



What Behavioral Economics tells us about Pay-for-Performance Design:

Lessons from a Medicaid Health Plan

March 4, 2015

Robert Moore MD MPH
Chief Medical Officer

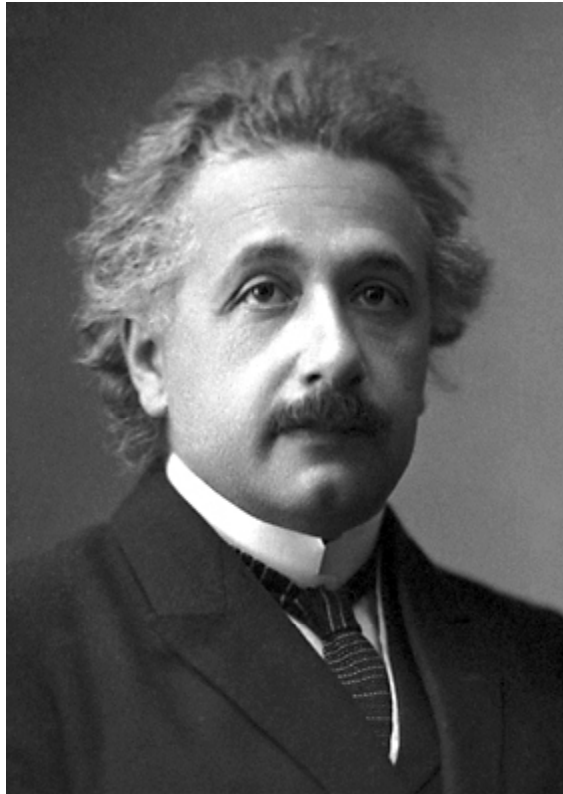
Using Incentives to Change Behavior

A Personal Journey

- Incentivizing front line staff: Health Center Medical Director
 - Incentivizing salaried clinicians
 - Responding to a Health Plan P4P program
- Switching sides: Health Plan CMO
 - How to design a better P4P system?
- What others have found
 - Some findings from the Behavioral Economics literature that shed light on P4P.

Opening Thoughts

Albert Einstein on P4P:



Unified
Newtonian physics,
thermodynamics, electromagnetism
and gravity

- “The only source of knowledge is experience”
- “A person who never made a mistake never tried anything new”

Incentivizing Salaried Clinicians

Community Health Clinic Ole

Federally Qualified Health Center in
Napa, CA



Medical Providers:

8 primary care physicians

2 physician assistants

5 nurse practitioners

Health Center income
dependent on volume of
face-to-face clinician visits



Case Study I: Productivity Incentives At Clinic Ole

Setting:

- **Pure salary, mission driven clinicians, no financial pressure**
 - Low productivity
 - Personalized care
 - Happy clinicians
- **Add pressure to meet community demand for services**
 - Top down efforts to increase productivity
 - Less happy clinicians

Case Study I: Productivity Incentives At Clinic Ole

Add productivity-based incentive

- **Structure of Incentive**
 - Per patient bonus for every patient over 10/shift seen on a given shift
- **Results**
 - Increased productivity (20% increase)
- **Analysis**
 - High productivity clinicians: faster, happier, less time with patients, less resentment of low volume clinicians
 - Low productivity clinicians: feel rushed, less happy
 - “Carefulness” suffers (hard to define)
 - “**Crowding out**” of altruistic incentives: retention affected

Unintended Consequences of Productivity Incentive

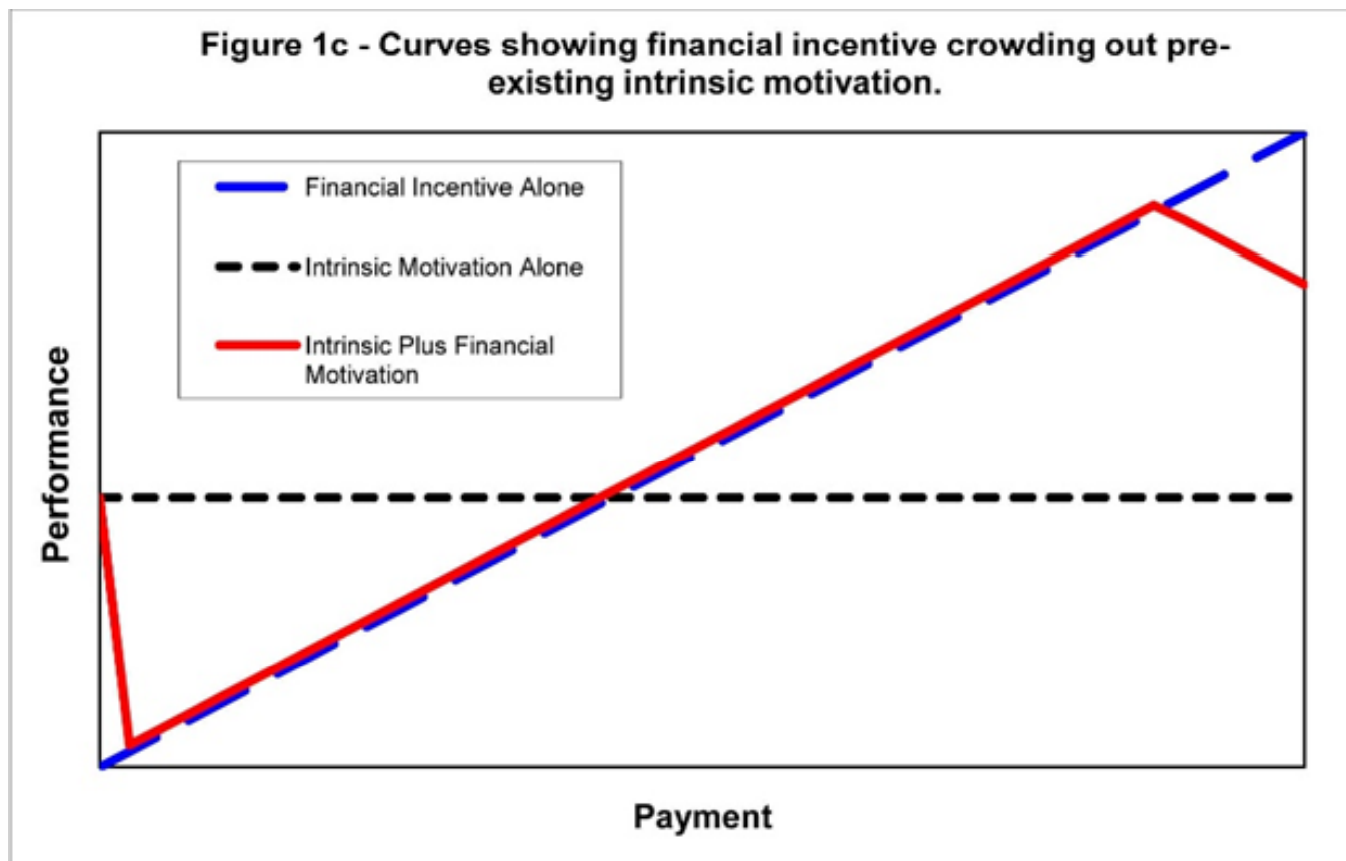
Lessons Learned:

- **A Productivity Incentive diverts energy away from time-consuming activities that improve quality:**
 - Addressing behavioral health
 - Motivational interviewing for self management of wellness and chronic disease
 - Employee wellness
 - Advance care planning discussions

Related Behavioral Economic Theory

Crowding Out: of altruistic incentives

- Several studies show that financial incentive payments can significantly and persistently decrease intrinsic motivation



Source: Woolhandler, Himmelstein, Ariely
"Will Pay for Performance Backfire? Insights from Behavioral Economics"
Health Affairs blog
2012

Case Study 2: Improving Quality at Clinic Ole

Background:

- Initial efforts to work with clinicians to improve quality

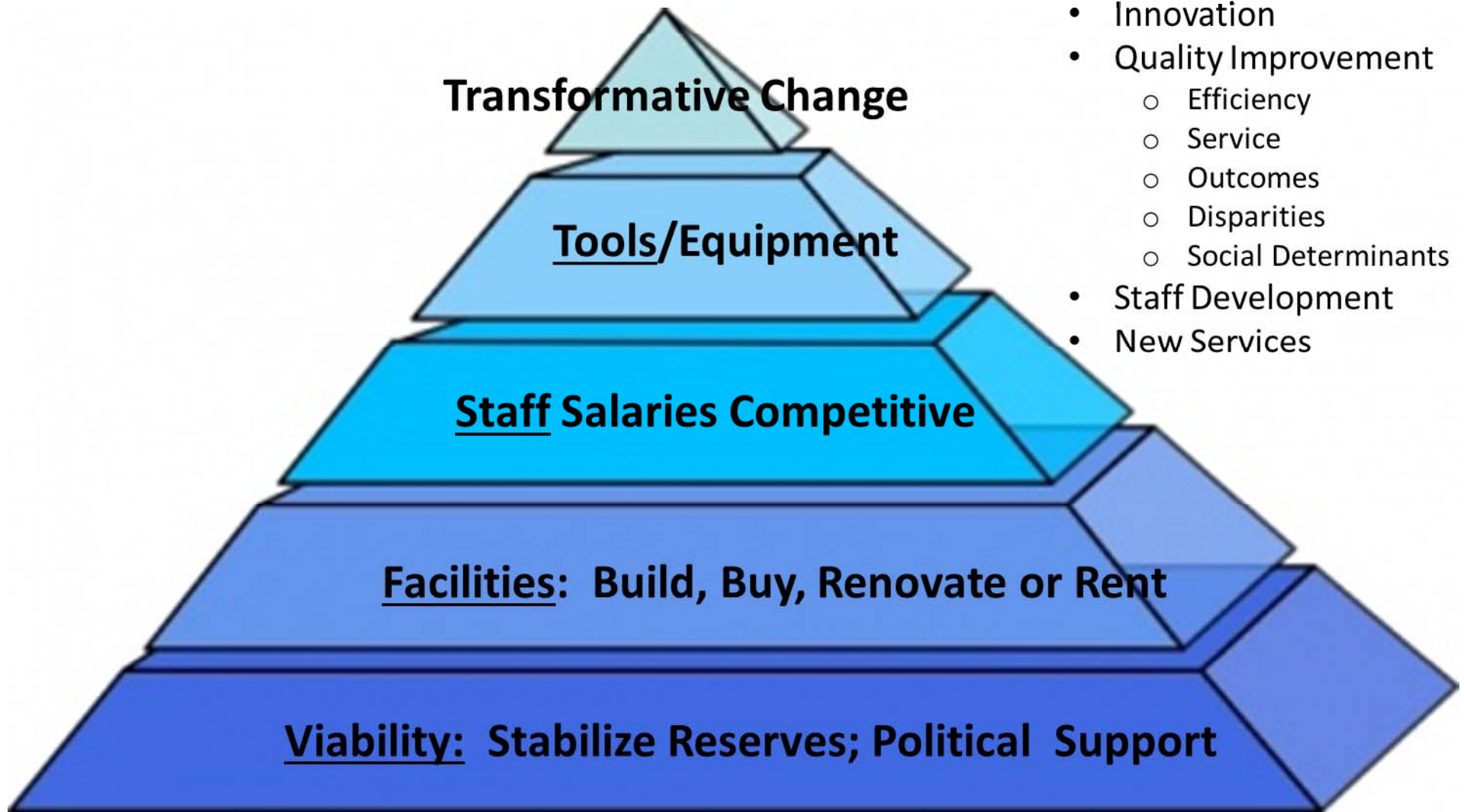
Findings

- Clinician time to do QI activities in conflict with:
 - Productivity incentive
 - Electronic health record implementation
 - Getting home at reasonable hour
 - Spending time to adequately address all prevention/chronic disease care issues.

Competing Priorities: other organizational priorities compete with energy devoted to QI activities.

What are these priorities?

Hierarchy of Needs for Medical Practice



Hierarchy of Needs for Medical Practice

Top Competing Priority:

- **Financial Stability:** Lack of this puts the very existence of the health center in jeopardy.
- **Is this true in general?**

Case Study 3

- Background:
 - Study of 35 emerging health center leaders in California (2014): If new funding were suddenly made available: what would you do with it?
- Findings:
 - If financial stability low: money put into reserves
 - If heavy competition and non-crisis reserves: money into customer service/better access

Case Study 4: Responding to Pay for Performance

How does the health center most effectively maximize revenue from a P4P program?

- What we did:
 - Carefully analyze whole P4P program
 - Hire staff to focus on population management:
 - Ordering labs due (LDL, HbA1c)
 - Scheduling visits for patients needing a service (Cervical cancer screening)
 - Completing quality reports
- Results:
 - Raises score without fundamentally changing clinician work or improving clinical outcomes

Consequence of Focus on Improving Measure

Analysis:

- Improving HbA1c scores by focusing effort on those patients closest to the threshold.
 - Goal < 7 ; focus on those with HbA1c 7.1 to 8, more likely to achieve goal as individuals and boost overall numbers.
 - The patients with HbA1c > 9 have much more benefit from intervention (with improved outcomes), but are less likely to help achieve the quality goal of more patients with HbA1c < 7 .

Consequence of Focus on Improving Measure

Related lessons from Behavioral Health Literature:

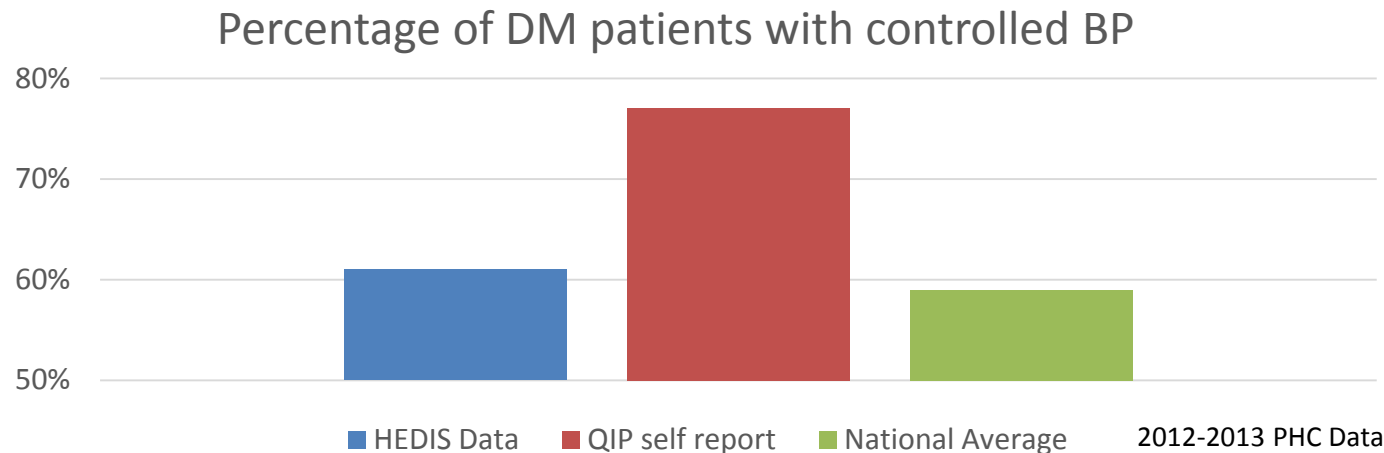
Gameability: The ability to improve scores without affecting underlying clinical outcomes

- May be due to efforts to ensure data is as complete as possible, and all supplementary activities that can be done without clinician are completed.
- May also be to slight “fudging” of data
 - experiments in chapter 11 of Ariely’s Predictably Irrational, where he shows that, on average, people cheat a little, while considering themselves honest)

Consequence of Focus on Improving Measure

Another example of **Gameability**

- Case study 5:
 - Data on BP control self reported data versus HEDIS audited data:
- Results:



Consequence of Focus on Improving Measure

Analysis: How do QI staff decide what to focus on?

- The amount of **increased funding** that would be available
- Divided by the **amount of work** required to make the numbers better. This, in turn, is related to an estimate the **ease** that interventions can improve scores

Size of P4P Incentive Relative to Visit-based Funding

Background:

Relationship between the size to the incentive and the effect at improving performance.

Case study 6:

Spurred by financial stresses caused by decreased productivity caused by Electronic Health Record

Implementation: focus on improved productivity

Results:

Decreased waiting time for appointment, faster phone answer time, more rapid cycle times. Side benefits of of improving access and patient satisfaction.

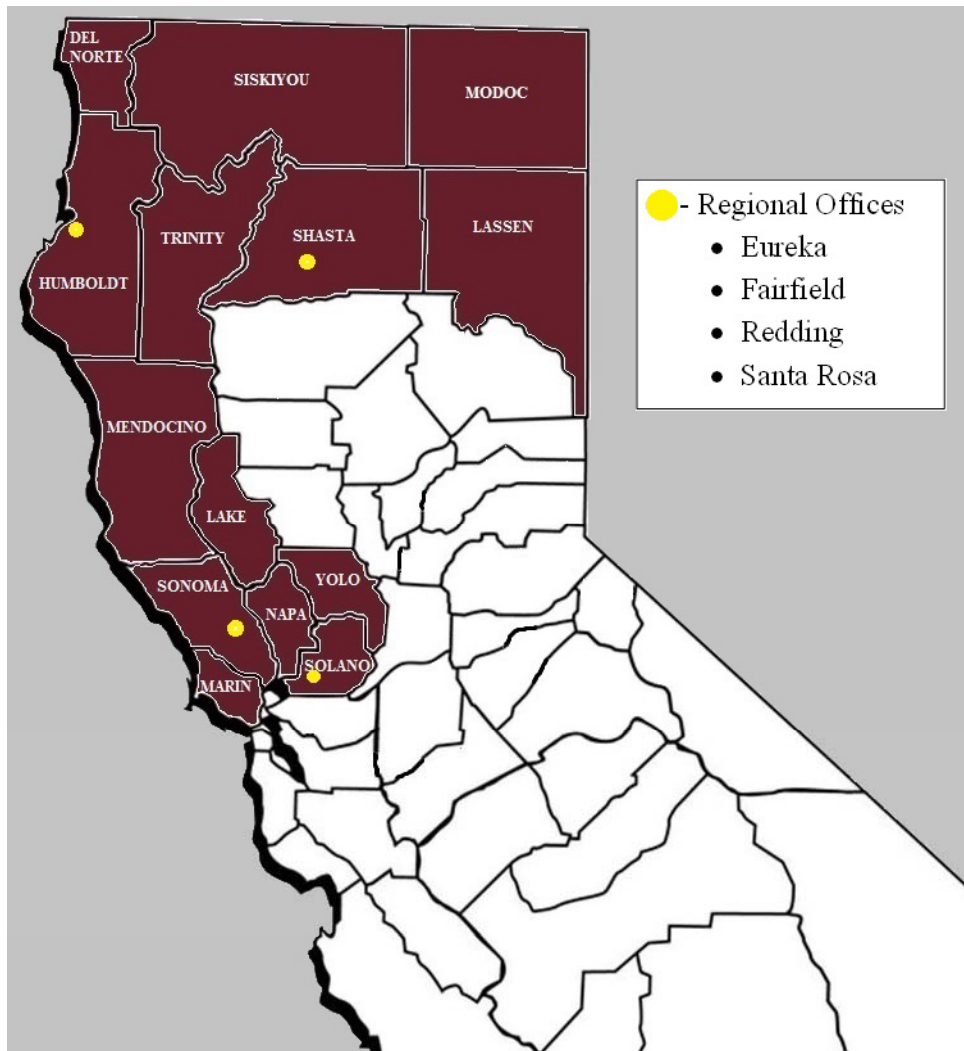
- Not metrics that are part of P4P

Improvements not Related to P4P

Analysis of case:

- Total of all P4P dollars is about 10% of the value of a single visit:
- It is rational to make the number of visits a priority in fiscally constrained environment
- Hierarchy of needs: financial survival takes precedence

Partnership HealthPlan of California



- **Mission:** *To help our members, and the communities we serve, be healthy*
- **Medi-Cal Managed Care:** County Organized Health System model
- **515,000 members**
- **14 counties**



Health Care Delivery at PHC: Primary Care

- Over 1500 **Primary Care Providers** in 220 organizations
- Member care provided:
 - 10% by Kaiser (full capitation)
 - 7% by Large Multispecialty Groups
 - 6% by Small Independent Practices
 - 77% seen by Organizations paid on a fixed, per-visit rate (Prospective Payment System)
 - Federally Qualified Health Centers (FQHC)
 - Rural Health Centers
 - Indian Health Service Health Centers

Primary Care Prospective Payment System (PPS)

Payment based on a per visit rate no matter how long or short the visit,

- if seen by an eligible clinician
- for a medically necessary visit
- at a PPS-eligible healthcare organization.

This is called the “PPS rate”

- Ranges from \$90 per visit to \$450 per visit
- Reason for variation: Includes allocation for overhead

Health Plan pays its usual rates, State pays the balance

Global payment by a Health Plan in the PPS setting is meaningless: PPS system trumps global payment, in terms of incentives

Permitted Alternatives to PPS

- Income outside the PPS process:
 - Quality Improvement Incentive Payments
 - Grants
- These may be paid by a Medicaid Managed Care Plan

P4P Effect on Different Provider Types

Central Question:

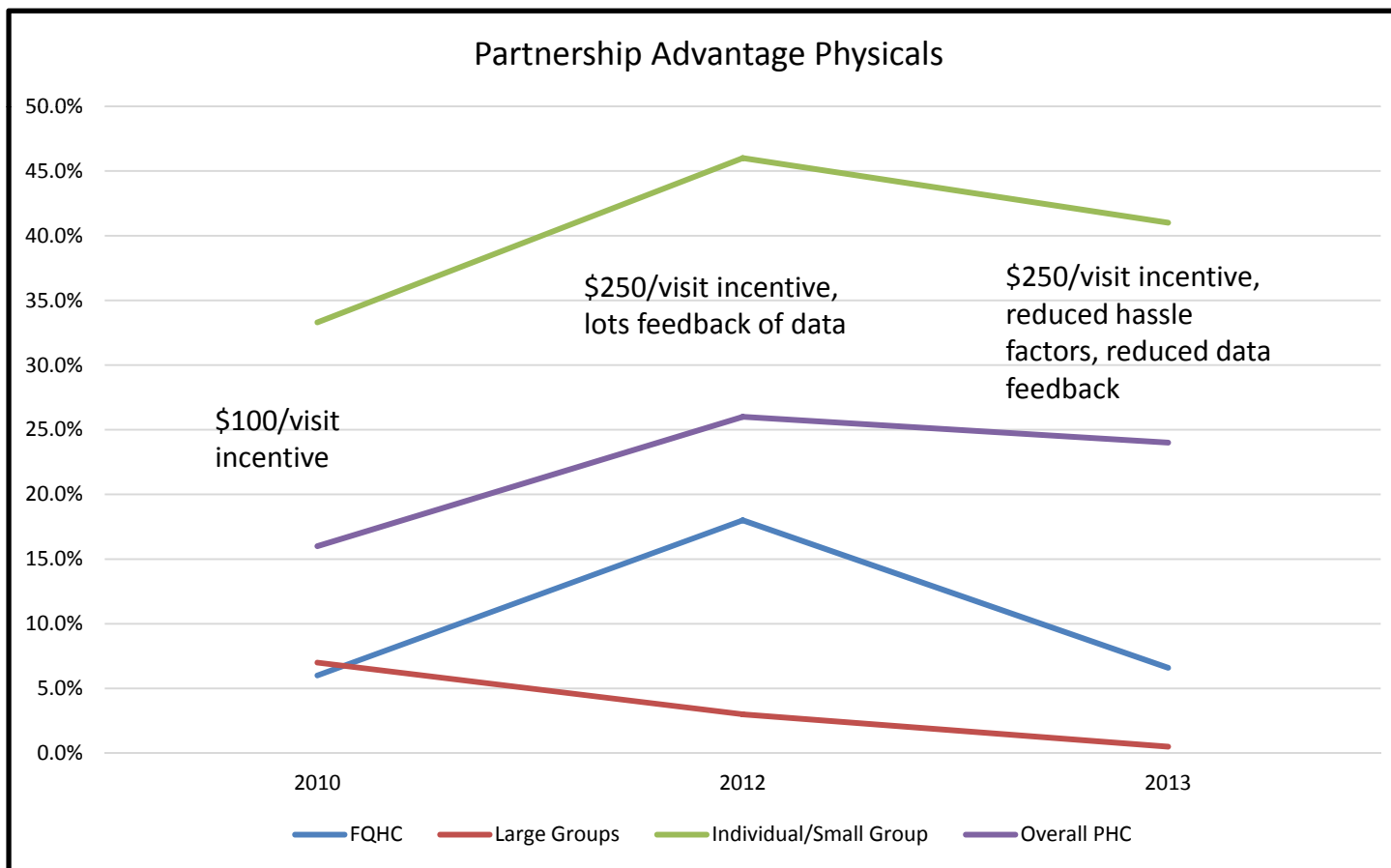
Given the different payment systems in place for different provider types, is there a differential effect of different types of P4P incentives?

Case 7: Incentives for Annual comprehensive visit for Members of Partnership Advantage Program

- High quality annual visit with specified data elements to meet Medicare Criteria and to capture data on total illness burden.
- Bonus payment (above usual visit fee) paid quarterly.

Effect of Practice Setting

Results:



Effect of Practice Setting

Lessons from this data:

1. Incentives based on payment per service work better for individual physicians/small groups, where the money goes directly to the physician
2. FQHCs have more responsiveness to payment per service than large medical groups, although the effect waned without attention.
3. When attention to the incentive program waned, performance worsened, in spite of reduction of hassle factors
4. (Data not shown): County clinic FQHC could not be incentivized to do any Physical exams; related to inability to develop alternative workflows to take advantage of extra funding without losing productivity
5. Most productive physician shares incentive with staff

Effect of Practice Setting

Why the difference?

1. **Directness of Link** between incentive payment and front line clinicians and other staff.

Support other studies:

- Cluster randomized trial of VA outpatient clinics showed that better performance on measures when individual clinicians are incentivized, instead of groups.

(Peterson et al 2013, JAMA: *Effects of Individual Physician-Level and Practice-Level Financial Incentives on Hypertension Care*)

Effect of Practice Setting

Why the difference?

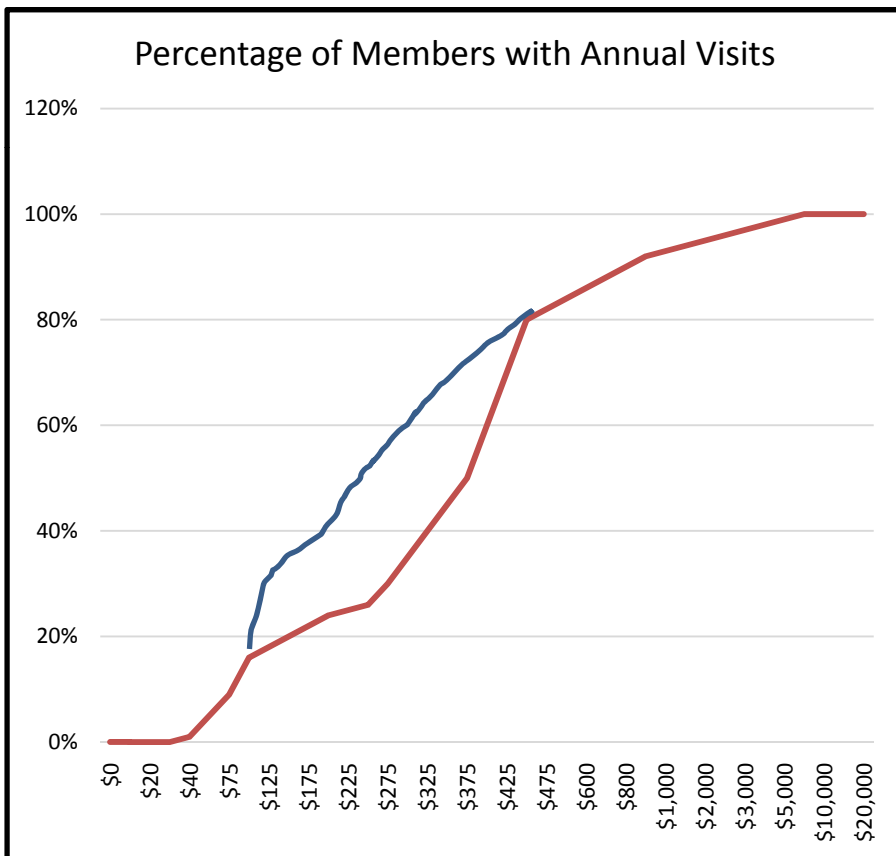
2. Size of the incentive (related to base pay)

- Increase in size of incentive increased the percentage of members that had an annual comprehensive visit, but not as much as anticipated.
- Increase from 16% to 26% of members with exams brought in additional \$341,000 (from improved HCG coding leading to better reimbursement), but cost the health plan an additional \$438,000.
- What is the relationship between increased funding and performance?

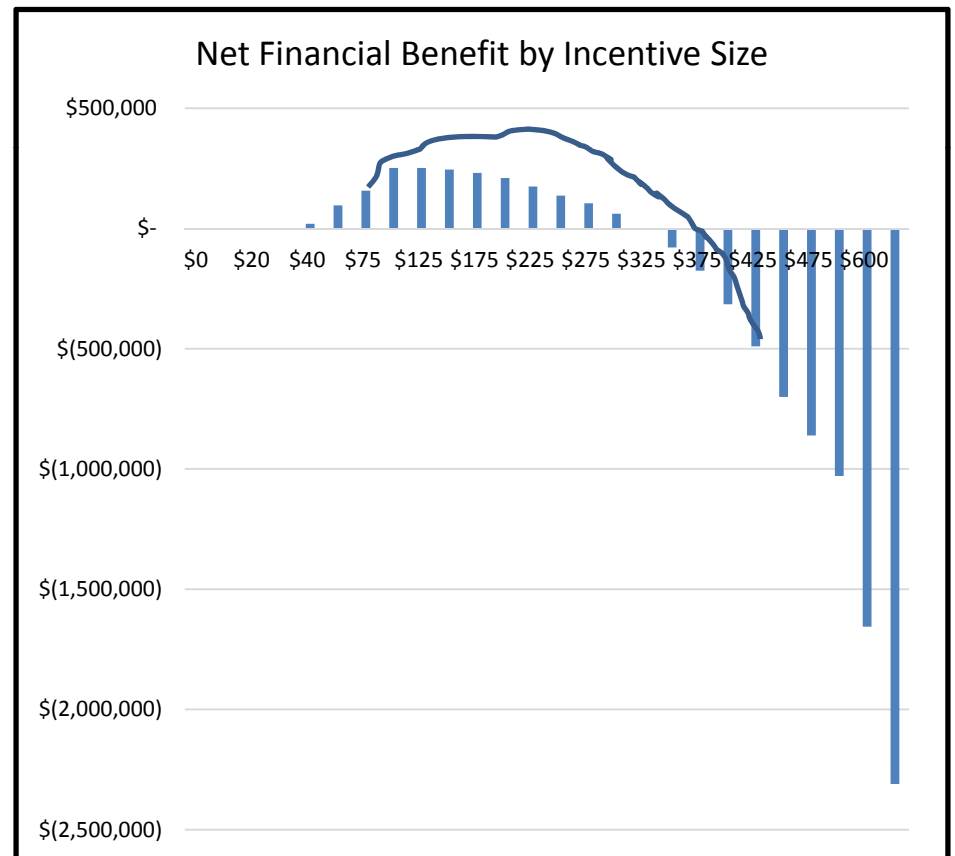
Effect of Incentive Size

Non-linear effect of incentive size,
with other variables being equal.

Optimizing Financial Benefit of P4P



Estimated from existing data and interviews



Extrapolating percentage, using PHC data

Effect of Practice Setting - II

Does this same effect of Practice Type hold true for other types of P4P?

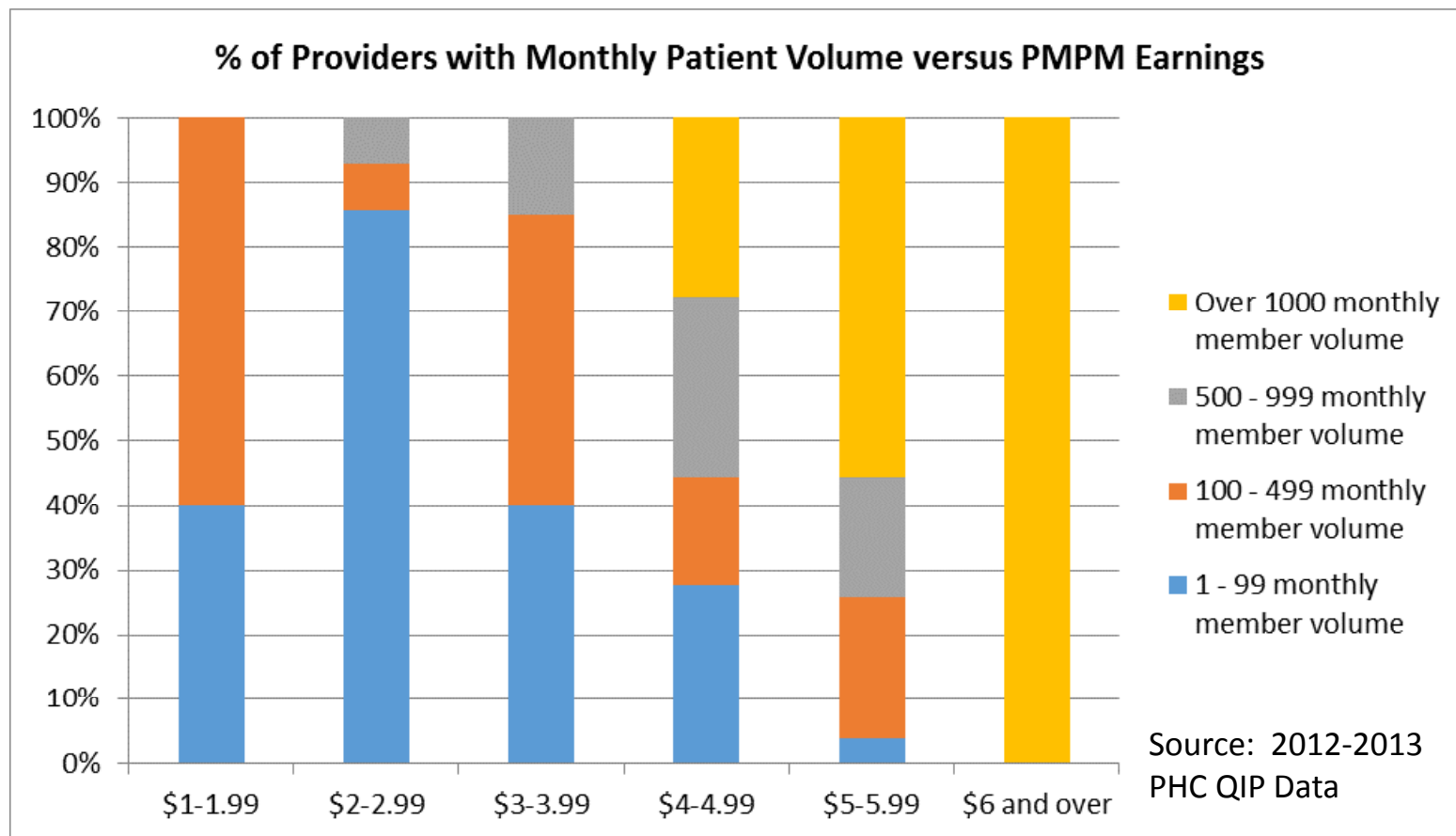
Case study 8: Performance on PHC's "fixed pool measurement set", based on number of members assigned.

- \$5 PMPM, all distributed
- Approximately 18 measures, each with points assigned to each, totalling 100 points
- Fixed pool measure set includes Clinical, Resource Use, Patient Experience, and Access measures.

Effect of Practice Setting - II

Results, part 1: Size of practice and performance

- *Opposite* of the effect seen in the Annual exam case



Effect of Practice Setting - II

Analysis: Why is performance worse for small practices in the fixed pool measures, but better for unit-of-service measures (like the Partnership Advantage Annual visit)?

A. The work of entering data:

1. Requires extra staff time to accomplish
2. Exceeds the operational ability of the office to handle
3. Outweighs the small benefit that would come from additional dollars
4. Competing priorities for the time of the support staff outweigh the pressure of the physician to focus time on achieving the goals

Which Measures Improve?

Testing this hypothesis

- If the difference in work involved was the key difference, would expect to see different global performance on measures, depending on the work need to improve them.

Case Study 9: Year over year comparison of results on P4P

- 2011-2012 year compared to 2012-2013
- Differences in measurement years: increased the thresholds for full payment
- Began implementation of pay for improvement

Measure Dependent Improvement

Results:

Table 1: Population level performance rates by clinical measure for provider sites submitting additional data

(See Appendix D for total denominators and numerators)

Measure	Total Rate 2011-2012	Total Rate 2012-2013
*Child BMI	80.7%	83.5%
*Adult BMI	76.7%	88.7%
Well Child	66.4%	72.3%
*Breast Cancer Screening	54.1%	64.3%
*Cervical Cancer Screening	64.3%	73.0%
HbA1c good control $\leq 9\%$	72.4%	72.4%
LDL good control <100	46.5%	46.3%
Blood pressure control	74.0%	76.7%

*Demonstrates statistical significance ($p < .05$) using Fisher's exact Test or a Chi-square with Yates correction

What distinguishes measures that **improve** from measures that **didn't**?

Measures that Improved

Analysis:

- Improved measures:
 1. Measures that are improved with entering of data
 2. *Measures that are improved by support staff ordering actions (all process measures)*
 - *Cervical Cancer screening*
 - *Breast Cancer screening*
 - *Ordering blood tests*
- Measures not improved: *require more comprehensive approach including activation of the patient to achieve. (DM, LDL, BP control)*
- *Illustrates the idea of **gameability** of measures*

Non-financial factors

In many cases, non-financial variables can overwhelm the effect of the financial incentive

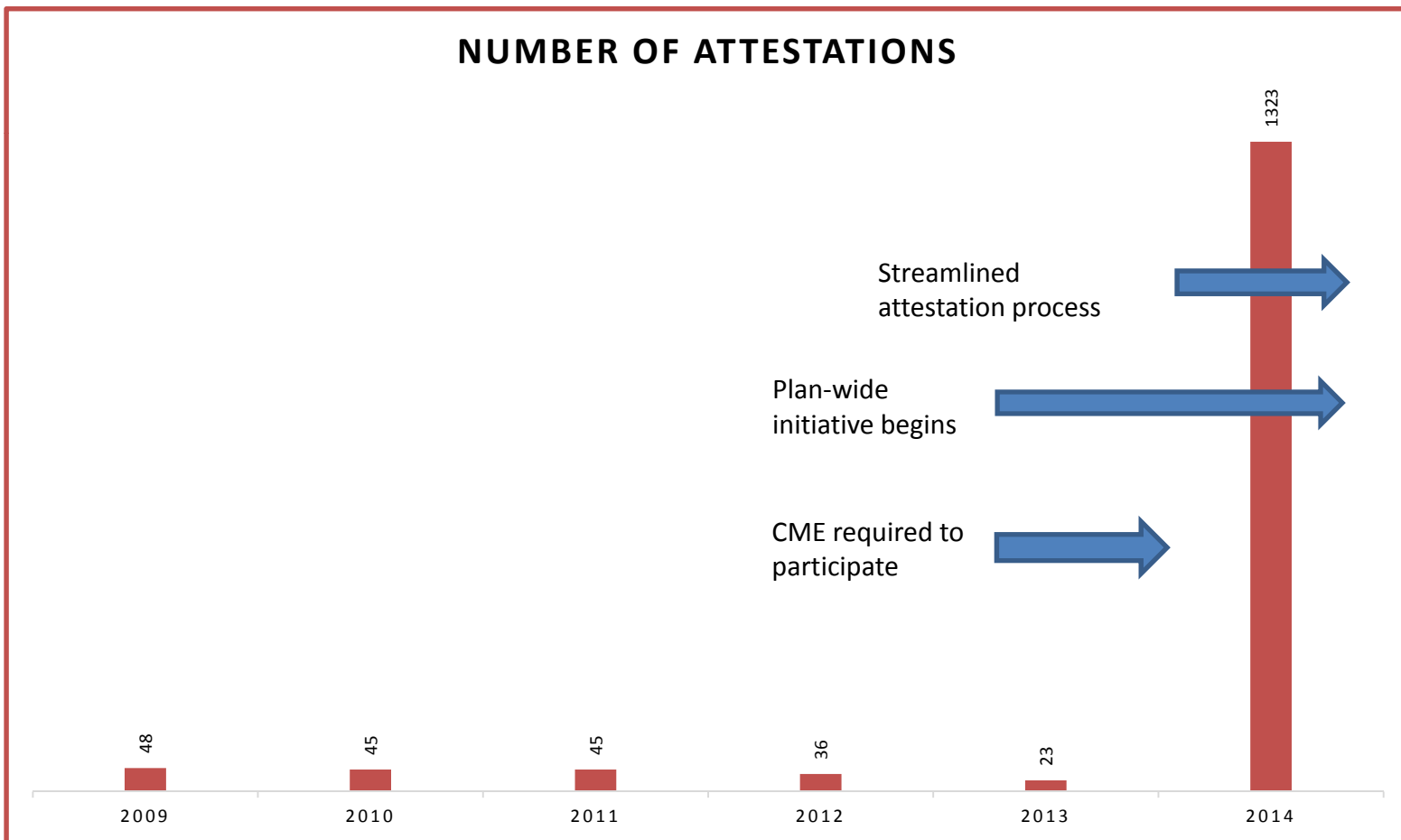
Case Study 10:

- Pay for performance measure for having a high quality conversation about Advance Care Planning
- Payment amount stable: \$100 per attestation submitted, per patient per year, up to maximum of 100 attestations per site. (total potential: \$10,000/site)
- Paid quarterly

Non-financial factors

Results:

- Quality Incentive for Advance Care Planning:



Non-financial factors

Analysis:

- Key variable is the **Hours of effort** put into the work of improving the measure.
- When work required increased, the number of attestations dropped, in spite of new plan wide initiative.
- When work required dropped, the number of attestations rose.
- Why did it increase so much?

Non-financial factors

Analysis:

Based on interviews of offices and health centers with large increases:

Key was **Leadership**:

- For many sites, the Leadership became dedicated to capturing the P4P dollars.
- For one site, the Leadership was dedicated to globally improving advance care planning systems at the health center. Side effect was improved score.

Case 11: Organizational Factors

Role of **Operational Strength**:

Strength of organizational operations and management in the organization. Includes Human Resource Management (HRM) and IT infrastructure.

- Case 11: Organization Self assessment of 15 Health Centers compared to quality outcomes. (48 questions; 9 domains)
- Findings:
 - Lower staff turnover most correlated with quality of care.
- Supportive evidence
 - Staff turnover at Baldrige award winning sites also low
 - Sophisticated Electronic Health record systems are important factors to improved quality in small practices in New York City.

Source:

-Bardach et al. 2013 JAMA: *Effect of Pay-for-Performance Incentives on Quality of Care in Small Practices with Electronic Health Records.*



Behavioral Economics

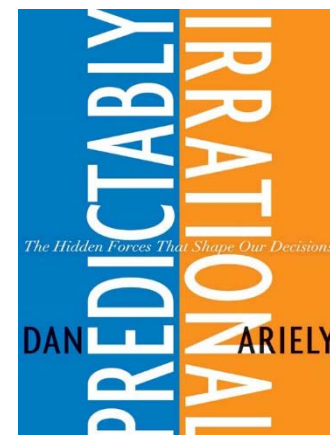
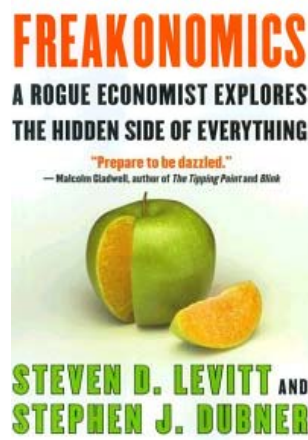
Definition: *Study of the effects of psychological, social, cognitive, and emotional factors on the economic decisions of individuals and organizations.*

- People do **not** act rationally based on the model of a purely “economic human” (Homo economicus), trying always to maximize their financial benefit.
- Human beings’ variation from “rationality” is experimentally provable and predictable.

Behavioral Economics

Popular Authors:

- Steven Levitt, University of Chicago Economist:
Co-Author of *Freakanomics*
- Dan Ariely, Duke University Psychologist:
Author of *Predictably Irrational*



Lessons from Behavioral Economics

Loss Aversion: Losses bite more than equivalent gains. (Dan Ariely: Predictably Irrational)

- Paying providers up-front and then asking for the money back if performance was not achieved would be expected to be more effective than paying afterwards.
- Psychological consequence: Severe negative feelings toward the health plan when they ask for the money back.

(Reference: Mehrotra et al. 2010 J Managed Care *Using the Lessons of Behavioral Economics to Design More Effective Pay-for-Performance Programs*)

The role of time

- Many studies show that incentives are stronger when they are more **closely related in time** to the behavior being rewarded.
- True in animals as well as humans. A variation on classical conditioning theory
- General principle of P4P: Link payment as closely as administratively possible to behavior being incentivized.
- At PHC, we structured Unit of Service measures to be paid quarterly, instead of yearly, partly to address this.
- Likely stronger effect for Individuals than for Organizations, which have a longer time frame for rewards.

Bureaucratic Complexity

Independent of the effort it takes to achieve a P4P goal, the **bureaucratic complexity** of the P4P program can reduce its effectiveness

- From interviews, this factor is at play with smaller practices in the PHC P4P program, who don't have QI staff who can take the time to understand the 100 page specifications document
- Implication: Need to balance simplicity with comprehensiveness of measure set and fairness

Source: Mehrotra et al. 2010, Am J Manag Care: Using the Lessons of Behavioral Economics to Design More Effective Pay-for-Performance Programs



Thresholds for achieving incentives

Goal Gradient. Different levels of goals for partial and full credit.

- Unachievable goals are unlikely to motivate effort
- Easily achieved goals are also unlikely to motivate effort
- Since different individual providers are starting from different places, need some gradient of goals to keep them motivated to improve.
- Continuous gradient may be less motivating (unproven).

Source: Mehrotra et al. 2010, Am J Manag Care: Using the Lessons of Behavioral Economics to Design More Effective Pay-for-Performance Programs



Using alternatives to money

Use of non-cash incentives.

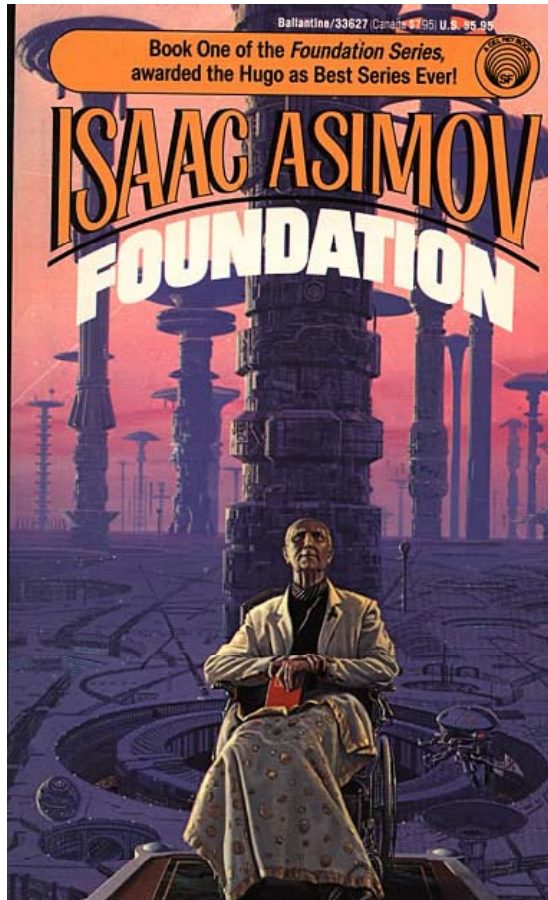
When money is not enough to incentivize providers and staff, sometimes using “prizes” of equal or lower value may have a greater effect.

- Examples: Professional Football Players and the Pro-Bowl
- Office staff: movie tickets; lunch for staff
- Clinicians or QI staff: scholarships to valued CME event

Source: Mehrotra et al. 2010, Am J Manag Care: Using the Lessons of Behavioral Economics to Design More Effective Pay-for-Performance Programs



How predictable is human behavior?



Hari Seldon:
Fictional Psychohistorian

Short digression on social science variables

Not physics variables with exact units that can cancel out (like $E=mc^2$)

Goals for understanding social science variables:

- Positive vs. negative correlation
- Linear vs. non-linear (if known)

Understanding Relationship of Variables

Positive Correlation

- Size of Incentive
- Leadership dedication to P4P
- Link between the measures and outcomes
- Link between incentive and front-line staff
- Financial stability
- Operational strength
- Loss aversion
- Use of goal gradient
- Use of non-cash incentives

Negative Correlation

- Crowding out of Intrinsic Rewards
- Competing priorities
- Work needed to raise scores
- Gameability
- Bureaucratic complexity

Behavioral Economics and P4P Architects

Four reasons that designers of P4P programs persist in valuing these programs although they have little evidence of benefit:

1. Ikea effect
2. Not invented here
3. Not understanding the psychology of behavioral economics



Source: Dan Ariely, 2011, The Upside of Irrationality

Lessons to Take Home

The number of inter-related variables that come to play in predicting the effectiveness of P4P mean that:

1. One cannot make any general statement about all of P4P based on any one set (or even a few sets) of data, looking at a few variables
2. In designing P4P, consider as many of the variables as possible, then carefully measure, quantitatively and qualitatively what happens, and try to understand it.

When building P4P programs, beware the Ikea effect!

Perfectly Designed

“Every system is perfectly designed to achieve the results it gets”. --Don Berwick

Corollary:

Every pay-for-performance program is perfectly designed to achieve the results that it gets.



Robert Moore, MD MPH

@Robert_Moore_MD

rmoores@partnershiphp.org

For details of Partnership's current Pay for Performance Program:

http://www.partnershiphp.org/Provider/MC_QIP.htm