**CAEECC-Hosted sub-Working Group on Underserved Customers**

**Meeting with UC Irvine Small-Medium Business (SMB) Research Team**

**December 17, 2020 10:30-12:00**

**Facilitators: Dr. Scott McCreary & Katie Abrams, CONCUR Inc**

**Meeting Summary[[1]](#footnote-1)**

# Attendees:

* *UC Irvine Research team*: Ed Coulson, Theo Love, Linda Tan, Matteo Zamparini, Carmen Liang
* *CAEECC members and other interested stakeholders*: Lara Ettenson, Anthony Kinslow II, James Dodenhoff, Kathleen Yip, Kaylee D’Amico, Marianne Bithell, Ted Howard, Aisha Cissna
* *Facilitators*: Dr. Scott McCreary and Katie Abrams

# Introductions & Agenda Review – *Scott McCreary*

Scott noted that the purpose of this meeting is for the UC Irvine Faculty and Students and Green Energy Economics research team to present early findings to the CAEECC Underserved Working Group. The research team anticipates finalizing their analysis in mid to late January 2020.

This timing and sequencing should enable full CAEECC Underserved Working Group to meet in early 2021 to review the analysis in depth and propose next steps for translating the analysis into policy recommendations to be presented to the CPUC for consideration.

# Present findings on SMB underserved research and analysis – *Professor Ed Coulson and Theo Love*

**Presentation highlights**

* Data sources and highlights
  + Used CEDARS to obtain claims and siteID, used Zipcode to cross-reference w/ other datasets (County Business Patterns and American Community Survey) to look at demographic and business information
    - Claim provides measure/project level info on program, site info, costs, savings, incentives, cost-effectiveness
    - Public[[2]](#footnote-2) dataset does not include site-specific details (city, zip, NAICS) –challenging to collate w/ other data sources
    - Data request provided unscrubbed claims for PGE, SCE, SCG, SDGE – allows zipcode analysis
  + **Initial claims analysis and data summary**
    - Summary: analyzed 505,500 unique claims from 2017-2019. Mean incentive $816, gross measure cost $2,292; 75% claims are small projects w/ some large projects $1.6M incentive; 56% zips in CA represented
    - Participation variable – challenging b/c PGE only PA to use “small commercial” tag
    - Investment flows – large barrier for participation is upfront cost; majority (56%) of claims have 99% incentive as percent of costs (likely direct install); otherwise wide range of incentive levels – varies by PA (SCE has highest % of cost covered by incentive; variation among IOUs: lower for SDGE and SCG; almost none for PGE – perhaps data tagging issue or no direct install program?)
    - Savings – still needs to be analyzed in depth, thus far have converted gas to kWh to account for fuel switching and make comparisons easier
  + County Business Patterns (CBP) and American Community Survey (ACS) regression analysis
    - **Initial analysis**:
      * Positive (higher % correlates w/ higher penetration): Hispanic, household size, income $50-75k
      * Negative (higher % correlates w/ lower penetration): number of establishments, population over 65, total payroll, African American
* **Next steps/future refinements**
  + ACS/CBP data: other income ranges, ethnic groups, and immigrant status
  + Claims data: will look at rebate amt, % cost covered by rebate, savings amount
  + Review correlations btw ACS/CBP data and claims data
  + Perform multivariate regression analysis
  + Collate w/ NAICS data
* **Overview of data challenges**
  + Upstream/downstream/midstream flags – SMB has only “upstream”, public has 3 flags and many claims have all three categories flagged simultaneously
  + Incentive issue – total gross incentive not in public data
  + Classification of residential/small commercial/commercial/large commercial/small industrial/agriculture is at the PA’s discretion
    - 1/3 of SMB data set is classified as small commercial
    - Notably, PGE is the only PA to use “small commercial” designation
  + Claims data challenges encountered: public data doesn’t include site info to compare to other data sources - or have clear documentation of how fields are derived on submission files (upstream issues, incentive issue, sectors and rate class issue); getting raw submitted data was difficult; submitted data doesn’t provide consistent mapping for rate classes or sectors. Facilitator note to research team: consider flagging these as potential areas to address in the future
  + ACS/CBP data challenges: requires zip-code level values; not all data is available for all zips

# Discussion and questions on research and analysis; initial suggestions for refining the analysis – *CAEECC members* – *CAEECC members*

CAEECC member and stakeholder questions are listed in italics below, with researchers’ responses in sub-bullets.

*Was the analysis able to differentiate between a Project vs a Measure? If a Project has multiple measures, are you able to determine that from the CEDARS Public Data?*

* Research is looking at site level not the project or measure level

*Am I understanding correctly that PG&E is the only PA to use “small commercial” designation in their dataset? That seems wildly unhelpful.*

* Yes, the other PAs tag “commercial” but PG&E tags “small commercial”. There must be a better way to tie claims data to rates; currently this is based on submitter (PA) discretion.

*Does the % of cost covered (slides 18-19) suggest that there’s a problem with the incentive process that makes it untenable or challenging for small commercial – can we infer that from this data?*

* The research team needs to look into more data, and the issue may speak to data issues.
* *A CAEECC member commented that PG&E is only IOU that runs DI through the Public local partnership program.*
* *Another member alluded to challenges with how data on local government partnerships projects (categorized as “Public”) appears in claims*

*Is household (slide 27) a proxy for SMB owners, or will you be using other proxy designations or data to capture ownership data?*

* Since claims data is at the zipcode level, we need to find another way to look at ownership data, for example demographic or payroll data

*Can the intercept (on slides 25 or 31) be used to determine a lack of service to groups?*

* Intercept indicated on these graphs shows 0 penetration level – i.e. where there is 0 % African American or Hispanic, which but this doesn’t tell us the average for that zipcode
* *A CAEECC member noted what’s interesting is that the intercept is the same for Hispanic and Native American, and we know that there are many zipcodes that have no Native American population*
* Multivariate analysis could give more insights, so next iteration may provide more details – which could show the effect of a population while holding the other variables constant. The research team confirmed they will using this approach in the next round of analysis.

*Can rurality status be included as an area of exploration/addition to the demographics data? Does the team have plans to incorporate that?*

* We may be able to get square mileage info and compare to density; applies one measure of rurality to entire level of zip level – we’ll give this thought, although it could be a challenge

*If the research team prefers presenting the final analysis in PowerPoint format, can an appendix be provided with relevant data details?*

* Yes an appendix with supporting data can be provided

*How is underserved defined here? Hypothesis is that certain customers aren’t being served by programs. Does “served” mean, for example, providing a thermostat incentive?*

* Does underserved mean SMB or businesses in underserved areas, so can measure by demographic criteria.
* The main way we’re looking at “served” is through participation, investment, and savings. There’s no perfect definition, but we’re finding statistically significant ways in which participation, investment, and savings vary.
  + *A CAEECC member noted that for Residential, we defined “underserved” by bottom quartile. Would that approach make sense here?*
  + We don’t know rate class – just if they participated, savings, and incentive amount – so this is a challenge.

# Next steps and timeline

Scott summarized the next steps as follows

* Research team to finalize first statistical analysis by 1/15
* Full UWG meeting in early February to review results and determine next steps

Lara reminded the group that the first question the working group is tasked with is determining whether there is sufficient data to support a hypothesis of being underserved that would warrant evaluation of program design and policies.

Theo expressed an interest in analyzing all commercial data. Lara offered to provide Theo with data if she’s able to obtain it from the CPUC. Theo specifically requested “Claims data with primary sector that doesn’t equal the residential sector.”

Scott and the CAEECC members expressed a final statement of appreciation to research team, especially the students helping propel the analysis behind the scenes.

1. Note this is an informal meeting summary, as the CAEECC process does not require formal meeting summaries for sub-working group meetings. Name attribution is provided in some places to assist in research/analysis followups [↑](#footnote-ref-1)
2. “Public” is used in the context of these notes, and the associated presentation to refer to publicly available data, *not* to data on the Public sector [↑](#footnote-ref-2)