Setting Pay for Performance Targets: Do Poor Performers Give Up?

Bryan Dowd, PhD
Roger Feldman, PhD
William Nersesian, MD

XXXII Jornados AES Economia de la Salud Bilbao, Spain
P4P

• **Pay for Performance (P4P):** financial incentives to encourage health care providers to improve their performance
• Performance usually means quality of care
• Most analyses of physician P4P programs have estimated *average treatment effects* across all participating physicians or physician practices
• Most studies show positive but modest average effects
Our Goal

• To estimate *heterogeneous treatment effects* of a P4P program directed at physician practices
  • Do poor performers try to improve?
  • Do practices above the performance target improve, stay put, or fall back?
• Motivation, prior research, our study setting, data, methods, results
• Implications for design of P4P programs
Why Do We Care?

Optimal Care for Patients in 128 Physician Practices in Minnesota, 2007

- Diabetes Care
- STD Screening
- Child with Sore Throat

Minimum
Maximum

0% 50% 100%
Prior Research

• Peterson (Annuals of Internal Medicine, 2006) reviewed 17 evaluations of P4P
  • 7 of 9 group-level studies showed improvements in quality

• Christianson, Leatherman, and Sutherland (2008) identified problems with existing studies:
  • small samples
  • lack of control groups
  • failure to understand the details of particular P4P programs
Rosenthal’s study

- Rosenthal and co-authors (*JAMA*, 2005) compared performance of 300 physician practices that got P4P incentive with 42 control practices in the same network
  - Rewards were based on meeting targets
  - Average effects: Cervical cancer screening improved significantly more in P4P practices
  - Heterogeneous effects: Practices that began at lowest levels of performance got the least rewards
  - Interestingly, these practices improved the most

- Most similar to our study
Goals of Our Study

• Like Rosenthal, we are interested in heterogeneous treatment effects
  – In particular, do poor performers simply give up when the P4P reward is based on meeting a performance target?

• Other goals:
  – Distinguish between random changes in performance versus changes caused by the P4P incentive
  – If we find non-random changes, are they related to the distance between the practice’s baseline performance and the P4P target?


Study Setting

• A health plan in the Minneapolis area offered a financial incentive to a physician network to increase the generic prescription rate (GPR) for outpatient pharmaceuticals for that plan’s enrollees
  – GPR = % of prescriptions written for generic drugs

• The network set targets for individual practices that varied by specialty and year (2004-2007)
  – Target rates varied between ~ 33% and 67%
  – In 2007 a practice with 8 physicians got $8,960 for meeting the target
  – Practices below the target got nothing
Study Setting

• We examined P4P incentives to improve performance

• We did not measure:
  – Changes in copayment tiers for generics, brand-name (sole source) drugs on the plan’s formulary, and name-brand drugs not on the formulary
  – Patents expire and new generic opportunities arise

• We captured these factors with annual time effects common to all physician practices
Data

- 54 physician group practices specializing in pediatrics, internal medicine, ObGyn, and family medicine
  - All practices were in ‘primary’ patient care
  - Results generalize only to such practices
- Average practice wrote 2,700 prescriptions (Rx) per year for enrollees from this health plan
  - We had 12 quarterly observations on each practice from 2nd quarter of 2004 through 1st quarter of 2007
  - We calculated GPR by quarter and compared it with the target
  - We also used differences over time (11 observations per practice)
‘Equilibrium’ GPR*

• In absence of P4P every practice has ‘equilibrium’ GPR* based on:
  – Patient mix (more or fewer generic drugs available for its type of patients)
  – Doctors’ awareness of generics and willingness to recommend generics
  – Technology to manage and monitor Rx

• Let GPR* = Xβ, where X are observed practice characteristics that may vary over time

• Deviating from GPR* is costly
Model of Moving Up

• Based on simple profit maximization with a threshold: practice moves up to target if Reward > Cost

• Why is moving up costly?
  – Patient resistance to trying generic drugs
  – Costly for physicians to learn about generics
  – Overcoming inertia of past prescribing patterns
  – Overcoming influence of ‘detailing’ (efforts by drug companies to promote brand-name products)
Model of Moving Up

Cost, Reward

Reward

Don’t Move  Move Up  Don’t Move

Cost of improving GPR (constant MC used for illustration)
‘Final’ GPR

\[ GPR^F_{i,t} = X_{i,t} \beta + (T - X_{i,t} \beta) \times D_{i,t} \]

\[ D_{i,t} = 1 \text{ if } \text{cutoff} \leq X_{i,t} \beta \leq T \]

- Denote the ‘final’ GPR under P4P as GPR^F
- Recall that all practices have ‘equilibrium’ GPR^* = X\beta
- Practices with \( X\beta > T \) and those with \( X\beta < \text{unobserved cutoff} \) are not motivated to deviate from their equilibrium
- Other practices want to move up to the target
- Next, we choose a function that approximates the unobserved cutoff
‘Final’ GPR Function

\[ GPR_{i,t}^F = X_{i,t} \beta + \delta_1 \text{BELOW}_{i,t-1} + (T - GPR_{i,t-1}) \times \delta_2 \text{BELOW}_{i,t-1} + u_{i,t} \]

- Proxy \( D = 1 \) if the practice was below the target last period
- We expect practices below the target last period will move up (\( \delta_1 > 0 \))
- But practices farther below the target last period will be less likely to move up (\( \delta_2 < 0 \))
- We also include the square of distance below \( T \) (not shown here)
Adjustment Process

• We embed the GPR^F function in a standard partial adjustment process:

\[ GPR_{i,t} - GPR_{i,t-1} = (1 - \gamma) \times \]

\[ (GPR_{i,t}^F - GPR_{i,t-1}) + e_{i,t} \]

• We estimate the speed of adjustment along with the parameters of GPR^F

\((1 - \gamma) = 1\) implies complete adjustment in one period

• We estimate the model with OLS because \(u\) and \(e\) are independent of explanatory variables
Descriptive Results

• The mean of 11 changes in GPR was +1.6 percentage points, but the change varied from -15 to +21 percentage points
• Overall, GPR increased by about 14 percentage points over the study period
• On average, 40% of practices met the target each quarter
## Average and (Target) GPR

<table>
<thead>
<tr>
<th>Specialty/Year</th>
<th>Internal Medicine</th>
<th>ObGyn</th>
<th>Pediatrics</th>
<th>Family Practice</th>
</tr>
</thead>
<tbody>
<tr>
<td>2004 (last 3 quarters)</td>
<td>48.15 (44)</td>
<td>41.81 (34)</td>
<td>46.71 (41)</td>
<td>50.09 (44)</td>
</tr>
<tr>
<td>2005</td>
<td>53.08 (54)</td>
<td>48.55 (44)</td>
<td>50.12 (46)</td>
<td>54.97 (54)</td>
</tr>
<tr>
<td>2006</td>
<td>58.72 (65)</td>
<td>53.35 (63)</td>
<td>58.80 (60)</td>
<td>60.12 (65)</td>
</tr>
<tr>
<td>2007 (first quarter)</td>
<td>62.54 (65)</td>
<td>58.60 (63)</td>
<td>64.48 (60)</td>
<td>63.24 (65)</td>
</tr>
</tbody>
</table>
## Main Regression Results

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(\text{Target} - \text{GPR}_{i,t-1}) \times \text{BELOW}$</td>
<td>0.3361**</td>
</tr>
<tr>
<td>$(\text{Target} - \text{GPR}_{i,t-1})^2 \times \text{BELOW}$</td>
<td>-0.0154**</td>
</tr>
<tr>
<td>$\text{BELOW}$</td>
<td>-0.2986</td>
</tr>
<tr>
<td>$-\text{GPR}_{i,t-1}$</td>
<td>0.2146**</td>
</tr>
</tbody>
</table>

* = significant at $\alpha = .05$
** = significant at $\alpha = .01$
Explanation

• Distance below the target has a curvilinear (inverted U) effect on moving up
  – Positive coefficient of distance and negative coefficient of squared distance

• Being BELOW the target last quarter does not have an independent effect on the change in GPR
  – Practices above T are just as likely to move up as those below T
Adjustment Process

- The coefficient of $-GPR_{i,t-1}$ provides an estimate of the speed of the adjustment process.
- $(1-\gamma) = 0.2146 \rightarrow$ about 1/5 of the adjustment to the long run occurs within 3 months.
- About 60% of the adjustment is complete within one year.
## More Regression Results

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>OBGYN</td>
<td>-1.8329*</td>
</tr>
<tr>
<td>Pediatrics</td>
<td>-0.7908</td>
</tr>
<tr>
<td>Family Practice</td>
<td>-1.0394</td>
</tr>
<tr>
<td>2005</td>
<td>-0.0926</td>
</tr>
<tr>
<td>2006</td>
<td>0.6222</td>
</tr>
<tr>
<td>2007</td>
<td>1.4332</td>
</tr>
<tr>
<td>Second quarter</td>
<td>-1.6961**</td>
</tr>
<tr>
<td>Third quarter</td>
<td>-0.09052**</td>
</tr>
<tr>
<td>Fourth quarter</td>
<td>0.7872*</td>
</tr>
</tbody>
</table>
In the long run: \[ GPR_{i,t} = GPR_{i,t-1} \]

- Set GPR before the P4P program at the practice’s equilibrium rate of generic prescribing, \( X\beta \)
- Vary the target and solve for GPR under the influence of the program
- Goal is to determine the maximum improvement and when poor performers give up
Long Run for GPR* = 40%

Max increase ~ 8.5 percentage points
Discussion

• Poor performers stop trying to improve only when the target is substantially higher (~23 percentage points) than the equilibrium rate of generic prescribing
  – Only one practice in our data is predicted to give up
• Targets will motivate poor performers unless they are set at unrealistically high levels
• But poor performers do not reach the target
  – They improve, but not enough to get rewards
  – What motivates them: Peer pressure? Agreeing with the goal of the target, even if they can’t reach it?
Effect of Target on Practices with Unexpected Poor Performance

• Suppose the practice’s equilibrium rate is at or above the target, but $\text{GPR}_{t-1}$ was below the target (unexpected bad luck)

• Does the target speed up the practice’s return to its equilibrium rate?

• We did an experiment like this and iterated until GPR converged to $\text{GPR}^*$

• Target speeds up the adjustment process
Speeding Up the Adjustment

Quarterly changes in GPR, with and without Target

![Graph showing quarterly changes in GPR, with and without Target.](image-url)
Implications for P4P Design

• Some argue that P4P targets should be lower for practices with low initial performance, otherwise, they will give up trying to improve.

• Others argue that setting low targets for this group would create ‘pushback’ from high performers.

• Our findings suggest this debate may not be necessary because low performers do not give up trying to reach a realistic target.

• Further study of incomplete improvement is warranted.
Thank You

• Copies of the paper are available from feldm002@umn.edu
• Paper is forthcoming in Health Economics