Appendix

In Section IV, we explain that our ability to observe variation in the incentives faced by teachers within a given month requires the assumption that lying is costly and that teacher also face costs in retroactively changing students’ attendance. In what follows, we build a structural model that describes the decision process a teacher must make when choosing to misrepresent a students’ attendance on a given day. Relative to the earlier assumptions, we assume here that not only is retroactively changing attendance costly, but that teachers simply cannot do it. As in Section IV, however, this is just a simplifying assumption that allows us to ignore the option to retroactively change attendance later in the month. Relaxing this assumption works against the empirical tests resulting from the model’s predictions.

A. A Dynamic Model of Teacher Behavior

The value of misrepresentation from the grain program derives solely through securing the grain for a child by making the child’s records appear as if the child has 80 percent attendance in a given month. This could be a warm glow from having given the child the grain or a bribe from the parents, or it could also be some peace of mind since parents of children who do not receive the grain often complain. We assume a constant cost to the teacher of misrepresenting attendance. The cost could be the psychic cost of lying, a probability of getting caught by principals or administrators, or a reputational cost: if students stopped attending school because they knew their teacher would ensure that they received the grain anyway, this may reflect negatively on the teacher.

Consider the problem faced by a teacher trying to decide whether or not to misrecord an absent child as present on a given day. The value of inflating a child’s attendance depends nonlinearly on the number of days left in the month, the student’s roster attendance this month, the teacher’s expectation of the child’s future attendance this month, and interactions of variables derived from these three. Early in the month, children whose attendance patterns will put them below 80 percent may make up their deficit, rendering an early misreport possibly wasted. Closer to the end of the month, a teachers’ misrepresentation is more likely to be decisive. How close a teacher can wait until the end of the month depends on the child’s attendance rate. For children who have not attended at all in a given month, teachers must start mis-recording attendance within the first twenty percent of days in the month to keep the child eligible for the
grain and misrecord many of the future days. For children on the cusp of meeting the 80 percent requirement, teachers can wait until the very last day to inflate attendance. Finally, for children in between these extremes, exactly how long a teacher can wait depends on how likely the child is to attend school in the remainder of the month. Students with high attendance rates are unlikely to require assistance allowing teachers to wait to use their discretion while poorly attending students would require early intervention to keep them eligible.

We posit a dynamic model of teacher behavior, where the teacher has to predict not only the child’s attendance on future days in the month but also her own behavior. Imagine it is the second to last day of the month. A child is absent, and the teacher must decide whether to misrepresent his attendance. From past experience, she predicts the probability that he attends on a given day is $a_t$; from her records, she notes that he has been marked as present $p_t$ times prior to today. The subscript $t$ denotes the day of the month, from 1 to $T$. Since there are only two observations left this month, she should not exaggerate his attendance if $p_t$ is less than or equal to $0.8T - 3$, because he has no chance of receiving the grain, or if $p_t$ is greater than or equal to $0.8T$, because he is already guaranteed the grain. If $p_t$ equals $0.8T - 1$, she can guarantee him the grain by misrepresenting the child’s attendance today or by waiting until the next day and misreporting his attendance if he is absent then. Since there is a cost to inflating attendance and waiting allows for the possibility that the child will be present, she should take the second option. She does not need to misrepresent today’s attendance since she will have the option to do so tomorrow. On the other hand, if $p_t$ is equal to $0.8T - 2$, the teacher must misrepresent his attendance today and if he is absent tomorrow, misrepresent it then as well. If she does not do so today, he is disqualified. Earlier in the month, the teacher makes a similar calculation with the additional condition that the expected number of times she will have to misrepresent attendance to guarantee the child the grain may be too costly.

To model the entire month, we build a system of finite Bellman equations specifying a value function that depends on the child’s attendance, allowing the teacher to misrepresent attendance if the child is absent. Let $c$ be the cost of exaggerating a child’s attendance on a single day, and let $G$ be the benefit to the teacher if a child receives the grain that month. Finally let $\beta$ be the discount rate. In the calculations that follow, we take the discount rate to be 1 since the length of a period is only a single day.
The Bellman equations for period T, the last day of the month, are straightforward. If the child is present, the value function is simply whether the child will receive the grain after taking into account that day’s observation. If the child is absent, the value function measures whether the child receives the grain, allowing the teacher to misrepresent that day’s attendance for a cost. Using superscripts P and A for present and absent, the value functions can be written as follows:

\begin{align*}
V_T^P(p_T) &= V(p_T + 1) \\
V_T^A(p_T) &= \max_{m_T} \{ -cm_T + V(p_T + m_T) \}
\end{align*}

where \( V(p) = \begin{cases} G & \text{if } \frac{p}{T} \geq 0.8 \\ 0 & \text{else} \end{cases} \)

where \( p_T \) is the number of past days the child attended school according to the roster and \( m_T \) is a choice variable that indicates whether the teacher misrepresents the child’s attendance.

On any previous day, the teacher must consider whether the child will attend school in the future, what costs she will incur misrepresenting the child’s attendance and whether the child will finally receive the grain. If the child is present on a given day, the Bellman equation is equal to the discounted expected value of the same equation for the following day; ‘expected’ because it depends on whether the child is present the following day. If the child is absent, the teacher gets to choose how to record his attendance, which impacts the parameters at time \( t + 1 \). Thus, the value functions can be written recursively as follows

\begin{align*}
V_t^P(p_t, a_t, d_t) &= \beta E(V_{t+1}(p_{t+1}, a_{t+1}, d_{t+1})) \\
&= \beta a_t V_{t+1}^P[p_t + 1, (a_t, d_t + 1)/(d_t + 1), d_t + 1] + \beta (1 - a_t) V_{t+1}^A[p_t + 1, (a_t, d_t + 1)/(d_t + 1), d_t + 1] \\
V_t^A(p_t, a_t, d_t) &= \max_{m_t, e_t} \{ -cm_t + E(V_{t+1}(p_{t+1}, a_{t+1}, d_{t+1})) \} \\
&= \max_{m_t, e_t} \{ -cm_t + \beta a_t V_{t+1}^P[p_t + m_t, a_t, d_t/(d_t + 1), d_t + 1] + \beta (1 - a_t) V_{t+1}^A[p_t + m_t, a_t, d_t/(d_t + 1), d_t + 1] \}
\end{align*}

where \( t \) takes on any integer value in \([1, T-1]\), \( a_t \) is the teacher’s updated perception of the probability the child attends school, \( d_t \) is the number of days elapsed in the year and \( p_t \) is as
defined above, but on day $t$. We keep track of $d_t$ in order to update the attendance probability, $a_{t+1}$. Note that the parameters at time $t + 1$ ($p_{t+1}, a_{t+1}$ and $d_{t+1}$) differ between the two equations.

Denote $V^A_t(m; p_t, a_t, d_t)$ to be the absent value function for a given $m$. This provides a convenient representation of the value of misrepresenting attendance when a child is absent. We define the incentive to inflating attendance on day $t$, $I_t$, as the difference between the value function when the teacher exaggerates the child’s attendance and when the teacher does not.

\[
I_t(p_t, a_t, d_t) = V^A_t(m = 1; p_t, a_t, d_t) - V^A_t(m = 0; p_t, a_t, d_t) = -c + E(V_{t+1}(p_t + 1, a_t, d_t, (d_t + 1), d_t + 1)) - E(V_{t+1}(p_t, a_t, d_t, (d_t + 1), d_t + 1))
\]

Solving the model in closed form is straightforward but tedious due to the large number of cases; instead we solve the model numerically.\footnote{We leave a structural estimation of this model to future work, focusing here on the implications of the model.} After first discussing the decision rule and the variation in the incentives to misrepresent attendance that arise from this model, we will describe the empirical specifications used to test for these patterns in the data.

Note that student behavior is an integral part of the teacher's problem; the teacher must predict whether the child will attend school in the future. We do not explicitly model the reputational cost, which could result in a response of the child’s attendance patterns to the teacher’s behavior itself. To justify this modeling decision, however, note that we do not need to assume that students do not respond to whether their teachers misrepresent attendance. The assumption that teachers do not anticipate this is sufficient, but also not necessary. Since all specifications will include teacher fixed effects or student fixed effects to account for unobservable variation in the cost of misrepresenting attendance, we simply have to assume that this reputation cost does not vary across different days we see the same student.

**B. The Decision Rule and Additional Variation in the Incentives**

Following the intuition above, teachers should only misrepresent attendance on days when the student needs that day’s attendance and perfect future attendance in order to earn the grain: $I_{T-n}$, the incentive on day $T - n$, is positive only when the child needs exactly $n+1$ days of future attendance, including that day, to make the threshold. If the child needs fewer than $n+1$
days, the teacher should wait, allowing for the possibility that the child attends enough that the teacher will not have to adjust his attendance. If the child needs more days, the child has already been disqualified. However, requiring n+1 days is not a sufficient condition. Early in the month and for a sufficiently large cost, $c$, the teacher may anticipate having to inflate attendance for a particular child too many times that she deems it too costly to award him the grain.

There is a fair amount of variation in $I_{T-n}$ besides this straightforward decision rule. For ease of exposition, we divide the large number of possible situations into four cases. The first is the only case in which $I_{T-n}$ can be positive: when the student needs exactly n+1 days of future attendance to make the 80% cut-off. In this case, variation in $I_{T-n}$ comes from the expected number of exaggerations necessary to guarantee the student the grain. The higher the number of future exaggerations necessary, the smaller the net benefit to the teacher of having awarded the child the grain. Thus, in this case, $I_{T-n}$ is decreasing in the number of days left in the month and increasing in the teacher’s perception of the child’s future attendance rate, $a_t$. Figures A1, A2 and A3 plot $I_{T-n}$ with respect to $a_t$, for n equal to 2, 3 and 5, respectively. The solid lines, representing this first case, show that $I_{T-n}$ is increasing in $a_t$. Note that in Figures A2 and A3, $I_{T-n}$ is negative for very low values of $a_t$ since it is too costly to give the child the grain.

In the second and third cases, $I_{T-n}$ is negative but does not vary with the number of days left or the child’s attendance rate. The second occurs when the student needs more than n+1 days of future attendance to make the 80% threshold. The third case occurs when the student has already passed the 80% threshold for the entire month (note: this can only happen close to the end of the month). The dotted line in Figure A1 shows that $I_{T-n}$ is constant in both these cases.

In the fourth case, when the student needs fewer than n+1 days of future attendance to make the 80% threshold, $I_{T-n}$ is negative because teachers should wait until the child hits the first (needs her help for certain) or third (already earned the grain) cases. Variation arises in $I_{T-n}$ in this case because of two opposing forces that derive from the probability the child will receive the grain and the likelihood the teacher will misrepresent the child’s attendance in the future. For a child who will earn the grain (either because he has a high attendance rate or because it is sufficiently cheap for the teacher to adjust her records), $I_{T-n}$ is decreasing in $a_t$. If the child is very likely to attend in the future, misrepresenting attendance today is very costly because the
teacher may never have needed to misrepresent his attendance. If the child is less likely to attend in the future, the teacher is simply mistiming the exaggerations. Denote $r$ to be the number of days of future attendance the child needs to earn the grain. The dashed line in Figure A1 demonstrates this force: a child for whom $c = 0.4$ and $r = 1$ is certain to earn the grain since the expected cost of misrepresenting attendance ($c$ multiplied by the probability the child is absent the next day) is less than the value of the grain. Thus, $I_{T-n}$ is decreasing in $a_i$. The two dashed lines in Figure A2 confirm this prediction: $I_{T-n}$ is decreasing in $a_i$ conditional on $r$. Note that the short dashed line ($r = 1$), is always below the long dashed line ($r = 2$). This is because the child is more likely to earn the grain himself, rendering the teacher’s action more often wasted when $r = 1$, than when $r = 2$. In other words, it is more likely the child will attend at least one of the future days than two of them.

However, an increase in $a_i$ also increases the probability the child will earn the grain (both on his own merits and because the teacher does not have to incur too many costs), which in turn increases $I_{T-n}$ because this probability rises faster when the child only needs $r - 1$ days of future attendance (the teacher having already adjusted one day’s records) than when he needs $r$ days of future attendance. For a child with a very low value of $a_i$, a teacher is no longer simply mis-timing the exaggerations because the child may not earn the grain at all, in which case the teacher’s action was wasted. Overall, an increase in $a_i$ increases the likelihood that the child could have earned the grain on his own; at low values of $a_i$ it becomes less wasteful to misrepresent a child’s attendance because it is less likely to be too costly to guarantee the child the grain and at higher values of $a_i$, the action becomes more wasteful since he would have earned the grain on his own. The two dash-dot combination lines in Figure A3 demonstrate this non-monotonic relationship between $a_i$ and $I_{T-n}$ for higher values of $r$. Note that, while $I_{T-n}$ is always negative, it is often greater than $-c$ since the child may eventually receive the grain.

C. Empirical Tests

To test whether teachers respond to variation in $I_{T-n}$, we estimate:

$$
\Pr(\text{misrepresent}_{ijt} = 1) = f(\alpha + \pi \cdot I_{T-n} + \nu_j + \epsilon_{ijt})
$$

(A1)
where \( I_{T-n} \) is taken directly from the numerical solution to the model described above. Recall that this calculation depends on the size of the parameter \( c \); we will test the sensitivity of our results to the calibration of \( c \). A positive \( \pi \) confirms that teachers misrepresent attendance in response to the grain, even when we control for teacher or student fixed effects.

In Table A1, we estimate specification (A1), demonstrating that teachers are more likely to misrepresent attendance for a particular child on a particular day when \( I_{T-n} \) is greater. The calculation of \( I_{T-n} \) comes from the numerical solution to the dynamic programming problem described above for each instance a child is absent. To perform this calculation, we must first calibrate the cost of exaggerating a child’s attendance on a given day, \( c \), and the benefit to the teacher if a child receives the grain, \( G \). Because each period is only a day, we assume the discount rate is equal to one. Since the values of \( c \) and \( G \) are important only relative to each other, we normalize \( G \) to 1 and vary the cost. We consider conservative lower and upper bounds for \( c \). Note that teachers are not willing to misrepresent attendance to make a child who rarely attends school eligible for the grain, suggesting a non-zero value for \( c \). For an upper bound, note that teachers are seen misrepresenting attendance three times within a single month for some children. This suggests a cost of misrepresenting attendance of less than 0.33. We calibrate the cost to equal 0.3 and test how sensitive our results are to this assumption.

The top panel of Table A1 estimates this specification using teacher-day fixed effects, while the bottom panel uses student fixed effects. These fixed effects ensure that the results are not driven by a teacher who misrepresents attendance for a particular child or for all children on a particular day. Both panels estimate a linear probability model, but the results are robust to employing a conditional logit model. Teachers respond strongly to variation in this net benefit to misrepresenting attendance, and this result is not sensitive to our calibration of the daily cost. Recall that the incentive to misrepresent attendance comes from a highly non-linear relationship between three variables: the number of days left in the month, the number of days of attendance the child needs to earn the grain and the child’s future attendance. Controlling linearly for these other variables (not shown) does not detract from the main effect; teachers respond to the nonlinearities. Since this variation could only come from the grain threshold, we conclude that teachers do misrepresent attendance in response to the grain distribution.
<table>
<thead>
<tr>
<th>Dependent Variable:</th>
<th>Misrepresent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost:</td>
<td>0.2 0.25 0.3 0.35 0.4</td>
</tr>
<tr>
<td>(1) (2) (3) (4) (5)</td>
<td></td>
</tr>
</tbody>
</table>

### Panel A

<table>
<thead>
<tr>
<th>Incentive to Misrepresent</th>
<th>0.129 ***</th>
<th>0.200 ***</th>
<th>0.268 ***</th>
<th>0.331 ***</th>
<th>0.386 ***</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0.027)</td>
<td>(0.028)</td>
<td>(0.031)</td>
<td>(0.033)</td>
<td>(0.035)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Additional Fixed Effects</th>
<th>Teacher XDay</th>
<th>Teacher XDay</th>
<th>Teacher XDay</th>
<th>Teacher XDay</th>
<th>Teacher XDay</th>
</tr>
</thead>
<tbody>
<tr>
<td>p-value of joint test of f.e.</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Num. of Obs.</td>
<td>24192</td>
<td>24192</td>
<td>24192</td>
<td>24192</td>
<td>24192</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.590</td>
<td>0.592</td>
<td>0.593</td>
<td>0.595</td>
<td>0.596</td>
</tr>
</tbody>
</table>

### Panel B

<table>
<thead>
<tr>
<th>Incentive to Misrepresent</th>
<th>0.040</th>
<th>0.080 **</th>
<th>0.119 ***</th>
<th>0.156 ***</th>
<th>0.186 ***</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0.036)</td>
<td>(0.039)</td>
<td>(0.041)</td>
<td>(0.043)</td>
<td>(0.044)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Additional Fixed Effects</th>
<th>Student</th>
<th>Student</th>
<th>Student</th>
<th>Student</th>
<th>Student</th>
</tr>
</thead>
<tbody>
<tr>
<td>p-value of joint test of f.e.</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Num. of Obs.</td>
<td>24192</td>
<td>24192</td>
<td>24192</td>
<td>24192</td>
<td>24192</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.675</td>
<td>0.675</td>
<td>0.675</td>
<td>0.675</td>
<td>0.675</td>
</tr>
</tbody>
</table>

Note: This table displays estimates of equation (A1). All observations are at the student-day level and are limited to days on which the child is absent. The dependent variable is whether or not the teacher marks the child as present despite the child's absence, while the primary independent variable is the net benefit a teacher receives from misrepresenting attendance. This value is estimated by solving the dynamic programming model numerically and calculating the change in the value function if the teacher inflates a child's attendance, IT. Different columns estimate this value assuming different daily costs to misrepresenting attendance. Panel B includes month fixed effects. All columns estimate a linear probability model; similar results are obtained with conditional logit. Robust standard errors clustered by teacher are in parenthesis. *** 1%, ** 5%, * 10%.
Figure A1: Sources of Variation in the Incentive to Inflate Attendance with 2 Days Left

Note: This graph plots the net benefit a teacher receives from misrepresenting attendance. This value is estimated by solving the dynamic programming model numerically and calculating the change in the value function if the teacher inflates a child’s attendance. This value is graphed against the teacher's perception of the child's future attendance rate. We assume a daily exaggeration cost of 0.4 and that 30 days have passed in the year in order to update the child's attendance rate.
Note: This graph plots the net benefit a teacher receives from misrepresenting attendance. This value is estimated by solving the dynamic programming model numerically and calculating the change in the value function if the teacher inflates a child's attendance. This value is graphed against the teacher's perception of the child's future attendance rate. We assume a daily exaggeration cost of 0.4 and that 30 days have passed in the year in order to update the child's attendance rate.
Figure A3: Sources of Variation in the Incentive to Inflate Attendance with 5 Days Left

Note: This graph plots the net benefit a teacher receives from misrepresenting attendance. This value is estimated by solving the dynamic programming model numerically and calculating the change in the value function if the teacher inflates a child's attendance. This value is graphed against the teacher's perception of the child's future attendance rate. We assume a daily exaggeration cost of 0.4 and that 30 days have passed in the year in order to update the child's attendance rate.