American Economic Journal: Economic Policy 2023, 15(3): 286–321
https://doi.org/10.1257/pol.20210208

The Costs of Misaligned Incentives: Energy Inefficiency and the Principal-Agent Problem†

By Joshua A. Blonz*

In many settings, misaligned incentives and inadequate monitoring lead employees to take self-interested actions. This paper identifies and quantifies the costs of this principal-agent problem in the context of an energy efficiency appliance replacement program. I show that contractors (agents) hired by the electric utility (the principal) increase their compensation by intentionally misreporting program data to authorize the replacement of nonqualified refrigerators. I estimate that each unqualified replacement reduces program benefits by $106 *and saves 30 percent less electricity than replacements that follow program guidelines. The same program without a principal-agent distortion would increase program benefits by $60 per replacement.* (*JEL* D82, L68, L94, L98)

The notion that economic agents respond to incentives is central to economics. The principal-agent problem is a canonical example of the power of incentives, in which self-interested agents may act against the interests of the principals. A standard formulation of the principal-agent problem involves a principal employing an agent to conduct a task where the principal cannot observe the agent's effort. In this setting, performance-based payments (rather than a fixed wage) offer the potential to align agents' and principals' incentives, by compensating agents proportionately to their output. Performance pay has been shown to increase productivity in cases where agents' output is easily observable (Lazear 2000; Shearer 2004), but there are many cases where the principal cannot observe output directly. When performance pay is used in these settings, agents have the incentive to undertake unobservable actions to maximize their compensation at the expense of the principal (Baker 1992).

* Federal Reserve Board (email: joshua.blonz@frb.gov). Naomi Feldman was coeditor for this article. The views in this paper are those of the author and do not necessarily represent the views or policies of the Board of Governors of the Federal Reserve System or its staff. I am deeply grateful to my advisor, Severin Borenstein, for his invaluable advice. I also thank Maximilian Auffhammer, Becka Brolinson, Fiona Burlig, Peter Christensen, Lucas Davis, Aluma Dembo, Meredith Fowlie, Sylvan Herskowitz, Sandile Hlatshwayo, Koichiro Ito, Erica Myers, Karen Palmer, Louis Preonas, Catherine Wolfram, and seminar participants at the NBER Summer Institute, AERE 2019, and the Energy Institute at Haas Energy Camp for their thoughtful and helpful comments. I thank the California Public Utilities Commission Public Advocates Office for providing data for this study.

† Go to https://doi.org/10.1257/pol.20210208 to visit the article page for additional materials and author disclosure statement(s) or to comment in the online discussion forum.

The challenge of imperfectly observed output is exacerbated if workers can exert effort on multiple dimensions or across many tasks (Kerr 1975; Holmström and Milgrom 1991), which applies to a wide range of settings including teacher pay based on test scores (Jacob and Levitt 2003), worker retraining programs (Courty and Marschke 2003), and commission-based sales (Tayan 2019). While compensation structures with misaligned incentives can result in losses, it has been difficult to measure their consequences, because agent actions are typically unobservable.

In this paper, I exploit a unique opportunity to observe agents taking self-interested actions at the expense of a principal, which enables me to estimate the costs of a principal-agent problem. In this empirical setting—an appliance upgrade program for low-income households—the principal does not directly monitor the agent, but I can observe how agents adjust their behavior in response to the level of incentive on one of their outputs. I develop a principal-agent framework to model how agents increase their compensation by disobeying guidelines set by the principal. Combining this modeling framework with empirical estimates of agents' misreporting behavior and outcome data, I find that the principal-agent problem turns an otherwise beneficial program into a program with negative net benefits.

I analyze data from 180,000 home energy efficiency retrofits conducted between 2009 and 2012 as part of the California Energy Savings Assistance (ESA) program, which provided free appliance replacements to low-income households. The electric utility Southern California Edison (SCE; the principal) hired contractors (the agents) to enroll households in the program, assess which upgrades each home was eligible for, and install upgrades.[1] By exploiting changes in program eligibility rules, I show that contractors intentionally misreported the age of refrigerators to increase the number of replacements they were able to provide, thereby increasing their compensation. I estimate that 13.8 percent of ESA enrollees were qualified for a refrigerator replacement, while 27.3 percent were determined to be eligible by the contractors. This implies that contractors misreported data for 13.5 percent of the households that participated in the program, resulting in about half of the refrigerator replacements conducted in the ESA program not qualifying based on program rules.

Next, I focus on how contract structure—which created variation in the financial incentive to misreport—affected the misreporting rate. Eleven of the 22 contractors in the ESA program conducted all the steps in the upgrade process, including enrolling households, determining refrigerator replacement eligibility, and (on a separate visit) installing the replacement refrigerators. I estimate that these "integrated-task" contractors had an average incentive to misreport of \$123 per household, which resulted in these contractors misreporting data for 19 percent of the households they enrolled. The other 11 contractors enrolled households and determined refrigerator eligibility, but were *not* responsible for refrigerator replacement. I estimate that these "separated-task" contractors had an average incentive to misreport of only \$25 per household, which resulted in these contractors misreporting data for 7.8 percent

[1] SCE runs the ESA program at the direction of the California Public Utilities Commission (CPUC). For simplicity, I refer to SCE as the principal because it directly administers the program, but the CPUC shares responsibility as the principal.

of the households they enrolled. In short, I find that the incentives created by the contractor compensation structure directly influenced misreporting and the size of the principal-agent distortion.

To evaluate how intentional misreporting affects program outcomes, I use data on monthly household-level electricity consumption before and after an upgrade. I combine this with model number data collected during assessments to classify refrigerator replacements conducted in the program as qualified or unqualified. I find that qualified replacements reduce household electricity consumption by 73 kilowatt hours (kWh) per month, whereas unqualified replacements save only 38 kWh per month. I then calculate the net benefits from qualified and unqualified refrigerator replacements. On average, qualified replacements increase net benefits by $60, while unqualified replacements *decrease* net benefits by $106. Contractor misreporting therefore substantially undermined the benefits of the ESA program.

Finally, I combine my model and empirical estimates to evaluate trade-offs between integrated-task contracts and separated-task contracts. Integrated-task contracts could be preferable if there were significant coordination benefits from having a single contractor conduct both initial assessments and physical replacements. However, the empirical results imply that the costs of misaligned incentives outweigh any potential coordination benefits. I find that integrated-task contracts reduce net benefits by an average of $12.32 more per assessment than separated-task contracts. The relative benefits of integrated-task contracts would need to be $95 per refrigerator replacement, or 135 percent above the average payment to contractors to conduct an assessment, in order to rationalize integrated-task contracts over separated-task contracts. The reduction in benefits from a separated-task contract is equal to 16 percent of the total electricity reduction benefits from the 25,000 refrigerator replacements conducted by integrated-task contractors, and demonstrates it would have been preferable to use the separated-task contract to implement the ESA program.

This paper makes three contributions. For one, to the best of my knowledge, this is the first paper to estimate the costs of a principal-agent distortion using detailed outcome data. The existing empirical principal-agent literature primarily identifies distorted agent behavior but has been unable to evaluate the costs of these distortions (e.g., Oyer 1998; Courty and Marschke 2004; Marion and West 2023).[2] This is because of insufficient outcome data and difficulty in identifying a valid counterfactual. In this paper, I am able to overcome these obstacles using policy-driven variation in agent incentives and electricity consumption data to estimate the costs of misaligned incentives between principals and agents. Of course, the estimates do not necessarily generalize to other contexts. However, my unique setting provides general insights on the role of contract structure in the principal-agent problem and specific insights on energy efficiency programs.

Second, I leverage variation in contract types to identify how contract structure influences both agent behavior and net benefits. Most of the previous work in this area of the contracts literature has focused on the impact of performance incentives

[2] See Prendergast (1999) and Holmström (2017) for a survey of the empirical principal-agent distortion literature.

compared with fixed wages (e.g., Asch 1990; Lazear 2000; Shearer 2004; Bandiera, Barankay, and Rasul 2007; Lewis and Bajari 2011, 2014).[3] In contrast, my results highlight how even within a given payment scheme, contract structure can distort incentives (all contractors in the study are paid a piece-rate wage) and reduce program benefits.

Finally, this is the first paper to identify and quantify the mechanism and incentives behind a principal-agent problem and demonstrate their impact on an energy efficiency program. The closest paper is Giraudet, Houde, and Maher (2018), which shows lower energy efficiency savings from home retrofits completed on Fridays. The authors attribute this to moral hazard and negative productivity shocks to workers at the end of the week. In my paper, I am able to directly observe agents misreporting and measure how that affects program outcomes. My findings add to a growing number of papers that have identified mechanisms behind the underperformance of energy efficiency programs and standards, including car manufacturers gaming of fuel economy standards (Reynaert and Sallee 2021), the landlord-tenant problem (Levinson and Niemann 2004; Davis 2012; Gillingham, Harding, and Rapson 2012; Myers 2020), heterogenous workmanship, and biased engineering models (Christensen et al. 2021).

Energy efficiency programs have become increasingly popular in the United States, with energy efficiency subsidies totaling \$8.2 billion in 2016 (CEE 2018). However, rigorous empirical analysis has consistently demonstrated that energy efficiency subsidies deliver lower energy savings than expected (Metcalf and Hassett 1999; Davis, Fuchs, and Gertler 2014; Graff Zivin and Novan 2016; Allcott and Greenstone 2017; Burlig et al. 2020; Fowlie, Greenstone, and Wolfram 2018). In this paper, I provide the first empirical evidence for a cause of the nonperformance of energy efficiency programs: the principal-agent problem. I am also able to quantify the relationship between realized energy savings and contract structure. The findings are of direct policy relevance. The CPUC, which is responsible for designing and monitoring the ESA program, learned of the potential principal-agent problems in the ESA program from my findings. A CPUC decision on the ESA program cited this paper and stressed the goal of reducing or eliminating the principal-agent problem I identify (CPUC 2019).

More broadly, energy efficiency is a critical component of most policies to mitigate carbon emissions and combat climate change (Loftus et al. 2015; Fowlie, Greenstone, and Wolfram 2018). While the ESA program's contract structure may exacerbate the principal-agent problem, many energy efficiency retrofit programs use contract structures similar to that of the ESA program. This suggests that the principal-agent problem likely contributes to lower-than-expected savings in many energy efficiency settings.

The rest of the paper is organized as follows: Section I describes energy efficiency policy and the ESA program. Section II introduces the model. Section III provides evidence of contractor misreporting. Section IV estimates the changes in electricity consumption from qualified and unqualified refrigerator replacements. Section V

[3] See Lazear (2018) for a survey of the performance-based compensation literature.

conducts a cost-benefit analysis, calculates a benefit-cost ratio, and calculates the benefits of an integrated-task contract using an expression derived in Section II. Section VI concludes.

I. Empirical Setting

A. *Energy Efficiency Policy*

Energy efficiency subsidy programs are typically considered a win-win policy intervention because energy efficiency investments are projected to reduce energy expenditures enough to more than offset their up-front costs, with the added bonus of reducing greenhouse gas emissions and local air pollution created by energy use. The majority of energy efficiency programs are mandated by state regulators, administered by the local electricity or gas utility, and funded by a surcharge on utility bills.

Policymakers give many justifications for subsidizing energy efficiency. Liquidity or credit constraints might prevent households from making an up-front investment that could take five to ten years to pay off. Reducing energy consumption can also reduce unpriced carbon and local air pollution externalities.[4] Landlords also may not have a strong incentive to invest in the efficiency of their rental units when the tenant pays the utilities (Davis 2012; Myers 2020).[5] There are also reasons why the government should not subsidize energy efficiency. Most electricity customers in the United States pay retail prices above the social marginal cost of electricity generation, which is a subsidy for the private adoption of energy efficiency (Novan and Smith 2018). Despite the subsidy for adopting energy efficiency caused by residential electricity pricing and the growing empirical evidence that energy efficiency programs are not cost-effective, energy efficiency remains a popular policy tool.[6]

Another policy tool used to increase energy efficiency is federal minimum efficiency standards. In the United States, most major appliances—including refrigerators—must satisfy a minimum level of efficiency. Before 1990, the average refrigerator consumed 127 kWh/month and represented one-quarter of total household electricity consumption (Energy Information Administration 1987). A similar refrigerator consumed only 65 kWh/month after the standards were tightened starting in 1993.[7] These binding standards have been responsible for progressively increasing energy efficiency, as the fleet of refrigerators is replaced over time.

[4] Unpriced externalities are primarily a concern in regions where the private cost is lower than the social marginal cost of electricity (Borenstein and Bushnell 2022).

[5] See Allcott and Greenstone (2012), Gerarden, Newell, and Stavins (2017), and Gillingham, Keyes, and Palmer (2018) for reviews of the policy justifications for energy efficiency programs and the research on their cost-effectiveness. Most of the programs are subsidized energy-efficient appliance replacements or building retrofits, which are designed to overcome barriers to energy efficiency adoption.

[6] For example, Fowlie, Greenstone, and Wolfram (2018), Allcott and Greenstone (2017), Burlig et al. (2020), and Christensen et al. (2021) all find a cost-benefit ratio meaningfully below 1.

[7] In Southern California, the average household uses 450 kWh/month, which means a refrigerator manufactured before 1993 consumed up to 20 percent of a household's electricity. See online Appendix A for a description of federal minimum efficiency standards.

B. *ESA Program*

The ESA Program provides free home energy efficiency upgrades to 400,000 low-income California households annually. I focus on the ESA program administered by SCE, which is one of the largest utilities in the United States and has about 100,000 ESA participants each year (SCE 2012b).[8] The program is mandated by the CPUC, which sets guidelines and funds the policy with a per-kWh charge on all electricity sold in California.[9]

Between 2009 and 2012, SCE hired 22 geographically dispersed contractors to assess who was eligible for the program and provide upgrades. The energy efficiency retrofits proceeded in the following steps: [10]

(Step 1) SCE reached out to eligible households through mail or phone calls. Interested households signed up to start the enrollment process.

(Step 2) SCE gave household contact information to an ESA contractor. The contractor scheduled a visit with the household to conduct a home assessment.

(Step 3) A contractor visited the household to verify that it was low income to complete the program enrollment. The contractor then assessed eligibility for refrigerator replacement and any other major upgrades.

(Step 4a) If a household was not eligible for any major upgrade, no action was taken and the contractor left.

(Step 4b) If a household was eligible for a major upgrade, the contractor provided up to five energy-efficient light bulbs and scheduled a second appointment to replace the refrigerator.[11]

(Step 5) The contractor provided the refrigerator replacement and any other major upgrades.

In this paper, I focus on the contractor behavior around the assessment and replacement of refrigerators, which were the most common major upgrade in the program. Between 2009 and 2012, 41,529 households received refrigerator replacements, which was 23.1 percent of households that enrolled in the ESA program. Refrigerator replacements totaled one-fourth of program expenditures and almost one-half of the program's projected energy savings (SCE 2012a). The contractor

[8] See online Appendix C1 for a description of program eligibility criteria, a full list of upgrades the ESA program provides.

[9] The utility does not earn a rate of return on the investments made through the ESA program (CPUC 2014).

[10] To supplement my understanding of the ESA program, I participated in a number of ride-alongs. Online Appendix C.2 discusses some of the insights I gained from directly observing the program.

[11] ESA program rules dictated that the contractor could install light bulbs only if the household qualified for a major upgrade.

assessed refrigerator eligibility in Step 3 by determining the year the refrigerator was manufactured.[12] A household was eligible for a free refrigerator replacement if its oldest working refrigerator was manufactured in 1992 or earlier. This eligibility cutoff was chosen because federal minimum efficiency standards tightened in 1993. As a result, replacing a refrigerator manufactured in 1992 or earlier is projected to save more electricity than replacing a refrigerator manufactured in 1993 or later.

There were two types of contractors that implemented the ESA program between 2009 and 2012. Half of the 22 contractor firms were able to conduct all five steps in the ESA program, from assessment to installing a major upgrade. These contractors, which I call "integrated-task" firms, were responsible for 50 percent of the assessments, and they conducted 100 percent of the refrigerator replacements in my data. The other 11 contractors, which I call "separated-task" firms, were able to conduct only Steps 1 through 4 of the ESA program—importantly, they did not replace refrigerators—and conducted the other 50 percent of the assessments. If a separated-task contractor determined a household was eligible for a refrigerator replacement, one of the 11 integrated-task firms provided that replacement.

Contractors were paid a fixed piece rate for every service they provided. If a household was eligible for a major replacement, contractors were paid \$25 to install five compact fluorescent light bulbs (CFLs) at the end of the assessment step. Contractors were paid \$224 to remove the old refrigerator for disposal and replace it with a new energy-efficient unit. Separated-task contractors, who never installed replacement refrigerators, received only the \$25 payment if CFLs were installed. Integrated-task contractors received both payments.[13] In most cases, a household only received CFLs when the contractor conducting the assessment determined that the dwelling was eligible for a refrigerator replacement. This program design gave separated-task contractors the incentive to find a household eligible for a refrigerator replacement so that they could install the CFLs and receive the \$25 payment.

While the ESA program did conduct audits during the period studied, these audits could not prevent unqualified refrigerator replacements resulting from contractor misreporting. The ESA program used audits to verify that contractors provided the replacements they billed to the program. The audits were completed after refrigerators were replaced, too late to check the refrigerator's manufacture year. Program guidelines set by the regulator did not require verification of the reported manufacture year, and SCE did not implement its own monitoring system. As a result, contractors could intentionally misreport that an ineligible household was eligible for a refrigerator replacement without much chance of being caught.

Contractors in the ESA program were paid a piece rate to incentivize them to conduct as many upgrades as possible (Research into Action 2011). This compensation structure, combined with the lack of monitoring, created the principal-agent problem where the principal wanted the contractors to replace only qualifying

[12] The manufacture year can be determined by looking at the serial number, which is usually on a plaque inside the main compartment. Contractors recorded information for all refrigerators in the dwelling.

[13] All households were given the same energy-efficient refrigerator. The new unit was meant to be approximately the same size as the one that was removed. Of refrigerator eligible households, 15.8 percent did not receive the replacement because of not wanting the replacement or logistical challenges. See online Appendix D.8 for more details.

refrigerators, but contractors had the incentive to misreport and provide ineligible replacements. Misreporting was particularly costly in this context, because the 1992 eligibility cutoff was chosen to maximize the savings from replacing the least efficient refrigerators.

II. Model

I consider a principal-agent model where the principal is an electric utility working under the supervision of the regulator, and the agent is a contractor hired by the utility. The electric utility may not directly care about the effectiveness of the ESA program. However, I assume the utility wants the program to function without misreporting to avoid reprimand or punishment from the regulator.[14] The principal hires an agent to provide services to a large number of households. Not all households qualify for the ESA program, and the principal cannot directly observe eligibility. Therefore, an agent must assess eligibility of customers on behalf of the principal before services can be provided. The contract proceeds in two steps. In the first step, "assessment," the principal hires an agent to assess n potential recipients' eligibility. In the second step, "service," the principal hires an agent to provide the service to participants who were deemed eligible in the assessment step.

Agents receive a piece-rate wage based on the number of assessments and services they provide. In the assessment step, the agent receives a fixed fee for each assessment she completes. If an agent reports that a potential recipient does not qualify, no services are provided. In addition to the flat fee, the agent receives a "referral payment" for each recipient she decides qualifies for the service.[15] In the service step, an agent receives a fixed payment for each service she provides. A portion α of potential recipients qualify for the service under program rules.[16] The principal and the agent have a shared but noisy expectation about the value of α. Subject to participation, the agent chooses a level of misreporting $Z \in [0,1]$ to maximize her expected payout, which is a static decision the agent makes at the beginning of the contract.

Separating the contract between different agents is costly. The agent must pay an "information acquisition cost," F, to complete her task during both the assessment and service steps.[17] In the integrated-task contract, the information acquisition cost is paid only once during the assessment visit because the same agent is responsible for both steps. The two agents cannot pass information in the separated-task contract, which requires the service provision agent to also pay the information acquisition cost for every recipient of the service. This extra information cost is the main disadvantage of a separated-task contract relative to an integrated-task contract. The

[14] I assume the regulator, which is the CPUC, wants to maximize program net benefits.

[15] In the ESA program, the referral payment is the $25 payment a contractor receives for installing five CFLs. The referral payment can take different forms in different settings.

[16] In the ESA program, α reflects the proportion of ESA-qualified households with a refrigerator manufactured in 1992 or earlier.

[17] In the ESA program, the information acquisition cost includes the costs associated with gaining the occupant's trust to schedule an appointment and enter the dwelling and learning about the occupants' appliances and their electrical system. During a ride-along, I observed a qualified refrigerator replacement that could not occur because the assessment contractor did not correctly note that the occupant did not have a grounded electrical socket.

principal's objective is to pick a contract structure that maximizes surplus knowing that the agent may intentionally misreport. The principal sets the payments to the contractors, which are driven by market rates.

The principal can observe only whether a contractor provided a service. To match the empirical setting I assume monitoring is not feasible or too costly for the principal to distinguish between a service provided to a qualified recipient and one resulting from false reporting in the assessment step. The principal is aware that the agent may misreport, but the only action the principal can take to punish an agent for suspected intentional misreporting is to cancel the contract with that agent for future assessments and services. For both types of contracts, I assume that agents are risk neutral and maximize expected contract value. By comparing the total expected contract surplus under each contract type, I derive the conditions under which the principal prefers the integrated-task contract.[18]

A. *Comparing the Two Contracting Approaches*

The principal's objective is to maximize net benefits by choosing the preferred contract structure for the setting. The derivation of the benefits of the integrated and separated contract types, Surplus_I and Surplus_S, is done in online Appendix B. To determine which contract the principal prefers, I compare the surplus in the two contract types to calculate the relative benefits of an integrated-task contract:

$$
(1) \quad E\left(\text{Surplus}_I - \text{Surplus}_S\right) = \overbrace{n b_2\left[n(1-\alpha)\left(Z_I^* - Z_S^*\right)\right]}^{\substack{(1)\\ \text{Costs from}\\ \text{increased unqualified service}\\ \text{provision with integrated contract}}}
$$

$$
\overbrace{-n(1-\alpha)\frac{1}{2}\left(\gamma_I Z_I^{*2} - \gamma_S Z_S^{*2}\right)}^{\substack{(2)\\ \text{Higher misreporting cost to agent}\\ \text{in integrated contract}}}
$$

$$
+ \underbrace{F\left[n(1-\alpha)Z_S^*\right]}_{\substack{(3)\\ \text{Information acquisition}\\ \text{cost in separated contract}\\ \text{for unqualified service provision}}} + \underbrace{Fn\alpha}_{\substack{(4)\\ \text{Information acquisition}\\ \text{cost in separated contract}\\ \text{for qualified service provision}}}.
$$

Equation (1) decomposes the factors that determine the net benefits of an integrated-task contract relative to a separated-task contract. I break it down into four main parts and describe each component. For brevity, I focus on describing the intuition behind each term.

Term (1) of equation (1) captures the costs of higher misreporting rates in the integrated-task contract. An important component of term (1) is the difference

[18] See online Appendix B.3 for a discussion of the assumptions in the model.

between the optimal misreporting rates, Z_I^* and Z_S^*, for the integrated- and separated-task contract.[19] Term (1) always reduces the relative value of an integrated-task contract. Term (2) is the difference in the misreporting cost to the agent between the integrated- and separated-task contracts. The γ_I and γ_S terms convert the misreporting rate into dollars for the integrated- and separated-task contracts. Term (2) always reduces the relative value of an integrated-task contract, because agents misreport more under the integrated-task contract, leading to a higher misreporting cost. Terms (3) and (4) capture the extra information acquisition cost that contractors pay in the separated-task contract. They show how information acquisition raises the cost of the separated-task contract for qualified *and* unqualified replacements, which increases the relative value of the integrated-task contract. In summary, the first two terms in equation (1) reduce the relative value of an integrated-task contract, and the last two terms increase the relative value of an integrated-task contract. The model does not predict which terms dominate or whether the integrated- or separated-task contract is preferred by the principal. In this paper, I apply the model to the ESA program to determine which contract type is preferred using empirical values derived in the following sections.

III. Intentional Misreporting

In this section, I identify and quantify the principal-agent distortion in the ESA program. I first look at how contractors intentionally misreport ineligible households as qualified for a refrigerator replacement. I use changes in program rules to develop a counterfactual against which I compare contractor behavior to identify intentional misreporting. I next use variation in contract type to examine how the incentive to misreport affects the size of the principal-agent distortion. Finally, I use the results to calculate the integrated-task and separated-task misreporting rates (Z_I and Z_S) used in the model.

A. *ESA Program Data*

I use two main datasets in this paper. First, I use confidential data from the SCE ESA program on assessments and installations. My main sample consists of 271,126 households that enrolled in the ESA program through SCE between January 2009 and December 2014.[20] The data include the reported refrigerator manufacture year and model number, along with information on the date of each visit, what upgrades or measures were installed, and which contractor performed each service.

B. *Overall Misreporting*

In this section, I investigate whether contractors intentionally misreport refrigerator manufacture years. Panel A of Figure 1 shows the distribution of refrigerator ages *reported* by contractors on the 180,105 assessments they conducted between January

[19] See online Appendix B.1 and B.2 for derivations of Z_S^* and Z_I^*.
[20] See online Appendix D.1 for details on the data cleaning process.

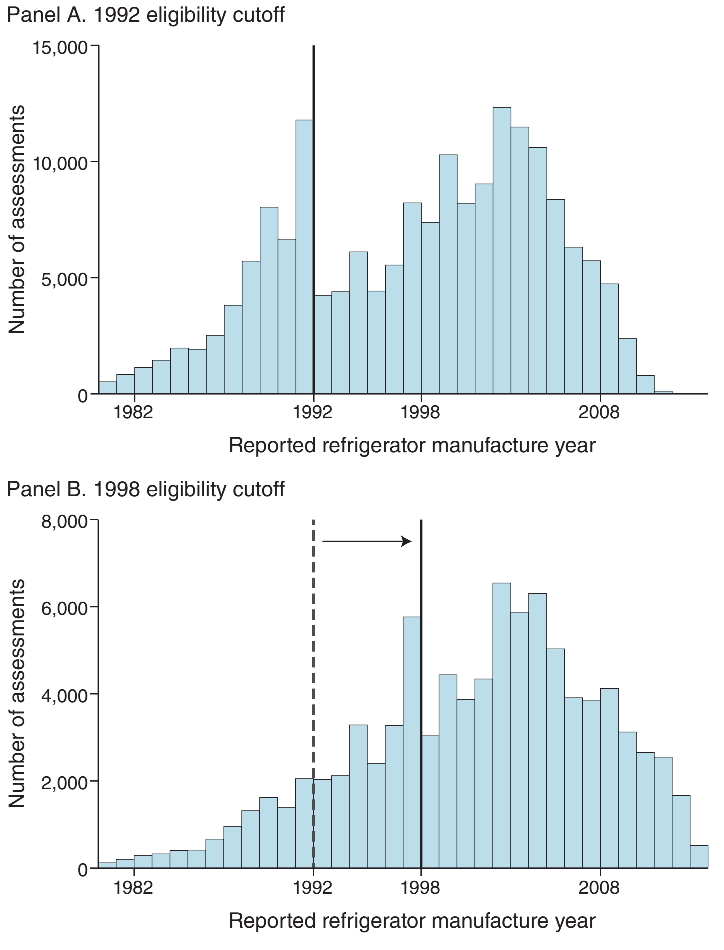

Figure 1. Reported Refrigerator Manufacture Years with 1992 and 1998 Eligibility Cutoffs

Notes: Panel A of this figure shows the reported refrigerator age from 180,105 assessments conducted between January 2009 and August 2012. Households are given free replacement refrigerators if their existing units are reported to be manufactured in 1992 or earlier. Of these assessments, 27.3 percent have reported refrigerator ages to the left of the 1992 cutoff, indicated by the solid vertical line. Panel B shows the reported refrigerator ages from 91,021 assessments conducted between September 2012 and December 2014. These households were eligible for replacements if their existing units were manufactured in 1998 or earlier, indicated by the solid vertical line. The dashed line indicates the previous 1992 refrigerator replacement eligibility cutoff. Both panels show bunching of reported refrigerator manufacturer years to the left of the eligibility cutoff.

2009 and the end of August 2012. Households to the left of the vertical line—those with refrigerators reported to be manufactured in 1992 or earlier—qualified for a free refrigerator replacement in the ESA program, while households to the right, with reported refrigerator manufacture year of 1993 or later, did not. Overall, contractors reported a qualified refrigerator in 27.3 percent of assessments.

Panel A shows substantial bunching to the left of the 1992 threshold. About 12,000 assessments had a reported refrigerator manufacture year of 1992 compared with just over 4,000 reported for 1993. Next, I test whether the bunching is due to contractor manipulation or reflects the true distribution of assessed refrigerators. It could be the case, for example, that households with pre-1993 refrigerators disproportionately enrolled in the program. To examine alternate explanations, I leverage a change in the ESA refrigerator replacement program. Starting in September 2012, the refrigerator eligibility cutoff year moved from 1992 to 1998. Panel B of Figure 1 shows the distribution of contractor-reported refrigerator ages for the 91,021 assessments conducted between September 2012 and December 2014. The solid vertical line indicates the new cutoff, and the dashed line represents the old 1992 cutoff. In panel B, the bunching to the left of the 1992 cutoff is no longer present. Instead, bunching appears to the left of the 1998 threshold, with almost 6,000 assessments having a reported refrigerator manufacture year of 1998, compared with just 3,000 manufactured in 1999. The bunching is smaller at the 1998 cutoff in panel B than the 1992 cutoff in panel A. The reduction in misreporting is likely because moving the cutoff caused a higher proportion of assessed households to be qualified for a refrigerator replacement, which could change a contractor's misreporting decisions.

One possible explanation for the bunching in panel A of Figure 1 is that low-income households with refrigerators manufactured in 1992 or earlier enrolled in the early phase of the ESA program at high rates knowing they could get free replacements. There are four reasons this is unlikely. First, the 1992 eligibility cutoff was not advertised.[21] A potential enrollee could have determined the refrigerator eligibility cutoff with some searching, but a large portion of the eligible population would need to have been informed about the cutoff to generate this bunching pattern.

Second, the change in bunching from the 1992 cutoff to the 1998 cutoff happened discontinuously in September 2012, the month the eligibility cutoff shifted. Figure 2 shows the timing of this changeover. The vertical axis reports the ratio of 1992 assessments to 1993 assessments conducted each month, which is a measure of bunching at the 1992 cutoff. The month the eligibility cutoff shifted to 1998, the assessments no longer exhibited this bunching to the left of the 1992 eligibility cutoff. It would have been difficult for enrolling households to learn about the changing program guidelines and adjust their behavior so rapidly, suggesting that it is contractors—who were aware of the eligibility shift in advance—adjusting their reporting behavior to the new guidelines.[22]

Third, if newly eligible households did respond to the eligibility shift shown in panel B, one would expect to see a jump in the number of 1993–1997 refrigerators. Instead, the only visible discontinuous increase is in refrigerators reported to be manufactured in 1998.

Fourth, the neighboring utility San Diego Gas and Electric (SDG&E), which runs an ESA program with the same 1992 eligibility cutoff for refrigerator replacements, does not exhibit the same bunching in assessments to the left of the 1992 eligibility cutoff. The only difference between the two programs is that SDG&E used a different contract structure that does not have the same incentives to misreport

[21] Online Appendix D.2 provides an example of an ESA brochure. It does not mention the refrigerator eligibility cutoff.

[22] Online Appendix D.3 discusses this changeover in more detail and formally estimates a regression discontinuity in time at the changeover date.

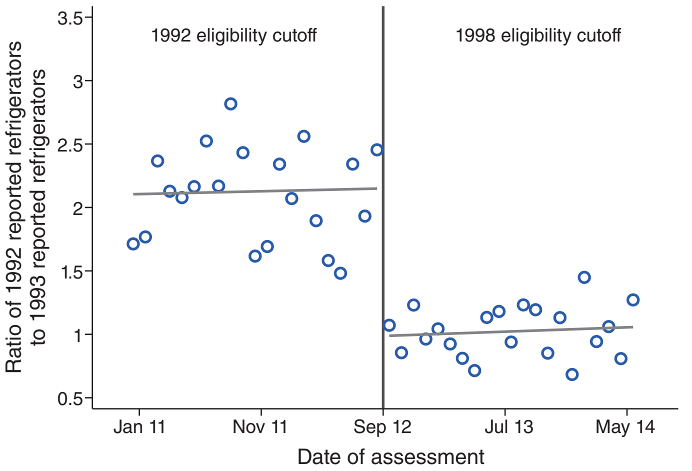

Figure 2. Contractor Reporting Behavior and the Change in Refrigerator Eligibility

Notes: This figure shows how contractors adjust their behavior in response to the eligibility cutoff shifting from 1992 to 1998. The vertical axis reports the ratio of 1992 assessments to 1993 assessments each month—a measure of bunching around the 1992 cutoff. Each dot represents about 200 assessments, and the horizontal lines are from a local linear regression described in online Appendix D.3. The discontinuity is significant with a coefficient of -1.16 and a standard error of 0.19.

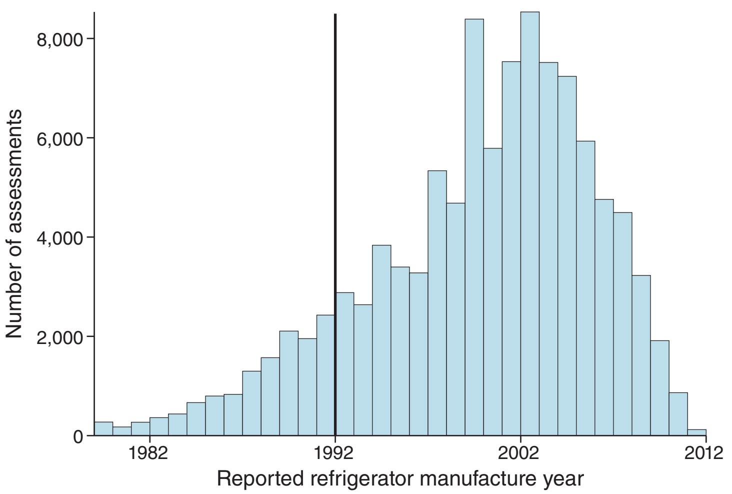

Figure 3. ESA Program Administered by Neighboring Utility SDG&E

Notes: This figure shows ESA assessment data from 106,179 assessments in the neighboring utility (SDG&E) between January 2009 and August 2012, where contractors did not have an incentive to intentionally misreport during the assessment step. The data were collected during the same time period as in panel A of Figure 1, but there is no bunching to the left of the 1992 eligibility threshold.

during the assessment step.[23] Both utilities were subject to the same rules from the regulator, and the eligible households live in similar regions. Figure 3 displays data from 106,179 assessments conducted during the same period by contractors hired by SDG&E. There is no bunching to the left of the 1992 eligibility cutoff in the SDG&E program, which shows that households with qualifying refrigerators did not select into that program. The lack of bunching in the neighboring SDG&E program suggests that the bunching in the SCE program in Figure 1 is caused by the incentives SCE contractors had to misreport.

C. *Estimating Misreporting*

The ESA program's change in eligibility requirements in September 2012 provides an opportunity to estimate the proportion of the ESA population with refrigerators eligible for upgrades at the time of the January 2009 to August 2012 assessments. Establishing a credible estimate for the proportion of eligible refrigerators is central to quantifying contractor misreporting. To generate this counterfactual, I use assessment data from 2013, after the refrigerator eligibility cutoff was moved to 1998 and contractors no longer had the incentive to misreport newer refrigerators as being manufactured in 1992 or earlier.[24] Figure 2 shows that starting in September 2012, the reported manufacture years no longer bunch around the 1992 cutoff, as was observed in the assessments collected between January 2009 and August 2012. Instead, contractors reported 1992 and 1993 refrigerators in equal proportions.

The 2013 assessment data will not reflect the 2009 to 2012 time period because some refrigerators will break and be replaced during that intervening year. In the 2013 assessment data, 11.6 percent of assessments recorded a household as having a refrigerator manufactured in 1992 or earlier, which reflects the refrigerator stock in 2013, but not in 2009 to 2012. If I were to use the raw 2013 assessment data, I would likely overstate the misreporting rate. To account for refrigerator retirements, I adjust the 2013 data by subtracting one from each reported manufacture year in the 2013 data, which shifts the full distribution down a year. This approach gives a counterfactual for the 2009 to 2012 time period of 13.8 percent of refrigerators manufactured in 1992 or earlier.[25] An alternate approach is to use the data from SDG&E shown in Figure 3, which shows that 12.8 percent of ESA households had a refrigerator manufactured in 1992 or earlier. While both are similar, I prefer the adjusted SCE data to reflect the refrigerator stock in SCE's service territory since the two programs were run differently and serve different populations.[26]

[23] SDG&E had one contractor firm that was responsible for administering the program and conducting all the assessments. Online Appendix D.4 provides a further discussion of assessment contractor incentives in the SDG&E ESA program.

[24] After the eligibility change, contractors have the incentive to report a refrigerator was manufactured in 1998 or earlier, which removes the incentive to misreport around the old 1992 cutoff. The 2013 data do not include any households that participated in the ESA program between 2009 and 2012. The ESA program blocked households from participating a second time unless their prior participation was before 2003.

[25] See online Appendix D.5 for a detailed discussion of the adjustment.

[26] A two-year adjustment to the 2013 data results in a counterfactual of 16.2 percent of households qualifying for a refrigerator replacement. I use the one-year adjustment because it is similar to the SDG&E estimate.

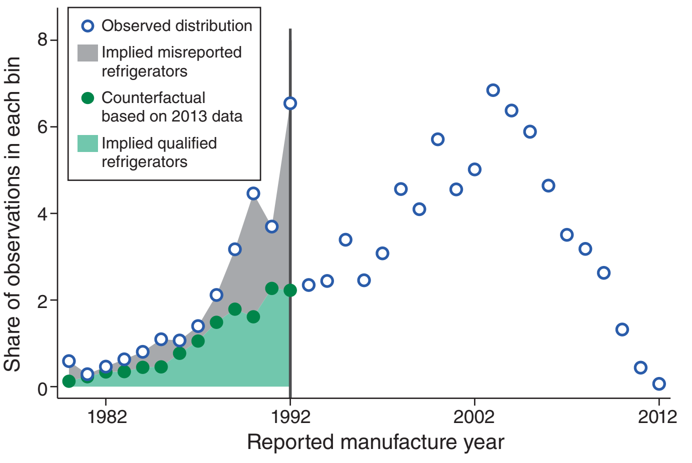

Figure 4. Reported Refrigerator Manufacture Years and Counterfactual

Note: This figure shows the reported refrigerator ages with blue hollow dots and the counterfactual based on 2013 data with solid green dots. The light green area below the green dots is the implied distribution of qualified refrigerators among ESA eligible households and is equal to 13.8 percent of assessments conducted between 2009 and 2012. The gray area represents the implied misreported refrigerators. It is the difference between the observed distribution and the counterfactual distribution and totals 13.5 percent of the assessments conducted. The counterfactual is calculated using data collected in 2013 when contractors did not have the incentive to misreport around the old 1992 eligibility threshold.

Figure 4 shows the observed distribution of the 2009 to 2012 data with hollow blue dots and the counterfactual distribution from the one-year adjusted 2013 data with solid green dots. The observed distribution is the same as panel A of Figure 1, with the vertical axis rescaled to the share of observations in each refrigerator manufacture year bin. The lower area, shaded light green, is the 13.8 percent of households in the ESA program that should have qualified for a replacement. This is in contrast to the 27.3 percent of households that contractors reported being eligible. The gray shaded area represents the implied misreported refrigerators, which total 13.5 of the assessments conducted between 2009 and 2012.

I test the statistical significance of the 13.5 percent misreporting rate using a *t*-test to estimate whether the proportion of assessments reporting a manufacture year of 1992 or earlier is the same in the observed and counterfactual distributions. I account for correlation within each firm by clustering at the firm level. Because there are only 22 firms in my sample, I use the wild cluster bootstrap, finding that the misreporting rate is significant at the 1 percent level (Cameron, Gelbach, and Miller 2008).[27]

The shape of the distributions in Figure 4 provide two additional insights. First, the majority of the misreporting was done with assessments indicating a manufacture year of 1989 to 1992. This could reflect contractors preferring to misreport

[27] I implement the t-test using a regression of a dummy variable equal to one if the assessment reported the refrigerator was manufactured in 1992 or earlier on a dummy of the observation being from the 2013 counterfactual period.

within a few years of the 1992 threshold.[28] Second, the calculated counterfactual appears to line up with the observed distribution on the right side of the 1992 cutoff. This is not a necessary condition for a reasonable counterfactual because the misreported refrigerators are drawn from the years above 1992, which could lead to a lower share of reported manufacture years in 1993 than the 1992 counterfactual. The smoothness through the 1992 cutoff suggests that contractors misreported refrigerators from many years above the 1992 cutoff.

This approach also allows for quantification of the misreporting without relying on statistical manipulation tests (e.g., McCrary 2008; Frandsen 2017).[29] I use the 2013 data to construct a counterfactual with few functional form assumptions or expectations about the shape of the misreporting. Figure 4 suggests contractors are misreporting refrigerators to be manufactured as far back as 1989. Statistical manipulation tests may not effectively capture this type of misreporting so far from the threshold and would likely undercount the misreporting. My counterfactual also captures contractors taking shortcuts and misreporting earlier refrigerators (e.g., pre-1990) as being manufactured in 1992, while statistical manipulation tests could count this behavior as additional misreporting.

Another advantage of using the 2013 data to calculate the counterfactual is that it accounts for unintentional misreporting that occurs due to error. If these errors are random and unrelated to the 1992 eligibility cutoff, they should, on average, occur at the same rate in the 2013 data. The error rate in the predicted distribution will be the same as in the reported distribution, allowing estimates of unintentional misreporting to be netted out. This allows me to attribute the estimated misreporting to the principal-agent problem and contractor incentives to maximize their compensation, not to errors in the assessment process.

D. *Model Numbers and Strategic Contractor Behavior*

SCE instructed contractors to record the refrigerator model number during an assessment. The model number data, when accurately reported, could be used to verify the reported refrigerator manufacture year.[30] However, contractors omit the model number 10.1 percent of the time. I examine the extent to which omitted model numbers are correlated with misreporting.

Panel A of Figure 5 shows the 161,955 assessments where contractors reported model numbers. There is less bunching to the left of the 1992 cutoff, with only 23.6 percent of households eligible for a replacement. The smaller distortion implies a misreporting rate of 9.7 percent, which is significant at the 1 percent level, and is 3.8 percentage points lower than the overall sample. This shows that contractors

[28] If a contractor was caught, a smaller difference between the true age and misreported manufacture year may make it easier to claim that the misreporting was accidental. I find anecdotal evidence of contractors preferring smaller misreporting jumps in a subsample where I can precisely identify misreporting behavior.

[29] These tests use smooth approximations of the variable of interest near a threshold to test for manipulation. They are commonly used to test the validity of regression discontinuity designs, but Frandsen (2017) can also be used to estimate the fraction of missing observations at the threshold. An RD approach yields a similar misreporting rate near the threshold.

[30] SCE has yet to use the model numbers for this purpose, but the potential for retrospective auditing could affect contractor behavior.

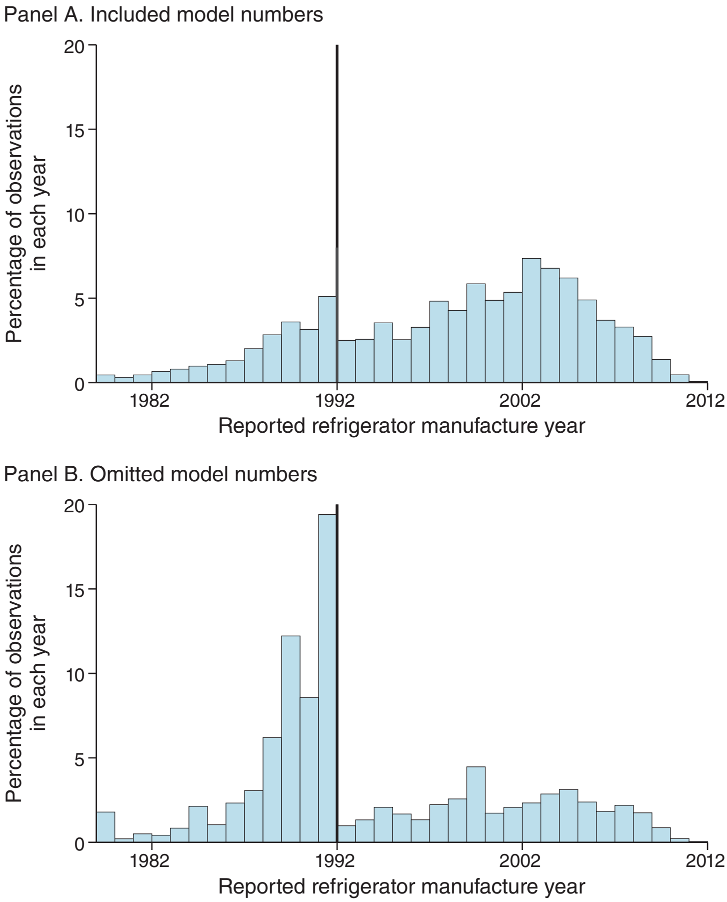

Figure 5. Assessments with Included versus Omitted Model Numbers

Notes: This figure shows the distribution of reported refrigerator manufacture years based on whether the contractor reported a model number on the assessment. Panel A shows the 161,955 assessments where contractors reported model numbers, which has an implied misreporting rate of 9.7 percent. Panel B shows the 18,150 assessments where contractors omitted the model numbers, which has an implied misreporting rate of 47.1 percent. The vertical line in both panels indicates the 1992 refrigerator replacement eligibility cutoff. The vertical axis reports the percentage of observations in each year to facilitate comparison across the panels.

are less likely to intentionally misreport an assessment when they report a model number.

Panel B shows the 18,150 assessments where contractors did not record model numbers. I find more bunching of assessments to the left of 1992 cutoff than in the full sample, with 60.9 percent of these assessments having a reported refrigerator manufacturer year before this date. The large distortion implies a misreporting rate of 47.1 percent, which is significant at the 1 percent level, and suggests that when contractors intentionally misreport, they are also more likely to omit the model numbers.[31] About 22 percent of refrigerator replacements in the ESA program were

[31] All the contractors engaged in this behavior; the results are not driven by a small number of contractors.

conducted based on assessments that did not report a model number, even though they represented only one-tenth of the assessments.

One explanation for the high level of implied misreporting in panel B is that the model numbers are missing from the refrigerators. Over time, model number labels could become hard to read, which would lead to more pre-1993 refrigerators having model numbers that are not visible. However, once the eligibility threshold shifted from 1992 to 1998, the percent of assessments with omitted model numbers that reported manufacture years before 1993 dropped from 60.9 to 18.3. This discontinuous 42.6 percentage point drop suggests that the change in the eligibility cutoff caused the contractors to adjust their model number omission behavior. The fact that contractors choose to disproportionately omit model numbers when misreporting the refrigerator manufacture years provides further evidence that they are trying to increase their compensation while minimizing the chance of being detected.[32]

E. *Contract Structure and Misreporting*

Next, I leverage the two types of contractors that SCE used to implement the ESA program to identify the intentional misreporting rates for the integrated- and separated-task assessments (Z_I and Z_S). Figure 6 shows all the assessments, split by whether the assessment was conducted by a separated-task contractor (i.e., one of the 11 firms that performs *only* assessments) or an integrated-task contractor (i.e., one of the 11 firms that *also* performs replacements). Panel A shows the distribution for 90,084 integrated-task contractor assessments, and panel B shows the distribution for 90,021 separated-task contractor assessments. The bunching of reported refrigerator manufacturer years is larger for the integrated assessments, with 32.9 percent of all assessments having a recorded refrigerator age of 1992 or earlier compared with 21.8 percent for the separated assessments. Both of these are higher than the predicted eligibility rate of 13.8 percent, suggesting that both types of contractors misreport. I find a misreporting rate of 19.0 percent for the integrated contracts, which corresponds to Z_I in the model.[33] The misreporting rate for separated contracts (Z_S) is less than half the size, at 7.8 percent, which confirms the assumption in the model that $Z_S < Z_I$ and reflects that separated-task contractors only receive $25 for misreporting.[34] Both of the misreporting estimates are significant at the 1 percent level, and I can reject that integrated- and separated-task contractors misreport at the same rate.

Households are assigned to contractors in a quasi-random manner. When SCE passes along the contact information of households that are enrolling in the SCE program to contractors (Step 2 in Section IB), it does so using a round-robin algorithm based on where the household is located and which contractors have available

[32] It does not appear that the occupant is playing a role in the contractor's misreporting decision. Contractors misreport at similar rates for owners and renters even though owners have a larger incentive to encourage misreporting.

[33] There is a range in misreporting rates across contractors. Some separated-task contractors misreport on over 15 percent of their assessments, while some integrated-task contractors misreport on less than 12 percent. I report the implied misreporting rate for each of the anonymized 22 contractors in online Appendix D.6.

[34] Separated-task contractors did not know who would conduct the upgrade, making it difficult to negotiate side payments. I verify the absence of side payments by examining the relationship between separated-task contractor misreporting and who conducts the upgrade, finding no pattern.

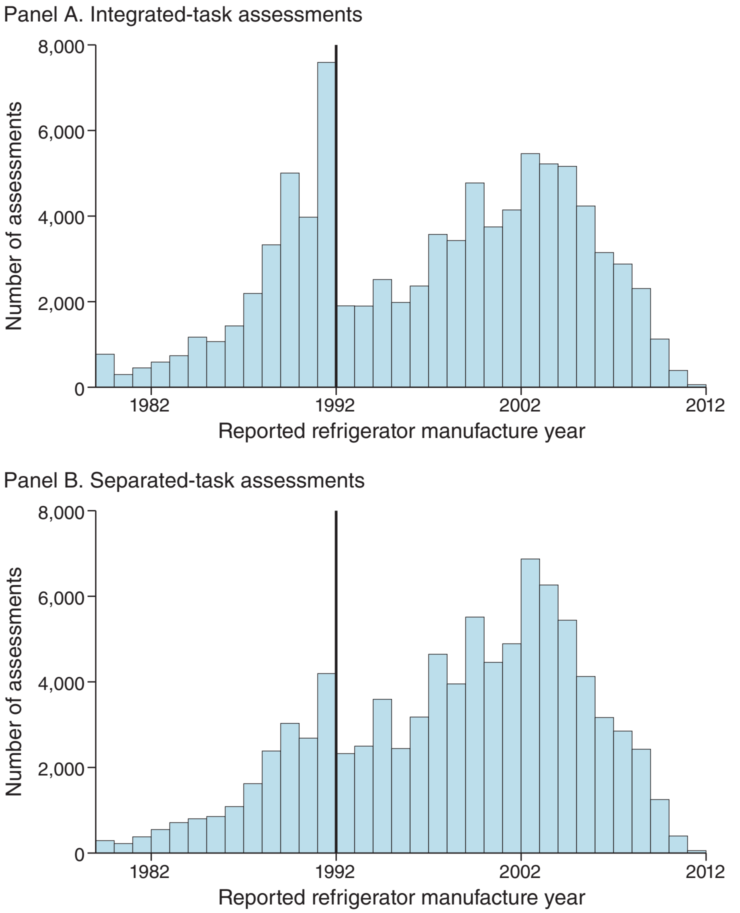

Figure 6. The Effect of Contract Type on Misreporting

Notes: This figure shows the reported refrigerator manufacture year by which type of contractor conducted the assessment. Panel A shows the 90,084 assessments that were conducted by integrated-task assessments. Integrated-task contractors have an implied misreporting rate of 19.0 percent. Panel B shows the 90,021 assessments that were conducted by separated-task contractors. Separated-task contractors have an implied misreporting rate of 7.8 percent. The vertical line in both panels indicates the 1992 refrigerator replacement eligibility cutoff.

capacity.[35] This assignment mechanism prevents contractors from targeting specific types of households for ESA upgrades.[36]

A potential concern is that contractors with the sophistication or propensity to misreport selected into the integrated-task contractor role. The observed differences in misreporting between integrated- and separated-task contractors could then be caused by selection. However, the ESA program did not allow contractors to pick their roles, mitigating selection concerns.

[35] The round-robin contract-assignment process is described in more detail in online Appendix D.7.

[36] I find that, conditional on being declared eligible during the assessment step, refrigerator replacements are completed at the same rate for integrated- and separated-task contractors. I also find that there are no major differences across contract types or whether the dwelling is rented versus owned. See online Appendix D.8 for a detailed breakdown.

To become a contractor in the ESA program, contractors filled out a form and added themselves to the ESA vendor list. There was no application or competitive solicitation. Many of the contractors in the ESA program already worked for existing SCE low-income service programs. When SCE needed more ESA services, they reached out to contractors on the vendor list and offered all potential contractors the same pricing.[37]

Once SCE selected an interested contractor to provide services through the ESA program, SCE assigned them a geographic service area and determined if they would be an integrated- or separated-task provider. SCE allocated responsibilities and the integrated- or separated-task contract status to contractors based on annual and regional ESA household participation targets provided by the CPUC.[38] The only eligibility requirement for contractors to install refrigerators was that a contractor had a general building license.[39]

The top-down assignment of contract type from SCE prevented contractors with a desire to misreport from choosing the integrated-task role and removed selection concerns. It also allowed for a comparison of misreporting rates between similar separated- and integrated-task contractors. I group contractors into three broad classes based on the services they offer *outside* of the ESA program. Online Appendix A7 provides a breakdown based on three broad classifications that I term "general low-income services," "utility subsidy contractors," and "general contractors."[40]

Breaking down contractors by the type of work they do provides two main insights. First, the misreporting rate varies across different types of contractors. Among general low-income service providers, separated- and integrated-task providers misreport at 17.5 percent and 36.2 percent, respectively, which is higher than the overall misreporting rates. Second, within all three contractor classifications, integrated-task contractors misreport at higher rates. Even if contractor altruism toward low-income households is responsible for some of the misreporting, the level of misreporting is still influenced by the level of incentive.

IV. Estimating Changes in Electricity Consumption

To quantify the cost of the principal-agent problem in the ESA program, I estimate the effect of refrigerator replacements on electricity consumption. I use model number data to separately identify the causal effect of a qualified and an unqualified refrigerator replacement on household electricity consumption. I use these results to show how intentional contractor misreporting reduces the electricity savings from refrigerator replacements in the ESA program.

[37] Most ESA contractors participated in the program before the 2009 to 2012 study window.

[38] Details on the SCE program structure and contractor participation were provided by the CPUC on background.

[39] Occasionally, a licensed electrician was needed to ground the refrigerator's outlet.

[40] I make these classifications using the contractor firm's name to search for other services they provide. Online Appendix D.10 shows these classifications by contract type and the number of assessments conducted and the implied misreporting rates by group.

A. *Electricity Consumption Data*

To estimate how a refrigerator replacement reduces electricity consumption, I combine ESA program data with confidential monthly household-level electricity consumption data. I merge model number data, reported by contractors during the assessment step, with external data on refrigerator characteristics to determine whether a refrigerator replacement was qualified or unqualified. From this model number data, I verify the refrigerator manufacture year for 57,847 households using exact matches.[41] I combine the ESA data from these households with monthly electricity consumption data from 2007 to 2014. I restrict the sample to households for which I have at least six months of consumption data before and after the ESA assessment was conducted, which removes households that moved shortly before or after enrolling in the ESA program and limits the sample to 38,008 households. My primary specification examines electricity usage in the two years before and after enrolling in the ESA program, leaving me with 1,648,915 observations.[42]

B. *Empirical Strategy*

This section describes the equation used to estimate the causal impact of a refrigerator replacement on household electricity consumption. The main specification is a panel fixed effects regression with a rich set of weather controls, household fixed effects, and time fixed effects that vary by region. Causal identification comes from comparing electricity use before and after a refrigerator replacement with electricity use by the control group that participates in the ESA program but does not receive a refrigerator replacement. In particular, I estimate

$$\text{(2)} \qquad \mathrm{Q}_{icmy} = \beta_1 \mathbf{1}\left[\text{refrigerator} = 1\right]_{imy} + \sum_u \beta_u \mathbf{1}\left[\text{upgrade} = \text{u}\right]_{imy} + CDD_{imy} + HDD_{imy} + \zeta_{myc} + \delta_{im} + \epsilon_{icmy},$$

where Q_{icmy} is the electricity consumption in kWh for customer i in climate zone c in month m and year y. Climate zones are designated by the California Energy Commission and represent geographic regions with similar energy usage and climates.[43] The variable $\mathbf{1}[\text{refrigerator} = 1]_{imy}$ is an indicator equal to 1 once a

[41] I could not match the model numbers to a large portion of the assessments because of incomplete or imprecisely recorded model numbers. I could not match refrigerators that did not have a reported model number. See online Appendix E.1 for a description of how I match model numbers and the effects of unmatched model numbers. I hand-verified a subset of the refrigerators to check how often the matching process produced errors and found an error rate of less than 1 percent. I also conduct a robustness check using a less conservative model matching approach in online Appendix F.5. I do not use the matched model numbers to calculate misreporting rates because, as shown in Section IIID, contractors are more likely to omit the model number when they are misreporting. Online Appendix E.4 compares the full sample of ESA households used in Section 3 to the subset used to estimate changes in electricity consumption.

[42] The average household in my sample has around 20 months of data before it enrolled in the program and the same amount after enrollment. The results are robust to including all available pre- and post-enrollment consumption data. See online Appendix E.2 for more details on the data cleaning.

[43] Climate zones are used to set building standards and other regulations. There are nine climate zones in SCE's service territory.

household's refrigerator was replaced by the ESA program. β_1, the coefficient of interest, represents the average change in monthly household electricity consumption in kWh caused by a refrigerator replacement. The variable $\mathbf{1}[\text{upgrade} = \text{u}]_{imy}$ is a separate indicator for the installation of each non-refrigerator upgrade that is conducted by the ESA program.[44] CDD_{imy} and HDD_{imy} are controls for monthly cooling degree days and heating degree days.[45] ζ_{myc} is a climate zone by month-of-sample fixed effect to control for contemporaneous regional shocks to electricity consumption. δ_{im} is a set of household-month-of-year fixed effects that flexibly controls for time-invariant household characteristics. Each household has a separate fixed effect for each month of the year (m) because household consumption varies seasonally throughout the year.

ϵ_{icmy} is the error term. The panel nature of this analysis makes the errors potentially correlated both within household over time and across households. I two-way cluster at the household and month-of-sample levels (Cameron, Gelbach, and Miller 2011). As a result, the standard errors are robust to both within-household and within-month correlation.

One potential concern in identifying β_1 is that households may have experienced other changes at the time of their refrigerator replacements that also could have affected electricity consumption, and could have resulted in changes in electricity consumption being misattributed to the refrigerator replacement.

The identifying assumption is that a refrigerator replacement is uncorrelated with electricity consumption conditional on the other ESA upgrades and fixed effects. Formally, this is written as $\text{cov}(\mathbf{1}[\text{refrigerator} = 1]_{imy}, \epsilon_{icmy} | X_{icmy}) = 0$, where X_{icmy} represents the other upgrades a household received, weather controls, and fixed effects. In other words, I assume households that received a refrigerator replacement had parallel trends in consumption to those that did not, and that these trends would have continued in the absence of an upgrade. To support this parallel-trends assumption, I use an event study in Section IVD and conduct robustness checks with household-specific time trends. I examine the pre-upgrade trends in electricity consumption by group in online Appendix F.1, finding similar patterns in pre-upgrade electricity usage across the treatment and control groups.[46]

C. *Sample Selection*

I report regression results for two different treatment groups: qualified replacements and unqualified replacements. I use model number data to classify replacements as qualified or unqualified based on the true manufacture year of the existing refrigerator at the time of the assessment. I classify a replacement as qualified or unqualified if I can verify its model number was manufactured before or after the 1993 eligibility cutoff. Table 1 shows average pre-period summary statistics.

[44] See online Appendix Table F11 for the non-refrigerator upgrades that were conducted.

[45] Weather data were obtained from MesoWest. See online Appendix E.3 for more details.

[46] A potential source of bias would be households with different unobserved time-varying trends in electricity consumption systematically selecting into the refrigerator replacement or control group. However, this is unlikely given the limited influence households have in the implementation of the ESA program and the analysis of pre-upgrade electricity consumption trends in Section IVD.

Table 1—Summary Statistics for ESA Participating Households in Regression Sample

Variable	Group 1 qualified replacements	Group 2 unqualified replacements	Group 3 no replacements (control)	*p*-value of difference: Group 1 versus Group 3	*p*-value of difference: Group 2 versus Group 3
Pre-period use (kWh/month)	476	455	437	<0.01	0.01
	(229)	(233)	(233)		
Average monthly CDD	95	98	91	<0.01	<0.01
	(55)	(53)	(53)		
Average monthly HDD	148	172	147	0.20	<0.01
	(64)	(70)	(63)		
Proportion renter	0.64	0.67	0.66	0.01	0.43
	(0.48)	(0.47)	(0.47)		
Proportion single-family home	0.68	0.69	0.66	<0.01	0.02
	(0.47)	(0.46)	(0.47)		
Proportion disabled resident	0.09	0.10	0.07	<0.01	<0.01
	(0.29)	(0.30)	(0.26)		
Proportion English speaker	0.56	0.50	0.48	<0.01	0.19
	(0.50)	(0.50)	(0.50)		
Households	3,715	1,261	33,032		
Observations	161,818	54,557	1,432,540		
Refrigerator manufacture year (model number verified)	pre-1993	1993–2012	1993–2012		

Notes: This table shows pre-period summary statistics broken down by the three groups used in the empirical analysis. Standard deviations are shown in parentheses. Group 1 consists of households that qualified for and received refrigerator replacements. Group 2 consists of households that were not qualified under ESA program rules to receive refrigerators but received replacement refrigerators because of intentional contractor misreporting. I differentiate between these two groups using the model numbers reported during the assessment step. Group 3 consists of control households that went through the ESA program and did not receive replacements because they did not qualify. The right two columns report p-values of the difference between each treatment group and the control group. CDD and HDD signify cooling and heating degree days, respectively. Pre-period use is the average monthly electricity consumption for all households in each group before they enroll in the ESA program.

Group 1 consists of the 3,715 households that received qualified replacements because their existing units were manufactured in 1992 or earlier.[47] Group 2 consists of the 1,261 unqualified households that received refrigerators because the contractors misreported during the assessment step. Group 3 is the control group, which consists of 33,032 households that went through the ESA program but did not receive refrigerators because the contractors correctly reported that their existing units were manufactured in 1993 or later.[48]

Table 1 shows statistically significant differences in observable characteristics between each of the treatment groups (Group 1 and Group 2) and the control group. In particular, households that receive qualified upgrades (Group 1) consumed significantly more electricity in the pre-ESA upgrade period than the control group. This difference in consumption is partially attributable to the fact that Group 1

[47] It is possible that a contractor could misreport both the refrigerator manufacture year and model number. This would lead me to classify an unqualified replacement in Group 1, which would bias downward the savings estimates from qualified replacements. I do not find any evidence of this behavior, but I cannot rule it out.

[48] In the primary specifications, I do not include the households that were deemed eligible (either qualified or unqualified) but ultimately did not receive an upgrade. The results are unchanged if those households are included.

households had older refrigerators that consumed more electricity. Group 2 also consumed more electricity than the control group, which is likely because these households live in warmer regions than the control group.[49] The difference in electricity usage, however, will not bias the estimate of β_1, because differences in levels are absorbed by the household-month-of-year fixed effects. I also directly control for cooling and heating degree days to account for weather. All other differences in time-invariant characteristics, such as proportion renters and English speakers, will also be absorbed by the household-month-of-year fixed effect.[50]

D. *Event Study: Support for Parallel Trends*

This section presents graphical results from an event study specification, which supports the parallel trends assumption. I estimate the event study design with the following equation:

$$(3)\quad Q_{icmy} = \sum_{q=-8}^{8} \beta_q \mathbf{1}\left[\text{quarter to enrollment} = q\right]_{imy} + \sum_{u} \beta_u \mathbf{1}\left[\text{upgrade} = u\right]_{imy} + CDD_{imy} + HDD_{imy} + \zeta_{myc} + \delta_{im} + \epsilon_{icmy},$$

where $\mathbf{1}[\text{quarter to enrollment} = q]_{imy}$ is an indicator in event time.[51] The household enrolls in the ESA program in quarter $q = 0$. Treatment effects are measured relative to the quarter before enrollment $q = -1$. The specification includes the eight quarters before and after enrollment. All other variables are the same as the panel fixed effects design shown in equation (2).

Figure 7 shows graphically—in two separate event study regressions—the causal effect of a qualified and an unqualified refrigerator replacement. The qualified replacement results compare households that received qualified replacements (Group 1) with those that did not qualify and did not receive replacements (Group 3). The unqualified replacement results compare households that received unqualified replacements (Group 2) with the same control group (Group 3). The horizontal axis is time in quarters before and after enrollment. Each dot represents the average change in monthly household kWh consumption compared with the consumption by the control group, and the vertical gray lines show 95 percent confidence intervals.

The treatment effect for both regressions in the months leading up to enrollment is flat and statistically indistinguishable from zero. This flat pre-treatment result suggests that households in the treatment group are similar to households in the control group, conditional on controls and fixed effects. It provides additional supporting evidence that this specification satisfies the parallel trends assumption, and that the estimates reflect the causal estimate of a refrigerator replacement.[52]

[49] Electricity use during the winter months is the same for Group 2 and the control group.

[50] Most of the differences in observable characteristics, while statistically different because of the large sample, are not economically meaningful.

[51] I use quarter of enrollment instead of quarter of refrigerator replacement because the quarter of enrollment better lines up the treatment group and the control group in event time.

[52] I use a placebo event study in online Appendix F.2 to show that enrolling in the program does not cause a change in electricity use for the control group.

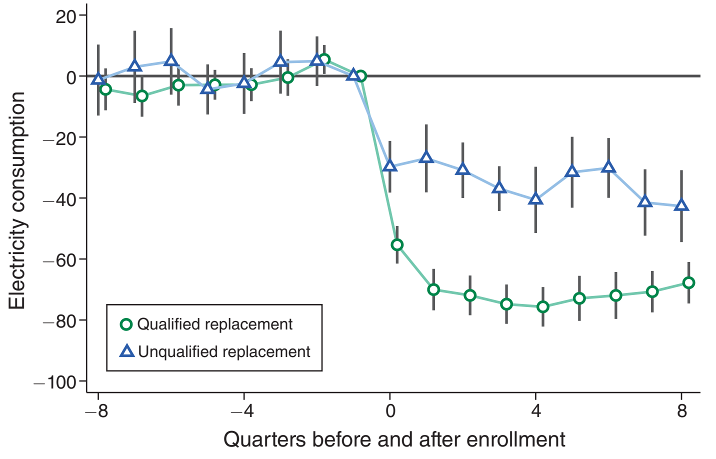

Figure 7. Electricity Savings from Qualified and Unqualified Refrigerator Replacements

Notes: This figure shows two event studies for refrigerator replacements, estimated using equation (3). The qualified replacement event study, indicated with circles, shows the electricity savings caused by qualified refrigerator replacements. The unqualified replacement event study, indicated with triangles, shows the electricity savings caused by unqualified refrigerator replacements. The results show that qualified replacements save about twice as much electricity as unqualified replacements. All results are in kWh per month and are relative to the quarter before enrollment. Both regressions control for non-refrigerator upgrades conducted in the ESA program and include weather controls, household-month-of-year fixed effects, and month-of-sample by climate-zone fixed effects. Standard errors are clustered at the household and month-of-sample levels.

The results show that a qualified refrigerator replacement provides statistically significant electricity reductions between 70 and 75 kWh/month, or around 16 percent of average monthly consumption. The quarter of enrollment in the ESA program ($q = 0$) has a smaller savings estimate than the months that follow. Not all refrigerator replacements are completed in the first three months after enrollment, so the full electricity savings are not realized until the next quarter. The savings estimates for the quarters after enrollment are flat, suggesting a stable long-run treatment effect for at least two years.[53]

The unqualified refrigerator replacement results (indicated with triangles) show that unqualified refrigerator replacements save about half as much electricity as a qualified replacement. Starting in 1993, all new refrigerators were subject to more stringent federal minimum efficiency standards, making them more efficient than similar refrigerators sold in 1992. In other words, the ESA program specifically excluded post-1992 refrigerators because of the timing of the standards. My results estimate that replacing a pre-1993 refrigerator saves substantially more electricity than replacing a post-1992 refrigerator. This means that contractor misreporting led to significantly smaller reductions in electricity consumption.

[53] I estimate the event study with a longer time horizon and find that the results are stable and persist for five years, the maximum I can test with the available data.

TABLE 2—THE EFFECT OF REFRIGERATOR REPLACEMENTS ON MONTHLY ELECTRICITY CONSUMPTION

	Qualified replacements (1)	Unqualified replacements (2)
Refrigerator replacement	−73.45	−38.02
	(2.68)	(3.20)
Cooling degree days	0.12	0.11
	(0.02)	(0.02)
Heating degree days	0.03	0.03
	(0.01)	(0.01)
Controls for other upgrades	Yes	Yes
Pre-period consumption (kWh/month)	448	445
Refrigerator replacements	3,715	1,261
Households	36,747	34,293
Observations	1,581,024	1,474,841

Notes: This table reports regression coefficients from two regressions using equation (2). The dependent variable in both regressions is monthly household electricity consumption. Refrigerator replacement is an indicator for a household's refrigerator being replaced. The refrigerator replacement coefficient in Column 1 shows that a qualified replacement causes a 73 kWh/month reduction in monthly consumption. Column 2 shows that an unqualified replacement causes a 38 kWh/month reduction in monthly consumption. Both column 1 and column 2 use the same control group of customers that went through the ESA program but did not qualify for refrigerator replacements. Both regressions control for non-refrigerator upgrades conducted in the ESA program and include weather controls, household-month-of-year fixed effects, and month-of-sample by climate-zone fixed effects. Standard errors are in parentheses and are clustered at the household and month-of-sample levels.

E. *Panel Fixed Effects Results*

Table 2 shows the results from estimating the effects of a refrigerator replacement using equation (2). Column 1 reports the changes in monthly electricity consumption from a qualified refrigerator replacement. The first row shows that a qualified refrigerator replacement causes a 73 kWh/month reduction in electricity consumption, or 16.4 percent of the average household's pre-upgrade electricity consumption.[54] These estimates are consistent with engineering predictions of the impacts of replacing a refrigerator manufactured in 1992 or earlier with a new energy-efficient model.[55]

Column 2 of Table 2 estimates the same regression for unqualified refrigerator replacements using sample Groups 2 and 3. I find that replacing a refrigerator manufactured in 1993 or later reduces consumption by 38 kWh/month.[56] The estimated savings are consistent with the engineering estimate prediction of replacing refrigerators manufactured after 1992 with new energy-efficient units. The savings from an unqualified replacement are half as large as and statistically different from the

[54] Online Appendix F.4 reports the individual coefficients for each of the ESA upgrades.

[55] There are no official ex ante energy savings engineering estimates for the ESA program because the projected savings are based on an ex post evaluation (Evergreen Economics 2013). As a result, there is no energy efficiency gap because the projected savings are from a retrospective analysis that likely includes qualified and unqualified refrigerator replacements. I calculate the engineering predictions using the average increase in minimum efficiency standards over time. See online Appendix A for more details. The similarity of the engineering projections to the estimated effects of a qualified replacement suggests that model numbers in Group 1 are accurately reported and that contractors are not misreporting both the manufacture year and the model number.

[56] With a less conservative model matching approach, I include about 14,000 more households and find similar savings estimates. See online Appendix F.5 for these results.

savings from a qualified replacement.[57] This finding shows how intentional contractor misreporting can reduce the electricity savings from a refrigerator replacement.

I conduct a number of robustness checks in online Appendix F, all of which support the above results. Online Appendix F.6 shows the results estimated with household-specific time trends, which yield quantitatively similar savings estimates. Online Appendix F.7 shows the main specification results for the subset of households that only received CFLs or refrigerators to verify that potential contractor misreporting on other major upgrades is not biasing the refrigerator savings estimates. The estimates in online Appendix F.7 are similar to the main results in Table 2. Online Appendix F.8 breaks down the results in columns 1 and 2 by the minimum efficiency standard of the refrigerator that was replaced. The results show that replacing older refrigerators saves more electricity. Replacing a refrigerator manufactured before 1990 saves 85 kWh/month. In contrast, replacing a refrigerator manufactured in 2001 or later reduces consumption by only 20 kWh/month.[58] Online Appendix F.3 considers the average energy savings of a refrigerator replacement in the ESA program by estimating equation (2) for a larger sample of 28,800 ESA refrigerator replacements where I cannot verify their manufacture year. I find that in the estimation sample, the average refrigerator saves 61 kWh/month, which shows that there is a mix of qualified and unqualified replacements being conducted in the program.[59]

V. Costs of Misreporting

A. *Benefits and Costs of Refrigerator Replacements*

In this section, I use the electricity savings estimates from Section IVE to calculate the social net benefit of a qualified and unqualified refrigerator replacement. I capture the key net benefits of refrigerator replacements in equation (4).[60] The net benefit measure is intended to capture the sum of all consumer and producer surplus and externality reduction benefits while subtracting the capital replacement costs and efficiency costs of funding the program through higher electricity bills. I assume that there are no income effects from a refrigerator replacement and that any remaining unmeasured benefits and costs are small.[61]

[57] See online Appendix Table F17 for the two effects estimated jointly.

[58] I also run a specification with the control group limited to households with refrigerators manufactured between 1993 and 1996. I find almost identical energy savings to Table 2, suggesting that households with newer refrigerators that did not receive replacements are not driving the results.

[59] The estimation sample does not include all of the 41,529 refrigerator replacements in the ESA program in the study window. Households were not included in the estimation sample if they did not have sufficient billing data. See online Appendix F.3 for a discussion.

[60] I do not include any additional costs for paying contractors to conduct unqualified replacements. If, for example, SCE had additional costs associated with paying contractors for fraudulent installations, then the net benefits of unqualified replacements would be lower.

[61] Homeowners and renters have the same energy savings from a refrigerator replacement, which suggests owners are not experiencing a substantial income effect from receiving a new refrigerator that could lead to increased energy consumption. See online Appendix G for a detailed discussion of the assumptions behind the net benefit calculation, including the potential for contractor market power. I assume the market for refrigerators is perfectly competitive and that there are no producer surplus benefits for refrigerator manufacturers from the ESA program.

Table 3—Net Benefits per Refrigerator Replacement

	Qualified replacement (1)	Unqualified replacement (2)
Electricity bill savings	\$414	\$288
Reduction in externalities	\$78	\$54
Electricity grid fixed costs not recovered	−\$166	−\$115
Capital replacement and labor costs	−\$266	−\$333
Distortion from raising funds	−\$0.03	−\$0.03
Net benefit per refrigerator replacement	\$60	−\$106

Notes: This table shows the net benefits of qualified and unqualified replacements broken down into the components in equation (4). All dollar values are net present values using a 3 percent annual real discount rate. The bottom row shows that the net benefits of a qualified replacement are \$60 and the net benefits of an unqualified replacement are −\$106.

$$\text{(4)} \quad \Delta\text{Net benefits} = \text{Electricity bill savings} + \text{Reduction in externalities} \\ - \text{Electricity grid fixed costs not recovered} \\ - \text{Capital replacement and labor costs} \\ - \text{Distortion from raising funds.}$$

Each component of the net benefits of a refrigerator replacement depends on household refrigerator replacement behavior in the absence of the ESA program.[62] I construct this non-ESA counterfactual using the ESA program assessment data, which allows for a nuanced characterization of when refrigerators would have been scrapped and replaced in the absence of the ESA program.[63] I find that 42 percent of qualified and 55 percent of unqualified refrigerator replacements would still be in operation 5 years after their replacement date in the absence of the ESA program.[64]

Table 3 provides a breakdown of the net benefits of qualified and unqualified refrigerator replacements in equation 4 as net present values using a 3 percent annual real discount rate.[65] Each component is calculated by comparing realized outcomes to the non-ESA counterfactual. The first row is the electricity bill savings of households that received replacements, which I calculate using the regression estimates in Table 2 separately for qualified and unqualified replacements. I find that qualified refrigerator replacements saved on average 3,677 kWh and unqualified refrigerator replacements saved on average 2,574 kWh. The discounted stream of electricity bill benefits saved households with qualified replacements \$414 and unqualified replacements \$288.

[62] There are many reasons why a household might not have replaced their old existing refrigerator despite potential financial benefits, including liquidity constraints and the landlord-tenant problem.

[63] A common assumption in the literature is that appliance replacement programs accelerate replacements by five years.

[64] The assessment data capture ESA-eligible households purchasing used refrigerators or keeping older refrigerators. See online Appendix G.1 for a discussion of household behavior absent the ESA program and how I calculate the counterfactual refrigerator scrappage and replacement.

[65] I find similar results using a range of discount rates in online Appendix G.4

The second row in Table 3 shows the dollar value for reducing CO_2 emissions and local air pollution externalities associated with the electricity reductions. Using a $50 social cost of carbon and local air pollution costs in California from Muller, Mendelsohn, and Nordhaus (2011), I calculate that qualified and unqualified replacements reduced externality damages by $78 and $54 over the lifetime of the replacements, with 84 percent of those benefits coming from CO_2 reductions. The third row shows the electricity grid fixed costs not recovered, reflecting how SCE collects revenue to pay for fixed transmission and distribution infrastructure costs. SCE bills residential customers per kWh, and there is no fixed monthly fee. As a result, when SCE sells less electricity because of a refrigerator replacement, it raises less revenue to cover its fixed electricity grid costs. To make up for the lost revenue, SCE will likely raise electricity prices in the long run. (Borenstein and Bushnell 2022).

The sum of the first three rows equals the total societal net benefit of the electricity reductions from a refrigerator replacement, which is equal to the electricity savings valued at the social marginal cost of electricity generation of $0.10/kWh (Borenstein 2012).[66] The total social net benefit of the electricity reductions for qualified and unqualified refrigerators equals $326 and $227, respectively, showing that replacing unqualified refrigerators provides substantially fewer electricity reduction benefits.

The fourth row in Table 3 shows the capital replacement and labor costs, which capture the capital and labor costs of scrapping an old, but working, refrigerator and moving the purchase of a replacement refrigerator to the present. I use the non-ESA counterfactual to determine when a household would replace their refrigerator in the absence of the ESA program.[67] For simplicity, I assume the ESA program provides the same refrigerator that households would eventually purchase when replacing their own refrigerators in the absence of the program.[68] Unqualified replacements have a higher capital replacement and labor cost because the unqualified refrigerators would have lasted longer before being replaced in the absence of the ESA program. The fifth row shows the distortion from increasing electricity prices a small amount to raise money to pay for the refrigerators. While important to account for, this cost is only $0.03 per refrigerator because electricity demand is inelastic.[69]

The last row in Table 3 shows the sum of the top five rows, and it shows the net benefits of qualified and unqualified refrigerator replacements. Qualified refrigerator replacements have a net benefit of $60 per replacement, while unqualified replacements have a net benefit of −$106 per replacement. The large difference between

[66] See online Appendix G.2 for a detailed discussion of the social marginal cost and electricity grid fixed costs not recovered.

[67] If contractors do not scrap old refrigerators, the net benefit calculation would depend on how the old refrigerator is used. Not removing the existing refrigerator is risky for a contractor because a post-upgrade audit conducted by the ESA program could catch them. I examine non-scrappage online Appendix G.3, finding no evidence.

[68] See online Appendix G.2 for a discussion of the capital replacement and labor cost calculation. See online Appendix G.3 for an alternate formulation where households in the counterfactual purchase refrigerators that met the minimum standards. It is possible that in the absence of the ESA program, households would purchase a replacement refrigerator that better fits their preferences than the standard ESA-provided model. If that were the case, then the net benefits of the ESA program would be lower for both qualified and unqualified replacements.

[69] In California, the price of electricity is higher than the social marginal cost (Borenstein and Bushnell 2022). Any price increase will inefficiently reduce electricity consumption. See online Appendix G.3 for a detailed discussion of the three costs associated with refrigerator replacement.

qualified and unqualified replacements shows high costs to the principal-agent problem in the ESA program.[70]

There are limitations to what a conventional cost-benefit calculation can measure. There may be non-energy benefits and costs of a refrigerator replacement that are not easily measured and are not included in the calculation in equation (4). Non-energy benefits include quieter operation, more consistent cooling, and better internal organization. Non-energy benefits are on average larger for qualified replacements, which are older and likely less desirable refrigerators. Non-energy costs can come from the ESA program providing the lowest-cost base model refrigerators without any extra features.[71] Many of these non-energy costs and benefits would have occurred in the absence of the ESA program when households eventually replaced their refrigerators, suggesting that they are not large enough to affect the cost-benefit analysis meaningfully.[72]

B. *Program Cost-Benefit Ratio*

The previous section calculates the net benefit of an individual qualified or unqualified refrigerator replacement. It is a useful comparison when determining the cost of misreporting, but it does not illustrate the overall impact of misreporting on the ESA refrigerator replacement program. In this section, I use a program benefit-cost ratio to show the effectiveness of the ESA program as implemented.

The ESA program replaced 41,529 refrigerators between 2009 and 2012. The results in Section IIIC suggest that a little less than half of the replacements were unqualified. Using these estimates, I find that the ESA refrigerator replacement program had a cost-benefit ratio of 0.93, meaning that for every dollar put into the program, there were $0.93 of benefits. Ideally, an energy efficiency program would have a cost-benefit ratio of at least 1, which would justify the program expenditure and make the program cost effective.

It is possible to make an out-of-sample calculation and consider what the refrigerator replacement program would look like if it conducted only qualified replacements. I find that conducting only qualified replacements would yield a benefit-cost ratio of 1.22, which is 29 percentage points higher than the actual program.[73] This large difference demonstrates the importance of the principal-agent problem and contract structure for cost-effectiveness.

[70] See online Appendix G.2 for a discussion of the assumptions used in each component of Table 3. I consider alternate counterfactual scenarios where refrigerators are replaced at different rates than predicted by the assessment data, finding similar results.

[71] In a small survey, 24 percent of respondents said their replacement was lower quality than their previous refrigerator (Research into Action 2011). See online Appendix G.3 for a discussion of how this could affect the net benefit calculations.

[72] Non-energy benefits and costs would not change the relative advantage of qualified refrigerator replacements over unqualified replacements.

[73] An ESA program that only provided unqualified replacements has a cost-benefit ratio of 0.68.

Table 4—Estimated and Derived Parameters to Calculate Benefits of Integrated-Task Contract

Parameter	Value	Origin of estimate
Predicted qualified eligibility rate (α)	0.138	Calculated from ESA data in Section IIIC
Separated misreporting rate (Z_S)	0.078	Calculated from ESA data in Section IIIE
Integrated misreporting rate (Z_I)	0.19	Calculated from ESA data in Section IIIE
Net benefits of unqualified replacement	\$106	Calculated in Section VA using regression results
Referral payment	\$25	ESA documentation and author's calculations
Refrigerator installation payment	\$98	ESA documentation and author's calculations
Integrated misreporting cost parameter (γ_I)	646	Calculated in Appendix G.7
Separated misreporting cost parameter (γ_S)	319	Calculated in Appendix G.7

Note: This table shows the parameters used to calculate the benefits of an integrated-task contract from equation (1). The parameters are derived empirically from program data and first-order conditions from the model.

C. *Calculating Benefits of Integration*

This section combines the empirical analysis done in Sections III and IV, the benefit-cost analysis in Section VA, and the model in Section II to estimate how the principal-agent problem interacts with contract type. I derive most of the key parameters empirically, and I take two parameters on the benefits to contractors of misreporting from ESA program documentation.

The first parameter is that contractors are paid \$25 to install five CFL light bulbs when they find a household eligible for a refrigerator replacement, which is the referral payment in Section IIA.[74] The second parameter is that contractors are paid \$224 for each refrigerator replacement, which covers the costs of installing the new unit along with removing and recycling the old unit in an environmentally appropriate manner (SCE 2012a). However, the contractor's expected profit of a refrigerator replacement is lower than \$224 for four reasons. First, when an old refrigerator is removed, the contractor must pay a disposal fee, which in the SCE service territory is \$27 (WM 2018). Second, there is a labor cost for providing the replacement, which I estimate at \$30 per replacement.[75] Third, only 84.2 percent of eligible households based on the assessments receive replacements.[76] Fourth, logistical constraints may prevent the integrated-task contractor who conducts the assessment from providing the replacement. The same integrated-task contractor does both steps 69 percent of the time, and contractors do not know which replacements they will be assigned to during the assessment step.[77] Accounting for all these factors, integrated-task contractors receive in expectation \$123 per refrigerator they report as eligible for replacement, while separated-task contractors receive only \$25.[78]

[74] SCE makes bulk purchases of refrigerators and CFLs, which it gives directly to the contractors to install. On the ride-alongs I participated in, I observed that it did not take contractors much time or effort to replace the CFLs once they had already conducted the home assessment.

[75] I base the labor cost estimate on the appliance delivery fee charged by major appliance retailers. In online Appendix G.5 I conduct robustness checks using labor costs of zero and \$60.

[76] See online Appendix D.8 for a discussion of why all eligible households do not receive refrigerator replacements. Integrated-task and separated-task contractors complete 84.9 and 83.0 of replacements in eligible households. The small difference suggests that contract structure is not affecting refrigerator installation completions.

[77] Much of this discrepancy is due to scheduling issues. See online Appendix G.6 for details.

[78] Integrated-task contractors receive \$25 to install five CFLs and \$98 in expectation for a refrigerator replacement.

Table 4 shows the parameters I use for the calculation of equation (1) and how I estimate them. I cannot measure the information acquisition cost, which is the benefits from having the same contractor responsible for both assessment and refrigerator replacement, with the available data.[79] Instead, I first use the model to calculate how large the information acquisition cost needs to be to justify an integrated-task contract. In the base case using the parameters displayed in Table 4, I estimate that the benefits would have to exceed \$95 for an integrated-task contractor to be preferred. Because contractors are paid only \$70 to conduct an assessment (Steps 3–4 in Section IB), it is unlikely that the information acquisition cost is \$95.

I next calculate the benefits of integration using a \$35 information acquisition cost. I find a benefit of integration of −\$12.32 per assessment. In total, integrated contract assessments reduce net benefits in the ESA program by \$1,109,000 across 90,084 assessments. The distortion introduced by using an integrated-task contract rather than a separated-task contract is equal to 16 percent of the total benefits of the refrigerator replacement program. The results show that separating the assessment and refrigerator replacement between two different contractors is beneficial.[80]

D. *Policy Fixes*

SCE continues to administer the ESA program under the supervision of the CPUC, which sets program guidelines but is not involved in the day-to-day implementation. The CPUC could choose not to renew the ESA program, but that is unlikely given that the program remains politically popular and the CPUC authorized it through 2026. The CPUC could also direct SCE to change the program implementation to mitigate misreporting. There are two main approaches to reducing the misreporting in the ESA program: monitor agent actions or fully separate the task of assessment and service provision between two agents.

After reading an earlier version of this paper, SCE decided to implement a monitoring system without a direct mandate from the CPUC, suggesting SCE wanted to fix the problem.[81] The monitoring system requires contractors to submit a picture of the serial number plaque for each assessed refrigerator. SCE plans to audit a small proportion of these pictures to verify that contractors are submitting accurate data. Photo-based monitoring should lead to a large drop in misreporting because contractors exhibited strategic reporting behavior in the existing program (e.g., Section IIID), suggesting that they will respond to the introduction of accountability and potential penalties.

Photo-based monitoring has disadvantages. First, agents can evade the monitoring by substituting the picture from a qualified refrigerator.[82] Second, contractors may shift their misreporting from refrigerators to other unmonitored measures. In the

[79] See online Appendix G.8 for a discussion of the information acquisition cost and the benefits of integration in ESA Program.

[80] The results are robust to different information acquisition costs. See online Appendix G.7 for details.

[81] SCE likely has an incentive to avoid conflict with the CPUC when possible because SCE requires regulatory approval for most policy changes.

[82] It would take a sophisticated contractor a significant effort for this type of misreporting because they would also have to manipulate the metadata (location, time, etc.) embedded in the picture they submitted.

ESA program, there is suggestive evidence that contractors are more likely to install upgrades that earn them higher profits. For example, only four firms are qualified to replace central air systems, giving them a larger incentive to find a household qualified for a replacement. I find that when one of these central-air-installing firms conducts an assessment, households in eligible regions receive a new central air system 30.3 percent of the time, compared with 23.3 percent of the assessments conducted by firms that are not qualified to install central air systems.[83] Any reallocation of misreporting may undermine some of the benefits of monitoring refrigerator installations.

Another approach to reduce overall misreporting is to fully separate the assessment from the service provision. In the ESA program, this could be accomplished with a modified separated-task contract where the assessment contractor does not have the $25 incentive to misreport. Full task separation would reduce the incentive to misreport on all upgrades, not just the ones that can be directly monitored. It could also introduce task coordination costs associated with having different firms conduct each step. It is not possible to determine if monitoring or task separation is better for the ESA program or other principal-agent problems without further research.[84]

VI. Conclusion

In this paper, I measure the costs of a principal-agent problem. In the empirical setting I study—an energy efficiency appliance replacement program—contractors intentionally misreport assessment data to provide unqualified refrigerator replacements and increase their compensation. This profit-seeking agent behavior has significant costs: I estimate that each unqualified replacement reduces net benefits by $106. In contrast, each replacement that follows program rules saves twice as much electricity and increases net benefits by $60. I find that the principal-agent problem makes an otherwise cost-effective refrigerator replacement program no longer be cost-effective.

These findings shed light on a key question in the energy efficiency literature: why do energy efficiency programs deliver lower savings than predicted? I provide the first evidence that the principal-agent problem may be an important part of the explanation. These findings are applicable to a wide range of energy efficiency policies, because similar incentives exist for contractors to maximize their income during program implementation. Other programs may use monitoring to avoid the misreporting problems observed in this paper, but monitoring every decision made by a contractor in an energy efficiency program can be expensive and impractical. Burlig et al. (2020) study a program that subsidizes energy-efficient capital upgrades, finding savings that were only 51 percent of ex ante expectations. In this type of program, utility-licensed contractors recommend upgrades from a menu of

[83] Households were only eligible for central air replacements in the hottest regions of SCE's service territory. The central-air-replacing firms have the ability to install refrigerators but install them at a lower rate than other firms. This behavior suggests that central-air-replacing firms may have shifted misreporting towards central air replacements, which earns them $2,000 in installation costs per replacement, and away from refrigerators.

[84] In online Appendix G.9 I outline how generalizable the results and the policy fixes in this paper are to other settings using the SANS conditions outlined in List (2020).

options. It is likely that they focus on installing upgrades that are more profitable for them to install but yield lower savings.

More broadly, I am able to quantify the costs of misaligned incentives between principals and agents. The principal-agent problem is a well-understood theoretical concept, but it has proved challenging to quantify its effects. Because asymmetric information is at the core of the principal-agent problem, it is difficult for the researcher to observe agent behavior and quantify the costs of self-interested agent actions. The existing literature primarily identifies how agents distort their behavior in a variety of contexts including commission-based sales behavior, teacher performance pay, and auto repair. The degree to which the observed distortions cause significant costs has generally remained an unanswered question. My research provides answers to this question in one setting, showing that self-interested agents can substantially reduce program benefits. Further research is necessary to learn how task assignment can better align the incentives of principals and agents in different settings. As this paper demonstrates, failing to consider the incentives of agents can have large costs and undermine policy goals.

REFERENCES

Allcott, Hunt, and Michael Greenstone. 2012. "Is There an Energy Efficiency Gap?" *Journal of Economic Perspectives* 26 (1): 3–28.

Allcott, Hunt, and Michael Greenstone. 2017. "Measuring the Welfare Effects of Residential Energy Efficiency Programs." NBER Working Paper 23386.

Asch, Beth J. 1990. "Do Incentives Matter? The Case of Navy Recruiters." *ILR Review* 43 (3): 89–106.

Baker, George P. 1992. "Incentive Contracts and Performance Measurement." *Journal of Political Economy* 100 (3): 598–614.

Bandiera, Oriana, Iwan Barankay, and Imran Rasul. 2007. "Incentives for Managers and Inequality among Workers: Evidence from a Firm-Level Experiment." *Quarterly Journal of Economics* 122 (2): 729–73.

Blonz, Joshua A. 2023. "Replication data for: The Costs of Misaligned Incentives: Energy Inefficiency and the Principal-Agent Problem." American Economic Association [publisher], Inter-university Consortium for Political and Social Research [distributor]. https://doi.org/10.38886/E166342V1.

Borenstein, Severin. 2012. "The Private and Public Economics of Renewable Electricity Generation." *Journal of Economic Perspectives* 26 (1): 67–92.

Borenstein, Severin, and James B. Bushnell. Forthcoming. "Do Two Electricity Pricing Wrongs Make a Right? Cost Recovery, Externalities, and Efficiency." *American Economic Journal: Economic Policy* 14 (4): 80-110.

Burlig, Fiona, Christopher Knittel, David Rapson, Mar Reguant, and Catherine Wolfram. 2020. "Machine Learning from Schools about Energy Efficiency." *Journal of the Association of Environmental and Resource Economists* 7 (6): 1181–217.

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller. 2008. "Bootstrap-Based Improvements for Inference with Clustered Errors." *Review of Economics and Statistics* 90 (3): 414–27.

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller. 2011. "Robust Inference with Multiway Clustering." *Journal of Business and Economic Statistics* 29 (2): 238–49.

Consortium for Energy Efficiency. 2018. *State of the Efficiency Program Industry*. Boston: CEE.

Christensen, Peter, Paul Francisco, Erica Myers, and Mateus Souza. 2021. "Decomposing the Wedge between Projected and Realized Returns in Energy Efficiency Programs." *Review of Economics and Statistics* 1–46.

Courty, Pascal, and Gerald Marschke. 2003. "Dynamics of Performance-Measurement Systems." *Oxford Review of Economic Policy* 19 (2): 268–84.

Courty, Pascal, and Geraldn Marschke. 2004. "An Empirical Investigation of Gaming Responses to Explicit Performance Incentives." *Journal of Labor Economics* 22 (1): 23–56.

California Public Utilities Commission. 2014. *Resolution E-4700*. San Francisco: CPUC.

California Public Utilities Commission. 2019. *Decision 19-06-022*. San Francisco: CPUC.

Davis, Lucas W. 2012. "Evaluating the Slow Adoption of Energy Efficient Investments: Are Renters Less Likely to Have Energy Efficient Appliances?" In *The Design and Implementation of U.S. Climate Policy*, edited by Don Fullerton and Catherine Wolfram, 301–16. Chicago: University of Chicago Press.

Davis, Lucas W., Alan Fuchs, and Paul Gertler. 2014. "Cash for Coolers: Evaluating a Large-Scale Appliance Replacement Program in Mexico." *American Economic Journal: Economic Policy* 6 (4): 207–38.

Energy Information Administration. 1987. *Residential Energy Consumption Survey*. Washington, DC: Energy Information Administration.

Evergreen Economics. 2013. *PY2011 Energy Savings Assistance Program Impact Evaluation*. Portland, OR: Evergreen Economics https://www.calmac.org/publications/2011_ESA_Impact_Evaluation_FINALES.pdf.

Fowlie, Meredith, Michael Greenstone, and Catherine Wolfram. 2018. "Do Energy Efficiency Investments Deliver? Evidence from the Weatherization Assistance Program." *Quarterly Journal of Economics* 133 (3): 1597–644.

Frandsen, Brigham R. 2017. "Party Bias in Union Representation Elections: Testing for Manipulation in the Regression Discontinuity Design When the Running Variable Is Discrete." In *Regression Discontinuity Designs Volume 38*, edited by Matias D Cattaneo and Juan Carlos Escanciano, 281–315. Bingley, U.K.: Emerald Publishing Ltd.

Gerarden, Todd D., Richard G. Newell, and Robert N. Stavins. 2017. "Assessing the Energy-Efficiency Gap." *Journal of Economic Literature* 55 (4): 1486–525.

Gillingham, Kenneth, Matthew Harding, and David Rapson. 2012. "Split Incentives in Residential Energy Consumption." *Energy Journal* 33 (2): 37–62.

Gillingham, Kenneth, Amelia Keyes, and Karen Palmer. 2018. "Advances in Evaluating Energy Efficiency Policies and Programs." *Annual Review of Resource Economics* 10: 511–32.

Giraudet, Louis-Gäetan, Sébastien Houde, and Joseph Maher. 2018. "Moral Hazard and the Energy Efficiency Gap: Theory and Evidence." *Journal of the Association of Environmental and Resource Economists* 5 (4): 755–90.

Holmström, Bengt. 2017. "Pay for Performance and Beyond." *American Economic Review* 107 (7): 1753–77.

Holmström, Bengt, and Paul Milgrom. 1991. "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design." *Journal of Law, Economics, and Organization* 7: 24–52.

Jacob, Brian A., and Steven D. Levitt. 2003. "Rotten Apples: An Investigation of the Prevalence and Predictors of Teacher Cheating." *Quarterly Journal of Economics* 118 (3): 843–77.

Kerr, Steven. 1975. "On the Folly of Rewarding A, While Hoping for B." *Academy of Management Journal* 18 (4): 769–83.

Lazear, Edward P. 2000. "Performance Pay and Productivity." *American Economic Review* 90 (5): 1346–61.

Lazear, Edward P. 2018. "Compensation and Incentives in the Workplace." *Journal of Economic Perspectives* 32 (3): 195–214.

Levinson, Arik, and Scott Niemann. 2004. "Energy Use by Apartment Tenants When Landlords Pay for Utilities." *Resource and Energy Economics* 26 (1): 51–75.

Lewis, Gregory, and Patrick Bajari. 2011. "Procurement Contracting with Time Incentives: Theory and Evidence." *Quarterly Journal of Economics* 126 (3): 1173–211.

Lewis, Gregory, and Patrick Bajari. 2014. "Moral Hazard, Incentive Contracts, and Risk: Evidence from Procurement." *Review of Economic Studies* 81 (3): 1201–28.

List, John A. 2020. "Non Est Disputandum de Generalizability? A Glimpse into The External Validity Trial." NBER Working Paper 27535.

Loftus, Peter J., Armond M. Cohen, Jane C. S. Long, and Jesse D. Jenkins. 2015. "A Critical Review of Global Decarbonization Scenarios: What Do They Tell Us about Feasibility?" *Wiley Interdisciplinary Reviews: Climate Change* 6 (1): 93–112.

Marion, Justin, and Jeremy West. 2023. "Socioeconomic Disparities in Privatized Pollution Remediation: Evidence from Toxic Chemical Spills." Unpublished.

McCrary, Justin. 2008. "Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test." *Journal of Econometrics, The Regression Discontinuity Design: Theory and Applications*, 142 (2): 698–714.

Metcalf, Gilbert E., and Kevin A. Hassett. 1999. "Measuring the Energy Savings from Home Improvement Investments: Evidence from Monthly Billing Data." *Review of Economics and Statistics* 81 (3): 516–28.

Muller, Nicholas Z., Robert Mendelsohn, and William Nordhaus. 2011. "Environmental Accounting for Pollution in the United States Economy." *American Economic Review* 101 (5): 1649–75.

Myers, Erica. 2020. "Asymmetric Information in Residential Rental Markets: Implications for the Energy Efficiency Gap." *Journal of Public Economics* 190: 104251.

Novan, Kevin, and Aaron Smith. 2018. "The Incentive to Overinvest in Energy Efficiency: Evidence from Hourly Smart-Meter Data." *Journal of the Association of Environmental and Resource Economists* 5 (3): 577–605.

Oyer, Paul. 1998. "Fiscal Year Ends and Nonlinear Incentive Contracts: The Effect on Business Seasonality." *Quarterly Journal of Economics* 113 (1): 149–85.

Prendergast, Canice. 1999. "The Provision of Incentives in Firms." *Journal of Economic Literature* 37 (1): 7–63.

Research into Action. 2011. *Low Income Energy Efficiency Program 2009-2010 Process Evaluation*. Portland, OR: Research into Action.

Reynaert, Mathias, and James M. Sallee. 2021. "Who Benefits When Firms Game Corrective Policies?" *American Economic Journal: Economic Policy* 13 (1): 372–412.

SCE, Southern California Edison. 2012. *Southern California Edison Company's (U 338-E) 2012 Annual Report for 2011 Low Income Programs*. Rosemead, CA: SCE.

SCE, Southern California Edison. 2012b. The Joint Utilities' 2011 Amended Report on CARE Eligibility and Penetration Rates. Rosemead, CA: SCE https://liob.cpuc.ca.gov/wp-content/uploads/sites/14/2020/12/3-20-12-ESAP-CARE-2012-2014-Joint-Parties-Amended-CARE-Eligibility-Estimates.pdf.

Shearer, Bruce. 2004. "Piece Rates, Fixed Wages and Incentives: Evidence from a Field Experiment." *Review of Economic Studies* 71 (2): 513–34.

Tayan, Brian, The Wells Fargo Cross-Selling Scandal. 2019. Rock Center for Corporate Governance at Stanford University Closer Look Series: Topics, Issues and Controversies in Corporate Governance No. CGRP-62 Version 2, Stanford University Graduate School of Business Research Paper No. 17-1. https://ssrn.com/abstract=2879102.

WM, Waste Management. 2018. *Lancaster Landfill & Recycling Center Service and Fee Schedule.* Antelope Valley, CA: WM.

Zivin, Joshua Graff, and Kevin Novan. 2016. "Upgrading Efficiency and Behavior: Electricity Savings from Residential Weatherization Programs." *Energy Journal* 37 (4): 1-23.