer effects. A separate concern when interpreting the effect of warnings received by own peers relates to direct or 'conventional' peer effects. As we showed in the first part of the paper, once an individual receives a warning, she reacts by failing less. 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

[17]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.

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 analyze if *warnings* of i's peers have an independent effect on i's likelihood to fail.[18] Second, we analyze whether child i's likelihood to fail 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:

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

$$Y_{iht} = \sum_{k=1}^{5} \alpha_k WS_{ht}^{k} + \sum_{n=-1}^{0} \beta_n PEERWARN_{ih,t+n} + \sum_{n=-1}^{0} \delta_n MaxPEERWARN_{ih,t+n} + \gamma X_{iht} + D_t + D_i + \epsilon_{iht} \tag{7}$$

where $PEERFAIL_{ih,t+n}$ is the fraction of i's classmates who are BFP recipients and fail to meet the 85 percent attendance threshold 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' failure we augment the specification with a one-month lag, because it is ex ante difficult to establish whether children learn about their peers' warnings in time to adjust their attendance in the current month or in the following month and because the effects may be persistent. We estimate equations (6) and (7) using a linear probability model and clustering the standard errors at the household level.

[18] This may lead to an underestimate of the "true" learning effect, if part of the learning about enforcement happens through observing one's peers' higher attendance.

5.2 Main results

Table 7 reports our main results on peers' warnings. The coefficient on the fraction of peers 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 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 failure would decrease by 13 percent in the same month and 27.5 percent in the following month.[19]

[Insert Table 7]

In column 2 we include the fraction of peers 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 failure to start decreasing before the arrival of a warning, hence we should see a negative coefficient on the lead of peers' warnings. Instead, the coefficient on the lead variable is a precisely estimated zero, lending support to our learning interpretation.

In column 3 we control for the fraction of peers who fail to meet the attendance threshold and find that both the contemporaneous and the lagged variable are positively correlated with own failure. At the same time, we still find a negative effect of peers' 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 7 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, i.e. 6 to 15 years old. It turns out that a large fraction of our original sample satisfies this condition,

[19] Appendix table A.8 presents the same results as table 7, but displays in addition the coefficients on the family's own warning stage, which we control for in all specifications.

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 failure rate, 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 the contemporaneous variable is also negative but insignificant. The findings in columns 5 and 6 lend further support to the hypothesis that people learn about the strictness of enforcement versus the possibility that common shocks may drive our results or that results are driven by the child responding solely to classmates' increased attendance: in fact the peers who receive warnings in column 5 and 6 are attending different grades or schools and are exposed to different shocks.

Interpretation and informational content

To gain further insights into the role played by the new information contained in warnings, we analyze how the effect of peers' 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. As for peers' warnings of the same stage where the family is, they 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 6]

In Table 6 we distinguish between peers who receive warning $1, 2, ...5$ and estimate four different regression equations, conditional on whether a family is currently in warning stage $1, 2, 3$ or 4. 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 6 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.

Finally, one remaining interpretation issue is the extent to which the behavioral responses we identify correspond to actual 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.[20] A first reaction to this is that even in this case our hypothesis that families learn about the strictness of enforcement and update their beliefs would be valid: it would be the action taken by the family after updating that would differ (e.g., persuading teachers instead of sending children to school). Secondly, our results on warnings received by peers –especially by siblings' peers– help address this point to some extent. As shown in Table 7, we find that individual *i*'s failure goes down when the classmates of *i*'s siblings get warned. Since the teacher of child *i* would not know about those children's warnings (in particular if the siblings attend different schools), this result is more likely explained in terms of actual child's attendance as

[20]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.

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 teachers during the day and each of them has to register attendance: bribing or convincing all of them may not be easy in the presence of generally high stakes.[21] 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 learn about and 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.

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 peers.

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

[21]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.

strictly enforce the rules.[22] 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.

[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).

[22] 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.

[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 ”, Forthcoming in *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 Enforcing Conditional Welfare Programs: Evidence from Brazil”, mimeo, Bocconi 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] 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.

[17] 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*.

[18] 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*.

[19] 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.

[20] 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*.

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

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

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

[24] 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,* 96(5).

[25] 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 Failure and Warning

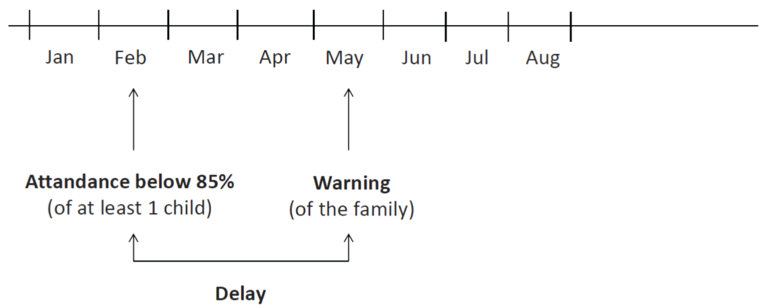


Figure 2: Response to warning

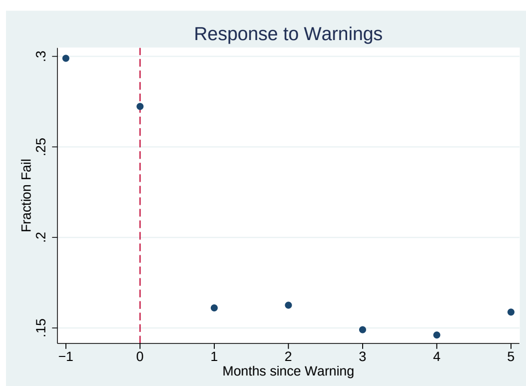


Figure 3: Response to warning, by delay

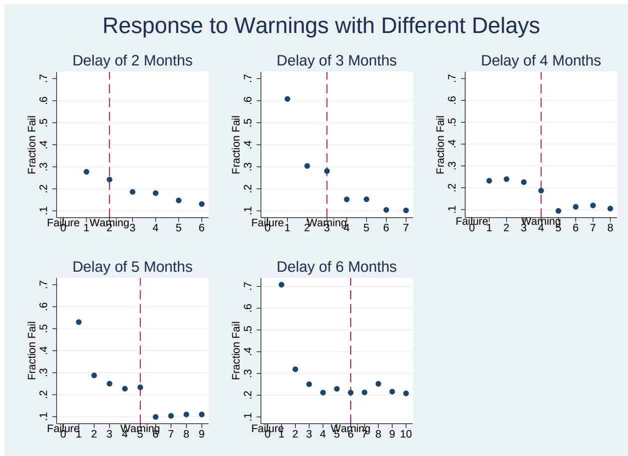


Table 1: Summary statistics

Variable	Observations	Mean	Std. Dev.
Own Warnings			
Failure to Comply	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
Peer Warnings			
Frac of Peers Warned	5914381	0.014	0.059
Frac of Peers Failed	5914381	0.010	0.046
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

Table 2: Effect of own warnings on attendance

Dependent Variable:	Noncompliance in a Given Month (Failure)	
	(1)	(2)
Warning Stage 1	0.0385***	-0.0624***
	(0.0000)	(0.0000)
Warning Stage 2	0.0632***	-0.1303***
	(0.001)	(0.001)
Warning Stage 3	0.0819***	-0.2050***
	(0.001)	(0.002)
Warning Stage 4	0.1099***	-0.2656***
	(0.002)	(0.003)
Warning Stage 5	0.1602***	-0.3477***
	(0.006)	(0.008)
Controls	Yes	Yes
Household FE	No	Yes
Time FE	Yes	Yes
No. Obs.	8,090,769	8,090,769
R-squared	0.02	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.

Table 3: Effect of own warnings, difference-in-differences

	Main result		Placebo	
	Definition 1	Definition 2	Definition 1	Definition 2
	Warning: t=0	Warning: t=0	Warning: t=0	Warning: t=0
	Before: t=-1,0	Before: t=-2,-1,0	Before: t=-1	Before: t=-2,-1
	Post: t=1	Post: t=1,2	Post: t=0	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 4: Effect of own warnings, random variation in timing

Dependent Variable:	**Noncompliance in a Given Month (Failure)**		
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 5: Learning about attendance

Dependent Variable:	**Noncompliance in a Given Month**		
Parents Informed in:	Writing	Meeting	Home Visit
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)
Controls (incl. Warning Stage)		Yes	
Household and Time FE		Yes	
No. Obs.		5,044,299	
R-Squared		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.

Table 6: Information content of peers' warnings

Dependent Variable:	**Noncompliance in Given Month (Failure)**			
Own Warning Stage:	1	2	3	4
Frac Peers Warned (WS 1)	-0.0123***	-0.0139	-0.0118	0.0318
	(0.003)	(0.009)	(0.016)	(0.028)
Frac Peers Warned (WS 2)	-0.0245***	-0.0203**	-0.0171	-0.0693
	(0.007)	(0.009)	(0.033)	(0.044)
Frac Peers Warned (WS 3)	-0.0188	-0.0308*	-0.0501**	-0.0392
	(0.012)	(0.017)	(0.021)	(0.051)
Frac Peers Warned (WS 4)	-0.0197	-0.0386	-0.0787*	-0.0590**
	(0.022)	(0.026)	(0.043)	(0.027)
Frac Peers 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 and 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 7: Effect of peers' warnings

Dependent Variable:	Noncompliance in a Given Month (Failure)					
Sample	*All Children*			*Children with Siblings*		
	Benchmark	Placebo	Control for Peers' Failure	Benchmark	Siblings' Peers	Siblings' Peers Other Schools
	(1)	(2)	(3)	(4)	(5)	(6)
Fraction of Peers 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 Peers 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 Peers Warned		0.0001				
		(0.002)				
Fraction of Peers Failed			0.2807***			
			(0.021)			
Lag of Frac of Peers Failed			0.0952***			
			(0.009)			
Max Frac of Siblings' Peers Warned					-0.0047**	
					(0.002)	
Lag Max Frac of Siblings' Peers Warned					-0.0076***	
					(0.002)	
Max Frac of Siblings' Peers Warned						-0.0013
(Only Sibs at Other Schools)						(0.002)
Lag Max Frac of Siblings' Peers 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.

ONLINE APPENDIX: NOT FOR PUBLICATION

Figures

Figure A.1: Frequency of delays

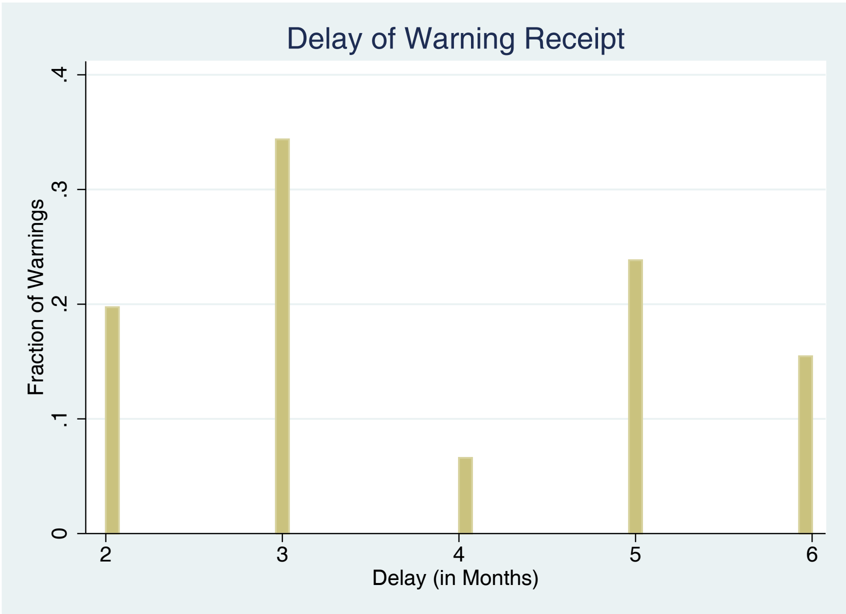
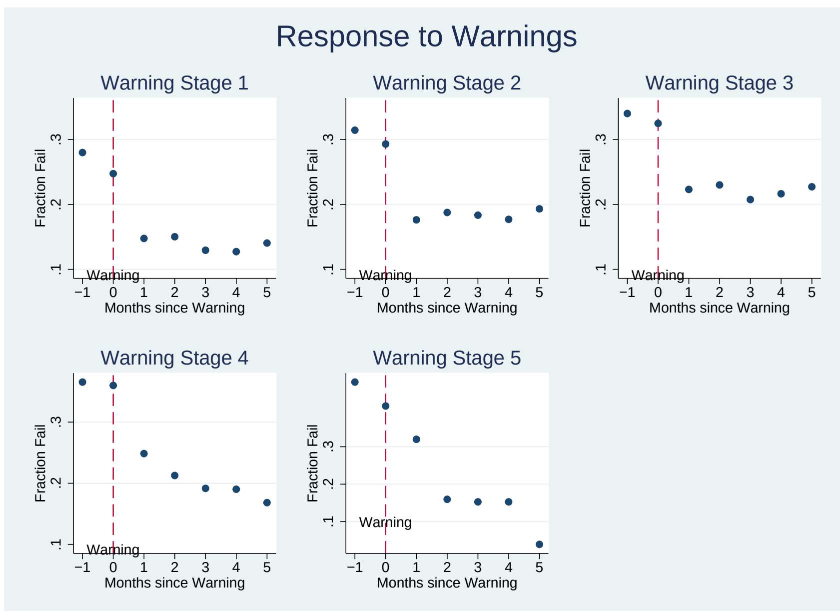


Figure A.2: Response to warning, by warning stage



Tables

Table A.1: 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.2: Summary statistics

Variable	Observations	Mean	Std. Dev.
Own Warnings			
Failure to Comply	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 Failed	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 Propoerty	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.3: Effect of own warnings, individual level

Dependent Variable:	**Noncompliance in a Given Month (Failure)**	
	(1)	(2)
Warning Stage 1	0.0216***	-0.0336***
	(0.0000)	(0.0000)
Warning Stage 2	0.0337***	-0.0657***
	(0.0000)	(0.001)
Warning Stage 3	0.0428***	-0.0970***
	(0.001)	(0.001)
Warning Stage 4	0.0569***	-0.1147***
	(0.001)	(0.002)
Warning Stage 5	0.0853***	-0.1366***
	(0.003)	(0.005)
Controls	Yes	Yes
Individual FE	No	Yes
Time FE	Yes	Yes
No. Obs.	16,056,547	16,056,547
R-squared	0.01	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. The second specification includes individual fixed effects.

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

Dependent Variable	Noncompliance in a Given Month (Failure)				
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: Learning about attendance

Dependent Variable:	Noncompliance in a Given Month (Failure)			
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.7: Effect of own warnings, by parental education

Dependent Variable:	**Noncompliance**
Warning Stage 1 * Educ Above Med	-0.0066***
	(0.001)
Warning Stage 2 * Educ Above Med	-0.0147***
	(0.002)
Warning Stage 3 * Educ Above Med	-0.0230***
	(0.004)
Warning Stage 4 * Educ Above Med	-0.0316***
	(0.007)
Warning Stage 5 * Educ Above Med	-0.0133
	(0.017)
Warning Stage 1	-0.0590***
	(0.001)
Warning Stage 2	-0.1238***
	(0.001)
Warning Stage 3	-0.1950***
	(0.002)
Warning Stage 4	-0.2516***
	(0.004)
Warning Stage 5	-0.3412***
	(0.009)
Controls	Yes
Household FE	Yes
Time FE	Yes
No. Obs.	7,511,834
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. All specifications include household fixed effects.

Table A.8: Effect of peer warnings

Dependent Variable:	Noncompliance in a Given Month (Failure)						
Sample		*All Children*				*Children with Siblings*	
	Own Warnings	Benchmark	Placebo	Control for Peers' Failure	Benchmark	Siblings' Peers	Siblings' Peers Other Schools
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Fraction of Peers 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 Peers 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 Peers Warned			0.0001				
			(0.002)				
Fraction of Peers Failed				0.2807***			
				(0.021)			
Lag of Frac of Peers Failed				0.0952***			
				(0.009)			
Max Frac of Siblings' Peers Warned						-0.0047**	
						(0.002)	
Lag Max Frac of Siblings' Peers Warned						-0.0076***	
						(0.002)	
Max Frac of Siblings' Peers Warned							-0.0013
(Only Sibs at Other Schools)							(0.002)
Lag Max Frac of Siblings' Peers 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.