Electronic Companion - "Infrastructure Planning for Electric Vehicles with Battery Swapping", by Mak, Rong and Shen

Proofs of Analytical Results

Proof of Proposition 1:
Define $g(x) = tx + \Phi^{-1}(\alpha)\sqrt{t}\sqrt{x}$. Note that both t and $\Phi^{-1}(\alpha)$ are nonnegative when α, the required service level is higher than 0.5. Without loss of generality, we can scale the coefficients in front of $\tilde{z}_l$, $\sum_{p\in P}\hat{\lambda}_{pl}Y_{jp}$ to 1 (and remove those equal to 0), for notational brevity. For each $l = 1, ..., L$, let F_l be the family of univariate distributions with mean μ_l, variance σ_l^2 and support $[\underline{z}_l, \bar{z}_l]$, and P_l be some distribution in this family. Note that the joint distribution $\mathbb{P}$ is a direct product $P_1 \times P_2 \times ...P_L$.

The moment problem can be rewritten as:

$$\sup_{\mathbb{P}\in\mathbb{F}} E_{\mathbb{P}}\left[g\left(\sum_{l=1}^{L}\tilde{z}_l\right)\right] = \sup_{(P_l\in F_l)_{l>1}} \sup_{P_1\in F_1} E_{(P_l)_{l=1...L}}\left[g\left(\sum_{l=1}^{L}\tilde{z}_l\right)\right].$$

To proceed, we make use of the following lemma, whose proof will be provided after we complete the proof of Proposition 1.

Lemma 2. *The following equality holds:*

$$\sup_{P_1\in F_1} E_{(P_l)_{l=1...L}}\left[g\left(\sum_{l=1}^{L}\tilde{z}_l\right)\right] = E_{(P_l)_{l>1}} \sup_{P_1\in F_1}\left[E_{P_1}\left[g\left(\tilde{z}_1 + \sum_{l>1}\tilde{z}_l\right)\right]\right].$$

Applying Lemma 2, we may further rewrite:

$$\sup_{\mathbb{P}\in\mathbb{F}} E_{\mathbb{P}}\left[g\left(\sum_{l=1}^{L}\tilde{z}_l\right)\right] = \sup_{(P_l\in F_l)_{l>1}} E_{(P_l)_{l>1}}\left[\sup_{P_1\in F_1} E_{P_1}\left[g\left(\tilde{z}_1 + \sum_{l>1}\tilde{z}_l\right)\right]\right]. \tag{24}$$

The next step requires the following lemma, whose proof will also be given after the proof of Proposition 1.

Lemma 3. *Consider a univariate random variable R, with given mean μ, variance σ^2 and support $[\underline{r}, \bar{r}]$. Denote the family of distributions satisfying the above by F and some particular member of this family by P. Then, the optimal solution to $\sup_{P\in F} E_P\left[g(R+K)\right]$, where K is some constant or some variable independent of R, is achieved by a two-point distribution with support points $\left(\frac{\mu\bar{r}-(\mu^2+\sigma^2)}{\bar{r}-\mu}, \bar{r}\right)$, with respective probabilities $\left(\frac{(\bar{r}-\mu)^2}{(\bar{r}-\mu)^2+\sigma^2}, \frac{\sigma^2}{(\bar{r}-\mu)^2+\sigma^2}\right)$.*

Applying Lemma 3, (24) implies that

$$\sup_{\mathbb{P}\in\mathbb{F}} E_{\mathbb{P}}\left[g\left(\sum_{l=1}^{L}\tilde{z}_l\right)\right] = \sup_{(P_l\in F_l)_{l>1}} E_{(P_l)_{l>1}}\left[E_{\check{z}_1}\left[g\left(\check{z}_1 + \sum_{l>1}\tilde{z}_l\right)\right]\right].$$

Suppose for some $k < L$,

$$\sup_{\mathbb{P}\in\mathbb{F}} E_{\mathbb{P}}\left[g\left(\sum_{l=1}^{L}\tilde{z}_l\right)\right] = \sup_{(P_l\in F_l)_{l>k}} E_{(P_l)_{l>k}}\left[E_{(\check{z}_i)_{i=1,...,k}}\left[g\left(\sum_{i=1}^{k}\check{z}_i + \sum_{l>k}\tilde{z}_l\right)\right]\right].$$

Then, applying a similar argument using Lemma 3 as done above, we can show that

$$\sup_{\mathbb{P}\in\mathbb{F}} E_{\mathbb{P}}\left[g\left(\sum_{l=1}^{L}\tilde{z}_l\right)\right] = \sup_{(P_l\in F_l)_{l>k+1}} E_{(P_l)_{l>k+1}}\left[E_{(\check{z}_i)_{i=1,\dots,k+1}}\left[g\left(\sum_{i=1}^{k+1}\check{z}_i + \sum_{l>k+1}\tilde{z}_l\right)\right]\right].$$

Then, by induction, we obtain the desired result.

Finally, we present the proofs of Lemmas 2 and 3.

Proof of Lemma 2: Observe that the optimal solution to $\sup_{P_1\in F_1} E_{(P_l)_{l=1\dots L}}\left[g\left(\sum_{l=1}^{L}\tilde{z}_l\right)\right]$ is a feasible solution to $\sup_{P_1\in F_1} E_{P_1}\left[g\left(\tilde{z}_1+\sum_{l>1}\tilde{z}_l\right)\right]$. Therefore, we have

$$\sup_{P_1\in F_1} E_{(P_l)_{l=1\dots L}}\left[g\left(\sum_{l=1}^{L}\tilde{z}_l\right)\right] \le E_{(P_l)_{l\neq 1}} \sup_{P_1\in F_1}\left[E_{P_1}\left[g\left(\tilde{z}_1+\sum_{l>1}\tilde{z}_l\right)\right]\right].$$

By Lemma 3, the optimal solution to $\sup_{P_1\in F_1} E_{P_1}\left[g\left(\tilde{z}_1+\sum_{l>1}\tilde{z}_l\right)\right]$ is achieved by the two-point random variable $\check{z}_1$ (recall that $\tilde{z}$'s are all mutually independent). Therefore, we have

$$E_{(P_l)_{l\neq 1}} \sup_{P_1\in F_1}\left[E_{P_1}\left[g\left(\tilde{z}_1+\sum_{l>1}\tilde{z}_l\right)\right]\right] = E\left[g\left(\check{z}_1+\sum_{l>1}\tilde{z}_l\right)\right].$$

Because $\check{z}_l$ is a feasible solution to $\sup_{P_1\in F_1} E_{(P_l)_{l=1\dots L}}\left[g\left(\sum_{l=1}^{L}\tilde{z}_l\right)\right]$, and it achieves an objective value equal to an upper bound of the optimal value, it is also the optimal solution. This completes the proof.

Proof of Lemma 3: For notational brevity, we will show the case for $K\equiv 0$. The case with nonnegative K can be proven by following the same steps trivially. The steps of the proof proceed as follows:

1. From the literature (e.g., [10]), for the problem $\sup_{P\in F} E_P\left[g(R)\right]$, it is sufficient to consider some distribution with three support points (r_1,r_2,r_3) with $r_1\le r_2\le r_3$ and corresponding probabilities (p_1,p_2,p_3) instead of all possible distributions in F.

2. By Lemma 5.1 of [1], one can find a two point distribution (z,r_3) with the same mean and variance as the three point distribution, where $r_1<z<r_2$.

3. Suppose that R is drawn from a family of two-point distributions with the same mean and variance, but possibly different supports. We provide Lemma 4 to show that $E[g(R)]$ is an increasing function of upper support point, as stated more formally in Lemma 4 below:

LEMMA 4. *Consider a pair of two-point distributions having the same mean and variance, with support points (r_1,r_2) and (s_1,s_2), where $r_1<r_2$ and $s_1<s_2$, and corresponding probabilities $(p,1-p)$, $(q,1-q)$ respectively. Then, $s_1>r_1$ implies that $s_2>r_2$, and the following:*

$$pg(r_1)+(1-p)g(r_2)\le qg(s_2)+(1-q)g(s_2)$$

4. Next, we provide Lemma 5 to show that given any three-point distribution, there always exists a two-point distribution with the same mean and variance, while having a higher expected value of $g(.)$.

LEMMA 5. *For a three point distribution with support points $r_1<r_2<r_3$ and corresponding probabilities of (p_1,p_2,p_3), we can always find a 2-point distribution (z,r_3), with corresponding probabilities $q,1-q$ and the same mean and variance such that:*

$$g(r_1)p_1+g(r_2)p_2+g(r_3)p_3<g(z)q+g(r_3)(1-q).$$

5. Finally, from the above steps, we show that the supremum in $\sup_{P\in F} E_P\left[g(R)\right]$ is achieved by a two point distribution with one point equal to the upper limit $\bar{r}$:

We know that solving the moment problem for general distributions is equivalent to solving the same problem over all three-point distributions with the given support, mean and variance. However, from Lemma 5, we know that any three-point distribution can be beaten by a two-point one satisfying the same mean and variance. Moreover, for two-point distributions, given mean and variance, $E[g(R)]$ increases with the higher support point by Lemma 4. Therefore, to maximize this expected value in the moment problem, we should set the higher support point as high as possible, until we hit the upper limit $\bar{r}$. Then, the other support point of the two-point distribution that meets the mean and variance requirements is given by $\frac{\mu\bar{r}-(\mu^2+\sigma^2)}{\bar{r}-\mu}$. If this two-point distribution is feasible (i.e., this lower support point lies in Ω), then it gives the optimal solution to the moment problem.

Therefore, we need to show that $\frac{\mu\bar{r}-(\mu^2+\sigma^2)}{\bar{r}-\mu} \geq \underline{r}$, which is equivalent to $\mu(\underline{r}+\bar{r})-\underline{r}\bar{r} \geq \mu^2+\sigma^2$ by rearranging terms.

For any two-point distribution with support points (x,y) and mean μ, the second moment is equal to $\mu(x+y)-xy$. Without loss of generality, we assume that $\underline{r}\leq x\leq \mu \leq y \leq \bar{r}$. We can see that $\mu(x+y)-xy \leq \mu(x+\bar{r})-x\bar{r} \leq \mu(\underline{r}+\bar{r})-\underline{r}\bar{r}$. If $\mu(\underline{r}+\bar{r})-\underline{r}\bar{r} < \mu^2+\sigma^2$, there is no feasible 2-point distributions with support points (x,y) such that its mean and variance is equal to μ and σ in $[\underline{r},\bar{r}]$.

From step 2, we show that, given some three-point distribution, we can find a two-point distribution with the same mean and variance. In addition, the higher support point of the latter equals the highest support point of the former, and the lower support point of the latter is strictly higher than the lowest support point of the former. Therefore, if $\mu(\underline{r}+\bar{r})-\underline{r}\bar{r} < \mu^2+\sigma^2$ and there exists no two-point distribution within the given support that has the given mean and variance, then there also exists no feasible three-point distribution within the support that gives the desired mean and variance. Combining this argument with the classical result that says three-point distributions are sufficient implies that our proposed two-point distribution must be feasible. Therefore, it is also optimal.

<u>Proof of Lemma 4:</u>

Note that if μ,σ,r_1,r_2 are given, then there is only one possible p satisfying the moment constraints. In particular:

$$p = \frac{r_2-\mu}{r_2-r_1}.$$

Conversely, given μ,σ,p, then r_1,r_2 are uniquely determined by the following:

$$\begin{aligned} r_1 &= \mu - \sqrt{\frac{1-p}{p}}\sigma \\ r_2 &= \mu + \sqrt{\frac{p}{1-p}}\sigma. \end{aligned}$$

It is clear that if μ,σ are given, both r_1 and r_2 are increasing in p. Therefore, if $s_1 > r_1$, we have $s_2 > r_2$. Then, our result is equivalent to that the expected value of the $g(.)$ function is increasing in p. We write down the expected value and take derivative with respect to p:

$$G(p) \equiv pg\left(\mu - \sqrt{\frac{1-p}{p}}\sigma\right) + (1-p)g\left(\mu + \sqrt{\frac{p}{1-p}}\sigma\right).$$

Consequently,

$$\begin{aligned}
G'(p) &= g\left(\mu-\sqrt{\frac{1-p}{p}}\sigma\right)+g'\left(\mu-\sqrt{\frac{1-p}{p}}\sigma\right)\frac{\sigma}{2p}\sqrt{\frac{p}{1-p}}\\
&\quad -g\left(\mu+\sqrt{\frac{p}{1-p}}\sigma\right)+g'\left(\mu+\sqrt{\frac{p}{1-p}}\sigma\right)\frac{\sigma}{2(1-p)}\sqrt{\frac{1-p}{p}}\\
&= g\left(\mu-\sqrt{\frac{1-p}{p}}\sigma\right)+g'\left(\mu-\sqrt{\frac{1-p}{p}}\sigma\right)\frac{1}{2(1-p)}\sqrt{\frac{1-p}{p}}\sigma\\
&\quad -g\left(\mu+\sqrt{\frac{p}{1-p}}\sigma\right)+g'\left(\mu+\sqrt{\frac{p}{1-p}}\sigma\right)\frac{1}{2p}\sqrt{\frac{p}{1-p}}\sigma\\
&= -g(r_2)+\frac{1}{2(1-p)}(\mu-r_1)g'(r_1)+g(r_1)+\frac{1}{2p}(r_2-\mu)g'(r_2)\\
&= -g(r_2)+g(r_1)+\frac{g'(r_1)+g'(r_2)}{2}(r_2-r_1)\geq 0.
\end{aligned}$$

Therefore, $g(.)$ is increasing in p and thus increasing in r_2.

Proof of Lemma 5:

By step 2, we can find a 2-point distribution (z,r_3) with corresponding probabilities $q, 1-q$ with same mean and variance such that $z>r_1$. Consider two cases:

Case 1: Suppose $p_3\geq 1-q$. We have

$$\begin{aligned}
&g(z)q+g(r_3)(1-q)-(g(r_1)p_1+g(r_2)p_2+g(r_3)p_3)\\
=\ &g(z)q-(g(r_1)p_1+g(r_2)p_2+g(r_3)(p_3-(1-q)))\\
=\ &g(z)-(g(r_1)p_1+g(r_2)p_2+g(r_3)(p_3-(1-q))+g(z)(1-q))>0.
\end{aligned}$$

The inequality is obtained by applying Jensen's inequality to the concave function $g(x)$, since $z=r_1p_1+r_2p_2+r_3(p_3-(1-q))+z(1-q)$ by $qz+(1-q)r_3=p_1r_1+p_2r_2+p_3r_3$ (equality of means).

Case 2: Suppose $p_3<1-q$. Then, $q<p_1+p_2$. We can define two new distributions:

1. (r_1,r_2) with respective probabilities $(l_1,l_2)=(\frac{p_1}{p_1+p_2},\frac{p_2}{p_1+p_2})$.
2. (z,r_3) with respective probabilities $(h_1,h_2)=(\frac{q}{p_1+p_2},\frac{1-q-p_3}{p_1+p_2})$.

Based on the above definitions, these two distributions have the same mean and variance. Moreover, $r_1<z<r_2<r_3$ by step 2. By Lemma 4, we have:

$$g(r_1)l_1+g(r_2)l_2<g(z)h_1+g(r_3)h_2$$

Recall that the expected value of $g(.)$ under the three-point distribution is given by:

$$\begin{aligned}
&g(r_1)p_1+g(r_2)p_2+g(r_3)p_3\\
=\ &(p_1+p_2)[l_1g(r_1)+l_2g(r_2)]+g(r_3)p_3\\
\leq\ &(p_1+p_2)[h_1g(z)+h_2g(r_3)]+g(r_3)p_3\\
=\ &qg(z)+(1-q)g(r_3).
\end{aligned}$$

This completes the proof.

Proof of Proposition 2:

We retain the definition $g(x)=tx+\Phi^{-1}(\alpha)\sqrt{t}\sqrt{x}$ used in the proof of Proposition 1. Again, without loss of generality, we scale the coefficients in front of $\tilde{z}_l$, $\sum_{p\in P}\hat{\lambda}_{pl}Y_{jp}$ to 1 (and remove those equal to 0), for notational brevity.

We first prove the upper bound. Let μ, σ and $[\underline{r},\bar{r}]$ be mean, standard deviation and support of $R=\sum_{p\in P}\lambda_p Y_{jp}$. An upper bound on the optimal objective value of the moment problem in (4) can be obtained by considering the following univariate moment problem: $\sup_{\mathfrak{P}\in\mathfrak{F}(\mu,\sigma,\underline{r},\bar{r})} E_{\mathfrak{P}}[g(R)]$, where $\mathfrak{F}(\mu,\sigma,\underline{r},\bar{r})$ is the family of univariate distributions with the given mean, standard deviation and support, and $\mathfrak{P}$ be some distribution in this family. This univariate problem gives an upper bound because any multivariate distribution of $\tilde{z}_l$'s, including the one that achieves the supremum in (4), can be projected to a univariate distribution in $\mathfrak{F}(\mu,\sigma,\underline{r},\bar{r})$. Note that not all univariate distributions can be projected to the family $\mathbb{F}$ of multivariate, independent distributions, and thus the upper bound is not necessarily sharp. By Lemma 3, this upper bound is achieved by a two-point distribution, leading to the following expected value:

$$t\mu+\Phi^{-1}(\alpha)\sqrt{t}\left[\sqrt{\bar{r}}-\frac{\bar{r}-\mu}{\sqrt{\bar{r}}+\sqrt{\mu-\frac{\sigma^2}{\bar{r}-\mu}}}\right]. \tag{25}$$

Next, we further bound this quantity above by the desired upper bound. Note that:

$$\begin{aligned}\mu &= \sum_{p\in P}\sum_{l=1}^{L}\hat{\lambda}_{pl}\mu_l Y_{jp}\\ \bar{r} &= \sum_{p\in P}\sum_{l=1}^{L}\hat{\lambda}_{pl}\bar{z}_l Y_{jp}\leq a'\sum_{p\in P}\sum_{l=1}^{L}\hat{\lambda}_{pl}\mu_l Y_{jp}=a'\mu\\ \sigma &\geq \frac{b'}{\sqrt{L}}\mu.\end{aligned}$$

In the above, $\sigma\geq b'\mu/\sqrt{L}$ because $\tilde{z}$'s are mutually independent. In this case, we know that $\sigma^2\geq b'^2\sum_{i=1}^{L}\mu_l^2$. One can show that $\sum_{i=1}^{L}\mu_l^2\geq\frac{1}{L}(\sum_{i=1}^{L}\mu_l)^2$. Therefore, we have $\sigma\geq\frac{b'}{\sqrt{L}}\mu$.

Now we set a new random variable U which has mean equal to μ, upper support point equal to $a'\mu$ and standard deviation equal to $b'\mu/\sqrt{L}$. U has the same mean, larger upper upper support point and smaller standard deviation compared to random variable R. From Lemma 4, the optimal objective value of $\sup_{\mathfrak{P}\in\mathfrak{F}(\mu,\sigma,\underline{r},\bar{r})} E_{\mathfrak{P}}[g(R)]$ is increasing in $\bar{r}$. It is also clear that the optimal objective value is decreasing in σ (e.g., from equation (25)). Therefore, $\sup_{\mathfrak{P}\in\mathfrak{F}(\mu,\sigma,\underline{r},\bar{r})} E_{\mathfrak{P}}[g(R)]\leq \sup_{\mathfrak{P}\in\mathfrak{F}(\mu,b'\mu,\underline{r},a'\mu)} E_{\mathfrak{P}}[g(U)]$. Substituting $b'\mu/\sqrt{L}$ and $a'\mu$ in place of σ and $\bar{r}$ in (25), we obtain the desired upper bound.

Next, we prove the validity of the lower bound. From Proposition 1, the supremal expected value in (4) is achieved by independent random variables, $\tilde{z}_l$'s, following two-point distributions. First, consider a new set of random variables, $\tilde{z}_l'$, that are independent and follow two-point distributions with support points $\left(\frac{\mu_l(a\mu_l)-(\mu_l^2+(b\mu_l)^2)}{a\mu_l-\mu_l}, a\mu_l\right)$, with respective probabilities $\left(\frac{(a\mu_l-\mu_l)^2}{(a\mu_l-\mu_l)^2+(b\mu_l)^2},\frac{(b\mu_l)^2}{(a\mu_l-\mu_l)^2+(b\mu_l)^2}\right)$. Note that these new random variables can be obtained by replacing the standard deviation by $b\mu_l$ and the upper support point by $a\mu_l$ in the definition of $\tilde{z}_l$.

Recall from Proposition 1 that the supremal expected value in (4) can be obtained by iteratively taking expectation over $\tilde{z}_l$ for each l, keeping the values of others unchanged. Then, recall from the proof of Lemma 3 that the expected value of $g(\tilde{z}_l+K)$ (over $\tilde{z}_l$) is increasing in $\bar{r}_l$ and decreasing in σ_l. This implies that $E[g(\tilde{z}_l'+K)]\leq E[g(\tilde{z}_l+K)]$. Repeating the same step for $l=1,...,L$, we can show that the supremal expected value in (4), which is given by $E[g(\sum_l\tilde{z}_l)]$, is bounded below by $E[g(\sum_l\tilde{z}_l')]$. On a side note, the definition of $\tilde{z}_l'$ does not require restrictions on the lower support point, except that it should be nonnegative (because of the square root term in the $g(.)$ function). Therefore, one can easily check that such a construction is possible as long as the condition $a\geq b^2+1$ holds.

Note also that the probability weights on the two support points of all the $\tilde{z}'_l$ variables are identical for all $l = 1, ..., L$. The next step is to show that, by restricting the two-point distributions to be perfectly positively dependent (i.e., all of them take the higher value at the same time and the lower value at the same time, possible only when the probability weights are identical), we obtain a lower expected value than when they are all independent. Consider $\tilde{z}'_{l_1}$ and $\tilde{z}'_{l_2}$, where $l_1 \neq l_2$. Conditioning on the event $\sum_{l \neq l_1, l_2} \tilde{z}'_l = K$, we are interested in the conditional expectation $E\left[g(\tilde{z}'_{l_1} + \tilde{z}'_{l_2} + K) \middle| K\right]$. Suppose $\tilde{z}'_{l_1}$ and $\tilde{z}'_{l_2}$ follow two-point distributions with support points (z_1^L, z_1^H) and (z_2^L, z_2^H) and (common) associated probabilities $(\hat{p}, 1-\hat{p})$, respectively. Then, if the two distributions are independent, the conditional expected value of interest is given by:

$$
\begin{aligned}
&\hat{p}^2 g(z_1^L + z_2^L + K) + (1-\hat{p})^2 g(z_1^H + z_2^H + K) + \hat{p}(1-\hat{p}) g(z_1^L + z_2^H + K) + (1-\hat{p})\hat{p} g(z_1^H + z_2^L + K) \\
&\geq \hat{p}^2 g(z_1^L + z_2^L + K) + (1-\hat{p})^2 g(z_1^H + z_2^H + K) + \hat{p}(1-\hat{p}) g(z_1^H + z_2^H + K) + (1-\hat{p})\hat{p} g(z_1^L + z_2^L + K) \\
&= \hat{p} g(z_1^L + z_2^L + K) + (1-\hat{p}) g(z_1^H + z_2^H + K).
\end{aligned}
$$

The inequality above holds by concavity of the $g(.)$ function. The last line is equal to the conditional expectation of interest when $\tilde{z}'_{l_1}$ and $\tilde{z}'_{l_2}$ are perfectly positively dependent, which implies that $\tilde{z}'_{l_1} + \tilde{z}'_{l_2}$ becomes a single univariate random variable following a two-point distribution. By applying this argument successively, we may obtain a lower bound by considering a single univariate random variable following a two-point distribution with support points $\left(\frac{a-(b^2+1)}{a-1}\mu, a\mu\right)$ and associated probabilities $\left(\frac{(a-1)^2}{(a-1)^2+b^2}, \frac{b^2}{(a-1)^2+b^2}\right)$. The resulting expected value gives the desired lower bound.

Proof of Proposition 3:

First, we rescale the random variable $\tilde{z}_l$ by defining $\tilde{s}_l = \frac{\tilde{z}_l}{\mu_l}$. Then, we have $E[\tilde{s}_l] = 1$ for all l. Moreover, $\tilde{s}_l$'s are nonnegative, independent and identically distributed random variables each with two-point distribution $\left(\frac{a-1-b^2}{a-1}, a\right)$ with probabilities $\left(\frac{(a-1)^2}{(a-1)^2+b^2}, \frac{b^2}{(a-1)^2+b^2}\right)$. Moreover, we define $w_l = \mu_l \sum_{p \in P} \widehat{\lambda}_{pl} Y_{jp}$. Since $Y_{jp}^2 = Y_{jp}$, we have $w_l = \mu_l \sum_{p \in P} \widehat{\lambda}_{pl} Y_{jp}^2$. Because $\sum_{p \in P} Y_{jp} \geq 1$, $\lim_{L \to \infty} \sum_{l=1}^L \widehat{\lambda}_{pl} \mu_l = \infty$ for all p, and $\lim_{L \to \infty} \widehat{\lambda}_{pl} \mu_l / \sum_{l=1}^L \widehat{\lambda}_{pl} \mu_l = 0$ for all p and l, it is clear that $\lim_{L \to \infty} \sum_{l=1}^L w_l = \infty$ and $\lim_{L \to \infty} \frac{w_l}{\sum_{l=1}^L w_l} = 0$. By the results of Jamison et al. [7], the sequence $\left[\sum_{l=1}^L w_l \tilde{s}_l\right] / \left[\sum_{l=1}^L w_l\right]$ converges to some constant κ in probability as $L \to \infty$.

Let $\beta(\epsilon, L) = P\left(\left|\left[\sum_{l=1}^L w_l \tilde{s}_l\right] / \left[\sum_{l=1}^L w_l\right] - \kappa\right| \leq \epsilon\right)$ for any $\epsilon > 0$ and L. That is, there is a probability $\beta(\epsilon, L)$ that $\left[\sum_{l=1}^L w_l \tilde{s}_l\right] / \left[\sum_{l=1}^L w_l\right] \in [\kappa - \epsilon, \kappa + \epsilon]$. Moreover, we have $0 \leq \tilde{s}_l \leq a$, by definition. Therefore, there is a probability of $(1 - \beta(\epsilon, L))$ that $\left[\sum_{l=1}^L w_l \tilde{s}_l\right] / \left[\sum_{l=1}^L w_l\right] \in [0, \kappa - \epsilon] \cup [\kappa + \epsilon, a]$. Then, by taking the smallest and largest possible values in the respective ranges, we obtain the following lower and upper bounds on the expectation:

$$
(\kappa - \epsilon)\, \beta(\epsilon, L) \leq E\left[\sum_{l=1}^L w_l \tilde{s}_l\right] / \left[\sum_{l=1}^L w_l\right] \leq (\kappa + \epsilon)\, \beta(\epsilon, L) + a(1 - \beta(\epsilon, L)).
$$

Since $\left[\sum_{l=1}^L w_l \tilde{s}_l\right] / \left[\sum_{l=1}^L w_l\right]$ converges to constant κ in probability, $\lim_{L \to \infty} \beta(\epsilon, L) = 1$ for any $\epsilon > 0$. By letting $L \to \infty$, we have $\kappa - \epsilon \leq \lim_{L \to \infty} E\left[\sum_{l=1}^L w_l \tilde{s}_l\right] / \left[\sum_{l=1}^L w_l\right] \leq (\kappa + \epsilon)$. On the other hand, because $E[\tilde{s}_l] = 1$ for any l, we have $\lim_{L \to \infty} E\left[\sum_{l=1}^L w_l \tilde{s}_l\right] / \left[\sum_{l=1}^L w_l\right] = 1$, which leads to $\kappa = 1$.

Furthermore, we have

$$\begin{aligned}\beta(\epsilon, L) &= \Pr\left(\left|\left[\sum_{l=1}^{L} w_l \tilde{s}_l\right] / \left[\sum_{l=1}^{L} w_l\right] - 1\right| \leq \epsilon\right)\\ &= \Pr\left(\left|\sqrt{\left[\sum_{l=1}^{L} w_l \tilde{s}_l\right] / \left[\sum_{l=1}^{L} w_l\right]} - 1\right| \left|\sqrt{\left[\sum_{l=1}^{L} w_l \tilde{s}_l\right] / \left[\sum_{l=1}^{L} w_l\right]} + 1\right| \leq \epsilon\right)\\ &= \Pr\left(\left|\sqrt{\left[\sum_{l=1}^{L} w_l \tilde{s}_l\right] / \left[\sum_{l=1}^{L} w_l\right]} - 1\right| \leq \frac{\epsilon}{\left|\sqrt{\left[\sum_{l=1}^{L} w_l \tilde{s}_l\right] / \left[\sum_{l=1}^{L} w_l\right]} + 1\right|}\right)\\ &\leq \Pr\left(\left|\sqrt{\left[\sum_{l=1}^{L} w_l \tilde{s}_l\right] / \left[\sum_{l=1}^{L} w_l\right]} - 1\right| \leq \epsilon\right)\end{aligned}$$

It indicates that there is at least a probability $\beta(\epsilon, L)$ that $\sqrt{\left[\sum_{l=1}^{L} w_l \tilde{s}_l\right] / \left[\sum_{l=1}^{L} w_l\right]} \in [1-\epsilon, 1+\epsilon]$. In addition, there is at most a probability of $(1-\beta(\epsilon, L))$ that $\sqrt{\left[\sum_{l=1}^{L} w_l \tilde{s}_l\right] / \left[\sum_{l=1}^{L} w_l\right]} \in [0, 1-\epsilon] \cup [1+\epsilon, \sqrt{a}]$. Using a similar argument as above, we can show that $\lim_{L\to\infty} E\left[\sqrt{\frac{\sum_{l=1}^{L} w_l \tilde{s}_l}{\sum_{l=1}^{L} w_l}}\right] = 1$. Finally, it is straightforward to see that $\lim_{L\to\infty} \overline{\Psi} = 1$, which completes the proof.

Proof of Proposition 4:
Define $g(x) = tx + \Phi^{-1}(\alpha)\sqrt{tx}$. We know that $g(x)$ is increasing function in x as $\alpha \geq 0.5$. Suppose that $g(x) \leq g_j$. Then it is equivalent to $\left(\sqrt{tx} + \frac{\Phi^{-1}(\alpha)}{2}\right)^2 \leq g_j + \frac{\Phi^{-1}(\alpha)^2}{4}$. Since $x \geq 0$, we know that $g(x) \leq g_j$ if and only if $x \leq \frac{\left(\sqrt{g_j + \Phi^{-1}(\alpha)^2/4} - \Phi^{-1}(\alpha)/2\right)^2}{t}$.

By the above argument and the assumption that $\sum_{l=1}^{L} \hat{\lambda}_{pl}\mu_l Y_{jp}$ is nonnegative, we have $P_{\mathbb{P}}\left(I_j(\mathbf{Y}) \leq g_j\right)$ is equivalent to $P_{\mathbb{P}}\left(\sum_{p\in P}\sum_{l=1}^{L} \hat{\lambda}_{pl}\hat{z}_l Y_{jp} \leq \hat{g}_j\right)$ where $\hat{g}_j = \frac{\left(\sqrt{g_j + \Phi^{-1}(\alpha)^2/4} - \Phi^{-1}(\alpha)/2\right)^2}{t}$

Proof of Lemma 1:
First, observe that the chance constraint is equivalent to the following:

$$\hat{\psi}_{1-\gamma}\left(-\sum_{p\in P}\pi_p\lambda_p + \sum_{j\in J} hI_j + T\sum_{j\in J} f_j X_j\right) \leq 0$$

where $\hat{\psi}_{1-\gamma}(.)$ is the value-at-risk of the expression inside the parentheses at the $1-\gamma$ quantile level. Because the CVaR value is always greater than or equal to the VaR value of the same random variable at the same quantile level, the CVaR constraint is more restrictive than the one on VaR, and thus the original chance constraint. This suggests that, by replacing the chance constraint with a CVaR constraint, the feasible region of the problem shrinks. As a result, the optimal solution to the new problem is feasible in the original one, and the optimal objective value of the new problem is a lower bound on that of the original problem.

Proof of Proposition 5:
For any positive number η, we know that $\sqrt{t\sum_{p\in P}\lambda_p Y_{jp}} \leq \frac{t\sum_{p\in P}\lambda_p Y_{jp}}{2}\eta + \frac{1}{2\eta}$. Note that this bound is tight, as the linear function touches the square root function at one tangent point. Then,

$$\psi_{1-\gamma}\left(-\sum_{p\in P}(\pi_p - ht)\lambda_p + \frac{h\Phi^{-1}(\alpha)t}{2}\sum_{j\in J}\eta_j\sum_{p\in P}\lambda_p Y_{jp} + \sum_{j\in J}\frac{h}{2\eta_j} + T\sum_{j\in J} f_j X_j\right)$$
$$\geq \psi_{1-\gamma}\left(-\sum_{p\in P}(\pi_p - ht)\lambda_p + h\Phi^{-1}(\alpha)\sum_{j\in J}\sqrt{\sum_{p\in P}\lambda_p Y_{jp}} + T\sum_{j\in J} f_j X_j\right).$$

To obtain a the tightest bound given $(\mathbf{X},\mathbf{Y})$ values, we need to solve problem (21), which is equivalent to:

$$\min_{\eta_j, O_j\geq 0}\psi_{1-\gamma}\left(-\sum_{p\in P}(\pi_p - ht)\lambda_p + \frac{h\Phi^{-1}(\alpha)t}{2}\sum_{j\in J}\eta_j\sum_{p\in P}\lambda_p Y_{jp}\right) + \sum_{j\in J}\frac{hO_j}{2} + T\sum_{j\in J} f_j X_j \tag{26}$$

subject to

$$\frac{1}{\eta_j} \leq O_j \tag{27}$$

Because $\eta_j, O_j \geq 0$, we can rewrite

$$\frac{1}{\eta_j} \leq O_j \Leftrightarrow \eta_j O_j \geq 1 \Leftrightarrow (\eta_j + O_j)^2 \geq 2 + \eta_j^2 + O_j^2 \Leftrightarrow \sqrt{2+\eta_j^2+O_j^2} \leq \eta_j + O_j.$$

Since $\psi_{1-\gamma}\left(-\sum_{p\in P}(\pi_p - ht)\lambda_p + \sum_{j\in J} h\Phi^{-1}(\alpha)t\eta_j\sum_{p\in P}\lambda_p Y_{jp}\right)$ can be bounded by an SOCP form in η_j when $(\mathbf{X},\mathbf{Y})$ is fixed, the whole problem can be bounded by solving an SOCP in η_j.

Proof of Proposition 6
We first show that the proposed problem gives an upper bound on the original CVaR value. By Proposition 5, we may obtain the following inequalities:

$$\sum_{n=1}^{N}\psi_{1-\gamma}\left(\sum_{l=1}^{L}\hat{c}_{ln}\tilde{z}_l + h\Phi^{-1}(\alpha)t\sum_{j\in J}\frac{\eta_{jn}}{2}\sum_{l=1}^{L}\tilde{z}_l\hat{y}_{jln}\right) + \sum_{j\in J}\frac{1}{2\eta_{jn}}\hat{d}_{jn}$$
$$\geq \sum_{n=1}^{N}\psi_{1-\gamma}\left(\sum_{l=1}^{L}\hat{c}_{ln}\tilde{z}_l + h\Phi^{-1}(\alpha)t\sum_{j\in J}\sqrt{\sum_{l=1}^{L}\tilde{z}_l\hat{y}_{jln}}\right)$$
$$\geq \psi_{1-\gamma}\left(\sum_{n=1}^{N}\sum_{l=1}^{L}\hat{c}_{ln}\tilde{z}_l + h\Phi^{-1}(\alpha)t\sum_{n=1}^{N}\sum_{j\in J}\sqrt{\sum_{l=1}^{L}\tilde{z}_l\hat{y}_{jln}}\right) \text{(By subadditivity of } \psi(.))$$
$$\geq \psi_{1-\gamma}\left(\sum_{n=1}^{N}\sum_{l=1}^{L}\hat{c}_{ln}\tilde{z}_l + h\Phi^{-1}(\alpha)t\sum_{j\in J}\sqrt{\sum_{n=1}^{N}\sum_{l=1}^{L}\tilde{z}_l\hat{y}_{jln}}\right) \text{(By subadditivity of } \sqrt{.})$$
$$= \psi_{1-\gamma}\left(-\pi_0\sum_{l=1}^{L}\hat{\lambda}_{0l}\tilde{z}_l - \sum_{p\in P}(\pi_p - ht)\sum_{l=1}^{L}\hat{\lambda}_{pl}\tilde{z}_l + h\Phi^{-1}(\alpha)\sum_{j\in J}\sqrt{t\sum_{p\in P}\sum_{l=1}^{L}\hat{\lambda}_{pl}\tilde{z}_l Y_{jp}}\right).$$

Note that $\psi(.)$, as a coherent risk measure, satisfies the subadditive property. To ensure that the expression in the square root is nonnegative, we impose nonnegativity constraints on $\hat{y}_{jln}$ variables. As the inequalities hold for all distributions in $\mathbb{F}$, they also hold for the supremum over all distributions in $\mathbb{F}$. Next, we show that this bound is tighter than the minimum of the N individual bounds. Let n^* be the index of the minimum of the N bounds. Then, we observe that this bound can be achieved by setting:

$$\begin{aligned}
\hat{c}_{ln^*} &= -\pi_0 \hat{\lambda}_{0l} - \sum_{p\in P} (\pi_p - ht)\hat{\lambda}_{pl}, \text{ for each } l = 1,...,L \\
\hat{d}_{jn^*} &= h, \text{ for each } j \in J \\
\hat{y}_{jln^*} &= \sum_{p\in P} \hat{\lambda}_{pl} Y_{jp}, \text{ for each } j \in J, l = 1,...,L, n = 1,...,N.
\end{aligned}$$

Therefore, the minimum of the N bounds corresponds to a feasible solution in the minimization problem. This suggests that the optimal solution gives a tighter upper bound.

List of Notation Used

Sets

$J =$ Set of candidate locations for swapping stations
$P =$ Set of inter-city travel paths
$Q =$ Set of subpaths longer than the tolerance distance for swapping

Parameters Related to Network Structure

$a_{jq} = 1$ if candidate location $j \in J$ is along subpath $q \in Q$
$b_{pq} = 1$ if subpath $q \in Q$ is part of origin-destination travel path $p \in P$

Cost, Revenue and Budget Parameters

$f_j =$ Fixed cost of building swapping station at site $j \in J$
$B =$ Budget to build swapping stations in Stage 1
$h =$ Cost of holding one battery at any swapping station
$\pi_0 =$ Exogenous variable profit earned per unit EV flow not requiring swapping
$\pi_p =$ Exogenous variable profit earned per unit EV flow on path p
$T =$ Target on return-on-investment

Parameters Related to Swapping Operations

$g_j =$ Maximum number of batteries that can be held at location $j \in J$ without exceeding electric load capacity
$t =$ Guaranteed minimum recharge time for batteries for a proportion α of customers
$\alpha =$ Proportion of swapping requests to be satisfied with batteries with a guranteed recharge time of t
$\epsilon_g =$ Maximum allowable probability that capacity requirement at a station be violated

Demand Parameters

$\lambda_0 =$ Flow rate of EVs along all paths without needing swapping
$\lambda_p =$ Flow rate of EVs along path $p \in P$
$\tilde{z}_l = l-$th primitive uncertainty term affecting demand
$\hat{\lambda}_{pl} =$ Coefficient of primitive uncertainty $\tilde{z}_l$ in linear flow rate function λ_p
$\mu_l =$ Mean of primitive uncertainty $\tilde{z}_l$
$\bar{z}_l =$ Maximum possible value of primitive uncertainty $\tilde{z}_l$
$\underline{z}_l =$ Minimum possible value of primitive uncertainty $\tilde{z}_l$

$\Sigma =$ Covariance matrix of primitive uncertainties
$\sigma_l =$ Standard deviation of primitive uncertainty $\tilde{z}_l$
Stage 1 Decision Variables
$X_j = 1$ if swapping station is built at site $j \in J$, 0 otherwise
$Y_{jp} = 1$ if EVs traveling origin-destinating path $p \in P$ visit station at site $j \in J$, 0 otherwise
$Z_{jq} = 1$ if EVs traveling along subpath $q \in Q$ visit station at site $j \in J$, 0 otherwise
$V_j =$ Worst case (supremum) expected number of batteries at location $j \in J$
$\gamma =$ Worst case (supremum) probability that the profit target is not met
Stage 2 Decision Variables
$I_j =$ Number of batteries held at station at location $j \in J$

Demand Model and Parameter Settings Used in Computational Tests

In the computational tests performed in Section 6, we constructed a realistic dataset based on the network of major freeways in the San Francisco Bay Area: I80, CA 84, CA 92, US 101, CA 237, I280, I580, I680 and I880, based on [4]. We define 316 segments based on the following steps. First, we define raw segments as sections between any adjacent exits and ramps connecting two freeways. Then, to shrink the problem size, we allow pairs of raw segments to be merged as long as (1) the length after merging does not exceed 1 mile and (2) neither raw segment considered is a connecting ramp. To define travel paths, we consider 53 cities in the Bay Area: San Francisco, Oakland, Emeryville, Berkeley, Richmond, El Cerrito, San Pablo, Pinole, Hercules, Vallejo, Fairfield, Vacaville, Woodside, Redwood City, Menlo Park, Newark, Fremont, Half Moon Bay, San Mateo, Hayward, Gilroy, San Martin, San Jose, Sunnyvale, Mountain View, Palo Alto, Belmont, Milbrae, South San Francisco, Brisbane, Sausalito, Marin City, Corte Madera, San Rafael, Novato, Pataluma, Cotati, Santa Rosa, Milpitas, Cupertino, Los Altos Hills, Hillsborough, San Bruno, Daly City, Livermore, Pleasanton, San Leandro, Walnut Creek, Pleasant Hill, Concord, Martinez, Benicia and Union City. We define the set of travel paths P as the shortest path connecting the segments nearest to the centers of every pair of cities in the list, excluding the pairs for which the shortest distance is shorter than half the one-charge travel range, which is assumed to be 80 miles in Sections 6.1 and 6.2, and 80, 100, 120 miles in Section 6.3. These values are within the typical range of current EVs. After the set P is defined, the set of subpaths Q are defined as illustrated in Figure 1. We assume that one out of every three segments (i.e., the exit at the end point) can be selected as a candidate station location. The resulting set J contains 101 candidate sites.

After defining the set of travel paths, we assume the following specific form of the general demand model. In particular, there are $L = |P| + 1$ primitive uncertainties, denoted by $\tilde{\lambda}_p, p \in P$ and $\tilde{\beta}$. The actual flow along path $p \in P$ is equal to $\lambda_p = \tilde{\lambda}_p + \rho_p \tilde{\xi}$. All the primitive uncertainties are mutually independent. In this model, all the demand flow rates are correlated through the inclusion of the common, zero-mean error term $\tilde{\xi}$, which represents an overall adoption factor for the entire market.

The mean, support and variance of $\tilde{\lambda}_p$ are given by μ_p, $[\underline{\lambda}_p, \bar{\lambda}_p]$ and σ_p, respectively. The forward and backward deviations [3] are denoted by ϕ_p and β_p, respectively. In Section 6, we generated these parameters as follows. First, the mean μ_p is generated using a gravity model, such that λ_p is proportional to the product of the populations of the origin and destination cities, divided by the squared path length. Such gravity models are widely used in transportation studies (see, e.g., [8]), and they reflect the travel behavior that more people travel out of and into cities with high population, separated by shorter distances. The flows, as computed by the gravity model, are scaled such that the total expected hourly flow ($\sum_{p \in P} \mu_p$) requiring swaps in the network is equal to 50, 100, 200, 400, 600, 800, 1000, 1200, 1400 in Section 6.1, and 600 in Section 6.2 and 6.3. These flows include all of those traveling on paths exceeding the travel range of the battery in one charge, and part (those traveling round trips) of those traveling on paths between 0.5 to 1 time the travel range.

The support parameters of the random variable $\tilde{\lambda}_p$ are assumed to be $\underline{\lambda}_p = 0.1\mu_p$ and $\bar{\lambda}_p = 2.5\mu_p$. The standard deviation σ_p is assumed to be $\sigma_p = 0.45\mu_p$. The forward and backward deviations are computed using the bounds introduced in Chen *et al.* [3].

We assume that all flows on paths exceeding the one-charge travel range (e.g., 80 miles) will require swapping. For paths between 0.5 to 1 times the travel range (e.g., 40 to 80 miles), EVs traveling round-trips will require swapping. We assume that, for all paths in this range, the proportion of EVs traveling round-trips is 20%. Finally, we consider the hourly flow of EVs not requiring swapping along all paths, denoted by $\tilde{\lambda}$. From a recent survey on the car usage pattern of potential EV adopters [9], it is suggested that the vast majority of vehicle-days monitored involve travels in the 0-50 mile range. Therefore, it is reasonable for us to assume that the vast majority of trips do not require swapping (i.e., not exceeding even half of the travel range). In particular, we scale the expected value of this flow, μ_0, to a value such that it occupies 90% of total expected vehicle flows, while the expected flows requiring swaps, as computed previously, will occupy 10%. Similar to the case of $\tilde{\lambda}_p$, the support parameters of the random variable $\tilde{\lambda}_0$ are assumed to be $\underline{\lambda}_0 = 0.1\mu_0$ and $\bar{\lambda}_0 = 2.5\mu_0$. The standard deviation σ_p is assumed to be $\sigma_0 = 0.45\mu_0$. The forward and backward deviations are computed using the bounds introduced in Chen *et al.* [3].

For the zero-mean adoption factor ξ, the mean, support and variance are given by 0, $[\underline{\xi}, \bar{\xi}]$ and σ_ξ, respectively. In Section 6, we assume that $\underline{\xi} = -0.9$, $\bar{\xi} = 1$ and $\sigma_\xi = 0.2$. The forward and backward deviations, denoted by ϕ_ξ and β_ξ, are again computed using the bounds introduced in [3]. Finally, we assume that $\rho_p = 0.08\mu_p$ for all $p \in P$, i.e., the impact of the common adoption factor ξ on the EV flow on path p is proportional to its expected flow.

The marginal exogenous variable profit per unit flow (π_p) values are determined as follows. Note that all our cost and revenue terms are annualized, and the flow rates are hourly. As the study in [9] suggests, only a small proportion of all cars (15% in their study) are on the road at the same time, even during weekday rush hours. Therefore, we may estimate the expected total number of registered EVs to be about 5 to 6 times $\sum_{p\in P\cup\{0\}} \mu_p$. Assuming that these vehicles are used for about 300 days a week, a reasonable estimate of the expected total vehicle-mile travelled by all registered EVs will be about $1500 \sum_{p\in P\cup\{0\}} \mu_p l_p$, where l_p is the length of travel path p. Then, assuming a consevative \$0.1 gross margin per vehicle-mile traveled (the cost of electricity is 2-3 cents per mile, while the cost of gasoline is 13 cents per mile and growing [6]), the value of π_p can be estimated to be \$$150 l_p$, where l_p is the length of path p. We conservatively assume that trips not requiring swapping has an average distance of one-fourth the one-charge travel range. Besides \$150 per mile of flow, we also consider \$75 per flow-mile in Section 6.1, and \$50 and \$100 per flow-mile in Section 6.3.

Next, we discuss the cost parameters. The swapping stations are estimated to cost \$500,000 apiece [5]. We annualize that figure to approximately \$50,000 per year and use it as the value of f_j in the computational tests. The initial budget B is assumed to be equivalent to the fixed cost to open 20 stations. The ROI target T is assumed to be 10%. The annual holding cost h for a battery is assumed to be \$2000 in Section 6.1 and 6.3, and \$1000, \$1500 and \$2000 in the study of technological advancements in Section 6.2. These numbers are based on the purchase cost of about \$10000 per battery, and an approximate life cycle of about 5 years.

Finally, we assume that the service guarantee maintained at swapping stations is that at least 0.95 probability (α), a battery will be recharged for at least 2 hours (t) in Section 6.1 and 6.3, and 1.5, 2, 2.5, 3, 3.5 hours in Section 6.2. The maximum allowable load at every station (g_j) is assumed to be 100 batteries that are recharged in parallel, and for simplicity, we assume that this load cannot be exceeded with any positive probability (i.e., $\epsilon_g = 0$).

Complete Formulation of Cost-Concerned Model

The complete formulation of the cost-concerned model, after all manipulations and approximations discussed in Section 4, is written below. As discussed in Chen and Sim [2], of the five bounds they proposed to approximate the worst-case CVaR value, the first and fourth ones typically dominate the other three. Therefore, we adopt only these two bounds in our computational studies.

$$\min \sum_{j\in J} \{f_j X_j + hV_j\}$$

subject to:

$$\begin{aligned}
Y_{jp} &\geq b_{pq} Z_{jq} \text{ , for each } j \in J, p \in P, q \in Q \\
\sum_{j\in J} a_{jq} Z_{jq} &\geq 1 \text{ , for each } q \in Q \\
Y_{jp} &\leq X_j \text{ , for each } j \in J, p \in P \\
\sum_{p\in P} t\mu_p Y_{jp} + \Phi^{-1}(\alpha)\bar{\Psi}\sqrt{\sum_{p\in P} t\mu_p Y_{jp}^2} &\leq V_j \text{ , for each } j \in J \\
r_{1j} + r_{4j} &\leq 0 \text{ , for each } j \in J \\
y_{10j} + \sum_{p\in P}(\bar{\lambda}_p - \mu_p) y_{1jp} + \bar{\xi} x_j &\leq r_{1j} \text{ , for each } j \in J \\
y_{40j} + \sqrt{-2ln(\epsilon_g)}\sqrt{\sum_{p\in P} u_{jp}^2 + \hat{x}_{4j}^2} &\leq r_{4j} \text{ , for each } j \in J \\
u_{jp} &\geq \phi_p y_{4jp} \text{ , for each } j \in J, p \in P \\
u_{jp} &\geq -\beta_p y_{4jp} \text{ , for each } j \in J, p \in P \\
\hat{x}_j &\geq \phi_\xi x_{4j} \text{ , for each } j \in J \\
\hat{x}_j &\geq -\beta_\xi x_{4j} \text{ , for each } j \in J \\
y_{10j} + y_{40j} &= \sum_{p\in P} \mu_p Y_{jp} - \hat{g}_j X_j \text{ , for each } j \in J \\
y_{1jp} + y_{4jp} &= Y_{js} \text{ , for each } j \in J, p \in P \\
x_{1j} + w_{4j} &= \sum_{p\in P} \rho_p Y_{jp} \text{ , for each } j \in J \\
X_j &\in \{0,1\} \text{ , for each } j \in J \\
Y_{jp} &\in \{0,1\} \text{ , for each } j \in J, p \in P \\
Z_{jq} &\in \{0,1\} \text{ , for each } j \in J, q \in Q.
\end{aligned}$$

Complete Formulation of Goal-Driven Model

The complete formulation of the goal-driven model, after all manipulations and approximations discussed in Section 4, is written below.

$$\min \sum_{n=1}^{N} \hat{\psi}_n + T \sum_{j\in J} f_j X_j$$

subject to:

$$\begin{aligned}
\sum_{j\in J} f_j X_j &\leq B \\
Y_{jp} &\geq b_{pq} Z_{jq} \text{ , for each } j \in J, p \in P, q \in Q \\
\sum_{j\in J} a_{jq} Z_{jq} &\geq 1 \text{ , for each } q \in Q
\end{aligned}$$

$$
\begin{aligned}
Y_{jp} &\leq X_j \text{ , for each } j \in J, p \in P \\
r_{1j} + r_{4j} &\leq 0 \text{ , for each } j \in J \\
r_{1j} &\geq y_{10j} + \sum_{p \in P} (\bar{\lambda}_p - \mu_p) y_{1jp} + \bar{\xi} x_j \text{ , for each } j \in J \\
r_{4j} &\geq y_{40j} + \sqrt{-2ln(\epsilon_g)} \sqrt{u_{j0}^2 + \sum_{p \in P} u_{jp}^2 + \hat{x}_{4j}^2} \text{ , for each } j \in J \\
u_{jp} &\geq \phi_p y_{4jp} \text{ , for each } j \in J, p \in P \\
u_{jp} &\geq -\beta_p y_{4jp} \text{ , for each } j \in J, p \in P \\
\hat{x}_j &\geq \phi_\xi x_{4j} \text{ , for each } j \in J \\
\hat{x}_j &\geq -\beta_\xi x_{4j} \text{ , for each } j \in J \\
y_{10j} + y_{40j} &= \sum_{p \in P} \mu_p Y_{jp} - \hat{g}_j X_j \text{ , for each } j \in J \\
y_{1jp} + y_{4jp} &= Y_{js} \text{ , for each } j \in J, p \in P \\
x_{1j} + w_{4j} &= \sum_{p \in P} \rho_p Y_{jp} \text{ , for each } j \in J \\
\hat{\psi}_n &= r_{10n} + r_{40n} + \sum_{j \in J} \frac{1}{2\eta_{jn}} \hat{d}_{jn} \\
\sum_{n=1}^{N} \hat{c}_{0n} &= -\pi_0 \\
\sum_{n=1}^{N} \hat{c}_{pn} &= -(\pi_p - ht) \text{ , for each } p \in P \\
\sum_{n=1}^{N} \hat{c}_{\beta n} &= -\sum_{p \in P} (\pi_p - ht) \rho_p \\
\sum_{n=1}^{N} \hat{d}_{jn} &= h \text{ , for each } j \in J \\
\sum_{n=1}^{N} \hat{y}_{jpn} &= Y_{jp} \text{ , for each } j \in J, p \in P \\
\sum_{n=1}^{N} \hat{y}_{j\beta n} &= \sum_{p \in P} \rho_p Y_{jp} \text{ , for each } j \in J \\
\hat{y}_{jpn} &\geq 0, \text{ for each } j \in J, p \in P, n = 1, ..., N \\
r_{10n} &\geq y_{100n} + (\underline{\lambda}_0 - \mu_0) \hat{y}_{100n} + \sum_{p \in P} \max\{(\bar{\lambda}_p - \mu_p) y_{10pn}, (\underline{\lambda}_p - \mu_p) y_{10pn}\} \\
&\quad + \max\{\bar{\xi} x_{0n}, \underline{\xi} x_{n0}\} \text{ , for each } n = 1, ..., N \\
r_{40n} &\geq y_{400n} + \sqrt{-2ln(\epsilon_g)} \sqrt{u_{00n}^2 + \sum_{p \in P} u_{0pn}^2 + \hat{x}_{40n}^2} \text{ , for each } n = 1, ..., N \\
u_{0pn} &\geq \phi_p y_{40pn} \text{ , for each } p \in P, n = 1, ..., N \\
u_{0pn} &\geq -\beta_p y_{40pn} \text{ , for each } p \in P, n = 1, ..., N \\
u_{00n} &\geq \phi_0 \hat{y}_{400n} \text{ , for each } n = 1, ..., N \\
u_{00n} &\geq -\beta_0 \hat{y}_{400n} \text{ , for each } n = 1, ..., N
\end{aligned}
$$

$$
\begin{aligned}
\hat{x}_{0n} &\geq \phi_\xi x_{40n} \text{ , for each } n=1,...,N \\
\hat{x}_{0n} &\geq -\beta_\xi x_{40n} \text{ , for each } n=1,...,N \\
y_{100n}+y_{400n} &= \sum_{p\in P\cup\{0\}} \hat{c}_{pn}\mu_p + h\Phi^{-1}(\alpha)t\sum_{j\in J}\frac{\eta_{jn}}{2}\sum_{p\in P}\mu_p\hat{y}_{jpn} \text{ , for each } n=1,...,N \\
\hat{y}_{100n}+\hat{y}_{400n} &= \hat{c}_{0n} \text{ , for each } n=1,...,N \\
y_{10pn}+y_{40pn} &= \hat{c}_{pn}+h\Phi^{-1}(\alpha)t\sum_{j\in J}\frac{\eta_{jn}}{2}\hat{y}_{jpn} \text{ , for each } p\in P, n=1,...,N \\
x_{10n}+w_{40n} &= \hat{c}_{\beta n}\rho_0+\sum_{j\in J}h\Phi^{-1}(\alpha)t\frac{\eta_j}{2}\hat{y}_{j\beta n} \text{ , for each } n=1,...,N \\
X_j &\in \{0,1\} \text{ , for each } j\in J \\
Y_{jp} &\in \{0,1\} \text{ , for each } j\in J, p\in P \\
Z_{jq} &\in \{0,1\} \text{ , for each } j\in J, q\in Q.
\end{aligned}
$$

Second-order Conic Reformulation of (21)

Given fixed values of $\mathbf{Y}$, problem (21) can be reformulated as follows. Note that $\eta_j, j\in J$ now become decision variables.

$$\min r_1+r_4$$

subject to:

$$
\begin{aligned}
r_1 &\geq y_{10}+(\underline{\lambda}_0-\mu_0)\hat{y}_{10}+\sum_{p\in P}\max\{(\bar{\lambda}_p-\mu_p)y_{1p},(\underline{\lambda}_p-\mu_p)y_{1p}\}+\max\{\bar{\xi}x,\underline{\xi}x\} \\
r_4 &\geq y_{40}+\sqrt{-2ln(\epsilon_g)}\sqrt{\sum_{p\in P}u_p^2+\hat{x}_4^2} \\
u_p &\geq \phi_p y_{4p} \text{ , for each } p\in P \\
u_p &\geq -\beta_p y_{4p} \text{ , for each } p\in P \\
u_0 &\geq \phi_0\hat{y}_{40} \\
u_0 &\geq -\beta_0\hat{y}_{40} \\
\hat{x} &\geq \phi_\xi x_4 \\
\hat{x} &\geq -\beta_\xi x_4 \\
y_{10}+y_{40} &= \pi_0\lambda_0-\sum_{p\in P}(\pi_p-ht)\mu_p+\sum_{j\in J}h\Phi^{-1}(\alpha)t\frac{\eta_j}{2}\mu_p Y_{jp}+\sum_{j\in J}\frac{hO_j}{2}+T\sum_{j\in J}f_jX_j \\
y_{1p}+y_{4p} &= -(\pi_p-ht)+\sum_{j\in J}h\Phi^{-1}(\alpha)t\frac{\eta_j}{2}Y_{jp} \text{ , for each } p\in P \\
\hat{y}_{10}+\hat{y}_{40} &= -\pi_0 \\
x_1+w_4 &= -\pi_0\rho_0-\sum_{p\in P}(\pi_p-ht)\rho_p+\sum_{j\in J}h\Phi^{-1}(\alpha)t\frac{\eta_j}{2}\sum_{p\in P}\rho_p Y_{jp} \\
\eta_j+O_j &\geq \sqrt{2+\eta_j^2+O_j^2} \text{ , for each } j\in J \\
\eta_j &\geq 0 \text{ , for each } j\in J \\
O_j &\geq 0, \text{ for each } j\in J.
\end{aligned}
$$

References

[1] Birge, J.R. and J.H. Dulá (1991). Bounding Separable Recourse Functions with Limited Distribution Information. *Annals of Operations Research*, 30(1), 277-298.

[2] Chen, W. and M. Sim (2009). Goal-Driven Optimization. *Operations Research*, 57(2), 342-357.

[3] Chen, X., M. Sim and S. Peng (2007). A Robust Optimization Perspective on Stochastic Programming. *Operations Research*, 55(6), 1058-1071.

[4] Department of Transportation, State of California (2010). *2009 Traffic Volumes on the California State Highway System*, California, `http://traffic-counts.dot.ca.gov/2009all/2009TrafficVolumes.htm`.

[5] Galbraith, K. (2009, May 13). Better Place Unveils Battery Swap Station. *New York Times*, Retrieved from `http://green.blogs.nytimes.com/2009/05/13/better-place-unveils-battery-swap-station/`.

[6] Green Options (2010). Introduction To Electric Cars And Electric Scooters. Retrieved from `http://www.greenoptions.com/wiki/introduction-to-electric-cars-and-electric-scooters`.

[7] Jamison, B., S. Obey and W. Pruitt (1965). Convergence of Weighted Averages of Independent Random Variable. *Probability Theory and Related Fields*, 4(1), 40-44.

[8] Ortúzar, J. de D. and L.G. Willumsen (2001). *Modelling Transport, Third Edition.* John Wiley & Sons, New York.

[9] Pearre, N.S., W. Kempton, R.L. Guensler and V.V. Elango (2011). Electric Vehicles: How Much Range is Required for a Day's Driving? *Transportation Research Part C*, 19(6), 1171-1184.

[10] Popescu, I. (2007). Robust Mean-Covariance Solutions for Stochastic Optimization. *Operations Research*, 55(1), 98-112.