

# Smart Building Energy Efficiency via Social Game: A Robust Utility Learning Framework for Closing-the-Loop

Ioannis C. Konstantakopoulos, Lillian J. Ratliff, Ming Jin, Costas Spanos, S. Shankar Sastry

**Abstract**—Given a non-cooperative, continuous game, we describe a framework for parametric utility learning. Using heteroskedasticity inference, we adapt a Constrained Feasible Generalized Least Squares (cFGLS) utility learning method in which estimator variance is reduced, unbiased, and consistent. We extend our utility learning method using bootstrapping and bagging. We show the performance of the proposed method using data from a social game experiment designed to encourage energy efficient behavior amongst building occupants. Using occupant voting data we simulate the game defined by the estimated utility functions and show that the performance of our robust utility learning method and quantify its improvement over classical methods such as Ordinary Least Squares (OLS).

## I. INTRODUCTION

Smart buildings are a fundamental component in constructing *smart cities*; their efficient design and operation enables flexibility and reliability in making urban spaces sustainable. It is well known that energy consumption of buildings, both residential and commercial, accounts for approximately 40% of all energy usage in the U.S. [1]. There have been many approaches to improve energy efficiency of buildings through control and automation, e.g., [2]–[7], as well as incentives and pricing, e.g., [8], [9].

The incentive for a building manager to encourage energy efficient behavior may be that they are accountable for the cost or are required to maintain an operational excellence measure. Beyond these motivations, demand response programs are being rolled out by utility companies and third-party solution providers—e.g. companies such as Ohmconnect—with the goal of correcting for improper load forecasting (see, e.g., [10], [11], [12]). In such a program, consumers enter into a contract in which they agree to change their demand when *demand response events* are called. The building manager may be required to keep this schedule. By capitalizing on new technological advances that enable smart building automation, the prescribed schedule can be met. Through automation and integration of the end-user, smart buildings play an integral role in creating a more sustainable and efficient *smart city*.

\*This work is supported by NSF (CNS-1239166) FORCES and the Republic of Singapore’s National Research Foundation through a grant to the Berkeley Education Alliance for Research in Singapore (BEARS) for the Singapore-Berkeley Building Efficiency and Sustainability in the Tropics (SinBerBEST) Program. BEARS has been established by the University of California, Berkeley as a center for intellectual excellence in research and education in Singapore.

\*\*The authors are with the Department of Electrical Engineering and Computer Sciences at the University of California, Berkeley, Berkeley CA 94720. {ioanniskon, ratliff1, jinming, spanos, sastry}@eecs.berkeley.edu

\*\*\*Ioannis C. Konstantakopoulos is a Scholar of the Alexander S. Onassis Public Benefit Foundation

Our approach to efficient building energy management focuses on *smart* office buildings that utilize building automation. We have designed and implemented a social game aimed at incentivizing occupants to modify their behavior. The impact is two fold: First, the overall energy consumption in the building is reduced and second, the building manager can effectively *close the loop* around the occupants. The latter allows for control of occupant behavior subject to comfort constraints thereby creating more flexibility and accuracy in meeting demand response scheduling.

At the core of our approach is the fact that we modeled the occupants as non-cooperative agents who play according to a *Nash equilibrium strategy*. This serves the purpose of modeling the agents as strategic entities who make decisions based on their own preferences in spite of others. The game theoretic framework both allows for qualitative insights to be made about the outcome of such selfish behavior and, more importantly, can be leveraged in designing *incentives* for modifying occupant behavior.

We designed a utility learning framework based on data generated from the social game experiment. Specifically, assuming a parametric form of utility function for each agent, we utilized Constrained Feasible Generalized Least Squares (cFGLS) to formulate a parameter estimation scheme in which the estimator variance is reduced, unbiased, and consistent. We explore wild bootstrapping, a powerful technique for asymptotic approximation of the bias and standard error of an estimator in a complex and noisy statistical model. Strengthened by the bootstrap estimators, we improve the parameter estimation scheme by using ensemble methods such as bagging. We showed that estimating agent utility functions via the proposed method based on the cFGLS results in a predictive model that outperforms several other standard techniques such as OLS and our preliminary estimation scheme in [13].

The rest of the paper is organized as follows. We begin in Section II by describing the social game and presenting the game formulation. Section III proposes the robust utility estimation approach. We present the results of our estimation scheme in Section IV. We conclude with some discussion and proposal for future work in Section V.

## II. SOCIAL GAME FRAMEWORK

In this section, we briefly describe the social game experimental setup and present a game theoretic framework for occupant decision making. We refer the reader to our previous works [13], [14] for a more detailed description of the former.

### A. Social Game

We have instrumented an office space in Cory Hall on the University of California, Berkeley campus with a heating, ventilation, and air conditioning (HVAC) system, automated lighting control (Lutron system<sup>1</sup>), plug-load metering and carbon dioxide sensors. The social game we designed and implemented consists of occupants voting according to their lighting usage preferences over comfort and productivity. We designed an online platform so that occupants can log their vote, view their points, and observe all occupants' consumption patterns and points. This platform stores past data that is used to estimate occupant behavior.

Occupants win points based on how energy efficient their vote is compared to others. The points are used to determine an occupant's likelihood of winning in a lottery. Specifically, occupants choose a value in the interval  $[0, 100]$  representing their vote for the dim level in their zone as well as neighboring zones. The lighting setting that is implemented in each zone is the average of all the votes weighted according to proximity to that zone. If occupants are not in the office, they are considered *absent* and are not allowed to vote. On the other hand, if they are present, they can *actively* vote according to their preferences or choose to *opt-out*, in which case their vote is taken to be the *default value*; we explored four different default settings,  $\{10, 20, 60, 90\}$ .

### B. Occupant Decision Making

In previous work [13], we modeled the interaction between the building manager and the occupants as a leader-follower(s) game. We designed the incentives by estimating the parameters using an online least squares framework and optimized the points and default value using the estimated utilities of the occupants. This work is based on the theoretical utility learning and incentive design framework presented in [15]. In the present work, we take a step back and re-examine the utility learning step using statistical methods that provide greater accuracy in the estimation and prediction of player decision-making. In future work, we will fold the new estimation scheme into the overall incentive design framework.

To describe the occupant game, let the number of occupants participating in the game be denoted by  $n$ . We model the occupants as utility maximizers having utility functions composed of two terms that capture the tradeoff between comfort and desire to win. We model their comfort level using a Taguchi loss function which is interpreted as modeling occupant dissatisfaction in such a way that it is increasing as variation increases from their desired lighting setting [16]. In particular, each occupant has the following *comfort* function:

$$\psi_i(x_i, x_{-i}) = -(\bar{x} - x_i)^2 \quad (1)$$

where  $x_i \in \mathbb{R}$  is occupant  $i$ 's lighting vote,  $x_{-i} = \{x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n\}$ , and

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad (2)$$

is the average of all the occupant votes and is the lighting setting which is implemented. Hence, this term measures the discomfort an occupant feels given that its vote is  $x_i$  and the state of the environment is actually  $\bar{x}$ . We acknowledge that an occupant may have some internal comfort level that is different than its vote; in particular, the occupant may realize that voting an extreme value pushes the average toward a more desirable setting. We set this type of *gaming* aside for the time being, and focus instead on the unknown preferences between *comfort* and *winning* resulting in a different kind of asymmetric information.

In addition, each occupant has the following *winning* function:

$$\phi_i(x_i, x_{-i}) = -\rho c(x_i)^2 \quad (3)$$

where  $\rho$  is the total number of points distributed by the building manager and  $c$  is a scaling factor which we set to  $10^{-4}$ . The points are distributed by the leader using the relationship

$$\rho = \frac{x_b - x_i}{nx_b - \sum_{j=1}^n x_j} \quad (4)$$

where  $x_b = 90$  is the baseline setting for the lights, i.e. the lighting setting that occurred before the implementation of the social game in the office. In our previous work [14] we modeled the function  $\phi_i$ , i.e. the desire to win, using the natural log of (4). We found that the form of  $\phi_i$  as defined in (3) provides a better estimation and prediction of all the occupant's behavior. It appears that it captures the occupants' perceptions about how the points are distributed and the value of the points as determined by each of the occupants more accurately.

Each occupant's utility function is then given by

$$f_i(x_i, x_{-i}) = \psi_i(x_i, x_{-i}) + \theta_i \phi_i(x_i, x_{-i}) \quad (5)$$

where  $\theta_i$  is parameter unknown to the leader, the source of asymmetric information in the leader-follower(s) game. Hence, the  $i$ -th occupant faces the following optimization problem:

$$\max_{x_i \in S_i} f_i(x_i, x_{-i}) \quad (6)$$

where  $S_i = [0, 100] \subset \mathbb{R}$  is the constraint set for  $x_i$ . The constraint set can be described as follows. Let  $h_{i,j}(x_i, x_{-i})$  for  $j \in \{1, 2\}$  denote the constraints on occupant  $i$ 's optimization problem. In particular, following Rosen [17], for occupant  $i$ , the constraints are

$$h_{i,1}(x_i) = 100 - x_i \quad (7)$$

$$h_{i,2}(x_i) = x_i \quad (8)$$

so that we can define  $\mathcal{C}_i = \{x_i \in \mathbb{R} \mid h_{i,j}(x_i) \geq 0, j \in \{1, 2\}\}$  and  $\mathcal{C} = \mathcal{C}_1 \times \dots \times \mathcal{C}_n$ . In this framework, the occupants are non-cooperative agents in a continuous game with convex constraints. We model their interaction using the Nash equilibrium concept.

**Definition 1:** A point  $x \in \mathcal{C}$  is a **Nash equilibrium** for the game  $(f_1, \dots, f_n)$  on  $\mathcal{C}$  if

$$f_i(x_i, x_{-i}) \geq f_i(x'_i, x_{-i}) \quad \forall x'_i \in \mathcal{C}_i \quad (9)$$

<sup>1</sup><http://www.lutron.com>

for each  $i \in \{1, \dots, n\}$ .

The interpretation of the definition of Nash is as follows: no player can unilaterally deviate and increase their utility. Additional constraints on the parameters  $\{\theta_i\}_{i=1}^n$  ensure that the game is a concave  $n$ -person game on a convex set.

*Theorem 1 (Rosen [17]):* A Nash equilibrium exists for every concave  $n$ -person game.

Define the Lagrangian of each player's optimization problem as follows:

$$L_i(x_i, x_{-i}, \mu_i) = f_i(x_i, x_{-i}) + \sum_{j \in A_i(x_i)} \mu_{i,j} h_{i,j}(x_i) \quad (10)$$

where  $A_i(x_i)$  is the active constraint set at  $x_i$ . The differential game form [18], [15] is given by

$$\omega(x, \mu) = [D_1 L_1(x, \mu_1)^T \cdots D_n L_n(x, \mu_n)^T]^T \quad (11)$$

where  $D_i L_i$  denotes the derivative of  $L_i$  with respect to  $x_i$ .

*Definition 2 (Ratliff [15]):* A point  $x^* \in \mathcal{C}$  is a differential Nash equilibrium for the game  $(f_1, \dots, f_n)$  on  $\mathcal{C}$  if  $\omega(x^*, \mu^*) = 0$ ,  $z^T D_{ii} L_i(x^*, \mu_i^*) z < 0$  for all  $z \neq 0$  such that  $D_i h_{i,j}(x_i^*)^T z = 0$ , and  $\mu_{i,j} > 0$  for  $j \in A_i(x_i^*)$ .

These conditions are sufficient for defining a local Nash equilibrium.

*Proposition 1 (Ratliff [13]):* A differential Nash equilibrium of the  $n$ -person concave game  $(f_1, \dots, f_n)$  on  $\mathcal{C}$  is a Nash equilibrium.

A sufficient condition guaranteeing that a Nash equilibrium  $x$  is isolated is that the Jacobian of  $\omega(x, \mu)$ , denoted  $D\omega(x, \mu)$ , is invertible [15]. We refer to such points as being *non-degenerate*.

### III. ROBUST UTILITY LEARNING

In this section we provide the theoretical formulation of our robust utility estimation method. The main contribution is the implementation of heteroskedasticity inference for correlated errors in the resulting utility regression models. In addition, we infer noise structures using residuals of the constrained least squares fitting.

#### A. Utility Estimation

The utility estimation problem is formulated as a convex optimization problem by using first-order conditions for Nash equilibria [15], [18]. At the observed Nash equilibrium the gradient of each occupant's Lagrangian should be identically zero. We define an  $\varepsilon$ -approximate differential Nash equilibrium as follows:

*Definition 3:* Given  $\varepsilon > 0$ , a point  $x^* \in \mathcal{C}$  is an  $\varepsilon$ -approximate differential Nash equilibrium for  $(f_1, \dots, f_n)$  if  $\omega(x^*, \mu^*) = \varepsilon$ ,  $z^T D_{ii} L_i(x^*, \mu_i^*) z < 0$  for all  $z \neq 0$  such that  $D_i h_{i,j}(x_i^*)^T z = 0$ , and  $\mu_{i,j} > 0$  for  $j \in A_i(x_i^*)$ . This is to essentially say that the first-order condition is *approximately* met.

We assume that each observation  $x^{(k)}$  corresponds to an  $\varepsilon$ -approximate Nash equilibrium where the superscript notation  $(\cdot)^{(k)}$  indicates the  $k$ -th observation. Thus, we can consider first-order optimality conditions for each occupants optimization problem and define a residual function capturing the amount of suboptimality of  $x_i^{(k)}$  [19], [20].

Define the residual of the stationarity and complementary conditions for occupant  $i$ 's optimization problem by

$$r_{s,i}^{(k)}(\theta_i, \mu_i) = D_i f_i(x_i^{(k)}, x_{-i}^{(k)}) + \sum_{j=1}^2 \mu_i^j D_i h_{i,j}(x_i^{(k)}) \quad (12)$$

and

$$r_{c,i}^{j,(k)}(\mu) = \mu_i^j h_{i,j}(x_i^{(k)}), \quad j \in \{1, 2\} \quad (13)$$

respectively.

Define  $r_s^{(k)}(\theta) = [r_{s,1}^{(k)}(\theta_1, \mu_1) \cdots r_{s,n}^{(k)}(\theta_n, \mu_n)]^T$  and  $r_c^{(k)} = [r_{c,1}^{(k)}(\mu_1) \cdots r_{c,n}^{(k)}(\mu_n)]^T$  where  $r_{c,i}^{(k)}(\mu_i) = [r_{c,i}^{1,(k)}(\mu_i) \ r_{c,i}^{2,(k)}(\mu_i)]$  and  $\mu_i = (\mu_i^1, \mu_i^2)$ .

Given the data from the occupants' actions, we solve the following convex optimization problem:

$$\begin{aligned} \min_{\mu, \theta} \sum_{k=1}^K \chi(r_s^{(k)}(\theta, \mu), r_c^{(k)}(\mu)) \\ \text{s.t. } \theta_i \geq \theta_{LB}, \mu_i \geq 0 \quad \forall i \in \{1, \dots, n\} \end{aligned} \quad (P)$$

where  $\theta_{LB}$  is a lower bound for the unknown parameters  $\{\theta_i\}_{i=1}^n$  that ensures the inferred game is concave and  $\chi: \mathbb{R}^n \times \mathbb{R}^{2n} \rightarrow \mathbb{R}_+$  is a nonnegative, convex penalty function satisfying  $\chi(z_1, z_2) = 0$  if and only if  $z_1 = 0$  and  $z_2 = 0$ , i.e. any norm on  $\mathbb{R}^n \times \mathbb{R}^{2n}$ , the inequality  $\mu_i \geq 0$  is element-wise. To determine  $\theta_{LB}$  we utilize the second derivative condition on players' utility functions; in particular, if  $D_{ii}^2 f_i(x) = -2(1 - 1/n)^2 - 2\theta_i c \rho < 0$  for each  $i$ , then the game is concave. Hence,  $\theta_i > -c^{-1} \rho^{-1} (1 - n^{-1})^2$  where the right-hand side is a negative non-increasing function of  $n$ . Thus, using a relaxation, concavity is ensured regardless of the number of players by setting  $n = 2$ , the minimum number of users in a non-cooperative game. Then, given fixed  $\rho$  and  $0 < \zeta \ll 1$ , the lower bound  $\bar{\theta}_{LB} = -0.3571 + \zeta$  will guarantee the estimated game is concave. However, the subgradient projection method applied to the gradient dynamics  $\dot{x} = [D_1 f_1(x)^T \cdots D_n f_n(x)^T]^T$  and the constraint set defined by (8) are known to converge to a differential Nash equilibrium of the constrained  $n$ -person concave game [21] and we know the differential Nash equilibrium is unique if the game Hessian

$$H = \begin{bmatrix} D_{11} f_1 & \cdots & D_{1n} f_1 \\ \vdots & \ddots & \vdots \\ D_{n1} f_n & \cdots & D_{nn} f_n \end{bmatrix} \quad (14)$$

is positive definite [18, Theorem 2]. This is automatically guaranteed for  $n \geq 4$  provided the constraint defined by  $\bar{\theta}_{LB}$  using  $\zeta = 10^{-2}$ ; this is straightforward to verify by determining the eigenvalues of  $H$  as  $n$  varies via the method described in [22]. Hence, we use a lower bound on the  $\theta_i$ 's in (P) that guarantees our the game is not only concave but has a unique differential Nash equilibrium. In particular, in (P) we set  $\theta_{LB} = -c^{-1} \rho^{-1} (1 - 4^{-1})^2 + \zeta = -0.8035 + \zeta$  for a given  $\rho$  and  $0 < \zeta \ll 1$ .

#### B. Robust Utility Estimation

Let  $K_i$  denote the number of data points for player  $i$ . Define the regressor  $X = \text{diag}(X_1 \cdots X_n)$  where  $X_i =$

$[(X_i^{(1)})^T \dots (X_i^{(K_i)})^T]^T$  and

$$X_i^{(k)} = \begin{bmatrix} D_i h_{i,1}(x_i^{(k)}) & D_i h_{i,2}(x_i^{(k)}) & D_i \phi_i(x_i^{(k)}) \\ h_{i,1}(x_i^{(k)}) & 0 & 0 \\ 0 & h_{i,2}(x_i^{(k)}) & 0 \end{bmatrix}, \quad (15)$$

the observation matrix  $Y = [Y_1 \dots Y_n]^T$  where  $Y_i = [-D_i \psi_i(x_i^{(1)}) \ 0 \ 0 \ \dots \ -D_i \psi_i(x_i^{(K_i)}) \ 0 \ 0]^T$ , the regression coefficient  $\beta = [\mu_1^1 \ \mu_1^2 \ \theta_1 \ \dots \ \mu_n^1 \ \mu_n^2 \ \theta_n]^T$ . Using the Euclidean norm on  $\mathbb{R}^n \times \mathbb{R}^{2n}$  for  $\chi$  in (P) leads to an OLS problem with inequality constraints:

$$\min_{\beta} \{ \|Y - X\beta\|_2 \mid \beta > \beta_{LB} \} \quad (P1)$$

where  $\beta_{LB} = [0 \ 0 \ \theta_{LB} \ \dots \ 0 \ 0 \ \theta_{LB}]^T$ . The above stated problem can be viewed as a classical multiple linear regression model with inequality constraints described by the data generation process

$$Y = X\beta + \epsilon, \quad \beta > \beta_{LB} \quad (16)$$

where  $\epsilon = (\epsilon_1, \dots, \epsilon_n)$  is the error term.

We assume heteroskedasticity [23, Chapter 5] and the data generation process has a nonspherical standard error  $\epsilon$ . With this general statistical model we are able to describe data generation processes in which the error terms do not have constant variance or are correlated. Therefore, the error, drawn from probability distributions with different variances and errors, can be autocorrelated or serial correlated. Mathematically the nonspherical errors are modelled by:  $\text{cov}(\epsilon|X) = G \succ 0$ ,  $G \in \mathbb{R}^{n_d \times n_d}$  where  $n_d$  is the total data points across all players. Given nonspherical standard errors constrained OLS (cOLS) estimator is biased; hence, it does not satisfy the Gauss–Markov theorem for Best Linear Unbiased Estimator (BLUE). However, using the following adaptation we can derive an unbiased estimator.

Multiplying (16) on the left with  $G^{-\frac{1}{2}}$  leads to the following constrained Generalized Least Squares (cGLS) statistical model:

$$(G^{-\frac{1}{2}}Y) = (G^{-\frac{1}{2}}X)\beta + (G^{-\frac{1}{2}}\epsilon), \quad \beta > \beta_{LB} \quad (17)$$

GLS estimators satisfy the BLUE property [23]. However, in many real regression applications, like the utility learning problem, we don't know the explicit form of  $\text{cov}(\epsilon|X) = G$ . We impose a heteroskedasticity structure on the covariance matrix  $G$ ; in particular,  $\hat{G} = \text{diag}(\tilde{K}, \dots, \tilde{K}) \in \mathbb{R}^{n_d \times n_d}$  where  $\tilde{K} \in \mathbb{R}^{K_i \times K_i}$ . As an example, if  $K_i = 2$ ,

$$\tilde{K} = \begin{bmatrix} \tilde{K}_{11} & \tilde{K}_{12} \\ \tilde{K}_{21} & \tilde{K}_{22} \end{bmatrix} \in \mathbb{R}^{2 \times 2} \quad (18)$$

where  $\tilde{K}_{11} = \frac{2}{n} \sum_{i=1}^{\frac{n}{2}} e_{2i-1}^2$ ,  $\tilde{K}_{22} = \frac{2}{n} \sum_{i=1}^{\frac{n}{2}} e_{2i}^2$ ,  $\tilde{K}_{12} = \tilde{K}_{21} = \frac{2}{n} \sum_{i=1}^{\frac{n}{2}} e_{2i-1} e_{2i}$  and the  $e_i$ 's are the residuals, namely the difference between observed and fitted values. This noise structure is widely used [23], [24].

After estimating  $\hat{G}$ , we substitute it to the cGLS estimators to get a one-step constrained Feasible Generalized Least Squares (cFGLS) or constrained Aitken estimators.

We iterate between the estimation of  $G$  and  $\beta$  either until convergence or for a fixed number iterations in order to prevent overfitting.

### C. Bagging

Given the modified GLS framework, we introduce *wild bootstrapping*, a technique of parametric bootstrapping that is consistent with heteroskedasticity inference and FGLS. In general, bootstrapping is a technique for asymptotic approximation of the bias and standard error of an estimator in a complex and noisy statistical model [23], [25].

We assume that  $E(Y|X) = X\beta$  but simultaneously we allow for heteroskedasticity by conditioning on the residual transformations that we imposed in the standard error structure. Wild or Weighted bootstrap is a technique that is consistent with our Heteroskedastic noise structure and the data generation process is given by

$$Y^* = X\beta_{cFGLS} + \Phi(e)\epsilon^* \quad (19)$$

where  $Y^* \in \mathbb{R}^{n_d \times 1}$  is the new response variables vector (pseudo-vector),  $\beta_{cFGLS} \in \mathbb{R}^{n_d \times 1}$  is the estimator from the constrained FGLS statistical model,  $\epsilon^* \sim N(0, I^{n_d \times n_d})$ ,  $e \in \mathbb{R}^{n_d \times 1}$  is the residuals vector given by  $e = Y - X\beta_{cFGLS}$  and  $\Phi(e) = \hat{G}^{\frac{1}{2}} = \text{diag}(\tilde{K}^{1/2}, \dots, \tilde{K}^{1/2}) \in \mathbb{R}^{n_d \times n_d}$  is a non linear transformation that maps from  $\mathbb{R}^{n_d \times 1}$  to  $\mathbb{R}^{n_d \times 1}$ . It is important to state that  $E(\Phi(e)\epsilon^*|X) = \Phi(e)E(\epsilon^*|X) = \Phi(e)E(\epsilon^*) = 0^{n_d \times n_d}$  and thus, we resample from i.i.d variables.

Using Wild bootstrapping, the empirical covariance matrix of  $\beta_{est}^*$  and the average of  $\beta_{est}^* - \beta_{est}$  are asymptotic approximations of the covariance matrix and bias, respectively. Moreover, the distribution of  $\beta_{est}^* - \beta_{est}$  is a good approximation of  $\beta_{est} - \beta$ .

The process can be described in two steps: First, we fit our cFGLS model. Then, we add Gaussian noise to the predicted values of the cFGLS statistical model and generate  $N$  replicates of *pseudo-data* which is, in turn, used that to fit bootstrap estimators by using ensemble methods to combine several *weak* cFGLS estimators. This improves accuracy since our data set is *small*.

Bagging in regression models and trees is a powerful technique similar to bootstrapping for reducing the overall variance [25]. The process is described as follows: Create  $N$  replicates of pseudo-data using wild bootstrapping, train a different model on each pseudo-data set and, finally, combine the resulting bootstrapped estimators by averaging. Bagging works efficient with high variance models and does not hurt the overall performance of the statistical model.

In the utility learning framework, we combine the estimators generated by cGLS using Bagging:

$$\theta_{Bagged} = \frac{1}{N} \sum_{s=1}^N \theta_{cFGLS}^s \quad (20)$$

where  $N$  is the pseudo-data replicates generated using Wild bootstrapping and  $\theta_{cFGLS}^s$  is the estimator using the  $s$ -th pseudo-data sample. We refer to the bagged estimates as bagged mega-learners since they combine a number of weak learners.

TABLE I

ROOT MEAN SQUARE ERROR (RMSE), MEAN ABSOLUTE ERROR (MAE) AND MEAN ABSOLUTE SCALED ERROR (MASE) OF FORECASTING USING BAGGED UTILITY LEARNERS VS COLS ESTIMATORS. MEAN ABSOLUTE SCALED ERROR (MASE) IS A NEW DEVELOPED MEASURE OF FORECAST ACCURACY [26] FOR COMPARING FORECAST ACCURACY ACROSS TIME SERIES. FORECASTING PREDICTS OCCUPANTS' DYNAMIC BEHAVIOR GIVEN THE DEFAULT LIGHTING SETTING IS 20 AND 10. ACTIONS CAN VARY BETWEEN 1 MINUTE TO SEVERAL HOURS DEPENDENT ON ACTIVITY LEVEL OF THE OCCUPANTS.

Default 20	Utility Learning Method	Bagged	cOLS
	RMSE	<b>8.31</b>	22.53
	MAE	<b>5.20</b>	18.35
	MASE	<b>2.08</b>	7.34
Default 10	Utility Learning Method	Bagged	cOLS
	RMSE	<b>8.69</b>	18.97
	MAE	<b>7.75</b>	15.35
	MASE	<b>3.02</b>	5.97

#### IV. UTILITY LEARNING RESULTS

In this section, we present results of the proposed robust utility learning method using data collected from the social game experiment for building energy management.

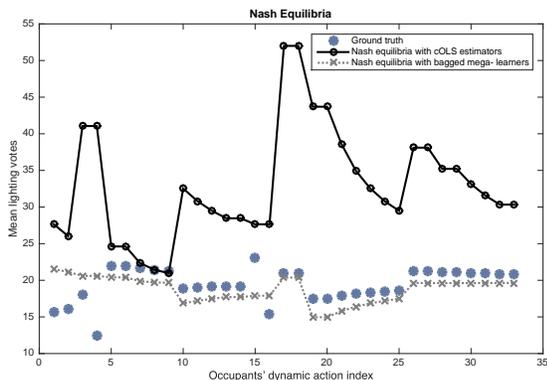


Fig. 1. The mean of the observed lighting votes for default lighting setting of 20 is depicted by the blue dots. The forecasting results via simulation of the occupant game using the cOLS and bagged mega-learner is indicated by the black line and the grey line respectively. On the  $x$ -axis we indicate the index of when a choice was made by one or more of the occupants (i.e. when the implemented lighting setting is changed); the time from one index to the next may be several minutes to hours depending on the activity of the occupants. Notice that the mean of the Nash equilibria of the simulated game using the bagged mega-learner estimates is approximately near the true mean where the cOLS estimates produce Nash equilibria with a large error.

Using occupant voting data we simulate the game defined by the estimated utility functions and show that the estimated model using our robust utility learning method significantly improves predictions for occupant behavior as compared to classical OLS (see Table I). Moreover, in Figure 1 we can see that our model using the estimated parameters obtained via bagging captures most of the variation in the true votes and provides significant improvement over predictions obtained via cOLS.

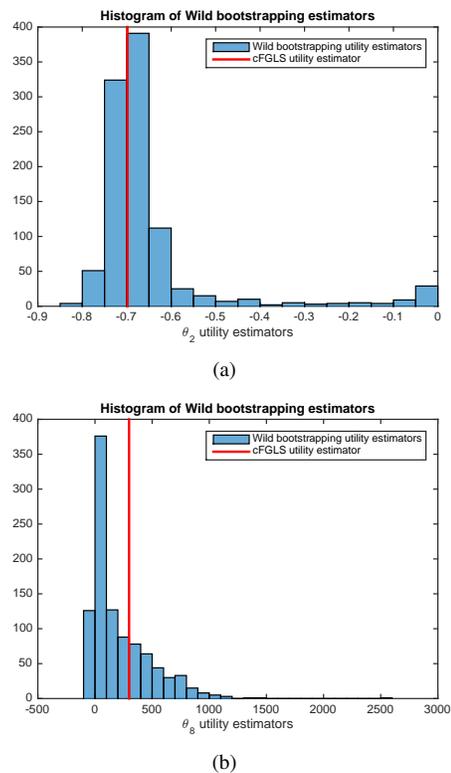


Fig. 2. The vertical red line represents the cFGLS estimator. The histogram depicts the estimates using the wild bootstrapping technique. Subfigure (a) depicts the unbiased estimator for occupant 2 where (b) depicts the biased estimator for occupant 8. In the unbiased estimator example, we have approximately a Gaussian distribution around the initial cFGLS estimator. In the bias example, the resulting distribution is not normally distributed around the cFGLS estimator, and thus the cFGLS estimator is biased.

In addition, we are able to approximate bias of the estimators for each  $\theta_i$ . In Figure 2, we show the resulting cFGLS estimators for two representative occupants obtained by replicates of data using wild bootstrapping. The particular occupants we selected represent players that prefer *winning* to comfort (occupant 8) and players that prefer *comfort* to winning (occupant 2). In Figure 2(b), we see that occupant's 8 parameter estimation is biased. On the other hand in Figure 2(a), we see that through heteroskedasticity inference we infer almost unbiased estimator for occupant 2. We believe this is largely due to the fact that there is significant variation in occupant 2's votes and little variation in occupant 8's votes, who voted zero most of the time.

In table II we present the cFGLS estimates, the bagged mega-learner estimates, and the bias for the most active occupants. For some occupants, the bias is not significantly reduced. This is likely due to the fact that some occupants were not nearly as active as others and hence, there was little variation in their data.

#### V. DISCUSSION AND FUTURE WORK

We presented a general framework for robust utility learning using a heteroskedasticity inference adaptation to cGLS. We applied this method to learn the utility functions of

TABLE II

CFGLS ESTIMATES, BAGGED MEGA-LEARNER ESTIMATES AND BIAS APPROXIMATION USING WILD BOOTSTRAPPING. IN BOLD, WE DENOTE THE OCCUPANTS WITH NEARLY UNBIASED ESTIMATORS. THESE ARE THE OCCUPANTS WITH THE LARGEST VARIATION IN THEIR VOTES.

User id	cFGLS	Bagged	bias
<b>2</b>	-0.69	-0.58	<b>0.11</b>
6	0.5	1.62	1.12
8	298.06	121.2	-176.86
14	337.52	151.26	-186.26
<b>20</b>	-0.8	-0.73	<b>0.07</b>

participants in the building energy management social game that we conducted. We were able to estimate nearly unbiased estimators for several occupant profiles and show that we significantly reduced the forecasting error of the occupants' actions over cOLS. Our robust framework can be extended to other choices of utility functions that incorporate different basis functions.

This framework enables us to close the loop around the building occupant and in effect, vary their behavior in order to meet, for instance, the requirements of a demand response program or to simply reduce over all consumption. Furthermore, such a framework is agnostic to the particular problem of energy efficiency in buildings and we believe it can be applied to any social game and could provide a useful tool in many experimental setups in smart city applications where learning decision making behavior is crucial.

There are several directions for future research. We are actively working employing a prior knowledge into the estimation of player utility functions. This leads to a penalized cFGLS model and has the potential to lead to more accurate forecasting. In addition, we are exploring a model for a mixture of different utility functions using hierarchical mixture of experts with softmax gates. On the social game experiment front, we are in the process of implementing new building energy management social game experiments, both small- and large-scale in Singapore and on the Berkeley campus.

## REFERENCES

- [1] J. McQuade, "A system approach to high performance buildings," *United Technologies Corporation, Tech. Rep.*, 2009.
- [2] A. Aswani, N. Master, J. Taneja, V. Smith, A. Krioukov, D. Culler, and C. Tomlin, "Identifying models of HVAC systems using semi-parametric regression," in *Proc. of the American Control Conf.*, 2012, pp. 3675 – 3680.
- [3] M. Boman, P. Davidsson, N. Skarmeeas, K. Clark, and R. Gustavsson, "Energy saving and added customer value in intelligent buildings," in *Third International Conference on the Practical Application of Intelligent Agents and Multi-Agent Technology*, 1998, pp. 505–517.
- [4] D. Bourgeois, C. Reinhart, and I. Macdonald, "Adding advanced behavioural models in whole building energy simulation: A study on the total energy impact of manual and automated lighting control," *Energy and Buildings*, vol. 38, no. 7, pp. 814 – 823, 2006, special Issue on Daylighting Buildings.
- [5] Y. Ma, G. Anderson, and F. Borrelli, "A distributed predictive control approach to building temperature regulation," in *Proc. of the American Control Conf.*, 2011, pp. 2089–2094.
- [6] F. Oldewurtel, A. Parisio, C. Jones, M. Morari, D. Gyalistras, M. Gworder, V. Stauch, B. Lehmann, and K. Wirth, "Energy efficient building climate control using stochastic model predictive control and weather predictions," in *Proc. of the American Control Conf.*, 2010, pp. 5100–5105.
- [7] T. Lovett, E. Gabe-Thomas, S. Natarajan, E. O'Neill, and J. Padgett, "'just enough' sensing to enlighten: A preliminary demonstration of sensing strategy for the 'energy literacy through an intelligent home energy advisor' (enliten) project," in *Proc. of the Fourth Inter. Conf. on Future Energy Systems*. New York, NY, USA: ACM, 2013, pp. 279–280.
- [8] J. Mathieu, M. Dyson, D. Callaway, and A. Rosenfeld, "Using residential electric loads for fast demand response: The potential resource and revenues, the costs, and policy recommendations," in *ACEEE Summer Study on Energy Efficiency in Buildings*, 2012.
- [9] M. Roozbehani, M. Dahleh, and S. Mitter, "Dynamic pricing and stabilization of supply and demand in modern electric power grids," in *First IEEE International Conference on Smart Grid Communications*, oct. 2010, pp. 543 –548.
- [10] M. Albadi and E. El-Saadany, "A summary of demand response in electricity markets," *Electric Power Systems Research*, vol. 78, no. 11, pp. 1989 – 1996, 2008.
- [11] J. L. Mathieu, P. N. Price, S. Kiliccote, and M. A. Piette, "Quantifying changes in building electricity use, with application to demand response," *IEEE Transactions on Smart Grid*, vol. 2, no. 3, pp. 507–518, 2011.
- [12] M. Lee, O. Aslam, B. Foster, D. Kathan, J. Kwok, L. Medearis, R. Palmer, P. Sporborg, and M. Tita, "Assessment of demand response and advanced metering," Federal Energy Regulatory Commission, Tech. Rep., 2013.
- [13] L. Ratliff, M. Jin, I. C. Konstantakopoulos, C. Spanos, and S. S. Sastry, "Social Game for Building Energy Efficiency: Incentive Design," in *Proc. of the 52nd Allerton Conf. on Communication, Control, and Computing*, 2014.
- [14] I. C. Konstantakopoulos, L. J. Ratliff, M. Jin, S. S. Sastry, , and C. Spanos, "Social game for building energy efficiency: Utility learning, simulation, and analysis," arXiv, Tech. Rep., 2014.
- [15] L. J. Ratliff, "Incentivizing efficiency in societal-scale cyber-physical systems," Ph.D. dissertation, University of California, Berkeley, 2015.
- [16] G. Taguchi, E. A. Elsayed, and T. C. Hsiang, *Quality engineering in production systems*. McGraw-Hill College, 1989.
- [17] J. B. Rosen, "Existence and uniqueness of equilibrium points for concave n-person games," *Econometrica*, vol. 33, no. 3, p. 520, 1965.
- [18] L. J. Ratliff, S. A. Burden, and S. S. Sastry, "On the characterization of local nash equilibria in continuous games," *IEEE Transactions on Automatic Control*, 2016 (to appear).
- [19] A. Keshavarz, Y. Wang, and S. Boyd, "Imputing a convex objective function," in *IEEE International Symposium on Intelligent Control*. IEEE, 2011, pp. 613–619.
- [20] L. J. Ratliff, R. Dong, H. Ohlsson, and S. S. Sastry, "Incentive design and utility learning via energy disaggregation," in *Proc. of the 19th World Congress of the Inter. Federation of Automatic Control*, 2014.
- [21] S. D. Flåm, "Solving non-cooperative games by continuous subgradient projection methods," in *System Modelling and Optimization*. Springer, 1990, pp. 115–123.
- [22] M. Gendreau, "On the location of eigenvalues of off-diagonal constant matrices," *Linear Algebra and its Applications*, vol. 79, pp. 99 – 102, 1986.
- [23] D. A. Freedman, *Statistical models: theory and practice*. cambridge university press, 2009.
- [24] D. A. Freedman and S. C. Peters, "Bootstrapping a regression equation: Some empirical results," *Journal of the American Statistical Association*, vol. 79, no. 385, pp. 97–106, 1984.
- [25] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning, Data Mining, Interference, and Prediction*. Springer, 2009.
- [26] R. J. Hyndman and A. B. Koehler, "Another look at measures of forecast accuracy," *International journal of forecasting*, vol. 22, no. 4, pp. 679–688, 2006.