Pacific Gas and Electric Company

Erik Jacobson
Director
Regulatory Relations

Pacific Gas and Electric Company
77 Beale St., Mail Code B10C
P.O. Box 770000
San Francisco, CA 94177

Fax: 415-973-7226

March 25, 2016

Advice 3698-G/4813-E
(Pacific Gas and Electric Company ID U 39 M)

Public Utilities Commission of the State of California

Subject: **Submission of High Opportunity Projects and Programs (HOPPs) Proposal - Residential Pay-for-Performance Program**

Purpose

The purpose of this Advice Letter (AL) is to submit a proposal to the California Public Utilities Commission (CPUC or Commission) to operate High Opportunity Projects and Programs (HOPPs) in compliance with the December 30, 2015 "Assigned Commissioner and Administrative Law Judge's Ruling Regarding High Opportunity Energy Efficiency Programs and Projects" (ACR). The ACR allows Program Administrators (PAs) to submit proposals for High Opportunity Programs to the Commission for expedited review, specifically, to the Commission's Energy Division via Tier 1 Advice Letters. (ACR, Paragraphs 1 and 2.)

PG&E plans to launch the Residential Pay for Performance (P4P) sub-program as a HOPP offering under the existing Residential Program. As explained below, the P4P program meets all of the requirements for HOPPs set forth in the ACR. Pursuant to Rule 5.1 of the Energy Industry Rules within General Order 96-B, PG&E designates this Tier 1 Advice Letter as effective pending disposition by the Energy Division. PG&E requests approval to be effective no later than April 15, 2016.

Background

On October 8, 2015, the Legislature enacted Assembly Bill (AB) 802, which amended Section 381.2 of the Public Utilities Code. New subsection (b) requires the Commission to authorize, by September 2016, electrical corporations or gas corporations to provide financial incentives, among other things, to increase the energy efficiency of existing buildings based on the reduction of metered energy consumption as a measure of energy savings. New subsection (c) states that "Effective January 1, 2016, electrical corporations and gas corporations are authorized to implement the provisions of subdivision (b) for high opportunity projects or programs." The idea behind HOPPs is to identify "high opportunity" interventions clearly within the ambit of legislative direction

before the Commission adopts a comprehensive program to provide incentives to improve the energy efficiency of existing buildings.
On October 30, 2015, the assigned Commissioner and Administrative Law Judge (ALJ) issued their scoping memorandum regarding energy efficiency “Rolling Portfolios” and established a process specifically for addressing “high opportunity programs or projects,” along with other aspects of AB 802.[1]

The ACR provides minimum standards for the development and implementation of HOPPs.[2] HOPPs may be funded from unspent funds in existing programs. There are no minimum requirement for expected savings for HOPPs. HOPPs may feature a variety of incentive structures, so long as the payment strategy reflect an accurate valutation of the savings. All HOPPs must incorporate a measurement and verification (M&V) plan, including the M&V protocols set out in the ACR. A key feature is that HOPPs proposals should emphasize measurement of the effects of interventions as detailed in Attachment A of the Ruling.

PAs are authorized to submit High Opportunity Program proposals with the documentation and specifications listed in the ACR. High Opportunity Project proposals are to be submitted through the CPUC Energy Division’s existing Custom Measure and Project Archive (CMPA) system.

This advice letter provides all of the material needed to meet the PA filing requirements, and addresses all the ACR’s preferred principles of HOPP program design. That is, PG&E’s HOPP:

(1) focuses on existing buildings,
(2) draws upon input from a diverse stakeholder group, the EM&V results and lessons learned from a similar offering, and best practice EM&V methods, and
(3) focuses on energy efficiency activities that are newly permissible under the statutory changes by considering all energy efficiency achievements, as measured at the customer’s meter, and by using a new intervention strategy and savings measurement regime.

PG&E’s HOPP proposal includes monthly program milestones that should lead to the enrollment of residential customers in September 2016, assuming that advice letter review and approval occur pursuant to the ACR’s procedure for review.

[1] “Assigned Commissioner and Administrative Law Judge's Ruling and Amended Scoping Memorandum Regarding Implementation of Energy Efficiency ‘Rolling Portfolios’ (Phases IIB and IIIA of R.13-11-005)” (Phase IIB/IIIA scoping memo).
[2] ACR, Paragraph 5.

Program Proposal

The residential P4P program proposal is summarized in Table 1 below. A detailed description of the P4P program is provided in Attachment A. The EM&V Plan for claiming energy savings is provided in Attachment B.

Table 1

Program Name:	Residential Pay-for-Performance
Proposal Type:	High Opportunity Program
Sector:	Residential
Brief Description:	Pacific Gas and Electric Company (PG&E) will offer a residential Pay-for-Performance (P4P) program based on a model originally described in Natural Resources Defense Council (NRDC)'s Phase II Workshop 3 comments, which was supported by PG&E and other stakeholders. Building off of NRDC's proposal, PG&E has worked with a broad stakeholder group to develop a framework intended to build a platform for scalable residential retrofits while minimizing administrative and implementation costs. This model seeks to more fully engage existing market actors like Property Accessed Clean Energy (PACE) loan providers, smart thermostat vendors, vertically integrated contractors, program implementers, and other businesses to advance and scale residential retrofits. This program will begin in 2016 with an initial enrollment period (IEP) of 2 years and annual incentive payments to Aggregator(s) one and two years after the initial interventions are performed. Aggregators are parties responsible for managing a portfolio consisting of numerous residential homes that receive energy efficiency interventions in an effort to maximize energy savings from those sites. The IEP will serve as an assessment period for the initial incentive design and evaluation strategies, during twhich PG&E will select several Aggregators through a competitive solicitation. The Aggregators will work directly with residential customers and contractors to achieve energy savings through retrofits in addition to operational and/or behavioral interventions. Aggregator payments will be determined based on gross energy savings through a PG&E facilitated weather normalized pre/post analysis of each participating customer's metered energy consumption. This measurement will be conducted through the CalTRACK system, a data analysis process which is under development with broad stakeholder input to provide a consistent measurement process across the state. The final details of the CalTRACK process will be submitted via a Program Implementation Plan (PIP) Addendum later in 2016, prior to the start of the IEP. Through this process, each home's usage will be measured individually and then added together to determine the aggregator's total portfolio performance. PG&E will pay each aggregator a set rate per therm and kWh based on their gross portfolio savings. PG&E will create and host a dashboard to display the performance of each portfolio of projects undertaken to increase market visibility into residential energy savings. The goal is to start with a simplified flat payment structure focused on gross savings.

	However once we establish the framework for this program, it can be expanded to emphasize other regional or state priorities such as locational savings, specific measures, time of use, and net savings. An additional incentive will be offered to aggregators for net savings during the IEP in order to promote management of this metric. While this additional incentive will be minimal for the purposes of initial enrollment, we intend for it to guide the market to focus on attainment of net savings. The early results of the IEP will help provide deeper insights into savings per measure and customer type and enable more sophisticated program metrics, allowing the further monetization of energy efficiency measures in the market. Future enrollment periods would be informed by this effort and include a price discovery mechanism to ensure the best value for rate payers. Additional discussion and further program details can be found in Attachment A.
Incentive Design:	PG&E will pay each Aggregator a set rate per therm and kWh for their delivered weather normalized gross portfolio savings (“payable savings”). Aggregators will be paid in two partial payments: one and two years post intervention; no up front payments will be made. An additional incentive will be offered to aggregators for net savings in order to promote management of this metric that follows the same cadence.
Measure Treatment:	Multiple measures; primarily retrofits such as heating, ventilation and air conditioning (HVAC) and insulation, also includes behavioral and operational measures.
Measurement technique:	Utilize CalTRACK, which is described in more detail in Attachment A, to perform pre/post intervention analysis of weather normalized metered consumption to determine gross payable savings.
EM&V methodology:	Pre/post intervention analysis of participant’s metered energy consumption compared to a matched pair control group through a quasi-experimental design approach. This methodology is described in more detail in Attachment B.
Proposed Budget:	2 years, $6M ($5M incentives) with the option to expand based on first year results.
Budget source(s):	Residential Energy Efficiency Programs
PG&E contact(s):	Primary Contact: Halley Fitzpatrick (hdf2@pge.com) Program Lead: Leif Magnuson (l3mz@pge.com) Policy Lead: Kate George (KEG9@pge.com) EM&V Lead: Brian Smith (B2SG@pge.com)

Protests

Anyone wishing to protest this filing may do so by letter sent via U.S. mail, facsimile or E-mail, no later than April 14, 2016, which is 20 days after the date of this filing. Protests must be submitted to:

CPUC Energy Division
ED Tariff Unit
505 Van Ness Avenue, 4th Floor
San Francisco, California 94102

Facsimile: (415) 703-2200
E-mail: EDTariffUnit@cpuc.ca.gov

Copies of protests also should be mailed to the attention of the Director, Energy Division, Room 4004, at the address shown above.

The protest shall also be sent to PG&E either via E-mail or U.S. mail (and by facsimile, if possible) at the address shown below on the same date it is mailed or delivered to the Commission:

Erik Jacobson
Director, Regulatory Relations
c/o Megan Lawson
Pacific Gas and Electric Company
77 Beale Street, Mail Code B10C
P.O. Box 770000
San Francisco, California 94177

Facsimile: (415) 973-7226
E-mail: PGETariffs@pge.com

Any person (including individuals, groups, or organizations) may protest or respond to an advice letter (General Order 96-B, Section 7.4). The protest shall contain the following information: specification of the advice letter protested; grounds for the protest; supporting factual information or legal argument; name, telephone number, postal address, and (where appropriate) e-mail address of the protestant; and statement that the protest was sent to the utility no later than the day on which the protest was submitted to the reviewing Industry Division (General Order 96-B, Section 3.11).

Effective Date

PG&E requests that this Tier 1 advice filing become effective on April 15, 2016, which is 21 calendar days after the date of filing.[3]

[3] ACR, p. 26.

<u>Notice</u>

In accordance with General Order 96-B, Section IV, a copy of this advice letter is being sent electronically and via U.S. mail to parties shown on the attached list and the parties on the service list for R.13-11-005. Address changes to the General Order 96-B service list should be directed to PG&E at email address PGETariffs@pge.com. For changes to any other service list, please contact the Commission's Process Office at (415) 703-2021 or at Process_Office@cpuc.ca.gov. Send all electronic approvals to PGETariffs@pge.com. Advice letter filings can also be accessed electronically at: http://www.pge.com/tariffs/.

/S/

Erik Jacobson
Director, Regulatory Relations

Attachments

cc: Service List R.13-11-005

CALIFORNIA PUBLIC UTILITIES COMMISSION

ADVICE LETTER FILING SUMMARY
ENERGY UTILITY

MUST BE COMPLETED BY UTILITY (Attach additional pages as needed)

Company name/CPUC Utility No. **Pacific Gas and Electric Company (ID U39 M)**

Utility type:	Contact Person: Yvonne Yang
☑ ELC ☑ GAS	Phone #: (415) 973-2094
☐ PLC ☐ HEAT ☐ WATER	E-mail: Qxy1@pge.com and PGETariffs@pge.com

EXPLANATION OF UTILITY TYPE	(Date Filed/ Received Stamp by CPUC)
ELC = Electric GAS = Gas PLC = Pipeline HEAT = Heat WATER = Water	

Advice Letter (AL) #: **3698-G/4813-E** **Tier: 1**

Subject of AL: **Submission of High Opportunity Projects and Programs (HOPPs) Proposal - Residential Pay-for-Performance Program**

Keywords (choose from CPUC listing): Compliance, Energy Efficiency

AL filing type: ☐ Monthly ☐ Quarterly ☐ Annual ☑ One-Time ☐ Other ______________________________

If AL filed in compliance with a Commission order, indicate relevant Decision/Resolution #: N/A

Does AL replace a withdrawn or rejected AL? If so, identify the prior AL: No

Summarize differences between the AL and the prior withdrawn or rejected AL: ____________________

Is AL requesting confidential treatment? If so, what information is the utility seeking confidential treatment for: No

Confidential information will be made available to those who have executed a nondisclosure agreement: N/A

Name(s) and contact information of the person(s) who will provide the nondisclosure agreement and access to the confidential information: ____

Resolution Required? ☐Yes ☑No

Requested effective date: **April 15, 2016** No. of tariff sheets: **N/A**

Estimated system annual revenue effect (%): N/A

Estimated system average rate effect (%): N/A

When rates are affected by AL, include attachment in AL showing average rate effects on customer classes (residential, small commercial, large C/I, agricultural, lighting).

Tariff schedules affected: N/A

Service affected and changes proposed: N/A

Pending advice letters that revise the same tariff sheets: N/A

Protests, dispositions, and all other correspondence regarding this AL are due no later than 20 days after the date of this filing, unless otherwise authorized by the Commission, and shall be sent to:

California Public Utilities Commission
Energy Division
ED Tariff Unit
505 Van Ness Ave., 4th Flr.
San Francisco, CA 94102
E-mail: EDTariffUnit@cpuc.ca.gov

Pacific Gas and Electric Company
Attn: Erik Jacobson
Director, Regulatory Relations
c/o Megan Lawson
77 Beale Street, Mail Code B10C
P.O. Box 770000
San Francisco, CA 94177
E-mail: PGETariffs@pge.com

Attachment A: Detailed Proposal for High Opportunity Program – Residential Pay for Performance

Section 1: General Program Description

Overview:
PG&E's Residential Pay-for-Performance (P4P) Program seeks to develop a scalable model for residential retrofits that leverages rapidly emerging market actors and products while minimizing administrative and implementation costs[1]. There are several improvements over our existing Home Upgrade program that this new P4P program design offers. The current Home Upgrade program uses a traditional Program Administrator implementer model with participating contractors and customer incentives based on a limited set of allowable measures. PG&E's P4P offering allows participation by more market actors and a high level of flexibility for them to choose the services and products that customers want and that achieve reduced energy consumption. Further, the current Home Upgrade program pays incentives based on predicted or deemed savings, which puts rate payer funding at risk when the savings don't materialize. The goal of the P4P approach is to limit risk by paying incentives only for energy savings that materialize at the meter.

After regulatory approval of this proposal, PG&E will release a competitive solicitation for the Initial Enrollment Period (IEP) for interested parties which will be overseen by a Peer Review Group (PRG) of non-monetarily interested parties to help ensure oversight and transparency to the competitive solicitation process. The solicitation will seek out parties who will directly or through a network of contractors perform energy efficiency interventions in customers' homes with the goal of maximizing measureable savings, referred to as Aggregators. Aggregators may consist of existing market actors such as Property Accessed Clean Energy (PACE) loan providers, smart thermostat vendors, vertically integrated contractors, program implementers or may be new entrants to the California market. These Aggregators will bid for funding through Power Savings Agreements (PSA); we anticipate multiple bids allowing us to test different approaches, geographies and measure mixes.

An initial $5M incentive budget will be allocated based on criteria defined in the solicitation. PG&E will only pay for Aggregator portfolio kWh and therm consumption reductions. This method reduces risk and costs by not paying for individual homes in the portfolio which have neutral or negative savings. Payments will be made annually one and two years after the initial intervention based on aggregators' total portfolio weather normalized metered savings. Performance is measured based on weather normalized energy consumption data obtained through PG&E's advanced metering infrastructure (AMI) and collected 12 months before and

[1] Examples of emerging market actors and products are PACE and other energy efficiency loans, Home Area Networks, smart thermostats and behavioral feedback devices and methods.

after interventions are performed. Additional incentive payments (“kickers”) are provided to aggregators that demonstrate net savings measured through the evaluation methodology detailed in Attachment B.

PG&E believes that energy efficiency incentives are needed to boost energy savings from existing and emerging market actors’ services and to optimize the deployment of new energy efficient products. Currently, these market actors either do not benefit directly from the energy savings they help their customers achieve or they are employing product solutions that do not maximize the energy savings their customers can achieve. This program will give these market actors a reason to focus on comprehensive projects and persistence of energy savings by incentivizing them directly for every kWh and therm they reduce. It will ensure maximization of customer trigger points and expenditures to improve the efficiency of existing residential buildings. Further, it will prepare market actors to successfully bid into future PG&E energy efficiency procurements by increasing their knowledge of what energy savings their interventions will achieve, and at what cost.

PG&E will continually assess the progress of the portfolios during the IEP and if success milestones are achieved additional funding may be made available. If the current program design does not achieve the goals set forth for the program, or the IEP yields lessons to improve the program structure, PG&E will issue a second enrollment period with modified requirements.

Payment and Savings Calculation Methods Overview:
The energy savings and associated performance payments will be made based on measured weather normalized usage reduction using the CalTRACK system. CalTRACK is the name we have adopted to describe the pre/post intervention energy usage data analysis system and process that is currently under joint development by the California Energy Commission (CEC), the California Public Utilities Commission (CPUC), and the Investor-owned Utilities (IOUs) with assistance from data analysis experts at DNV-GL, Energy Savvy, Olivine, and the Department of Energy, National Renewable Energy Laboratory (DOE-NREL). The goal of the CalTRACK effort is to determine a standardized process for measuring residential energy savings.

The initial version of the CalTRACK system will be completed by mid-2016, prior to the launch of P4P. When completed, all CalTRACK methods and source code will be open source so that aggregators and other entities involved in, or managing similar programs can run their own analysis using standardized methods and common computational algorithms. The CalTRACK system will access smart meter data via the PG&E Share My Data platform allowing for near real time transmittal of usage data.

Utility claimable energy savings will be determined using quasi-experimental design practices. In order to perform this evaluation, we will require Aggregators currently active in PG&E territory to provide existing customer data to help establish baselines for claimable savings and program influence (see Appendix B for additional details).

Stakeholders[2] have expressed a strong interest for the program to drive net savings as well as gross payable savings. With this in mind, PG&E will offer Aggregators an additional incentive per kwh and per therm for net savings measured through the evaluation methods defined in Appendix B. However, this incentive will be initially nominal as it is included primarily to inform the market of the importance of delivering net savings while managing costs and risks to market actors and ratepayers. Learnings from this IEP will inform how quickly, or if, we can move to a net savings payment model in the future.

Measure Treatment:
The P4P program is designed to offer maximum flexibility for retrofit options coupled with operational and behavioral interventions. As a result, there is no list of required eligible measures. However we will require aggregators to report intervention tactics and associated implementation dates to inform the evaluation process. Retrofit measures will likely include traditional items such as insulation, air sealing, HVAC replacement, water heating, windows, pool products, large appliances and hardwired lighting fixtures. Operational and behavioral items may include, but are not limited to, connected devices, engagement and feedback applications, HVAC and water heating setting adjustments, and ensuring equipment is meeting manufacturer designed performance metrics. This program will not include retro-commissioning and as a result will not require maintenance plans be adopted.

Participants must sign up through the Aggregator acknowledging participation and inability to participate in other incentive offerings. This will include releasing the incentive payment and usage data to the aggregator. The Aggregator will be responsible for providing customers with insights into their energy usage patterns and striving to maximize the savings achieved. Customers with solar PV or who add solar PV while enrolled must provide verifiable production data to calculate energy savings and to allow for paying incentives to Aggregators for that site.

PG&E will measure weather normalized gross savings for two years after the intervention and at least one year after aggregator payments cease. This will give greater insights into measure performance, persistence and long term savings claims.

Sharing Best Practices and Lessons Learned:
PG&E will use CalTRACK to create a dashboard to display the performance of each portfolio of projects undertaken by each aggregator so that others can track the program's performance. The dashboards will be updated quarterly based on the analysis of the past quarter's weather normalized meter data. PG&E will perform a process evaluation to identify opportunities for program improvement and expansion. Annual reports will include findings from the process evaluations along with internal analysis and program updates. Additionally, data from Program will be compiled, aggregated and published in reports on the California Measurement Advisory Council (CALMAC) website to enable further market growth.

Stakeholder Engagement:
In order to develop the details of the program, PG&E worked closely with numerous stakeholders including NRDC, TURN, ORA, Sustainable Spaces, PACE providers, contractors and others with longstanding participation in residential energy efficiency markets. PG&E plans

[2] Office of the Rate-payer Advocates (ORA), The Utility Reform Network (TURN)

to continue this stakeholder engagement through the IEP to share results, ideate improvements and plan for the second enrollment period.

Section 2: Background and Program Drivers

The residential sector represents 31% of electricity consumption and 44% of total natural gas consumption within PG&E's service territory[3]. Existing single family buildings comprise 76% the sector's consumption[4]. To meet the State's ambitious goals established in legislation such as Senate Bill (SB) 350, AB 32, and other energy efficiency and carbon reduction goals, we must be able to scale our interventions in this segment.

Alignment with State Goals and HOPPs Requirements:

PG&E's P4P program is in close alignment with AB 802, the Strategic Plan, AB 758, Market Transformation and other key state objectives:

- **Assembly Bill (AB) 802:** The program will directly align by paying incentives based on metered savings from all retrofit, behavioral and operational savings, which will encompass to code and above code opportunities.
- **California Long-Term Energy Efficiency Strategic Plan:** This program can achieve energy savings in a manner that directly addresses several of the Strategic Plan's stated goals. This includes broad goals such as reducing greenhouse gas emissions in a rapid and low cost manner[5] and creating a way for utilities to stimulate market transformation levering non-utility actors to push the market[6]. In addition, it aligns with specific residential goals 2 and 3 detailed in Section 2 of the plan[7]. These goals focus on embracing bundled multi-measure energy efficiency approaches, as well as more efficient plug load products and customer behavioral elements.
- **AB 758 and California's Existing Buildings Energy Efficiency Action Plan:** This program directly supports Strategy 3.2.2 as it employs performance-based incentives to support savings realization and persistence, in tandem with finance mechanisms[8].
- **Market Transformation:** Recently the IOUs led the development of a Home Upgrade Market Transformation plan[9]. This plan defined strategic market transformation initiatives as those intended to both establish and meet the goal to save energy by changing market structures and consumer behavior. The report concluded that for such an effort to succeed, existing programs (e.g., PACE loan programs, IDSM, smart homes, and other such efforts) will need to be coordinated to change the culture and operation of California's existing residential market. This proposal leverages these insights and is

[3] Ref: 2015 Integrated Energy Policy Report (IEPR), Docket # 15-IEPR-01, CEC-100-2015-001-CMF, http://www.energy.ca.gov/2015_energypolicy/index.html

[4] California Existing Buildings Energy Efficiency Action Plan. September 20015. Pg. 11.

[5] *California Long-Term Energy Efficiency Strategic Plan* (2011), Section 1, page 2.

[6] *California Long-Term Energy Efficiency Strategic Plan* (2011), Section 1, page 5.

[7] "*Develop partnerships for innovative financing programs, such as performance contracts … Design implement, monitor and continuously improve financial products and programs for whole house energy efficiency and renewable energy retrofits.*" *California Long Term Energy Efficiency Strategic Plan (2011), Section 2, page 11*

[8] *California's Existing Buildings Energy Efficiency Action Plan* (2015), Section 3, page 75.

[9] A *Comprehensive Strategic Market Transformation (SMT) Plan for a Home Upgrade Program SMT Initiative* (2015). Navigant Consulting. www.calmac.org

designed to pull together the disparate market actors to drive scalable residential energy efficiency.

In addition to alignment with legislation, Strategic Plan and Market Transformation goals, this proposal exemplifies many other desired aspects defined in the CPUC guidance for High Opportunity Programs and Projects (HOPPs):

- **New and innovative design, partnerships, concepts or measure mixes**: This program allows technical innovations, leverages innovative intervention and market strategies, and is designed to provide scalable savings with less ratepayer funds than existing segment offerings. The program grants Aggregators the freedom to tailor a mix of interventions based on customer needs, granting them unparalleled flexibility to introduce new measures to help customers save energy as long as these measures lead to persistent measureable consumption reductions, measured at the customer's meter. Aggregators will bear the risk and costs associated with marketing and implementing the program. PG&E will only pay for measured savings.
- **Scalable:** The program is designed to leverage existing and emerging market actors to bring residential retrofits to scale. One set of potential aggregators is PACE financiers. PACE programs in California have driven over twice the volume and triple the private investments in energy efficiency projects than the Energy Upgrade California® Home Upgrade (Home Upgrade) program during a comparable time period[10]. By monetizing energy savings from PACE projects, we can entice the PACE programs to drive greater and more persistent savings, higher project volumes and increase energy savings transparency.
- **Brings buildings to or above code**: By incentivizing aggregators for gross savings, the P4P program promotes measures that bring building up to and above code. The business model of both aggregators and their contractors is to focus on equipment replacement versus repair due to the higher sale or loan price. This program will build upon this by financially motivating aggregators and contractors to identify early replacement opportunities for equipment, and to maximize the installation of heating ventilation and air conditioning (HVAC) equipment and shell measures. Additionally, this methodology promotes high quality installations through savings visibility, ensuring that customers obtain their anticipated savings.
- **Includes operational and behavioral**: The customer intervention starts with retrofit measures and then allows for unlimited operational and behavioral steps to maximize realization of measurable gross savings.

Historic Challenges with Scaling Residential Retrofits:

While PG&E has been able to generate over 14,000 upgrades since the inception of the Home Upgrade program, the program's continued growth is limited by administrator costs and lack of customer flexibility. However, in order to meet the state energy efficiency and carbon reduction

[10] California Alternative Energy and Advanced Transportation Financing Authority (CAEATFA) website http://www.treasurer.ca.gov/caeatfa/pace/activity.asp

goals, PG&E is exploring ways to scale the efficiency in existing buildings while minimizing ratepayer subsidies. PG&E's P4P program is designed to solve a number of challenges:

- **Program flexibility:** While only 17% of customers are interested in comprehensive retrofits, 66% have expressed interest in completing at least two interventions[11]. The current Home Upgrade program has reached just 0.4% of the eligible single family homes, indicating a large untapped market for incremental upgrades and savings in residential buildings. By allowing inclusion of retrofit measures and unlimited behavioral and operational interventions (BROs), this offering will allow aggregators to reach the untapped potential in both sets of motivated customers.
- **Financial assistance:** The principal reason interested homeowners cited for not completing the Home Upgrade program is the high cost of an upgrade[11]. Despite this, offerings such as PACE have 46% higher average job costs and rapidly growing participation rates. As of February 2016, PACE in California, which started about a year after the Home Upgrade program, has completed over 36,000 projects totaling more than $810M in financing[12]. While the PACE market is growing quickly, the focus has not been on delivering measurable persistent energy savings because PACE lenders make money from bundling and securitizing the loans – the more loans they sell, the more they profit. The energy savings play no direct role in their profits, only indirectly to the extent state and local policies allow PACE programs to operate based on the assumption of energy savings. This offering will create a new, direct energy savings focus for PACE and other potential aggregators by creating a new cash flow and business model for these providers; projects that deliver measurable results will become more profitable, directly aligning market incentives with energy policy goals. This Program enables price discovery, allowing Aggregators and utilities better value energy efficiency interventions and the costs associated with achieving them.
- **Capturing emerging savings opportunities:** Home Area Network (HAN) and thermostat device manufacturers have developed effective marketing channels and products and are experiencing rapid growth in the residential energy efficiency marketplace[13]. The market for connected devices such as thermostats and home energy management systems is increasing customers' energy efficiency opportunities through control, operational and behavioral savings. However, our existing residential energy efficiency programs are limited in their ability to similarly engage customers to drive further adoption or increase the savings potential of these products.
- **Incentivize only realized savings:** Studies show as many as 30% of Home Upgrade participants either have neutral or negative savings[14]. This is due to a variety of causes,

[11] SBW Consulting, Inc. 2013. 2010–2012 PG&E Whole House Retrofit Program Phase II Process Evaluation Study. San Francisco: Pacific Gas & Electric Co.
[12] California Alternative Energy and Advanced Transportation Financing Authority (CAEATFA) website http://www.treasurer.ca.gov/caeatfa/pace/activity.asp
[13] Technavio's analysts forecast the global smart grid HAN market to grow at a CAGR of 15.1% over the period 2014-2019. Accessed on 2/23/2016 at http://www.marketresearchstore.com/report/global-smart-grid-home-area-network-market-34060
[14] 2010–2012 PG&E Whole House Retrofit Program PHASE II PROCESS EVALUATION STUDY – PGE0302.04 12/31/2013

including negative behavior changes and lack of financial motivations to focus on persistent energy savings. Incentivizing aggregators for measured energy savings encourages more measure bundles that achieve significant energy savings at the meter. By allowing interventions to include operational and behavioral savings, we incentivize customers and installation contractors to participate in an on-going relationship which will lead to greater persistence of savings.

Section 3: Program Metrics

Program Goals and Objectives:

1. Allow aggregators to determine the mix of interventions that is most attractive to customers and can lead to significant energy savings beyond what is currently available in residential offerings.
2. Establish a scalable model for the residential energy efficiency market by incentivizing privately financed market actors (aggregators) to deliver measureable energy savings.
3. Determine whether this platform can increase residential energy savings at less cost to ratepayers compared to current residential energy efficiency programs.
4. Demonstrate how a simpler, more transparent method to determine savings using weather normalized meter consumption data is more effective at enticing privately financed market actors to participate in rate payer funded programs and achieve greater energy savings.
5. Monetize energy savings from residential buildings and build a foundation for a model that can successfully transition to grid-tied procurement in order to effectively respond to demand side procurement needs in the future.

Program performance metrics or non-resource objectives and success criteria (see also Program Theory and Logic Model section below for further details)**:**

Goal	Metric	Logic Model Box	Target for Initial Enrollment Period (IEP): Years 1-2	Logic Model Box	Target for Second Enrollment Period: Years 3-5
Develop Scalable Business Models	Participating customers	I	2,100 / year	O, T	Triple IEP
	Participating aggregators and contractors	B	3-5 aggregators, 50 active contractors (>5 jobs / year)	N	Triple IEP
	Non-incentive costs	N	< 20% of total costs	S	< 16% of total costs
	Total cost per home	N	< $1,500/home	S	< $1,200/home
	Savings	R	4.83 GWH, 4.7 KW, 0.945 therms	R	Triple IEP
	Competing ESPs			P	50
Data Availability and Transparency	Transparent aggregator portfolio savings	K, N	Provided quarterly		
	Monetize savings	N	Aggregators able to bid into auction by 2018-2019		

Proposed Program Timeline:

To launch the Program in 2016, PG&E proposes an accelerated timeline.

Date	Milestone	Dependencies or Potential Delays
March 2016	PG&E submits Advice Letter	
April 2016	CPUC reviews and approves Advice letter	Potential protests or request for additional information
May 2016	PG&E establishes PRG and solicitation perimeters	Interest and availability of non-monetarily interested parties
May/June 2016	PG&E opens competitive solicitation	Agreement on solicitation perimeters
July 2016	PG&E closes aggregator solicitation and selects aggregators	Multiple aggregator proposals received
August 2016	PG&E and aggregators sign PSAs	Agreement on PSA terms
Sept. 2016	Customer enrollment period begins	Completion of aggregator enrollment

Program Performance:
The main goal of this effort is to build a more scalable and cost effective intervention method for addressing residential retrofits. We anticipate achieving equivalent or higher savings than Home Upgrade program with less ratepayer funded costs. However, due to the high incremental measure cost (total project cost) of these types of activities, PG&E anticipates that the total resource cost (TRC) will continue to be on the same order of magnitude as PG&E's 2015 Home Upgrade Sub-Program TRC of 0.34[15]. PG&E believes the more appropriate cost-effectiveness test for determining the success of this Program to be the program administrator cost (PAC) test. The PAC test, which we predict to be greater than 2.0, will be used to measure the success of the IEP.

As noted in the HOPPs Ruling[16], the full measure cost is used to determine the cost effectiveness of measures when using an existing condition baseline. For this program we propose adjusting the baseline relative to a comparison group (see Attachment B). When energy savings baseline adjustments are made, corresponding cost adjustments should also be applied to maintain the integrity of cost effectiveness calculations.

Section 4: Program Budget, Incentives and Energy Savings

Program Budget:

	2016	2017	2018	2019	Total
Administration	$60,000	$60,000	$60,000	$60,000	$240,000
Direct Implementation	$140,000	$140,000	$140,000	$140,000	$560,000
Incentive	$0	1,250,000	$2,000,000	$1,750,000	$5,000,000
Savings Measurement	$50,000	$50,000	$50,000	$50,000	$200,000
Total Initial Enrollment Period Budget	**$250,000**	**$3,500,000**	**$1,250,000**	**$2,000,000**	**$6,000,000**

Savings Targets:

	2016	2017	2018	2019	Total
Electric Savings (GWH)	0	0.966	1.932	1.932	4.83
Demand Reduction (MW)	0	0.945	1.89	1.89	4.725

[15] This represents initial calculations which may differ slightly than official annual reporting figures.
[16] December 30, 2015 "Assigned Commissioner and Administrative Law Judge's Ruling Regarding High Opportunity Energy Efficiency Programs and Projects" (ALJ Ruling). Page 13.

Gas Savings (MM Therms)	0	0.189	0.378	0.378	0.945

Savings and Budget Assumptions:
Savings targets are 4.83 GWH, 4.725 MW and 0.945 therms. These were set based on results from the 2010-2012 Impact Evaluation which determined that gross savings for PG&E's Advanced Home Upgrade (AHU) program were 460 KWH, 0.45KW and 90 therms[17]. This would be equivalent to 6% electric savings and 16% gas savings per home. While PG&E believes this program will have a different savings profile due to the inclusion of operational and behavioral savings, it is the best point of reference given the similar program goal to advance residential retrofits.

In the 2013-2015 program cycle, the PG&E AHU program averaged between 1,800-2,100 upgrades annually. Given the 2 year enrollment period, PG&E estimates 4,200 customer enrollments for the P4P program. PG&E will pay for portfolio savings annually for 2 years. PG&E proposes to continue to measure and claim the savings for at least another 1-3 years to better inform future program design and properly value measures with a longer useful life.

Year 1:	Aggregators enroll new customers
Year 2:	Aggregators enroll new customers; PG&E pays aggregators for portfolio savings of all year 1 participants; PG&E claims net savings from year 1 participants
Year 3:	PG&E pays aggregators for portfolio savings of all year 1 and year 2 participants; PG&E claims net savings from year 1 and year 2 participants
Year 4:	PG&E pays aggregators for portfolio savings of all year 2 participants; PG&E claims net savings from year 1 and year 2 participants
Year 5:	PG&E claims net savings from year 1 and year 2 participants; PG&E uses year 5 savings persistence data to define the actual claimable persistent savings period

We will price the aggregator incentives in a manner to promote measures with longer useful life, such as HVAC equipment and shell measures, during the IEP. To capture the savings associated with measures with a longer useful life, we will be measuring savings 1 year after aggregator payments have ceased. Since aggregator interventions will cease after incentives are no longer available, this data point will be used to define the persistent savings for the 15 year assumed useful life of the intervention.

Incentive Design:
PG&E will pay each aggregator a set rate per therm and kWh annually one and two years after the initial intervention for their portfolio savings; no upfront payments will be made. PG&E will make aggregator incentive payments in two forms (1) $/kwh and $/per therm for gross savings and (2) an additional 5-10% will be provided for net savings measured through PG&E's

---

[17] 2010-2012 Whole House Retrofit Impact Study (2014), DNV GL - Energy. http://www.calmac.org/publications/CPUC_WO46_Final_Report.pdf

evaluation methodology. The program will not pay for demand savings (KW) in the initial enrollment period, but PG&E will consider how to integrate incentives for demand reduction in future periods.

PG&E proposes paying $0.80/kwh and $1.80/per therm for gross savings for the IEP. However, PG&E would like to conduct further stakeholder engagement to confirm the appropriateness of this rate. Should this rate change before the program begins, PG&E will submit a PIP Addendum or Advice Letter as appropriate. This incentive rate is based roughly on realized average net savings claims identified in the 2010-2012 Impact Evaluation, as well as the current rebates offered through AHU. However, we have prioritized electric savings to further target the inclusion of plug loads and operational and behavioral savings.

A principal goal of the IEP is to determine the ideal price to pay for savings in this model. We believe these rates are protective of the ratepayer and significant enough to encourage leading market actors into the program. As mentioned above, incentive rates will be re-evaluated and adjusted in future offerings, following the price discovery learnings in this period.

Payable Savings:
The move to pay for performance presents a new challenge. Aggregators need a savings calculation method that is known upfront and stable throughout the program period, or at least replicable by them using an established, transparent and easily replicable methodology. Otherwise, the risk for them to make the needed upfront investments is too great, limiting their participation and jeopardizing the desired scalability of the program.

PG&E will use this initial enrollment period to test and establish the best method to determine gross savings. During the IEP, we propose to pay aggregators based on a simple pre/post weather normalized savings method. This simple method will be used to determine the payments for 90+% of aggregator portfolio savings. An additional 5-10% will be paid for net savings calculated according to the methods described in Appendix B.

To help PG&E select the best method for determining gross savings during the IEP, we propose that the CalTRACK technical workgroup test the better known and canonical methods proposed in Appendix B as well as newer methods that take advantage of the precision and timeliness of Advanced Meter Infrastructure (AMI) data. Our findings will determine gross savings for the purpose of calculating net claimable savings in the IEP and will determine how we calculate gross savings for the purpose of calculating payable savings under future offerings, provided it is stable, transparent and easily replicable by aggregators.

Claimable Savings:
PG&E’s savings claims will be determined by the estimated net energy savings that are attributable to the program. Estimated net energy savings result in the best estimate of the incremental benefit of the program and are is used in benefit/cost calculations such as the Total Resource Cost (TRC) Test.

We examined five options for estimating net savings. Based on an analysis presented in the EM&V document in Appendix B, PG&E recommends that the following three options be used for estimating the net savings. Note that using more than one method for estimating net savings is consistent with the enhanced level of rigor specified in the California Energy Efficiency Evaluation Protocols.

1. Quasi-Experimental Design. This design, the non-equivalent comparison group design, is an alternative to a randomized control trial. This method effectively adjusts the energy savings baseline and further normalizes the savings estimate for factors beyond weather.
2. Self-Report Approach. This approach involves the estimation of gross impacts that is adjusted using a self-reported net-to-gross ratio (NTGR).
3. P4P Versus PACE Loan Program (or similar). We also considered the use of the non-equivalent comparison group design to test the hypothesis that P4P, by allowing aggregators to determine the mix of interventions that is most attractive to customers, including behavioral, operational and retrofit activities and paying them based on verified energy savings, can lead to significant energy savings above the existing PACE Loan Program (or similar).

In Appendix B, PG&E provides details on all five options considered to ensure that our methodological choices are transparent. Please note that the details of this EM&V plan cannot be finalized until the Energy Division approves both the general approach and the EM&V budget.

Section 5: Program Theory and Logic Model

Logic models go hand-in-hand with program theory; Rosenberg and Hoefgen state: "program logic models are graphic representations of the causal links between program activities, short-term responses to those activities among market actors, and longer-term market effects"[18]. The elements used to describe or represent a logic model include inputs, activities, and outputs, which in combination loosely form a program process theory, short-term outcomes (sometimes called initial, proximal, or immediate outcomes), mid-term outcomes (sometimes called intermediate or proximal outcomes), and long-term outcomes (sometimes called distal outcomes or impacts), which are intended to represent a program impact theory[19,20,21,22]. In these logic models, activities are the actions undertaken to bring about a desired end, outputs are the immediate results of an action, and outcomes are the anticipated changes that occur directly or indirectly as a result of inputs, activities, and outputs.

The P4P concept is built upon a series of hypothesized causal linkages between program activities, outputs, and intended program outcomes that are depicted in the program logic model as illustrated in Figure 1**Error! Reference source not found.**. The development of this logic model is based on two sources of information:

---

[18] Rosenberg, M. and L. Hoefgen. 2009. *Market Effects and Market Transformation: Their Role in Energy Efficiency Program Design and Evaluation*. Prepared for the California Institute for Energy and Environment. p. 48. Available at: http://uc-ciee.org/downloads/mrkt_effts_wp.pdf
[19] Donaldson, S. I. 2007. *Program Theory-Driven Evaluation Science*. New York, NY: Lawrence Erlbaum
[20] Donaldson, S. I., & Lipsey, M. W. 2006. "Roles for Theory in Contemporary Evaluation Practice: Developing Practical Knowledge." *In:* I. Shaw, J. C. Greene, & M. M. Mark (Eds.), *The Handbook of Evaluation: Policies, Programs, and Practices* (pp. 56-75). London, UK: Sage.
[21] Lipsey, M. W., Rossi, P. H., & Freeman, H. E. 2004. *Evaluation: A Systematic Approach* (7th ed.). Thousand Oaks, CA: Sage.
[22] Patton, M. Q. 2008. Utilization-Focused Evaluation (4th ed.). Thousand Oaks, CA: Sage.

- Prior theory and research (e.g., consumer behavior, evaluations of similar programs, etc.)
- Implicit theories of those close to the program (e.g., PG&E program managers, experience of CPUC-ED and its consultants, and PG&E EM&V staff and its consultants),

Importantly, the P4P concept is one of the first programs of its type aimed at longer-term market transformation in the State of California and beyond through an intervention strategy with a midstream emphasis. As a result, outcomes of the program are expected to occur over different time frames involving different market actors.

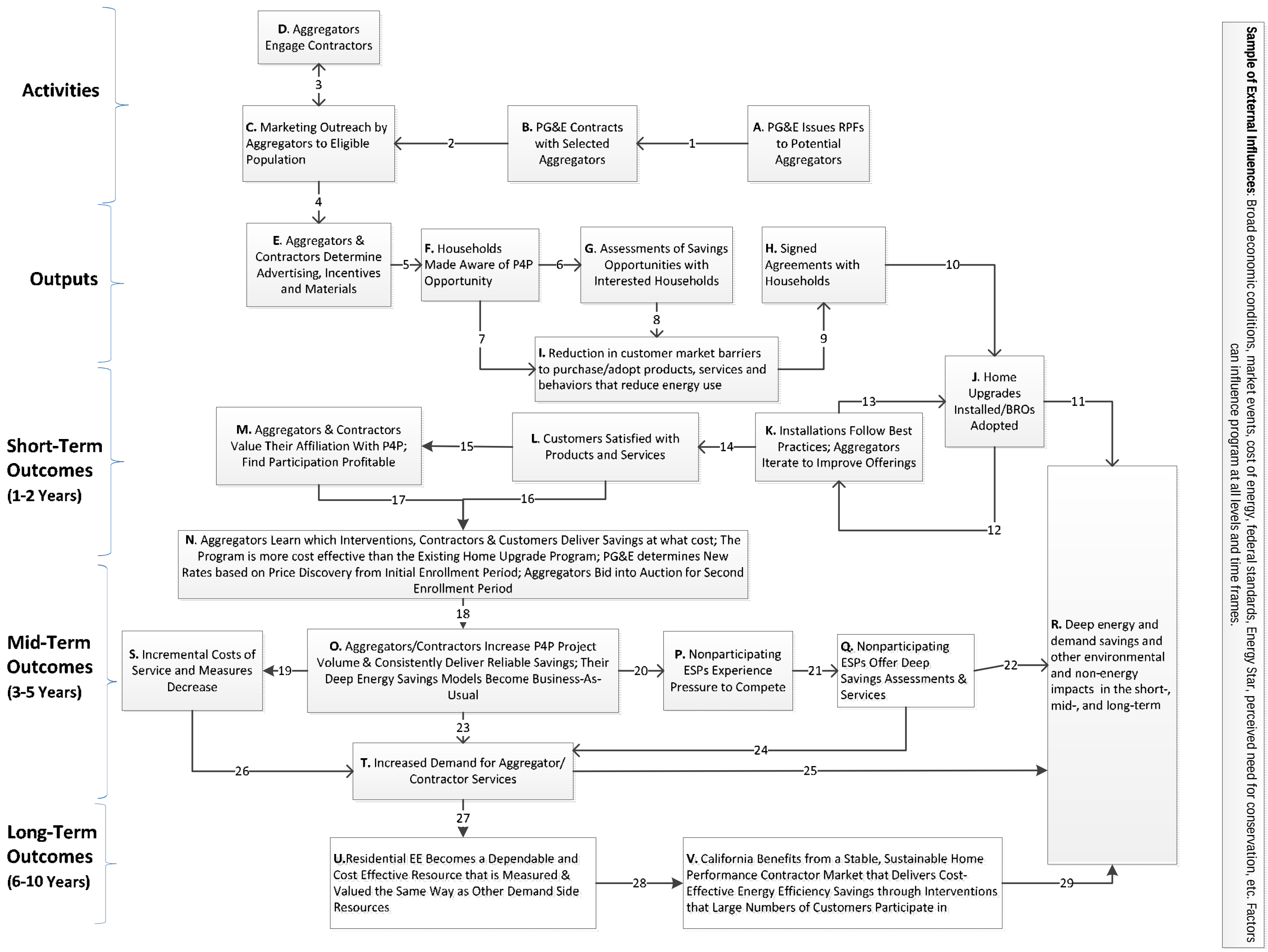
P4P Logic Model Description:
As depicted in Figure 1, initial program activities are aimed at issuing an RFP to potential aggregators **(A)**, signing contracts with selected aggregators **(B)**, aggregators conduct marketing and outreach to eligible population and **(C)** and aggregators engage with contractors **(D)**. Next, aggregators and contractors determine the advertising, incentives and materials needed to recruit and educate customers **(E)** and a subset of the eligible population are made aware of the P4P opportunity **(F)**. As a result, a subset of the aware households agrees to an energy savings assessment **(G)**. As a result of **F** and **G**, barriers to purchasing/adopting energy efficient products, services and behaviors are reduced **(I)** leading to signed agreements with interested households **(H)**. Next, home upgrades are installed and BROs are adopted **(J)** leading to deep energy and demand savings and other environmental and non-energy impacts in the short-term **(R)**.

Aggregators work with contractors to make sure that their work conforms to best practices and adjustments are made to customer offerings, as needed to ensure energy savings are being realized based on feedback from the CalTRACK portfolio tracking dashboard and household level analytics that aggregators are conducting on their own **(K)**. As a result, customers are satisfied with products and services **(L)** which leads to aggregators and contractors valuing their affiliation with P4P **(M)**.

Over time, the aggregators learn which interventions, contractors and customers deliver savings at what cost; PG&E determines new rates based on price discovery from initial enrollment period and uses other insights gleaned from process evaluations to develop a refreshed second enrollment period program design; aggregators bid their $/kWH and $/therm savings into an auction for second enrollment period **(N)**. As a result, aggregators/contractors increase P4P project volume & consistently deliver reliable savings; their deep energy savings models become business-as-usual **(O)**. This increased volume decreases the incremental costs of services and measures **(S)**. The increased volume **(O)** and decreased costs **(S)** leading to increased demand for aggregator/contractor services **(T)** leading to deep energy and demand savings and other environmental and non-energy impacts in the mid-term **(R)**. Nonparticipating energy service providers (ESPs) experience pressure to compete with P4P aggregators **(P)** eventually motivating them to also offer deep energy savings assessments and services **(Q)** leading to deep energy and demand savings and other environmental and non-energy impacts in the mid-term **(R)**.

The increased demand in the mid-term **(T)** leads to residential energy efficiency becoming a more dependable and cost-effective resource that is measured and valued the same way as other demand-side resources **(U)**. As a result, California benefits from a stable, sustainable home performance contractor market that delivers cost-effective energy efficiency savings through interventions that large numbers of customers participate in **(V)** leading to deep energy and demand savings and other environmental and non-energy impacts in the long-term **(R)**.

Figure 1: P4P Logic Model



Attachment B:
Evaluation, Measurement and Verification Plan for the PG&E Residential Pay-for-Performance Program: Claimed Savings

Prepared by the Pacific Gas and Electric Company

March 25, 2016

Contents

1. Evaluation, Measurement and Verification (EM&V) Overview

The EM&V plan for the Residential Pay-for-Performance Program ("P4P") presented in this document is the result of substantial collaboration with professionals with years of experience in the field of energy efficiency program evaluation. It has been reviewed by program administrators and other stakeholders. We believe that the result is a workable plan that balances the competing desires for accurate estimates of net program savings, generalizability of the results and efficient program administration. For now, PG&E recommends that this EM&V plan serve as a general framework until a detailed EM&V plan can be developed based on the types of customers who actually join P4P and the measures and practices they adopt. PG&E also recognizes that this general EM&V framework, the more detailed EM&V plan, and an evaluation budget must be approved by the ED. We note that the details of this EM&V plan cannot be finalized until the ED approves both the general approach and the EM&V budget.

AB802 provides for the simple estimation of savings based on the difference in normalized annual consumption from the pre to the post period. For this HOPPs program, PG&E plans to *claim estimated net energy savings.*[1] Estimated net energy savings results in the best estimate of the *incremental* benefit of the Program and is used in benefit/cost calculations such as the Total Resource Cost (TRC) Test. Moreover, estimated net energy savings is the preferred basis for assessing whether program administrators have met their energy savings goals which are a key input in the calculation of utility earnings for the administration of energy efficiency programs.

We examined five methods for estimating net savings and propose the following three:

1) **Quasi-Experimental Design.** This design, the non-equivalent comparison group design, is an alternative to a randomized control trial.
2) **Self-Report Approach.** This approach involves the estimation of gross impacts based on the pre-to-post difference in normalized annual consumption that is adjusted using a self-reported net-to-gross ratio (NTGR).
3) **P4P Versus PACE Loan Program**. This approach uses a second non-equivalent comparison group design to test the hypothesis that P4P, by allowing aggregators to determine the mix of interventions that is most attractive to customers (including behavioral, operational and retrofit activities) and paying aggregators based on energy savings, can lead to significant energy savings greater than those obtained by the existing PACE Loan Program.

We recommend both the first and second options for estimating the net savings that PG&E will claim. Note that using more than one method for estimating net savings is consistent with the *enhanced* level of rigor specified in the *California Energy Efficiency Evaluation Protocols: Technical, Methodological, and Reporting Requirements for Evaluation Professionals* (The

[1] Net savings are defined as "The total change in load that is attributable to the utility DSM program. This change in load may include, implicitly or explicitly, the effects of free-drivers, free-riders, state or federal energy efficiency standards, changes in the level of energy service and natural change effects" (*California Energy Efficiency Evaluation Protocols: Technical, Methodological, and Reporting Requirements for Evaluation Professionals*. The TecMarket Works Team, 2006, pp. 233-234).

TecMarket Works Team, 2006, p. 36). The third option will assess the relative effectiveness of the P4P versus the PACE Loan Program and may be used to inform future program offerings.

In the sections that follow, we describe the three recommended methods and assess the advantages and disadvantages of each. Subsequently, we discuss the other methods that we considered but ultimately rejected in order to make the reasons for our methodological choices transparent.

2. Method 1: Quasi-experimental Design

We recommend a quasi-experimental design to estimate net impacts. Of the many quasi-experimental designs, the one we examined is the *non-equivalent comparison group design* (Shadish, Cook and Campbell, 2002). The difference between a non-equivalent comparison group design and a randomized encouragement design is that, in a quasi-experimental design, eligible customers are *not* randomly assigned to the treatment and comparison groups (i.e., a *participant* is any eligible customer who was encouraged by a P4P aggregator and a *nonparticipant* is any member of the eligible population who was not encouraged by a P4P aggregator). That is, the aggregators decide, on some nonrandom basis, which customers they will target. Members of the treatment group are defined as eligible households that were exposed through various means to the P4P opportunity by aggregators. Members of the comparison group are defined as eligible households that were not exposed in any way to the P4P opportunity by aggregators. A subset of those exposed to the P4P opportunity will self-select into the P4P. The counterfactual inference depends on a non-equivalent comparison group deliberately chosen to have maximum pretest similarity to the treatment group on as many observed characteristics as possible (e.g., climate zone, size of home, age of home, prior energy usage) and on other particular features that the researcher believes will be particularly salient threats to internal validity.

2.1. Major Threat to Internal Validity

In a non-equivalent comparison group design, the main threat to internal validity[2] is self-selection bias. Self-selection bias occurs when groups exposed to treatments non-randomly may differ in ways that mimic what the treatment might achieve (Shadish, Cook and Leviton, 1991). There is considerable evidence that nonrandom assignment often (*but not always*) yields different results than random assignment does (Chalmers et al., 1983; Colditz, Miller and Mosteller, 1988; Lipsey and Wilson, 1993; Mosteller, Gilbert and McPeek, 1980; Wortman, 1992), more so when participants self-select into conditions than when others make the decision (Heinsman and Shadish, 1996; Shadish, Matt, Navarro and Phillips, 2000; Shadish and Ragsdale, 1996) – so self-selection should be avoided if possible. But if such a situation cannot be avoided for reasons such as those given in Section 2, econometricians and statisticians over the years have also devoted an enormous amount of effort to developing strategies to mitigate self-selection.

We digress here to note that self-selection has been given considerable attention over the last 30 years in the evaluation of energy efficiency programs in California due to the fact that, with the

[2] Internal validity is the basic minimum without which any experiment is uninterpretable: did in fact the experimental treatments make a difference in this specific experimental instance? (Campbell and Stanley, 1963, p.5)

exception of evaluations of neighbor comparison ("Opower")-type programs, nearly all the evaluations that relied on billing analysis to estimate net impacts have been based on quasi-experimental designs. Since the early 1980s, most billing analyses aimed at estimating net savings used some form of analysis of covariance (Huitema, 2011) to control for the *observed* differences between the treatment and comparison groups. Efforts to address the biasing effects of *unobserved* differences using inverse Mills ratios began at least as early as the late 1980s. Since then, Train (1993) and Goldberg and Train (1995), using simulated datasets, demonstrated that failing to correct for self-selection can overestimate net savings, but that there are effective strategies to reduce this bias substantially. Finally, the use of quasi-experimental designs has been allowed in both sets of California EM&V protocols (PG&E et al., 1996; The TecMarket Works Team, 2006) as long as evaluators made methodologically-sound efforts to address self-selection.

Below, based on a relatively limited review of the literature, we provide a series of strategies to improve internal validity primarily by addressing self-selection. Before finalizing this evaluation plan, PG&E proposes that a more comprehensive review of the more recent literature regarding strategies for addressing self-selection should be conducted.

2.2. Strategies to Mitigate Self-Selection

There are number of strategies to strengthen this quasi-experimental approach by mitigating self-selection bias and bring the results closer to results that would be produced by a true experimental design. Below, we recommend a number of these strategies that should be considered in evaluating the P4P evaluation method.

Use of Internal Controls. Assignment can often be controlled in other ways than by random methods. Nonrandom comparisons to an internal rather than external control can sometimes yield more accurate results (Aiken et al., 1998; Bell et al., 1995; Heinsman and Shadish. 1996; Shadish and Ragsdale, 1996). Internal controls are drawn from the same pool of participants (i.e., from students in the same school or class or from all program applicants). External controls are drawn from patently different pools (e.g., patients in different treatment settings) and are presumed to have less in common. Drawing on members of the P4P-eligible population will serve as our internal controls.

Joint Use of a Pretest and a Comparison Group. The joint use of a pretest and a comparison group makes it easier to examine certain threats to validity. Because the groups are nonequivalent by definition, selection bias is presumed to be present. The pretest allows exploration of the possible size and direction of that bias. For example, we will match treatment and comparison group households on historical monthly kWh consumption. Note that while adding a pretest to a design helps assess selection biases and attrition as sources of observed effects, adding repeated pretests of the same construct on consecutive occasions prior to treatment helps reveal maturational trends and detect regression artifacts. However, the extent to which the pretest can render self-selection implausible depends on the size of any selection bias and the role of any *unmeasured* variables that cause selection and are correlated with the outcome. The absence of pretest differences in a quasi-experiment is never proof that selection bias is absent.

Modelling Approaches. As noted earlier, attempting to correct for self-selection bias is essential in any observational study. To the extent that the differences between the two groups can be observed, variables that represent those differences can be addressed by first by using internal controls to form a comparison group and then matching the two groups on an observed characteristic. For example, we could match treatment and comparison group households on monthly kWh consumption. But matching on a single variable such as pre-monthly kWh consumption is no guarantee the selection bias has been adequately addressed (Shadish, Cook and Campbell, 2002) since there might be more than one variable that plays a role in explaining why households chose to self-select into the program. In such a case, treatment and comparison group households can be matched on propensity scores, the predicted probability of being in the treatment (versus comparison) group from a logistic regression equation. The logistic regression reduces each household's set of covariates to a single propensity score, thus making it feasible to match or stratify on what are essentially multiple variables simultaneously. Another approach is to enter the propensity score as an additional covariate into the regression model. Of course, the most difficult issue to address is the differences between participants and non-participants that are unobserved and unobservable. To mitigate both overt and hidden bias, a variety of approaches that attempt to take advantage of recent developments in statistics and econometrics will be explored:

1. Sample selection models (e.g., Heckman's two-step estimator (1978, 1979); treatment effect model (Green, 2003); instrumental variables estimator (Wooldridge, 2002)
2. The propensity score matching model (Rosenbaum and Rubin, 1983, 1985; Hansen and Klopfer, 2006; Guo and Fraser, 2014)[3]
3. Matching estimators and synthetic controls (Abadie and Imbens (2002, 2006)
4. Propensity score analysis with nonparametric regression (Heckman et al., 1997, 1998)

Other Strategies. In addition, the very nature of billing analysis allows evaluators to avoid a host of other problems that plague any experiment. Two of these are listed below.

- ***Clearly Defined Post Period***. The major reason for assessing any posttest after the treatment is to eliminate the ambiguity about the temporal precedence of cause and effect. In conducting a billing analysis, we have participation dates and create dead bands around these participation dates to clearly separate the pre from the post period, i.e., the monthly post kWh measurements clearly comes after the treatment.

- ***Lack of Reactivity***. The very nature of measuring kWh consumption using electricity meters means that customers cannot react to the fact that they (i.e., their households) are being measured (Rosnow and Rosenthal, 1997).

Again, before finalizing this evaluation plan, PG&E proposes that a more comprehensive review of the more recent literature regarding strategies for addressing self-selection should be conducted.

---

[3] Note that propensity scores cannot remove hidden biases except to the extent that unmeasured variables are correlated with the measured covariates used to compute the propensity score

2.3. The Regression Model

To estimate net savings, a pooled, fixed-effects, time-series/cross-sectional (panel) regression model that incorporates the treatment and comparison groups could be estimated. The treatment and comparison groups would be matched on key variables such as consumption level (not just overall, but month-by-month similarity), available customer demographics (especially income and education), dwelling unit type, geography (ZIP code, if feasible), and energy end uses. Any *observed* differences in the composition of the treatment and comparison groups can be controlled statistically[4]. Equation 1 illustrates one possible specification.

$$ADC_{it} = \alpha_i + \delta_m + \beta_1 Post_t + \beta_2 Treatment_i \cdot Post_t + \beta_3 HDD_{it} + \beta_4 CDD_{it} + \sum \beta_k X_i + \varepsilon_{it} \quad (1)$$

Where:

ADC_{it}= Average daily consumption (kWh or therms) for household *i* at time *t*

α_i= Household-specific intercept

δ_m = 0/1 Indicator for each time interval *m*, time series component that track systematic change over time

β_1= Coefficient for the change in consumption between pre and post periods

β_2= Coefficient for the change in consumption for the treatment group in the post period compared to the pre period and to the control group. This is the basis for the net savings estimate.

β_3= Coefficient for HDD

β_4= Coefficient for CDD

Post = dummy variable for pre (Post=0) and post (Post=1)

Treatment = dummy variable for treatment (Treatment=1) and control (Treatment=0)

HDD_{it}= Sum of heating degree-days (e.g., base 65 degrees Fahrenheit)

CDD_{it}= Sum of cooling degree-days (e.g., base 75 degrees Fahrenheit)

β_k = A vector of k coefficients that reflect the energy change associated with a one unit change in the k^{th} explanatory variable

X_i = A vector of explanatory variables (i.e., covariates), such as changes in occupancy or square footage, for the i^{th} factor

ε_{it} = Error

To obtain the final estimate of net savings per participant, the coefficient β_2 is then multiplied by the total number of P4P participants and divided by the participation "take rate" among the

[4] Inverse Mills ratios interacted with δ_m will also be explored as a way to control for *unobserved* differences between the treatment and comparison groups.

treated or encouraged. The final specification of this model will depend on which strategies for addressing self-selection are used and the availability of various covariates. These savings would be tracked over a three-year period.

In addition to net energy savings, the P4P will also produce net peak demand reductions that will be claimed by the IOUs. Consistent with DEER, we will determine the electric demand impacts of measures using the average kWh reduction over a 9-hour window. The nine-hour window is from 2 p.m. to 5 p.m. over a three-day "heat wave" that is determined for each climate zone. The three-day demand periods for the new weather data were chosen based on these criteria:

- Occurs between June 1st and September 30th,
- Does not include weekend days or holidays (based on 2009),
- Has the highest value for
  - average temperature over the three-day period +
  - the average temperature from noon to 6 p.m. over the three-day period +
  - the peak temperature over the three-day period.

The treatment group and the comparison group will be compared with respect to kW demand during peak periods based on AMI data. A regression model similar to Equation 1 will be specified.

We will also explore the use of a simpler method described in Equation 2:

$$Net\ Peak\ Demand\ Reduction = \frac{CF \times kWh_{net}}{Hours_{Peak}} \qquad (2)$$

where

CF= the fraction of the peak demand of a population that is in operation at the time of system peak[5].

kWh_{net} = the average net kWh savings per household

$Hours_{Peak}$= The number of hours in the summer on peak period

3. Method 2: Self-Report Approach to Estimating Net Savings

There will always be some uncertainty surrounding the estimates of net savings derived from the application of any individual methodology. One way to reduce uncertainty would be to obtain a second estimate of net savings by multiplying the estimated gross savings[6] by a net-to-gross ratio (NTGR) estimated using the self-report approach (SRA). The method for estimating gross savings is discussed below, followed by a discussion of the self-report method for estimating the

[5] Northeast Energy Efficiency Partnerships (NEEP) defines it as, "The ratio of the average hourly demand during a specified period of time of a group of electrical appliances or consumers to the sum of their individual maximum demands (or connected loads) within the same period." (NEEP 2011).

[6] Gross savings defined as "The change in energy consumption and/or demand that results directly from program-related actions taken by participants in the DSM program, regardless of why they participated" (The TecMarket Works Team, 2006, p. 227).

NTGR. Based on findings from Method 1, PG&E may elect to pursue Method 2 to verify or augment the validity of the results.

3.1. Proposed Methods for Estimating Gross Savings

Since the intervention will result in savings over time, gross savings must also be tracked over time. To track gross savings over time, two approaches will be used, one for the short-term (i.e., first year) and one for the longer-term (multi-year). The short-term method uses future participants as a comparison group since it is a more rigorous approach for controlling for exogenous changes (such as self-selection), under certain assumptions to be discussed later. For example, the comparison group for the first program year will be the participants in the second program year. However, since we are also interested in estimating gross savings for more than one year, the comparison group composed of participants in the second program year would no longer be useful for successive years. Instead, we propose that the participants in the third program year be used as the comparison group in estimating the savings in the second year of the program, and so on. This "rolling comparison group" design would work so long as the program design remains relatively stable over time and provided that there is an adequate amount of pre-program consumption data (at least one year) available for future participants to serve as points of comparison. If either condition (that is, a stable program design and sufficient historic data from the comparison group) is not met, then another approach will be used to identify a comparison group composed of a random sample of eligible households. Both the short-term and long-term methods are described below. Note that the methods described in the following sections would be used to estimate savings for each participant cohort for each year of their participation.

3.1.1. Estimation of Short-Term Gross Savings

Gross savings will be estimated in a manner consistent with AB802 and IPMVP[7] Option C which allow for an existing conditions baseline in estimating gross savings. The method recommended is based on the two-stage approach described in Chapter 8 of the Uniform Methods Project[8].

Stage 1. Individual Premise Analysis

A third-party selected by PG&E and approved by Energy Division will perform the following activities:

1. Fit a premise-specific degree-day regression model (as described in Step 1, below) separately for the pre- and post-periods.
2. For each period (pre- and post-) use the coefficients of the fitted model with normal-year degree days to calculate the normalized annual consumption (NAC) (defined below) for that period.
3. Calculate the difference between the pre- and post-period NAC for the premise (i.e., ΔNAC).

[7] International Performance Measurement and Verification Protocol (IPMVP) available from the Efficiency Valuation Organization at http://evo-world.org/en/
[8] Uniform Methods Project: Methods for Determining Energy Efficiency Savings for Specific Measures (http://energy.gov/eere/about-us/ump-protocols)

Step 1. Fit the Basic Stage 1 Model

$$E_m = \mu + \beta_H H_m + \beta_C C_m + \varepsilon_m \quad (3)$$

$E_m =$ Average consumption per day during interval m

$H_m =$ Specifically, $H_m(\tau H)$, average daily heating degree days at the base temperature(τH) during meter read interval m, based on daily average temperatures on those dates

$C_m =$ Specifically, $C_m(\tau C)$, average daily cooling degree days at the base temperature(τC) during meter read interval m, based on daily average temperatures on those dates

$\mu =$ Average daily baseload consumption estimated by the regression

$\beta_{H,}\beta_C =$ Heating and cooling coefficients estimated by the regression

$\varepsilon_m =$ Regression residual.

Step 2. Apply the Stage 1 Model
To calculate NAC for the pre- and post-installation periods for each premise and timeframe, we combine the estimated coefficients μ, β_H, and β_C with the annual normal-year or typical meteorological year (TMY) degree days H_0 and C_0 calculated at the site-specific degree-day base(s), τH and τC. Thus, for each pre- and post-period at each individual site, we use the coefficients from Equation 7 for that site and period to calculate the weather-normalized annual consumption (NAC) (see Equation 4). This example puts all premises and periods on an annual and normalized basis.

$$NAC = \mu * 365 + \beta_H H_0 + \beta_C C_0 \quad (4)$$

The same approach can be used to put all premises on a monthly basis and/or on an actual weather basis.

Step 3. Calculate the Change in NAC
For each site, the difference between pre- and post-program NAC values (ΔNAC) represents the change in consumption under normal weather conditions. For *future participants* who are used as a comparison group to current participants, these same three steps are followed.

Stage 2. Cross-Sectional Analysis
Next, the cross-sectional model in Equation 5 is estimated incorporating both current and future participants.

$$\Delta NAC_j = \beta + \gamma I_j + \varepsilon_j \quad (5)$$

$I_j =$ 0/1 dummy variable, equal to 1 if customer j is a (current-year) participant, 0 if customer j is in the comparison group composed of future year participants.

$\beta, \gamma =$ Coefficients determined by the regression model

$\varepsilon_j =$ Regression residual.

From the fitted equation:

- The estimated coefficient γ is the estimate of mean savings.
- The estimated coefficient β is the estimate of mean change or trend unrelated to the program.

The coefficient β corresponds to the average change among the comparison group, while the coefficient γ is the difference between the comparison group change and the participant group change. That is, this regression is essentially a difference-of-differences formulation and can be accomplished outside of a regression framework as a difference of the two mean differences. More complex models that include other available premise characteristics can be included that can improve the extrapolation of the billing analysis to the full population. Total P4P first-year annual savings are calculated by multiplying the difference between the comparison group change and the participant group change by the number of participating households.

There are two approaches to using future participants as a comparison group, the *full-year approach* and the *rolling specification*. Chapter 8 of the UMP observes that, although using the full-year comparison group specification is simple, it requires data from farther back in time. The rolling specification, however, allows data from a more-compressed timeframe to be used, as it uses a rolling pre- and/or post-period across the current program year. We will explore both approaches.

There is one important concern about the use of future participants as a comparison group. The implicit baseline is that, absent the intervention, the future participants did nothing (either measures or behaviors) that would affect their energy use substantially in the period of time that the current participants received the P4P. However, it is possible that *some* of the future participants might have installed a number of other measures or engaged in other behaviors that affected their energy use during the same period of time as the current participants, resulting in a lower estimate of gross savings (that is, an estimate of savings that is somewhere between gross and net). To address this possibility, we propose conducting telephone interviews with 70 future P4P participants to determine the extent to which the future participants might have installed a number of other measures or engaged in other behaviors during the same period of time as the current participants. The results of these interviews will aid in the interpretation of the results as being gross estimates, or as being estimates that lie somewhere between gross and net, and perhaps suggest an approach to adjusting the estimates upward to make them better represent gross savings.[9]

Once savings and uncertainty bands are estimated for the population of treated sites, sites will be grouped according to several factors and analyzed in aggregate. Factors by which accounts are grouped and analyzed include, but are not limited to:

- Project aggregator
- Measures included

[9] PG&E notes that the use of future participants in a comparison group as a way to control for exogenous changes in the estimation of gross savings will actually result in an estimate of net savings for low income programs. The rationale for this conclusion is that, not only are future participants in low income programs extremely unlikely to have the financial means to adopt measures during the evaluation period, but members of the general eligible low income population are equally unlikely to as well.

- Weather station
- Typical annual usage

As we noted earlier, since we are interested in estimating gross savings for more than one year, the first year's comparison group (composed of participants in the second program year) would no longer useful for the second year. Instead, we propose that the participants from the third program year be used in estimating the savings from the second year of the program, and so on. This process could work so long as the program remains stable over time and provided that the future participants have an adequate amount of pre-program consumption data to serve as points of comparison. If both conditions are not met, a pooled fixed-effects approach, discussed next in Section 3.1.2, will be used.

3.1.2. Estimation of Long-Term Gross Impacts

The longer-term method will involve a pooled fixed-effects approach. The pooled approach addresses exogenous change without the inclusion of a separate comparison group. In this model, participants who received a measure installation during a certain time interval serve as a steady-state comparison for other participants in each other time interval. Almost all observations include premises that are still in their pre-installation period ***and*** premises that are in their post-installation period, so the effect of post- versus pre- is estimated to control for exogenous trends.

The basic structures of the site-level and the second-stage consumption data model are effectively combined in the pooled approach. All monthly participant consumption data (both pre- and post-installation) are included in a single model. This model has:

- A site-level fixed-effect component (analogous to the site-level baseload component) and average overall heating and cooling components
- A post-installation indicator variable capturing the change in the post-installation period.

The recommended pooled model is illustrated in Equation 6.

$$ADC_{it} = \alpha_i + \delta_m + \beta_1 Post_t + \beta_2 HDD_{it} + \beta_3 CDD_{it} + \sum \beta_k X_i + \varepsilon_{it} \quad (6)$$

Where:

ADC_{it}= Average daily consumption (kWh or therms) for household i at time t

α_i= Household-specific intercept

δ_m = 0/1 Indicator for each time interval *m*, time series component that track systematic change over time

β_1= Coefficient for the change in consumption between pre and post periods

β_2= Coefficient for HDD

β_3= Coefficient for CDD

Post = dummy variable for pre (Post=0) and post (Post=1) participation in P4P

HDD_{it}= Sum of heating degree-days (e.g., base 65 degrees Fahrenheit)

CDD_{it}= Sum of cooling degree-days (e.g., base 75 degrees Fahrenheit)

β_k = A vector of k coefficients that reflect the energy change associated with a one unit change in the kth explanatory variable

X_i = A vector of explanatory variables (i.e., covariates), such as changes in occupancy or square footage, for the ith factor

ε_{it} = Error

An additional set of variables will be included to explain variation in consumption over time for reasons other than the central installation variable. That is, these variables will attempt to capture the effects of economic, historical, social, and weather conditions that could not be explicitly modeled. Examples of variables that could be included are:

- Real per capita personal income provided quarterly by MSA
- California unemployment rate
- California consumer price index
- Aggregate residential consumption: It is reasoned that electricity consumption over all PG&E residential premises would vary with economic and other historical conditions. During recessions, consumption will decrease, and when the economy is good, electricity use will increase. Aggregate monthly consumption for all members of the eligible P4P population will be calculated and incorporated into the regression model.
- Monthly dummies

3.1.3. Normalization

For P4P, there is no need to normalize the gross savings to account for different baseline assumptions for equipment that is replaced on burn-out since P4P assumes that all installations will be early replacement. This is consistent with the treatment of Energy Upgrade California Advanced Home Upgrade program. While this makes the existing conditions the appropriate baseline for estimating first-year annual savings, the lifecycle gross savings must be adjusted to account for early replacement in calculating the TRC. However, a method for adjusting these regression-based lifecycle savings has not yet been identified. An approach developed for the New York State Department of Public Service by Ridge, Jacobs, Tress and Hall (2011) is one possibility.

3.1.4. Double Counting

Finally, because participants are able to participate in other PG&E programs (except for Energy Upgrade California), we will conduct an analysis to avoid any double counting of savings potentially claimed by other PG&E measures using the standard program tracking database.

3.2. Self-Report Net-To-Gross Ratio

The Self-Report Approach (SRA) method will be consistent with *Guidelines for Estimating Net-To-Gross Ratios Using the Self Report Approach* (Ridge, Keating and Megdal, 1997). The methods and instrument contained in the *Joint Simple Net of Free-Ridership and Participant Spillover Self-Report Survey Battery* (Residential Net-To-Gross Ratio Working Group, 2008) will be customized to address the unique characteristics of P4P. A stratified sample will be designed so that customers with the largest estimated savings will be overrepresented. As noted below in Section 7.4, participating residential customers will be interviewed twice, immediately after the treatment and one year later. It is critical that the self-report survey be conducted as part of the first survey to minimize the problem of recall.

If this SRA based estimate of net savings is reasonably close to the regression based estimate, then one is reassured that the regression-based estimate is sufficiently accurate. This approach is referred to in the literature as triangulation, which provides redundant or confirmatory measurement (Scriven, 1991).

3.3. Sample Design

A sample design is not required in estimating P4P gross savings because the analysis will be performed on the full, relevant program population with sufficient pre- and post-treatment consumption data. Given this, there is no sampling error, i.e., there are no confidence intervals around the estimates of short-term and long-term gross savings.

4. Method 3: P4P Versus PACE Loan Program

Similar analyses will be conducted using an alternate comparison group comprised of customers who have participated during the initial enrollment period in the PACE Loan Program. The results of this analysis will test the hypothesis that the P4P, by allowing aggregators to determine the mix of interventions that is most attractive to customers, including behavioral, operational and retrofit activities and paying them based on verified energy savings, can lead to significant energy savings above the existing PACE Loan Program.

To estimate net savings, a pooled, fixed-effects, time-series/cross-sectional (panel) regression model (similar to Equation 1) that incorporates the participants in the P4P *and* the PACE Loan Program will be estimated. The main difference is that the treatment variable in Equation 1 would represent whether one participated in the P4P or the PACE Loan Program.

5. Alternative Method Considered: Random Encouragement Design

We also considered—but ultimately rejected— the use of a random encouragement design (Cappers, 2014). In a randomized encouragement design (RED), a homogeneous group of customers are divided randomly by a third party into two groups whereby one group of eligible customers is "encouraged" to take up the treatment (but some may not do so) and another group of eligible customers is not encouraged. The evaluation of the treatment effect in such a design necessitates including both the customers who actually took up the treatment and those who did not within the encouraged group. In aggregate, this "treatment" group can be compared against a randomly-drawn control group from the eligible customer population, which would likewise be comprised of those who, if given the offer of treatment, would accept it as well as those who would reject the offer. This randomly-drawn control group from the eligible customer population is therefore, in expectation, an unbiased counterfactual[10] to the behavior of the treatment group.

[10] The counterfactual is the result that would have been expected had the intervention not been implemented.

Specifically, we considered three different evaluation designs:

1) **Randomize Control Trial.** This design involves randomly assigning eligible PG&E customers to participate in the P4P (treatment) and randomly assigning eligible PG&E customers to not participate in the P4P (control).
2) **Full Eligible Population Random Encouragement Design.** This Random Encouragement Design (RED) involves randomly assigning all eligible PG&E residential customers to one of two groups. The *treatment* group is encouraged to participate in the P4P. The *control* group is not allowed to participate in the pilot for two years.[11]
3) **Partial Eligible Population Random Encouragement Design.** This RED is a variation on the full approach described above but is less invasive to the operation of the program because a portion of eligible customers are untouched by the experiment (the "business as usual" group). The remaining customers are enrolled in the experiment and will be assigned to either the *treatment* or to the *control* group as outlined in the full design described above.

In the course of the development of this evaluation plan, we realized RCT and RED designs might not be feasible *or* desirable for P4P. The key limitations of these two designs are listed below:

1. **Feasibility**. In a RCT design, to randomly assign eligible customers to the P4P means that PG&E customers would be mandated to participate in the P4P, which is impossible. The only reason such a design is feasible for such program as OPower is that customers are not asked to opt in, i.e., every eligible household receives the monthly energy report.

2. **External validity.** In RCT designs that involve random assignment to treatment and control groups are at a slight disadvantage when it comes to external validity[12]. Mohr (1995) concludes: "Because they demand enough control to be able to assign subjects to treatments at random, they make it more difficult to employ typical subjects and natural or representative setting; the randomization often upsets natural groupings and setting and leads to the selection of atypical subjects simply because they are easy or convenient or at least possible to randomize" (p. 97). External validity for this evaluation is critical since a key component of the underlying theory of the P4P is that the market actors should be allowed the flexibility to implement the program using their best professional judgement. A fair test of this program design component would be to give control of the marketing and targeting of the program to the aggregators who are supposed to implement the program. For PG&E program staff to impose their definition of the eligible market means that the results of this evaluation will be less generalizable to a scaled-up future program in which the aggregators have full control of the marketing and targeting of the program.

---

[11] A true experimental design isn't possible since PG&E cannot mandate that a random sample of eligible customers actually participate in the P4P and that a random sample of eligible customer cannot participate in the P4P.
[12] The issue of external validity concerns the extent to which one may safely generalize the conclusions derived from an evaluation.

3. **Customer equity.** In both RCT and RED designs, a significant portion of eligible customers would be denied any benefits of participating in the program for two years.

4. **Ability to manage aggregator marketing behavior.** In RED designs, a given aggregator might not agree that the PG&E list of eligible customers assigned to the treatment group is optimal. As a result, they might supplement this list with households that they believe have greater savings potential and higher probability of participating. This of course would compromise the randomness of our design, effectively turning it into a quasi-experimental design.

5. **Ability to attract aggregators.** In RED designs, aggregators might be too risk averse to sign a contract that requires them to market only to PG&E-identified households that they believe are a sub-optimal group of households, or that limits their ability to use the targeting approaches they see as being the most effective (such as neighbor referral or geographic targeting approaches that may be incompatible with assignment approaches used in a RED design).

6. **Statistical power requirements.** In RED designs, sample size requirements are greater than the sample size requirements for a true experimental design. The power analysis[13] used to estimate the sample size must take into consideration that the number of households required to obtain a given level of statistical power in a RED increases by a factor of $1/c^2$ where c is defined as the share of treatment group households that participate in the program (Cappers, 2014). Such a large sample size might not be possible for a pilot program in which few customers might be expected to participate.

7. **Maintaining the integrity of the design.** In RED designs, the implementation can be challenging. PG&E, in close collaboration with the aggregators, would need to agree on the definition of the eligible population in order to improve the external validity of the design. This definition would probably be broader than the eligible population defined by any one aggregator since it must include unique customer types that each of the aggregators might prefer to target.[14] Aggregators would then be instructed to encourage only those assigned by PG&E to the treatment group and to create a database of all these encouraged households. Aggregators would be supplied on an on-going basis with random samples of the eligible population which they must approach since *all* members of the eligible population must be encouraged by aggregators not just a subset of those that they might prefer to target. Only when each sample is exhausted, could an aggregator request another sample. Those assigned to the control group would not be allowed to opt into the Pilot for two years. Maintaining the integrity of this design requires clear communication among all parties, effective management of samples of those eligible for treatment, and discipline on the part of 1) the aggregators to market only to those

[13] The statistical power of a study translates into the probability that the study will lead to the correct conclusion (i.e., that it will detect the effects of treatments (Murphy and Myors, 1998).
[14] Note that agreement among PG&E and the aggregators regarding the definition of the eligible population could help to mitigate (not eliminate) the first concern.

assigned to the treatment group and 2) PG&E to deny treatment to those control group households that might seek to participate.

6. Process Evaluation

A process evaluation is defined as: *A systematic assessment of an energy efficiency program for the purposes of (1) documenting program operations at the time of the examination, and (2) identifying and recommending improvements that can be made to the program to increase the program's efficiency or effectiveness for acquiring energy resources while maintaining high levels of participant satisfaction* (TecMarket Works 2004). A process evaluation is particularly important for new programs in which the mechanics of implementing the program are relatively new and untested. To gather the necessary data, telephone interviews will be conducted with the following:

1. P4P staff
2. Participating aggregators
3. Nonparticipating energy service providers
4. Participating residential customers
5. Residential customers who were marketed to but chose not to participate

Below, for each of these five groups, we describe the general topics to be covered, the targeted number of interviews, the targeted level of confidence and precision when sampling is used, and the frequency with which interviews will be conducted.

6.1. Program Administrator Staff

In-depth interviews will be conducted with two Program Administrator staff members. They will be interviewed by telephone twice each year, covering such topics as:

- Substantial deviations from original program design and the reasons why
- Ideas for improvement of program design and delivery
- Ideas for making the program more scalable
- Perceptions of whether participating aggregators prefer the P4P approach to more traditional PG&E energy efficiency programs.

6.2. Participating Aggregators

In-depth interviews will be conducted with staff members of each participating aggregator. They will be interviewed by telephone twice each year, covering such topics as:

- Ideas for improvement of program design and delivery
- Ideas for making the program more scalable
- Perceptions of whether they prefer the P4P approach to more traditional PG&E energy efficiency programs.
- PG&E management of the P4P
- Whether program requirements were onerous
- Whether they were paid in a timely manner

6.3. Nonparticipating Energy Service Providers (ESPs)

In-depth interviews will be conducted with a random sample of 70 nonparticipating energy service providers (ESPs) with the expectation of achieving the 90 percent level of confidence plus or minus 10 percent. They will be interviewed by telephone once each year, covering such topics as:

- Awareness and knowledge of the P4P
- Interest in participating in the P4P
- Barriers to participating in the P4P

6.4. P4P Participants

Interviews with 70 participating residential customers will be conducted twice, immediately after the treatment and one year later. Within each aggregator, the sample might be further stratified by size of expected savings. The 90 percent level of confidence plus or minus 10 percent has been targeted. Interviews will address such topics as:

- Perception of risk of not achieving energy savings
- Program satisfaction
- Spillover
- Ideas for improvement of program design and delivery
- Other household changes (e.g., increased occupancy, addition of energy using equipment) from the pre to the post period that might have affected energy use
- Use of web portal to track the energy use of their household
- Sharing of their P4P experience with their friends and neighbors
- Participant non-energy benefits (e.g., increased comfort)

6.5. Nonparticipating Eligible Residential Customers

Seventy customers who were encouraged/marketed to but who decided not to participate in the P4P will be interviewed once each year. The sample might be further stratified by CEC climate zone. The 90 percent level of confidence plus or minus 10 percent has been targeted. Interviews will address such topics as:

- Barriers to participating in the P4P
- Interest in future participation in the P4P
- Awareness and knowledge of the P4P

7. Early EM&V

It is critical that PG&E monitor (using aggregator dashboards) post-intervention consumption data for each participating household in order to determine if the observed ex post savings match the expected ex ante estimates. If savings are less than expected, PG&E can explore possible causes and take corrective action. For example, QA/QC audits might discover that a particular aggregator is not performing quality installations of HVAC units and could benefit from additional training. Or, interviews with samples of participants might reveal that some occupants are changing their behavior (e.g., turning up their thermostat in the winter or turning it down in the summer). In such cases, the non-energy benefits could be clearly documented or participants would be reminded that such “take back” will reduce their expected bill savings.

8. P4P Performance Metrics

A summary of the key performance metrics that will be tracked over the course of the P4P is provided in Table 1. These metrics are informed by the logic model provided in the HOPPs submission for P4P.

Table 1. Program Performance Metrics to Be Tracked

Goal	Metric	Logic Model Box	Target for Initial Enrollment Period (IEP): Years 1-2	Logic Model Box	Target for Second Enrollment Period: Years 3-5
Develop Scalable Business Models	Participating customers	I	2,100 / year	O, T	Triple IEP
	Participating aggregators and contractors	B	3-5 aggregators, 50 active contractors (>5 jobs/year)	N	Triple IEP
	Non-incentive costs	N	< 20% of total costs	S	< 16% of total costs
	Total cost per home	N	< $1,500/home	S	< $1,200/home
	Savings	R	4.83 GWH, 4.7 KW, 0.945 therms	R	Triple IEP
	Competing ESPs			P	50
Data Availability and Transparency	Transparent aggregator portfolio savings	K, N	Provided quarterly		
	Monetize savings	N	Aggregators able to bid into auction by 2018-2019		

Note that some of these indicators are simple (e.g., the number of participating customers) while other are more complex and will require more work to operationalize. To the extent possible each will also have to be transformed into SMART (**s**pecific, **m**easureable, **a**mbitious, **r**ealistic, and **t**ime-bound) objectives.[15] For each, decisions must also be made regarding the frequency of data collection (e.g., monthly, quarterly, annually, every third year, etc.) as well as the targeted level of accuracy and precision. Finally, both the Energy Division and the IOUs must collaborate in making all of these decisions.

[15] Poister, Theodore H. (2003). *Measuring Performance in Public and Nonprofit Organizations*. San Francisco, CA: Jossey-Bass.

9. Establishing Evaluation Data Requirements

Table 2 summarizes the data requirements and sources to support the impact and process evaluations of the P4P. The evaluators will prepare a written request to PG&E and the aggregators.

Table 2. Data and Sources

Data	P4P Aggregator	PG&E	Third Parties/ Telephone Surveys	California Energy Commission (CZ 2010)	PACE Implementor*
Aggregator ID	X				X
Unique site ID	X				X
Customer Account	X				X
Customer contact information (name, mail address including zip code, telephone number, and e-mail address)	X				X
Measures installed	X				X
Estimated savings per site	X				X
Date(s) of measure installation(s)	X				X
Begin dates and end dates for each billing cycle		X			
12 months of pre-treatment kWh consumption and 12 months of post-treatment kWh consumption		X			
Read type (indicating estimated and other non-actual reads)		X			
Variables required to merge consumption data with program tracking data (e.g., account numbers)		X			
Location information or other link to CZ 2010 weather data		X		X	
Customer tenancy at the premise (the tenancy starting and ending dates)		X			
Household characteristics from program-tracking database e.g., number and age of occupants, income, education) for members of treatment group		X			
Household characteristic data for members of the comparison group		X	X		

* PACE Loan Program implementors must provide all the premise information in the table for a matched set of PACE standard offer loan participants

Household characteristics data for members of the comparison group require some further details. Household characteristics data for members of the comparison group would be obtained from PG&E which has contracts with two third party providers. Such variables as age, household income, owner/renter, home square footage, year home was built, household size, language, length of residence, number of adults, and presence of children can be obtained for a very large percent of households in the comparison group. It is assumed that the same data will be collected from households in the treatment group by the aggregators. Another possibility is to collect the same data through telephone interviews with a random sample of comparison group households. This would be a more expensive option and the number of households from which we would be able to collect these data would be smaller due to budget constraints. The smaller sample would also reduce the statistical power of any regression models.

10. Attachment A: References

Abadie, A. and G. W. Imbens. (2002). Simple and bias-corrected matching estimators [Technical Report]. Department of Economics, University of California, Berkeley. Retrieved March 6, 2016, from http://ideas.repec.org/p/nbr/nberte/0283.html.

Abadie, A. and G. W. Imbens. (2006). Large sample properties of matching estimators for average treatment effects. *Econometrica*, 74, pp. 235-267.

Agnew, Ken and Mimi Goldberg. (2013). Whole-Building Retrofit with Consumption Data Analysis Evaluation Protocol. In U.S. Department of Energy. *Uniform Methods Project: Methods for Determining Energy Efficiency Savings for Specific Measures* (pp. 8-1 – 8-35) (http://energy.gov/eere/about-us/ump-protocols)

Aiken L. S. and S. G. West, D. E. Schwalm, J. L. Carroll and S. Hsiung. (1998). Comparison of a randomized and two quasi-experimental designs in a single outcome evaluation: Efficacy of a university-level remedial writing program. *Evaluation Review*, 22, pp. 207-244.

Campbell, Donald T. and Julian C. Stanley. (1963). *Experimental and Quasi-Experimental Designs for Research*. Chicago, Ill.: Rand McNally College Publishing Company.

Cappers, Peter. (2014). *Summary of Utility Studies: Smart Grid Investment Grant Consumer Behavior Study Analysis*. LBNL Paper LBNL-6248E (http://escholarship.org/uc/item/4rv446b)

Goldberg, M. L. And K. Train, (1995) Net Savings Estimation: An Analysis of Regression and Discrete Choice Approaches. Submitted by Xenergy Inc. To CADMAC Subcommittee on Base Efficiency.

Gou, Shenyang and Mark W. Fraser, (2015). *Propensity Score Analysis: Statistical Methods and Applications*. Los Angeles, CA: SAGE Publications.

Greene, W. H. (2003) Econometric Analysis (5th Ed.). Upper Saddle River, N.J.: Prentice Hall.

Hansen, B. B. and S. O. Klopfer. (2006). Optimal Full Matching and Related Designs via Network Flows. Journal of Computational and Graphical Statistics, 15, pp. 1-19.

Heckman, J. J. (1979). Sample Selection Bias as a Specification Error. Econometrica, 47(1), pp. 153-162.

Heckman, J. J. (1978). Dummy Endogenous Variables in a Simultaneous equations system. *Econometrica*, 46(1), pp. 931-960.

Heckman, J. J., H. Ichimura, and P. E. Todd. (1997). Matching as an Econometric Evaluation Estimator. *Review of Economic Studies*, 65, pp. 261-294.

Heckman, J. J., H. Ichimura, and P. E. Todd. (1998). Matching as an Econometric Evaluation Estimator: Evidence from evaluating a job training programs. *Review of Economic Studies*, 64, pp. 605-654.

Huitema, Bradley E. (2011). *The Analysis of Covariance and Alternative: Statistical Methods for Experiments, Quasi-Experiments, and Single-Case Studies*. Hoboken, New Jersey: John Wiley & Sons, Inc.

IPMVP: *International Performance Measurement and Verification Protocol*: *Concepts and Options for Determining Energy and Water Savings Volume I*, EVO-10000-1.2007, Efficiency Valuation Organization. (http://www.evo-world.org/en/products-services-mainmenu-en/products-ipmvp-mainmenu-en)

Johnson, Aaron C., Marvin B. Johnson and Rueben C. Buse. (1987). *Econometrics: Basic and Applied*. New York: Macmillan Publishing Company.

Johnson, Dr. Katherine. (2011). *A Supplement to the New York State Evaluation Guidelines: Update 2011*. Prepared for the New York State Energy Research and Development Authority and New York State Evaluation Advisory Group

Northeast Energy Efficiency Partnership (NEEP). (March 2011). *Glossary of Terms: A Project of the Regional Evaluation, Measurement and Verification Forum*. http://neep.org/uploads/EMV%20Forum/EMV%20Products/EMV%20Glossary%20of%20Terms%20and%20Acronyms%20-%20Version%202%20FINAL.pdf

PG&E, San Diego Gas & Electric, Southern California Edison, Southern California Gas, California Energy Commission, Division of Ratepayer Advocates and the Natural Resources Defense Council. (1996). *Procedures for the Verification of Costs, Benefits, and Shareholder Earnings from Demand-Side Management (DSM) Programs*. As adopted by the California Public Utilities Commission Decision 93-05-063 and Revised February 1996 Pursuant to Decision 94-05-063, 94-10-059, and 95-12-054.

Residential Net-To-Gross Ratio Working Group. (2008). *Residential/Small Commercial SR NTG (Revised Final—4/19/2008)*. Prepared for the Energy Division of the California Public Utilities Commission.

Ridge, Richard, Pete Jacobs, Harvey Tress, and Nick Hall. (2011). *One Solution to Capturing the Benefits of Early Replacement: When Approximately Correct Is Good Enough*. Presented at the International Energy Program Evaluation Conference.

Ridge, Richard, Ken Keating, Lori Megdal, and Nick Hall. (2007). *Guidelines for Estimating Net-To-Gross Ratios Using the Self Report Approach*. Prepared for the California Public Utilities Commission.

Ridge, Richard, Mike Baker, Nick Hall, Ralph Prahl, and William Saxonis. *Gross Is Gross and Net Is Net: Simple, Right?* Originally presented at the International Energy Program Evaluation Conference in August, 2013 and republished in *Progress in Industrial Ecology – An International Journal, Vol. 8, 2013.*

Rosenbaum, P. R. and D. B. Rubin. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika, 70, pp. 41-55.*

Rosenbaum, P. R. and D. B. Rubin. (1985). Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *American Statistician*, 39, pp. 33-38.

Scriven, Michael. (1991). *Evaluation Thesaurus*. Newbury Park, CA: SAGE Publications.

TecMarket Works. (2004). *The California Evaluation Framework.* Prepared for the California Public Utilities Commission and the Project Advisory Group http://www.calmac.org/publications/California_Evaluation_Framework_June_2004.pdf

The TecMarket Works Team. (2006). *California Energy Efficiency Evaluation Protocols: Technical, Methodological, and Reporting Requirements for Evaluation Professionals*. Directed by the CPUC's Energy Division, and with guidance from Joint Staff.

Train, Kenneth. (1993). *Estimation of Net Impact from Energy Conservation Programs*. Submitted to the Southern California Edison Company.

Wooldridge, J. M. (2002). *Econometric Analysis of Cross Section and Panel Data*. Cambridge: MIT Press.

PG&E Gas and Electric
Advice Filing List
General Order 96-B, Section IV

AT&T
Albion Power Company
Alcantar & Kahl LLP
Anderson & Poole
Atlas ReFuel
BART
Barkovich & Yap, Inc.
Bartle Wells Associates
Braun Blaising McLaughlin & Smith, P.C.
Braun Blaising McLaughlin, P.C.
CPUC
California Cotton Ginners & Growers Assn
California Energy Commission
California Public Utilities Commission
California State Association of Counties
Calpine
Casner, Steve
Cenergy Power
Center for Biological Diversity
City of Palo Alto
City of San Jose
Clean Power
Clean Power Research
Coast Economic Consulting
Commercial Energy
Cool Earth Solar, Inc.
County of Tehama - Department of Public Works
Crossborder Energy
Davis Wright Tremaine LLP
Day Carter Murphy
Defense Energy Support Center
Dept of General Services
Division of Ratepayer Advocates
Don Pickett & Associates, Inc.
Douglass & Liddell
Downey & Brand
Ellison Schneider & Harris LLP
Evaluation + Strategy for Social Innovation
G. A. Krause & Assoc.
GenOn Energy Inc.
GenOn Energy, Inc.
Goodin, MacBride, Squeri, Schlotz & Ritchie
Green Power Institute
Hanna & Morton
International Power Technology
Intestate Gas Services, Inc.
Kelly Group
Ken Bohn Consulting
Leviton Manufacturing Co., Inc.
Linde
Los Angeles County Integrated Waste Management Task Force
Los Angeles Dept of Water & Power
MRW & Associates
Manatt Phelps Phillips
Marin Energy Authority
McKenna Long & Aldridge LLP
McKenzie & Associates
Modesto Irrigation District
Morgan Stanley
NLine Energy, Inc.
NRG Solar
Nexant, Inc.
ORA
Office of Ratepayer Advocates
OnGrid Solar
Pacific Gas and Electric Company
Praxair
Regulatory & Cogeneration Service, Inc.
SCD Energy Solutions
SCE
SDG&E and SoCalGas
SPURR
San Francisco Water Power and Sewer
Seattle City Light
Sempra Energy (Socal Gas)
Sempra Utilities
SoCalGas
Southern California Edison Company
Spark Energy
Sun Light & Power
Sunshine Design
Tecogen, Inc.
Tiger Natural Gas, Inc.
TransCanada
Troutman Sanders LLP
Utility Cost Management
Utility Power Solutions
Utility Specialists
Verizon
Water and Energy Consulting
Wellhead Electric Company
Western Manufactured Housing Communities Association (WMA)
YEP Energy