Price of Uncertainty in Multistage Stochastic Power Dispatch

Junjie Qin and Ram Rajagopal

Abstract— **The price of uncertainty measures the impact of uncertainty in the total cost of power dispatch in networks with significant penetration of renewables. It has been shown to exist in problems with deterministic prices and two stage markets. The existence of this price in multistage dispatch problem has remained an open question. This paper demonstrates the existence of the price of uncertainty in multiple stage stochastic power dispatch problems with stochastic prices. The existence proof requires a careful derivation of the structure of the optimal dispatch for this scenario. The paper concludes with various examples and applications of the price of uncertainty.**

I. Introduction

Power systems with deep penetration of renewable generation face significant costs to mitigate the increase in supply variability. In the existing power grid, generators are progressively scheduled limited by the lead times they require to ramp up outputs. This scheduling process requires utilizing sequentially updated forecasts of loads and renewable generator outputs. It also relies on prices decided according to the marginal cost of generation at each stage. Since stage by stage schedules are unknown in advance, uncertainty in prices at different periods needs to be considered. Various forms of stochastic and robust dispatch methods have been proposed as potential approaches to reduce energy and reserve capacity costs [1], [2], [3], [4], [5], [6]. Other settings including storage [7] and generator ramping constraints [8] have been developed as well. In common to all these approaches is the significant role of uncertainty in determining both the optimal control decisions and the total expected cost of the process. Therefore it is important to *benchmark the impact of uncertainty* in the power dispatch process.

The *price of uncertainty* has been proposed as a fundamental measure of performance of a stochastic dispatch procedure [5], [9]. It captures the idea that the additional cost of the dispatch process due to uncertainty is *linear* in the standard deviation of the forecast of the net of all loads and generators. The price of uncertainty has been shown to exist under mild assumptions in two-stage dispatch problems with prices assumed to be deterministic. Stochastic prices have been considered in [10] in a two stage setting, but the existence of this price was not verified. In this paper we demonstrate the existence of the price of uncertainty in a general setting with multiple stages and stochastic prices. In particular, we provide a more appropriate definition of this price as the partial variation of the expected cost with respect to forecast standard deviation. Calculating the price of uncertainty requires a formal derivation of the optimal dispatch control for sequential allocation with stochastic prices and multiple stages, significantly generalizing the results in [2]. Besides proving as a useful benchmark to measure the performance of arbitrary dispatch procedures and improvements in forecasts, the price of uncertainty also is a very useful structural result. We show its applicability, for example in renewable aggregation design problems where the existence of this price enables decoupling of uncertainty consideration from optimal selection of participating generators.

The paper is organized as follows. Section II formulates the stochastic control problem for multistage dispatch with forecast revisions and stochastic prices. Section III then analyzes the structural property of the problem, and derives closed form solution to the optimal dispatch policy and the cost-to-go function. The notion of price of uncertainty is then formalized for this setting in Section IV, where conditions for constant price of uncertainty are established together with its evaluation in various settings. Section V concludes the paper.

II. Problem Formulation

We consider the problem faced by a system operator, who needs to dispatch power production from a sequence of forward markets to meet an uncertain net demand. We call the time when the demand is serviced the *delivery time*. These markets are organized at different times; for instance, a day ahead of, an hour ahead of, and 15 minutes ahead of the delivery time. Two effects differentiate these markets from each other:

- More information about the net demand becomes available in later markets. As such, the forecast of the net demand is expected to improve as the delivery time approaches.
- Shorter lead times (that is, time for the generators to ramp up) are allowed in later markets. This restricts the participation of those cheap but slow generators, such as coal generators, into markets closer to the delivery time. Consequently, the price of later markets can be higher.

The system operator aims to strike a balance between *economy* and *reliability*. Here the reliability is defined in terms of supplying enough power to meet the demand. In each market, the system operator dispatches energy at a fixed price, which is random and becomes realized at the beginning of the market. We characterize the system reliability with a

J. Qin is with Institute for Computational and Mathematical Engineering, Stanford University, Stanford, CA, 94305. J. Qin is supported by the Satre Family Fellowship. Email: jqin@stanford.edu

R. Rajagopal is with the Department of Civil and Environmental Engineering, Stanford University, Stanford, CA, 94305. Email: ramr@stanford.edu

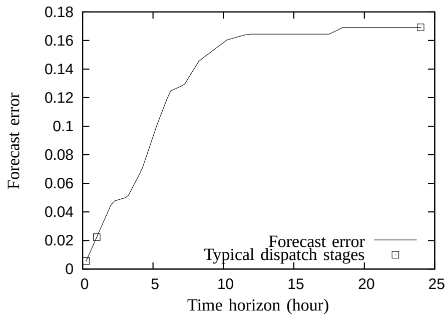

Fig. 1: Percentage root mean squared forecast error v.s. forecast horizon for a single wind farm. Data are obtained from Iberdrola Renewables.

risk measure, which is a function of the power imbalance at the delivery time. Thus the balance is achieved by utilizing all the information that is presently available at each of the markets, and taking a sequence of dispatch decisions to minimize an expected sum of the economic cost and systematic risk. The mathematical details of the the decision problem are provided in the following subsections.

A. Information Updates and Statistical Models of the Demand and Prices

Variable power generation is easier to forecast in a shorter forecast horizon. Figure 1 shows a typical error curve for wind forecasting for a single wind farm [2]. Intraday markets have been advocated and implemented in some countries or regions, in part to leverage the improved forecast quality for corrective dispatch actions.

In this paper, we explicitly model the information updates (or forecast revision) process. Let the sequence of markets (or stages) be indexed by $t = 1, \dots, T$, with the delivery time be indexed by $T+1$. Evidently, the information that is available at market t is a subset of that available at market $t+1$, for $t = 1, \dots, T$. This effect is captured using a version of the "onion-layer peeling" information update models developed for inventory control problems [11]. In particular, we model the forecast revision process for the net demand as an additive process, that is, the forecast of the net demand is updated as

$$\widehat{d}_{t+1} = \widehat{d}_t + \epsilon_{t+1}, \quad t = 1, \dots, T, \tag{1}$$

where $\widehat{d}_t$ is the stage-t forecast of the net demand, and ϵ_{t+1} is the forecast correction contributed by the newly available information during stage t. Note that because at the delivery time $T+1$, the net demand d is realized, we have $\widehat{d}_{T+1} = d$. This indicates that, if ϵ_t's are zero-mean and mutually independent for all t , then the $\{\widehat{d}_t\}$ process is a sequence of unbiased estimators of d with decreasing mean square errors, *i.e.*,

$$\mathbb{E}\left[\left(d - \widehat{d}_t\right)^2\right] = \mathbb{E}\left[\left(\sum_{\ell=t+1}^{T+1} \epsilon_\ell\right)^2\right] = \sum_{\ell=t+1}^{T+1} \mathrm{Var}(\epsilon_\ell)$$

decreases as t approaches $T+1$.

The prices for each unit of power in the future stages are usually not known ahead of time. Thus we model the price at each stage t as a random variable p_t, which is realized at the beginning of stage t. [1] Since the realization of the price p_t is part of the additional information revealed at the beginning of stage t, it may have an effect in the forecast update of stage t. More specifically, we assume that the price and forecast update at stage t have a joint probability density function (pdf) $f_t(\epsilon_t, p_t)$. For convenience in exposition, we assume (ϵ_t, p_t) and (ϵ_s, p_s) are independent for different stages s and t. Remark 3 discusses the relaxation of this assumption.

B. Decision Problem

The decision making problem can be described as follows:

- Before the first stage (for each delivery time), the system operator obtains the statistical information about the forecast corrections and prices, possibly from historical data. Such information contains $f_t(\epsilon_t, p_t)$ for $t = 1, \dots, T+1$.
- At the beginning of each stage t, $t = 1, \dots, T$, new information since the last decision is collected; quantities ϵ_t and p_t are revealed to the system operator. The forecast update (1) is performed. (For convenience, we write $\widehat{d}_1 = \widehat{d}_0 + \epsilon_1 = \epsilon_1$.)
- The system operator makes a dispatch decision $g_t \ge 0$ at the cost of $p_t g_t$.
- The accumulated power x_t is updated as
$$x_{t+1} = x_t + g_t, \quad t = 1, \dots, T. \tag{2}$$
- At the terminal stage $T+1$, $d = \widehat{d}_{T+1} = \widehat{d}_T + \epsilon_{T+1}$ and p_{T+1} are realized. A cost for imbalance $J(d - x_{T+1}; p_{T+1})$ is incurred.

This decision problem can be cast as a stochastic control program:

$$\begin{aligned}
\text{minimize} \quad & \mathbb{E}\left[\sum_{t=1}^{T} p_t g_t + J(d - x_{T+1}; p_{T+1})\right] && \text{(3a)}\\
\text{subject to} \quad & x_{t+1} = x_t + g_t, && \text{(3b)}\\
& \widehat{d}_{t+1} = \widehat{d}_t + \epsilon_{t+1}, && \text{(3c)}\\
& g_t \ge 0, && \text{(3d)}
\end{aligned}$$

where $x_1 = 0$, and we aim to find a control policy at each stage t that maps the information available at the stage to the optimal dispatch g_t. Here J is penalty function characterizing the systematic risk (or financial loss) due to the unmet demand $z = d - x_{T+1}$. We impose the following technical conditions on J:

A1 Convexity: For each $p \in \mathbb{R}$, the function $J(z; p)$ is convex in z, and thus is continuous and differentiable Lebesgue almost everywhere (a.e.).

A2 Integrability: $J(z + \epsilon_t; p_t)$ is integrable for all $z \in \mathbb{R}$, that is, $\mathbb{E}|J(z + \epsilon_t; p_t)| < \infty$, and differentiable a.e. with respect to the probability measure induced by the distribution of (ϵ_t, p_t), $t = 1, \dots, T+1$.

[1] We choose not to use a detailed information update model for the price process as the randomness in the price is usually not a primary concern for the system operator.

An example of primary interest in the power system literature is the Value of Loss Load (VOLL) penalty

$$J(d-x_{T+1};p_{T+1})=p_{T+1}(d-x_{T+1})_+,$$

where $(u)_+=\max(u,0)$. The penalty characterizes the cost of serving (or curtailing) the unfulfilled demand $d-x_{T+1}$ at a stochastic price p_{T+1}. Its expectation sometimes is referred to as a measure of operational risk of the power system. Although we are working mostly with the function J before taking expectation, we note that, as our assumptions on J is mild, upon taking expectation, such form can incorporate many other risk measures studied in the finance and portfolio management literature (*cf.* coherent risk measures discussed in [12]).

III. Optimal Dispatch

Observing the additive structure of the updates in x_t and $\widehat{d}_t$, and the fact that the last-stage cost only depends on $d-x_{T+1}$, we can re-write (3) in the following form:

$$\begin{aligned} \text{minimize} \quad & \mathbb{E}\left[\sum_{t=1}^{T} p_t g_t + J(z_{T+1};p_{T+1})\right] && \text{(4a)}\\ \text{subject to} \quad & z_{t+1}=z_t-g_t+\epsilon_{t+1}, && \text{(4b)}\\ & g_t\geq 0, && \text{(4c)}\end{aligned}$$

where $z_t=\widehat{d}_t-x_t$, $t=1,\dots,T+1$, is the (expected) residual net demand after accounting for the accumulated power x_t.

From the theory of stochastic dynamic programming, the following sequence of the functions are the key to solving (4):

Definition 1: The cost-to-go functions are defined via the Bellman's recursion:

$$J_{T+1}(z_{T+1};p_{T+1})=J(z_{T+1};p_{T+1}), \qquad \text{(5a)}$$

$$J_t(z_t;p_t)=\min_{g_t\geq 0}\Big\{p_tg_t+\mathbb{E}[J_{t+1}(z_t-g_t+\epsilon_{t+1};p_{t+1})]\Big\}, \qquad \text{(5b)}$$

for $t=1,\dots,T$, where the expectation is taken over ϵ_{t+1} and p_{t+1}. The objective of (5b) is defined as the state-action cost-to-go function[2], *i.e.*, for stage $t=1,\dots,T$, we have

$$Q_t(z_t,g_t;p_t)=p_tg_t+\mathbb{E}[J_{t+1}(z_t-g_t+\epsilon_{t+1};p_{t+1})]. \quad (6)$$

For each fixed stage t and realization p_t, the cost-to-go function summarizes the minimum expected cost from the current stage with the given residual net demand z_t to the end of the decision process. The optimal dispatch can then be solved from

$$g_t^\star(z_t;p_t)=\operatorname*{argmin}_{g_t\geq 0} Q_t(z_t,g_t;p_t).^3 \qquad (7)$$

Equations (5), (6), and (7) can be thought of as a characterization of the optimality condition for (4). However, they do not directly lead to an efficient algorithm for computing the optimal dispatch, because, for example, using (5b) to obtain the function J_T from J_{T+1} would require solving an optimization for each possible $(z_T;p_T)$ pair, whereas the total number of such pairs is clearly infinite. Thus a closer investigation of the problem is necessary. We start with a convexity result about the cost-to-go functions.

Lemma 1: For each stage $t=1,\dots,T+1$ and each fixed p_t, the cost-to-go function J_t is convex in z_t. Similarly, for each stage $t=1,\dots,T$ and each fixed p_t, the state-action cost-to-go function Q_t is convex in (z_t,g_t).

Proof: By a set of standard arguments, see, for example, [13] and references therein. ■

A direct consequence of this result is the following structural characterization of the optimal dispatch policy $g_t^\star(z_t;p_t)$:

Lemma 2: For each stage $t=1,\dots,T$, there exists an adjustment $\Delta_t(p_t)$ which is independent of z_t, such that the optimal dispatch takes the form

$$g_t^\star(z_t;p_t)=(z_t-\Delta_t(p_t))_+. \qquad (8)$$

Proof: Consider the Bellman's recursion at stage t, in which for each fixed z_t and p_t, the following optimization is solved

$$\begin{aligned}&\min_{g_t\geq 0}\{p_tg_t+\mathbb{E}[J_{t+1}(z_t-g_t+\epsilon_{t+1};p_{t+1})]\}\\ &=p_tz_t+\min_{y_t\leq z_t}\{-p_ty_t+\mathbb{E}[J_{t+1}(y_t+\epsilon_{t+1}:p_{t+1})]\}, \quad (9)\end{aligned}$$

where the identity follows from a change of variable $y_t=z_t-g_t$. Let $\Delta_t(p_t)=\operatorname{argmin}_{y_t\in\mathbb{R}} h(y_t)$, where $h(y_t)=-p_ty_t+\mathbb{E}[J_{t+1}(y_t+\epsilon_{t+1};p_{t+1})]$, and notice that h is convex in view of Lemma 1. Thus the constrained minimizer of y_t in (9) is of the form $y_t^\star=\min(\Delta_t(p_t),z_t)$. Indeed, the claim obviously holds when $\Delta_t(p_t)\leq z_t$. When $\Delta_t(p_t)>z_t$ and $y_t^\star=z_t$, suppose there is another $y_t'<y_t^\star<\Delta_t(p_t)$ such that $h(y_t')<h(y_t^\star)$. It follows that there exists a $\theta\in(0,1)$ such that $y_t^\star=\theta y_t'+(1-\theta)\Delta_t(p_t)$. By convexity of h, we have

$$\begin{aligned}h(y_t^\star)&=h(\theta y_t'+(1-\theta)\Delta_t(p_t))\leq\theta h(y_t')+(1-\theta)h(\Delta_t(p_t))\\ &\leq\theta h(y_t')+(1-\theta)h(y_t')=h(y_t'),\end{aligned}$$

thus a contradiction. Recalling the relation between y_t and g_t, we conclude that $g_t^\star(z_t;p_t)=(z_t-\Delta_t(p_t))_+$. ■

This result is intuitive: whenever feasible, the optimal dispatch is the expected residual net demand z_t plus a price-dependent adjustment $(-\Delta_t(p_t))$ (We will see in later examples that $\Delta_t(p_t)$ is likely to be negative.) The reserve $(-\Delta_t(p_t))$ is optimally selected to hedge against the uncertainty. Such structural results are in particular appealing when the prices are deterministic.

Remark 1 (The case of deterministic prices): If the $\{p_t\}$ sequence can be forecast accurately ahead of time, then $\Delta_t(p_t)\equiv\Delta_t$ and may be computed offline.

We proceed to give a closed-form characterization of these adjustments. To that end, the following functions turn out to be pivotal for the development of the analytical solutions to (4):

[2]The terminology is borrowed from the literature of Markov decision processes and reinforcement learning.

[3]When the minimizer is not unique, an arbitrary minimizer is selected. We use the same convention in the sequel.

Lemma 3: Recursively, from backward, define

$$K_T(y) = \mathbb{E}[J'(y + \epsilon_{T+1}; p_{T+1})], \tag{10a}$$

$$K_t(y) = \mathbb{E}\Big\{p_{t+1}\mathbf{1}(K_{t+1}(y+\epsilon_{t+1}) \geq p_{t+1}) + K_{t+1}(y+\epsilon_{t+1})\mathbf{1}(K_{t+1}(y+\epsilon_{t+1}) < p_{t+1})\Big\}, \tag{10b}$$

for $t = 1, \dots, T-1$, where $J'(y;p)$ is the derivative of J with respect to y.[4] Let $K_t^{\min} = \inf_{y\in\mathbb{R}} K_t(y)$ and $K_t^{\max} = \sup_{y\in\mathbb{R}} K_t(y)$. Define, for $t = 1, \dots, T$,

$$\widetilde{\Delta}_t^p = \widetilde{\Delta}_t(p_t) = \begin{cases} K_t^{-1}(p_t) & \text{if } K_t^{\min} \leq p_t \leq K_t^{\max}, \\ -\infty & \text{if } p_t < K_t^{\min}, \\ +\infty & \text{if } p_t > K_t^{\max}, \end{cases} \tag{11}$$

where $K_t^{-1}(p) = \inf\{y : K_t(y) \geq p\}$, and

$$G_T(y) = \mathbb{E}[J(y + \epsilon_{T+1}; p_{T+1})], \tag{12a}$$

$$G_t(y) = \mathbb{E}\Big\{ \Big[p_{t+1}(y + \epsilon_{t+1} - \widetilde{\Delta}_{t+1}^p) + G_{t+1}(\widetilde{\Delta}_{t+1}^p)\Big] \times \mathbf{1}(K_{t+1}(y+\epsilon_{t+1}) \geq p_{t+1}) + G_{t+1}(y+\epsilon_{t+1})\mathbf{1}(K_{t+1}(y+\epsilon_{t+1}) < p_{t+1})\Big\}, \tag{12b}$$

for $t = 1, \dots, T-1$. The following claims hold for each $t = 1, \dots, T$:

(a) $\widetilde{\Delta}_t^p$ is well-defined;
(b) $G_t(y)$ is convex and differentiable;
(c) $G_t'(y) = K_t(y)$;
(d) $K_t(y)$ is nondecreasing, so $K_t^{\min} = \lim_{y\to-\infty} K_t(y)$, and $K_t^{\max} = \lim_{y\to+\infty} K_t(y)$.

Proof: See appendix. ■

Different from the cost-to-go functions $\{J_t\}$ defined in (5), the sequences of functions $\{K_t\}$ and $\{G_t\}$ can be evaluated without optimization. Given the joint density of (ϵ_t, p_t)'s, many efficient simulation or numerical integration schemes can be utilized to compute these $\{K_t\}$ and $\{G_t\}$ functions. We now state the optimal dispatch and cost-to-go functions in terms of the quantities defined in Lemma 3.

Theorem 1: For $t = 1, \dots, T$, the adjustment in (8) is the same as $\widetilde{\Delta}_t^p$ defined in (11), that is, the optimal dispatch is

$$g_t^\star(z_t; p_t) = (z_t - \Delta_t(p_t))_+ = (z_t - \widetilde{\Delta}_t^p)_+, \tag{13}$$

and the cost-to-go function is

$$J_t(z_t; p_t) = \begin{cases} p_t(z_t - \Delta_t(p_t)) + G_t(\Delta_t(p_t)) & \text{if } \Delta_t(p_t) \leq z_t, \\ G_t(z_t) & \text{if } \Delta_t(p_t) > z_t. \end{cases} \tag{14}$$

Proof: See appendix. ■

We provide an illustrative example to demonstrate the use of Lemma 3 and Theorem 1.

Example 1 (Two-stage with VOLL penalty): We consider an instance of the problem with two stages. In the first stage (*e.g.*, day-ahead market), a net demand $z_1 = \widehat{d}_1 - x_1 = \widehat{d}_1 > 0$ is predicted. In the second stage (*e.g.*, real-time market), the actual net demand $d = \widehat{d}_2 = \widehat{d}_1 + \epsilon_2$ is realized. The VOLL penalty is used, *i.e.*, $J(z_2) = p_2(z_2)_+$, where p_2 is random in the first stage. Then we have

$$K_1(y) = \mathbb{E}[p_2\mathbf{1}(y + \epsilon_2 \geq 0)]$$

with $K_1^{\min} = 0$ and $K_1^{\max} = \mathbb{E}p_2$. Note that if $p_1 < K_1^{\min} = 0$, in the first stage it is optimal to dispatch infinite amount of power; if $p_1 > K_1^{\max} = \mathbb{E}p_2$, it is expected to be cheaper to dispatch power in the real time and thus there is little incentive to dispatch any power in the first stage. The case of most practical interest is $0 \leq p_1 \leq \mathbb{E}p_2$. In this case, we have $\widetilde{\Delta}_1^p = K_1^{-1}(p_1)$ and so the optimal first stage dispatch is $g_1^\star(z_1; p_1) = (\widehat{d}_1 - K_1^{-1}(p_1))$.

To further crystallize the intuition for the optimal adjustment $\widetilde{\Delta}_t^p$ and connect our results to earlier results in the literature, we consider a special case of Example 1 where the prices are deterministic.

Example 2 (Two-stage with deterministic prices): In the same setting as Example 1, we further assume all prices are known ahead of time. Let $\overline{F}_{\epsilon_2}(\cdot) = 1 - F_{\epsilon_2}(\cdot)$ be the complementary cumulative distribution function of ϵ_2. Then $\widetilde{\Delta}_1^p = -\overline{F}_{\epsilon_2}^{-1}(p_1/p_2)$, and $g_1^\star(z_1; p_1) = (\widehat{d}_1 + \overline{F}_{\epsilon_2}^{-1}(p_1/p_2))$. Note that if the distribution of ϵ_2 is such that $\mathbb{P}(\epsilon_2 \geq 0) = \mathbb{P}(\epsilon_2 \leq 0)$, then $\overline{F}_{\epsilon_2}^{-1}(p_1/p_2) \geq 0$ whenever $p_1/p_2 \leq 1/2$. The condition $p_1/p_2 \leq 1/2$ is usually the case when the second stage price corresponds to the marginal cost of using fast ramping generators to supply the shortfall, whose marginal cost can easily be ten times higher than that of the first stage slow generators.

We give several remarks regarding the evaluation of the optimal policy in general cases and the relaxation of the inter-temporal independence assumption.

Remark 2 (Computation for general cases): We focus our analytical examples on two-stage cases, as in these cases, simple closed-form results can obtained and they are usually easy to be interpreted. However, the optimal dispatch procedure for the general multistage and stochastic price case, as developed in Lemman 3 and Theorem 1 can be implemented for numerical solutions to the dispatch problem as numerical integration is a mature field with many well-developed software routines. We also hint that an efficient implementation may exploit the fact that the integration in ϵ_t terms can be expressed as a convolution integral, which allows efficient implementations.

Remark 3 (Removing the independence assumption): The sequences $\{K_t\}$ and $\{G_t\}$ provide sufficient information in computing the optimal dispatch and optimal cost-to-go functions. Under the inter-temporal independence assumption, evaluating G_t (for instance) from G_{t+1} requires an two dimensional integration, with respect to the realizations of $(\epsilon_{t+1}, p_{t+1})$ and weighted by the joint density $f_{t+1}(\epsilon_{t+1}, p_{t+1})$. One can obtain analytical results similar to that in this section without the inter-temporal independence assumption. In this case, instead of iterated integration, evaluating G_t would require integration with respect to $\{(\epsilon_s, p_s) : t+1 \leq s \leq T+1\}$ weighted by their joint density conditional on the observed realization of $\{(\epsilon_s, p_s) : s < t+1\}$. As such, the computation is much

[4] At points where J is non-differentiable (which form a set of measure zero), we can define J' to be any of J's subderivative at the point.

more cumbersome. In practice, the independence assumption is widely used also because obtaining the joint density for the entire process $\{(\epsilon_s, p_s) : s \leq T+1\}$ requires high order information and may not be feasible from historical data.

IV. Price of Uncertainty

The term *price of uncertainty* (PoU) is coined in [5] as a key quantity to characterize the dependence between the integration costs of variable energy resources and the level of uncertainty in the system. In a simple setting (that of Example 2 with the forecast correction being Gaussian), it is shown in [5] that the *expected* integration cost, defined as the difference between the optimal cost of (4) and the expected clairvoyant cost conditional on the first stage forecasts [5], is a linear function of the standard deviation of the forecast error, when the uncertainty in the system is small. Here we take an alternative approach to reduce the last requirement; we consider the *realized* integration cost, which replaces the expected clairvoyant cost in the previous definition by its actual realized value. Because the actual clairvoyant cost is independent from the ability to forecast the demand and prices in forward markets, the PoU defined using the realized integration cost is the same as the sensitivity of the optimal cost of (4) to the uncertainty level. We will show later that this treatment not only simplifies the calculation, but also leads to consistent results as in [5] and at the same time is obtained without invoking the assumption that the uncertainty level is small.

A. Existence in Two-Stage Dispatch

We start by formalizing these notions for a two-stage problem with the following additional requirement on the function J:

A3 Positive homogeneity: For all $a \geq 0$, $J(ay) = aJ(y)$.

Note that convexity and positive homogeneity are common requirements on "good" risk measures. (Coherent risk measures also satisfy translation invariant and monotonicity conditions which we do not assume here.)

Furthermore, based on discussions in the previous section, we rule out degenerate cases by requiring the support of p_t be a subset of $[K_t^{\min}, K_t^{\max}]$, that is,

$$f_t(\epsilon, p) = 0, \text{ if } p \notin [K_t^{\min}, K_t^{\max}], \tag{15}$$

for any ϵ and all t. Then the two-stage price of uncertainty can be defined as follows.

Definition 2 (Price of uncertainty, two-stage case): For a two-stage problem with a fixed $z_1 = \widehat{d}_1$, let the forecast correction be $\epsilon_2 = \sigma E$, where E has zero mean and unit standard deviation. The optimal dispatch cost is $J_1(\sigma; \widehat{d}_1, p_1) \triangleq J_1(\widehat{d}_1; p_1)$, and the price of uncertainty is defined as

$$p^{\text{PoU}}(\sigma, \widehat{d}_1, p_1) \triangleq \partial J_1(\sigma; \widehat{d}_1, p_1)/\partial\sigma.$$

In the sequel, we often omit the dependence of p^{PoU} on certain parameters whenever appropriate.

[5]The clairvoyant cost is the minimum cost of (4) assuming all stochastic elements are known ahead of time.

The next result gives a necessary and sufficient condition for the PoU to be a constant independent of σ (or the integration cost is a linear function of σ); formulas for the PoU are also provided.

Theorem 2: Let $K_{E,1}(y) \triangleq \mathbb{E}[J'(y+E;p_2)]$ and $G_{E,1}(y) \triangleq \mathbb{E}[J(y+E;p_2)]$ be the K-function and G-function defined as in Lemma 3 for the normalized forecast correction E. Then the following statements are true for each p_1 and $\widehat{d}_1$:

(a) The PoU is independent of σ and the integration cost is affine in σ if and only if

$$\widetilde{\Delta}_1^p \triangleq \sigma K_{E,1}^{-1}(p_1) \leq \widehat{d}_1, \tag{16}$$

in which case, the PoU is

$$p^{\text{PoU}}(p_1) = -p_1 K_{E,1}^{-1}(p_1) + G_{E,1}\left(K_{E,1}^{-1}(p_1)\right), \tag{17}$$

and the optimal dispatch cost is

$$J_1(\sigma; \widehat{d}_1, p_1) = p_1 \widehat{d}_1 + \sigma p^{\text{PoU}}(p_1). \tag{18}$$

We refer to this case as the *constant PoU* case as our primary focus is on the dependence of PoU on the uncertainty level σ, and p_1 is known at the beginning of stage 1.

(b) If the condition (16) fails, *i.e.*, $\widetilde{\Delta}_1^p > \widehat{d}_1$, then

$$p^{\text{PoU}}(\sigma, \widehat{d}_1) = G_{E,1}\left(\widehat{d}_1/\sigma\right) - \frac{\widehat{d}_1}{\sigma} K_{E,1}\left(\widehat{d}_1/\sigma\right). \tag{19}$$

Proof: See appendix. ■

B. Examples

We illustrate the use of Theorem 3 and discuss the intuitions behind it through two simple examples.

Example 3 (Condition of constant PoU for Example 2): Using the setup of Example 2, we have that the PoU is constant whenever

$$\widehat{d}_1 \geq \sigma K_{E,1}^{-1}(p_1) = -\sigma \overline{F}_E^{-1}(p_1/p_2). \tag{20}$$

It can be seen that under this condition, the unconstrained minimizer of g is no smaller than 0, that is, no nonlinear effect of the thresholding on g is introduced. Assuming $\widehat{d}_1 > 0$ (the renewable penetration is less than 100%), $\mathbb{P}(E > 0) = \mathbb{P}(E < 0)$, the above condition can be written as

$$\begin{cases} \frac{\sigma}{\widehat{d}_1} \geq -\frac{1}{\overline{F}_E^{-1}(p_1/p_2)} & \text{if } p_2 > 2p_1, \\ \frac{\sigma}{\widehat{d}_1} \leq -\frac{1}{\overline{F}_E^{-1}(p_1/p_2)} & \text{if } p_2 < 2p_1, \end{cases}$$

where the case $p_2 = 2p_1$ reduces to $\widehat{d}_1 \geq 0$ which holds by assumption. This indicates when the price in second stage is sufficiently higher ($p_2 > 2p_1$), the PoU is always a constant. When the second stage price is such that $p_2 < 2p_1$, the PoU is constant only if the (normalized) uncertainty level $\sigma/\widehat{d}_1$ is smaller than a certain threshold. Figure 2 indicates that for a dominant portion of the parameter regions, PoU is constant in the uncertainty level.

Example 4 (Expressions of PoU for Example 2): For the constant PoU case (when (20) holds), by (17),

$$\begin{aligned} p^{\mathrm{PoU}}(p_1) &= p_1 \overline{F}_E^{-1}(p_1/p_2) + G_{E,1}\left(-\overline{F}_E^{-1}(p_1/p_2)\right) \\ &= p_1 \overline{F}_E^{-1}(p_1/p_2) + p_2 \mathbb{E}\left[\left(E - \overline{F}_E^{-1}(p_1/p_2)\right)_+\right] \\ &= p_2 \mathbb{E}\left[E;\ E \geq \overline{F}_E^{-1}(p_1/p_2)\right], \end{aligned}$$

and in the Gaussian case, the last expression further simplifies to $p_2 f_E\left(\overline{F}_E^{-1}(p_1/p_2)\right)$, where f_E is the pdf of E. In the other case where PoU is not a constant, (19) reduces to

$$\begin{aligned} & p^{\mathrm{PoU}}(\sigma, \widehat{d}_1) \\ &= p_2 \mathbb{E}\left[(\widehat{d}_1/\sigma + E)_+\right] - \frac{\widehat{d}_1}{\sigma}\mathbb{E}\left[\mathbf{1}(\widehat{d}_1/\sigma + E \geq 0)\right] \\ &= p_2 \mathbb{E}[E; E \geq -\widehat{d}_1/\sigma], \end{aligned}$$

and similarly when the forecast correction is Gaussian, the last expression reduces to $p_2 f_E(-\widehat{d}_1/\sigma)$. Figure 2 depicts the values of the PoU for various p_1/p_2 and $\sigma/\widehat{d}_1$ ratios. These results are consistent with that in [5].

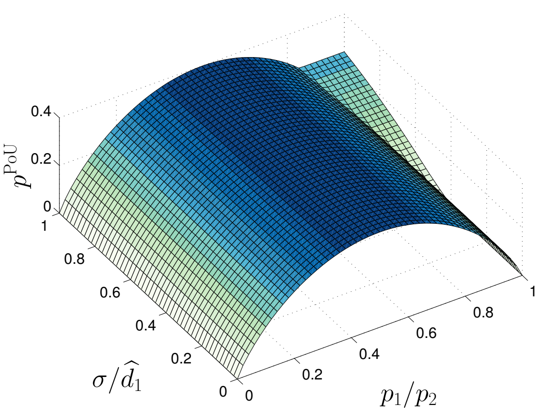

Fig. 2: Values of the PoU for Gaussian forecast errors and $p_2 = 1$. The PoU values are computed for various p_1/p_2 and $\sigma/\widehat{d}_1$ values. PoU is not constant in the uncertainty level only in the area where both ratios p_1/p_2 and $\sigma/\widehat{d}_1$ are close to 1.

Example 5 (Placement of a single forward market): One important application is to decide at which time to place a single forward market. Theoretical guidance about such decision may be obtained by studying the relation between the PoU and the lead time, denoted by τ, of the market. Selecting a lead time that has a small PoU may facilitate renewable integration; in this case, increasing the renewable penetration (which in turn increases σ) results in relatively small growth of system cost due to the increased uncertainty. Mathematically, suppose one has the empirical relation characterizing how the forecast error standard deviation, denoted by $\sigma(\tau)$, and the price for each unit of power generation in the forward market, denoted by $p(\tau)$, depend on the lead time τ. We expect $\sigma(\tau)$ to be non-decreasing and $p(\tau)$ to be non-increasing. Substituting these functions in the expressions of PoU in Theorem 2, one obtains an explicit relation between PoU and the lead time.

Example 6 (Separation between forecast and uncertainty): In applications that involve the aggregation of diverse and stochastic power consumption, under mild assumptions, it has been shown that the level of uncertainty σ usually depends primary on the level of aggregation (*e.g.*, the number of users that have been aggregated) [14]. PoU, in this context, captures the sensitivity of the system cost conditioning on an optimal stochastic dispatch to the uncertainty level. The optimal cost formula (18) for the constant PoU case is extremely useful in such applications, as it allows optimization of design choices, such as deciding the optimal aggregation group size and the selection of users without the need to solves a stochastic control program.

C. Existence in Multiple Stage Dispatch

We proceed to generalize these results to multi-stage settings. The following definition is analogous to Definition 2

Definition 3 (Price of uncertainty, general case): For general multi-stage problem with a fixed $z_1 = \widehat{d}_1$, let the forecast correction be $\epsilon_t = \sigma E_t$, $t = 2, \ldots, T+1$, where $[E_2, \ldots, E_{T+1}]$ is a reference forecast correction vector. The optimal dispatch cost is $J_1(\sigma; \widehat{d}_1, p_1) \triangleq J_1(\widehat{d}_1; p_1)$, and the price of uncertainty is defined as

$$p^{\mathrm{PoU}}(\sigma; \widehat{d}_1, p_1) \triangleq \partial J_1(\sigma; \widehat{d}_1, p_1)/\partial\sigma.$$

The key to calculate the PoU in multi-stage settings is the following lemma regarding scaling properties of function K_t and G_t:

Lemma 4: Let the K-functions and G-functions for the reference forecast correction, denoted by $\{K_{E,t}\}$ and $\{G_{E,t}\}$, be defined via (10) and (12), respectively, with ϵ_t replaced by E_t. Given J is positive homogeneous, we have

$$G_t(y) = \sigma G_{E,t}(y/\sigma), \tag{21}$$

$$K_t(y) = K_{E,t}(y/\sigma), \tag{22}$$

and

$$\widetilde{\Delta}_t^p = \sigma K_{E,t}^{-1}(p_t) \tag{23}$$

for $t = 1, \ldots, T$.

We summarize our results for the PoU in multi-stage setting in the following theorem:

Theorem 3: For each stage $t = 1, \ldots, T$, the cost-to-go depends on the uncertainty level in the form

$$J_t(\sigma; z_t, p_t) = \begin{cases} p_t(z_t - \sigma K_{E,t}^{-1}(p_t)) + \sigma G_{E,t}(K_{E,t}^{-1}(p_t)) & \text{if } \sigma K_{E,t}^{-1}(p_t) \leq z_t, \\ \sigma G_{E,t}(z_t/\sigma) & \text{if } \sigma K_{E,t}^{-1}(p_t) > z_t, \end{cases}$$

where $\{K_{E,t}\}$ and $\{G_{E,t}\}$ are defined for the reference problem and are independent of σ. Replacing $K_{E,1}$ and $G_{E,1}$ in Theorem 2 by that in Lemma 4, then all results in Theorem 2 hold for the multistage case.

Proof: See Appendix. ■

V. CONCLUSIONS

In this paper, we formulate the stochastic control problem of multistage risk limiting dispatch with forecast revisions and stochastic prices. The structure of the problem is studied, based on which the closed form solutions of the optimal control and cost-to-go are obtained. These formula are then used to give analytical characterization of the price of uncertainty in a general setting. Necessary and sufficient conditions for the price of uncertainty to be independent of the uncertainty level have been established. Simple numerical examples demonstrate that in a two stage setting and with Gaussian forecast error, the region where the price of uncertainty is constant is large. Formula for the price of uncertainty are also developed. In future work, we present efficient numerical algorithm for the general problem setting, and investigate the possibility of including forecast updates for prices.

REFERENCES

[1] P. Varaiya, F. Wu, and J. Bialek, "Smart operation of smart grid: Risk-limiting dispatch," *Proceedings of the IEEE*, vol. 99, no. 1, pp. 40–57, 2011.

[2] R. Rajagopal, E. Bitar, P. Varaiya, and F. Wu, "Risk-limiting dispatch for integrating renewable power," *International Journal of Electrical Power & Energy Systems*, vol. 44, no. 1, pp. 615–628, 2013.

[3] M. Ilic, L. Xie, and J. Joo, "Efficient Coordination of Wind Power and Price-Responsive Demand Part I: Theoretical Foundations," *IEEE Transactions in Power Systems*, vol. 26, no. 4, pp. 1875–1884, 2011.

[4] A. Conejo, M. Carrin, and J. Morales, *Decision making under uncertainty in electricity markets*. Springer Verlag, 2010, vol. 153.

[5] R. Rajagopal, D. Tse, and B. Zhang, "Network risk limiting dispatch: Optimal control and price of uncertainty," accepted*, IEEE Trans. on Automatic Control*, 2012.

[6] J. Warrington, P. Goulart, S. Mariethoz, and M. Morari, "Policy-Based Reserves for Power Systems," *IEEE Transactions on Power Systems*, vol. 28, no. 4, pp. 4427–4437, Nov 2013.

[7] J. Qin, H. Su, and R. Rajagopal, "Storage in risk limiting dispatch: Control and approximation," in *American Control Conference, 2013*, 2013.

[8] J. Qin, B. Zhang, and R. Rajagopal, "Risk limiting dispatch with ramping constraints," in *IEEE International Conference on Smart Grid Communications (SmartGridComm)*, 2013, pp. 791–796.

[9] J. Nair, S. Adlakha, and A. Wierman, "Energy procurement strategies in the presence of intermittent sources," Caltech Technical Report, Tech. Rep., 2012.

[10] E. Bitar, R. Rajagopal, P. Khargonekar, K. Poolla, and P. Varaiya, "Bringing wind energy to market," *IEEE Transactions on Power Systems*, vol. 27, no. 3, pp. 1225–1235, 2012.

[11] S. P. Sethi, H. Yan, and H. Zhang, "Peeling layers of an onion: Inventory model with multiple delivery modes and forecast updates," *Journal of Optimization Theory and Applications*, vol. 108, no. 2, pp. 253–281, 2001.

[12] S. Y. Chun, A. Shapiro, and S. Uryasev, "Conditional value-at-risk and average value-at-risk: Estimation and asymptotics," *Operations Research*, vol. 60, no. 4, pp. 739–756, 2012.

[13] J. Qin, H. Su, and R. Rajagopal, "Risk limiting dispatch with fast ramping storage," *Submitted to IEEE Transactions on Automatic control*, 2012. [Online]. Available: http://arxiv.org/abs/1212.0272

[14] S. Patel, R. Sevlian, B. Zhang, and R. Rajagopal, "Aggregation for Load Servicing," *ArXiv e-prints*, Dec. 2013.

[15] D. Bertsekas, "Stochastic optimization problems with nondifferentiable cost functionals," vol. 12, no. 2, pp. 218–231, 1973.

APPENDIX

We prove all claims in Lemma 3 and Theorem 1 via backward induction, together with the following additional claim for each stage:

Proposition 1: For each $t = 1, \dots, T$ and all $y \in \mathbb{R}$, $y + \epsilon_t \geq \Delta_t(p_t)$ if and only if $K_t(y + \epsilon_t) \geq p_t$.

Base case: For stage T, $G_T(y)$ is convex as J is convex and it is differentiable such that

$$G'_T(y) = \frac{d\mathbb{E}[J(y + \epsilon_{T+1}); p_{T+1}]}{dy} = \mathbb{E}[J'(y + \epsilon_{T+1}; p_{T+1})];$$

see *e.g.*, [15] for more details. Thus we have $G'_T(y) = K_T(y)$. It follows that K_T is nondecreasing, with $K_T^{\max}$ and $K_T^{\min}$ defined as in Lemma 3. By intermediate value theorem, $\widetilde{\Delta}_T = K_T^{-1}(p_T)$ is well-defined whenever $K_T^{\min} \leq p_T \leq K_T^{\max}$. Consider the stage-$T$ optimization. Using the change of variable discussed in the proof for Lemma 2, we aim to solve the optimization

$$p_T z_T + \min_{y_T \leq z_T} \{-p_T y_T + \mathbb{E}[J(y_T + \epsilon_{T+1} : p_{T+1})]\}, \quad (24)$$

whose objective is convex and its derivative with respect to y_T is $-p_T + K_T(y)$. If $p_T < K_T^{\min}$, $-p_T + K_T(y)$ is non-decreasing for $y \in \mathbb{R}$; if $p_T > K_T^{\max}$, $-p_T + K_T(y)$ is non-increasing for $y \in \mathbb{R}$; if $K_T^{\min} p_T \leq K_T^{\max}$, by the first-order optimality condition, $K_T^{-1}(p_T)$ is the unconstrained minimizer of (24). Thus on extended reals $\mathbb{R} \cup \{\pm\infty\}$, one can summarize the unconstrained minimizer of (24) in the form of (11). Per Lemma 2 and using $y_T = z_T - g_T$, the optimal last stage dispatch is of the form $g_T^\star(z_T; p_T) = (z_T - \Delta_T(p_T))_+ = (z_T - \widetilde{\Delta}_T^p)_+$ as claimed. Then depending on whether $z_T - \widetilde{\Delta}_T^p \geq 0$, J_T has the form

$$J_T(z_T; p_T) = \begin{cases} p_T(z_T - \Delta_T(p_T)) + G_T(\Delta_T(p_T)) & \text{if } \Delta_T(p_T) \leq z_T, \\ G_T(z_T) & \text{if } \Delta_T(p_T) > z_T. \end{cases}$$

For the statement in Proposition 1, we have that

$$\begin{aligned} \{y + \epsilon_T \geq \Delta_T(p_T)\} &= \{y + \epsilon_T \geq K_T^{-1}(p_T)\} \cup \{p_T < K_T^{\min}\} \\ &= \{K_T(y + \epsilon_T) \geq p_T\} \cup \{p_T < K_T^{\min}\} \\ &= \{K_T(y + \epsilon_T) \geq p_T\}. \end{aligned}$$

Inductive step: Suppose all claims hold for stage $t + 1$. In particular, $J_{t+1}(z_{t+1}; p_{t+1})$ has the form as in (5b) and $\mathbf{1}(K_{t+1}(y + \epsilon_{t+1}) \geq p_{t+1}) = \mathbf{1}(\Delta_{t+1}(p_{t+1}) \leq z_{t+1})$. It is easy to check that, by definition of G_t in (12),

$$G_t(y) = \mathbb{E}[J_{t+1}(y + \epsilon_{t+1}; p_{t+1})], \quad (25)$$

and so is convex by Lemma 1. Again by [15] [6] and using $G'_{t+1}(y) = K_{t+1}(y)$, we have

$$G'_t(y) = \mathbb{E}\Big\{ p_{t+1}\mathbf{1}(K_{t+1}(y + \epsilon_{t+1}) \geq p_{t+1}) + K_{t+1}(y + \epsilon_{t+1})\mathbf{1}(K_{t+1}(y + \epsilon_{t+1}) < p_{t+1}) \Big\},$$

that is, $G'_t(y) = K_t(y)$. Therefore $K_t(y)$ is nondecreasing. The rest of the claims in Lemma 3 and the claim in Proposition 1 follow from the same arguments as in the

[6] One can alternatively check this calculation using Leibniz's rule.

base case. For the claims in Theorem 1, consider stage-t optimization

$$\min_{g_t\ge 0}\{p_t g_t+\mathbb{E}[J_{t+1}(z_t-g_t+\epsilon_{t+1};p_{t+1})]\}$$
$$=p_t z_t+\min_{y_t\le z_t}\{-p_t y_t+G_t(y_t)\},$$

where the identity follows from the usual change of variable and (25). Repeating the arguments in the base case with $G_t(y_t)$ in the place of $\mathbb{E}[J(y_T+\epsilon_{T+1};p_{T+1})]=G_T(y_T)$ concludes the thesis.

Proof of Lemma 4:

By induction. For the base case, we have

$$\begin{aligned}G_T(y)=&\mathbb{E}[J(y+\sigma E_{T+1};p_{T+1})]\\=&\sigma\mathbb{E}[J((y/\sigma)+E_{T+1};p_{T+1})]\\=&\sigma G_{E,T}(y/\sigma).\end{aligned}$$

Since $J(ty;p)=tJ(y;p)$, on the set that J is differentiable in its first argument, we have $t\partial J(ty;p)/\partial y=t\partial J(y;p)/\partial y$, and so $J'(ty;p)=J'(y;p)$. As a result,

$$\begin{aligned}K_T(y)=&\mathbb{E}[J'(y+\sigma E_{T+1};p_{T+1})]\\=&\mathbb{E}[J'((y/\sigma)+E_{T+1};p_{T+1})]\\=&K_{E,T}(y/\sigma).\end{aligned}$$

Under the assumption (15), $\widetilde{\Delta}^p_T$ satisfies

$$K_t(\widetilde{\Delta}^p_T)=K_{E,T}(\widetilde{\Delta}^p_T/\sigma)=p_t,$$

and so $\widetilde{\Delta}^p_T=\sigma K^{-1}_{E,T}(p_T)$. Now suppose the claims hold for the $(t+1)$th stage, we have

$$\begin{aligned}K_t(y)=\mathbb{E}\Big\{&p_{t+1}\mathbf{1}(K_{t+1}(y+\sigma E_{t+1})\ge p_{t+1})\\&+K_{t+1}(y+\sigma E_{t+1})\mathbf{1}(K_{t+1}(y+\sigma E_{t+1})<p_{t+1})\Big\}\\=\mathbb{E}\Big\{&p_{t+1}\mathbf{1}(K_{E,t+1}(y/\sigma+E_{t+1})\ge p_{t+1})\\&+K_{E,t+1}(y/\sigma+E_{t+1})\mathbf{1}(K_{E,t+1}(y/\sigma+E_{t+1})<p_{t+1})\Big\}\\=K_{E,t}&(y/\sigma).\end{aligned}$$

Similarly,

$$\begin{aligned}G_t(y)=&\mathbb{E}\Big\{\mathbf{1}(K_{E,t+1}(y/\sigma+E_{t+1})\ge p_{t+1})\\&\times\Big[p_{t+1}(y+\sigma E_{t+1}-\sigma K^{-1}_{E,t+1}(p_{t+1}))+G_{t+1}(\sigma K^{-1}_{E,t\ +1}(p_{t+1}))\Big]\\&+G_{t+1}(y+\sigma E_{t+1})\mathbf{1}(K_{E,t+1}(y/\sigma+E_{t+1})<p_{t+1})\Big\}\\=&\sigma\mathbb{E}\Big\{\mathbf{1}(K_{E,t+1}(y/\sigma+E_{t+1})\ge p_{t+1})\\&\times\Big[p_{t+1}(y/\sigma+E_{t+1}-K^{-1}_{E,t+1}(p_{t+1}))+G_{E,t+1}(K^{-1}_{E,t+1}(p_{t+1}))\Big]\\&+G_{E,t+1}(y/\sigma+E_{t+1})\mathbf{1}(K_{E,t+1}(y/\sigma+E_{t+1})<p_{t+1})\Big\}\\=&\sigma G_{E,t}(y/\sigma).\end{aligned}$$

The same arguments as for the base case give $\widetilde{\Delta}^p_t=\sigma K^{-1}_{E,t}(p_t)$, thus all claims indeed hold for all t.

Proof of Theorem 2 and Theorem 3:

We prove the general case (Theorem 3); the two-stage results (Theorem 2) follow from the general case as a corollary.

Notice that the sequences $\{K_{E,t}\}$ and $\{G_{E,t}\}$ are defined for the reference problem and thus are independent of σ. Therefore, substituting the results from Lemma 4 into the expression of J_1 in Theorem 1, we have

$$J_t(\sigma;z_t,p_t)=\begin{cases}p_t(z_t-\sigma K^{-1}_{E,t}(p_t))+\sigma G_{E,t}(K^{-1}_{E,t}(p_t)) & \text{if } \sigma K^{-1}_{E,t}(p_t)\le z_t,\\ \sigma G_{E,t}(z_t/\sigma) & \text{if } \sigma K^{-1}_{E,t}(p_t)>z_t,\end{cases}$$

and in particular, for the first stage

$$J_1(\sigma;\widehat{d}_1,p_1)=\begin{cases}p_1(\widehat{d}_1-\sigma K^{-1}_{E,1}(p_1))+\sigma G_{E,1}(K^{-1}_{E,1}(p_1)) & \text{if } \sigma K^{-1}_{E,1}(p_1)\le \widehat{d}_1,\\ \sigma G_{E,1}(\widehat{d}_1/\sigma) & \text{if } \sigma K^{-1}_{E,1}(p_1)>\widehat{d}_1.\end{cases}$$

If $\sigma K^{-1}_{E,1}(p_1)\le\widehat{d}_1$, taking derivative of J_1 with respect to σ gives

$$p^{\mathrm{PoU}}(\sigma;\widehat{d}_1,p_1)=-p_1K^{-1}_{E,1}(p_1)+G_{E,1}\left(K^{-1}_{E,1}(p_1)\right),$$

which does not depend on σ. Re-writing the expression of J_1 in terms of PoU, we have, in this case,

$$J_1(\sigma;\widehat{d}_1,p_1)=p_1\widehat{d}_1+\sigma p^{\mathrm{PoU}},$$

that is J_1 is linear in σ. If $\sigma K^{-1}_{E,1}(p_1)>\widehat{d}_1$, taking derivative gives

$$p^{\mathrm{PoU}}(\sigma;\widehat{d}_1,p_1)=G_{E,1}\left(\widehat{d}_1/\sigma\right)-\frac{\widehat{d}_1}{\sigma}K_{E,1}\left(\widehat{d}_1/\sigma\right),$$

which completes the proof.