prepared for
Joint Inadvertent Interchange Taskforce
Resources Subcommittee, Operating Committee
North American Electric Reliability Council

OBJECTIVES:

1. Create a reasonably efficient economic framework, beyond a mere "barter" mechanism, for valuating/paying for Balancing Authorities' inadvertent, but not for explicitly targeting/evaluating inadvertent/frequency. Only an economic framework is compatible with the current, market behavior of entities. Otherwise commercial entities who become BAs can benefit unfairly in the market from gross economic inefficiencies in exclusively bartering BAs' inadvertent.

2. Recommend a revised economic-assessment framework for valuating/paying for entities' energy imbalance that does not target/evaluate it and that is compatible with a reasonably efficient framework for valuating/paying for BAs' inadvertent, and that is therefore reasonably efficient itself.

CRITERIA:

1. *The "framework" principle*. Such frameworks should be open to market pricing, without themselves prescribing the actual market mechanisms that set prices and that make these frameworks full economic/market mechanisms. It is appropriate for NERC to demonstrate the marketizability, compatibility with markets, or "market interface" of such mechanisms.

2. *The neutrality principle*. Such frameworks should be sensitive to system conditions but only pay/charge for service provided without explicitly evaluating performance relative to some target. They should be open to performance evaluation mechanisms without being themselves actual mechanisms that explicitly set targets and assess performance.

3. *The separation principle*. Inadvertent interchange and energy imbalance each consists of two things: energy, on the one hand, and control/deviation/reliability, on the other. In other words, inadvertent and energy imbalance are "unscheduled energy" which is two things: (i.1) the "energy" part, and a related (i.2) transmission congestion (loading) component, and (ii) the "unscheduled" aspect. The unscheduled part is the "inconvenience" factor, "hassle" factor, or degree of suddenly needing the energy.

(i.1) The energy part is separately and readily valuated/paid-for like all other energy in the existing energy market/system, for example at the hourly spot market price. Homogeneous integrated systems can even continue bartering the "energy" portion between themselves.

(i.2) The transmission loading component is the value of congested transmission, in addition to the value of unconstrained energy, needed to enable inadvertent to flow through a constraint. It also refers to the cost, in addition to the value of unconstrained energy, of having to specifically locate inadvertent to avoid congestion while the inadvertent is deployed to correct frequency. Load following (one way on the expensive side of the constraint, and in the opposite direction on the cheap side of the constraint) is sufficient to deploy to relieve congestion caused by inadvertent and protect the line from overheating. In a region that exempts inadvertent from transmission congestion management, a value of zero is assigned to the transmission loading component.

(ii) The "unscheduled" part, since it is what directly relates to frequency which is a system-wide characteristic (a "social choice" in economists' parlance), can and should be subject to a common valuation/payment metric/standard enforced by NERC on the BAs through mechanisms amenable to marketization. The hassle factor may be hard to visualize or point to separately, but it is separated out and valuated in markets in the concept of "options". The value of a stock option, for example, incorporates both these factors: how high a stock (call) option is valued depends on both "how low" the "exercise" price is for the possible purchase of the stock itself, and how "frequently" the stock would exceed that price, or the "volatility" of the stock's price, or the likelihood of the exercise price's being reached. Stock options trade off of the price volatility of the stock market and, so, serve to dampen that volatility.

Thanks to the separation of the "unscheduled" part from the "energy" in unscheduled energy, energy can be paid or charged for independently of whether it was good or bad for frequency. But the profitability of the energy part depends on the impact of frequency in the following framework of "dual pricing":

Dual pricing of unscheduled energy
Ambiguity along the diagonal

Revenue/Expense

Unscheduled part $\boldsymbol{b} > 0$ / energy part $U > 0$	good	bad	
			receive
sold	$p_e U + p_b \boldsymbol{b}$	$p_e U - p_b \boldsymbol{b}$	pay
bought	$-p_e U + p_b \boldsymbol{b}$	$-p_e U - p_b \boldsymbol{b}$	

The separation principle is already applied in the energy and transmission markets to valuating/charging for congested energy, separating it into the uncongested "energy" part and the "congestion" or congested-"transmission" part. Transmission congestion is triggered by the "energy" part of unscheduled energy, not by the "unscheduled" part. The transmission congestion would be included in the unscheduled Energy part in a market design that has combined transmission congestion and energy pricing. If the market has a design that has kept the energy price separate from the transmission congestion pricing, then the transmission congestion price of the "unscheduled" part would be an independent component of the pricing. This would result in a three part price that includes: 1) an Energy price component to recover the base energy price, 2) a Transmission Use component that captures the costs associated with the use of the transmission system, and 3) a Frequency Control Component that captures the costs associated with maintaining interconnection frequency under shared control. --Howard Illian

4. *The "covariance" principle*. The unscheduled aspect of inadvertent and energy imbalance should be valuated (as payment or receipt) depending on how much it affects ("covaries" with) frequency and in which direction, good or bad. This principle is the essence of $CPS1$ adopted by NERC (for BAs). $CPS1$ recognizes that it is overkill to valuate an individual's (a BA's) "unscheduled" performance at any moment independently of how much it hurt or helped frequency, in other words independently of the performance of the system at that moment. Unscheduled is good or bad depending on being counter-directional or co-directional to frequency deviation. A given bad individual performance is worse if the system's performance (frequency change) is very bad, and not so bad if the system's performance is just marginally bad. [While an individual's behavior can be measured as contributing to or harming the system's and compensated as such, a system's behavior would in a performance evaluation mechanism be different degrees of "bad" outside of some range $|\boldsymbol{e}|$, within which it is good.] If one entity's unscheduled deficit offsets another entity's unscheduled surplus, and both entities are lucky enough that frequency has not deviated, the unscheduled frequency components should not be paid or charged for. "Covariance" means that you simply multiply the individual's behavior ("deviation") "times" the system's behavior (frequency deviation) and you get some payment/valuation for the product from the numeraire applied to the units. For payment/valuation purposes the system and the individual can cancel out each other's behavior and at other times magnify each other's behavior. As frequency control is not a morality play, you're not paid or charged for deviating if you're unlucky or lucky enough that your deviance was masked by the combined actions of all the others. We recognize and punish someone's "theft" provided it wasn't offset by somebody else's charity, or punish somebody's charity for being too much provided it wasn't

offset by somebody's theft. This is terrible personal morality but great for allocating responsibility for frequency deviation.

This principle is not new as applied to frequency control; it also has been applied to the calculation of transmission congestion. The only difference is that transmission congestion is a binary instead of a continuous function. Frequency has a magnitude equal to the frequency error, while transmission has congestion representation that is binary. Therefore, the covariance is with transmission that is either congested as represented by a 1 or transmission that is un-congested as represented by a zero. Transmission congestion also has a sign to indicate whether or not the unscheduled energy reduces congestion or increases congestion, good vs. bad. --Howard Illian

5. *The time-averaging assessment principle*. Because of the random aspect of the "unscheduled" component (versus the "energy" component) of inadvertent or energy imbalance, and the rapid variability, behavior should be paid or charged for not at each data-acquisition moment (ultimately in seconds) but, say, monthly as a time-average over some reasonable period, say over the previous month to as long as over the previous year. In other words, schemes that seek to pay/charge for inadvertent or unscheduled energy instantaneously and deliberately, like some kind of AGC, may pose serious management/programming and stability challenges. Even if, out of practicality, we start out by measuring behavior hourly, rather than sub-hourly, "unscheduled" remains "hassle" factor and we don't want to further hassle ourselves by compensating/charging and settling at every single measurement interval—as that would only compound our inconvenience. Like the "covariance" principle, the time-averaging principle is also embodied in $CPS1$.

It is economically rational to treat random events on a statistical averaging, or fuzzy, basis, in the same way that risk is priced for insurance where you don't apply caution instant by instant. Time averaging is appropriate in this case because a metric for the "unscheduled" part of unscheduled energy is a valid representation of risk, the probability of interconnection failure. On an electric system where each time period is independent from every other time period because there is no inventory, only risk and dollars can be averaged over time. While BAs' averaging of energy values by payment-in-kind has masked this risk, separating out and averaging the "unscheduled" component would eliminate the need to change the payment-in-kind for Inadvertent energy. --Howard Illian

The market price of a unit of the unscheduled part would be the price of time-average units and would be determined up until and including assessment time as the bad units settle on a price to pay for the good units, and the good units get transferred to the holders of the bad units. The options transactions mentioned in comment 12. would be settlements transacted in advance of final, after-the-fact assessment of any residual bad unscheduled parts. The unit price an entity pays for an option serves as the price of the entity's actual bad unscheduled part that the option exercise corrects, as opposed to an after-the-fact price for that bad unscheduled part that the entity would have had to pay up until and including final assessment time had that entity not corrected that bad unscheduled part by exercising an option. Options are a particularly appropriate representation for the unscheduled part because they represent risk. --Howard Illian

SOLUTION:

(1) Implement a payment/valuation framework for BAs' inadvertent on the basis of hourly energy and average clock-hour frequency error data, and compatible with an allowed-for payment/valuation framework for entities, until such a time as the metered data becomes infra-hourly and CPS1 performance targeting and evaluation can be and are implemented by NERC.
(2) Allow for a payment/valuation framework for entities' energy imbalance, compatible with CPS1, but based on hourly-metered data and eventually infra-hourly metered data after CPS1 is implemented.

COMMENTS:

1. *Frequency-support/bias must be excludable from a behavior metric applicable to both BAs and entities*. Entities are not NERC-jurisdictional and therefore have no NERC conceived obligation for frequency control. Since entities are not assigned bias or a frequency-support obligation like BAs are, entities'

behavior must be valuated/paid for based on energy imbalance alone, without reference to bias. For comparability and performance-neutrality purposes, so should BAs.

2. *Premature to apply $CPS1$ to BAs.* $CPS1$ is conceived on the basis of ACE infra-hour intertie-error and infra-hour frequency-error data for assessing BAs' frequency support performance, especially with respect to the covariance/tolerance $\boldsymbol{e}^2$. JIITF has decided that a framework based on infra-hour data is not immediately feasible, partly because of still difficult data-resolution incompatibilities. Since only hourly-metered energy data and clock-hour average frequency data are available, it would be futile and defeat the purpose of $CPS1$ to try to implement the $CPS1$ performance evaluation metric now, based on this data.

3. *Outside $CPS1$, measurement of individual behavior relative to frequency remains.* While the $CPS1$ performance evaluation metric will not be used, the "covariance" essence of the $CPS1$ metric may be kept for payment/valuation purposes, meaning that in a payment/valuation metric we can preserve the relativity of individual behavior to frequency/system behavior. A simple expression of that is multiplying the individual BA's inadvertent by frequency deviation, and that's the numerator $I_i \times \overline{\Delta F}_h$ left in Howard Illian's equation (9)

$$CPS1_{60} : AVG\left(\frac{I_i \times \overline{\Delta F}_h}{-10\overline{B}_{i,h}}\right) \leq \boldsymbol{e}_h^2 - AVG\left(\overline{\Delta F}_h^{\,2}\right) \qquad (9)$$

of a hypothetical $CPS1$ based on hourly inadvertent-based AIE instead of infra-hourly intertie-error based ACE. [Cohn Inadvertent Decomposition of AIE doesn't make $CPS1$ any more usable, as Cohn Decomposition just reduces the bad of the over- or under- provision of frequency support, magnifies the good of frequency support if it is overprovided, and reduces the good of frequency support if it is underprovided, the more so the bigger the BA's share of the interconnection.] And it would contravene the "covariance" principle (criterion 4) above to use AIE alone, and we don't want to target performance by, say, setting $AIE = 0$ for BAs. Since we are aiming only to compensate/charge for behavior and not to evaluate performance relative to some target, we should drop the frequency support obligation part $-10B_i\overline{\Delta F}$ of

$$AIE = (NI_A - NI_S) - 10B_i\overline{\Delta F} \; .$$

With no $CPS1$-intrinsic $\boldsymbol{e}$ either, we are left with I_i and $\overline{\Delta F}$ to work with to get a metric of "unscheduled" performance for BAs, just as for entities.

4. *Purely "unscheduled" aspect of unscheduled "energy" is the energy's contribution to frequency deviation.* The simplest, most direct way to express the "unscheduled" part (vs the "energy" part) of inadvertent or energy imbalance, according to criterion 3 (the separation principle) above, is to express the inadvertent's or energy imbalance's impact on frequency. In other words, we need first to discover the metric that converts the frequency deviation into the BA's inadvertent or into the entity's energy imbalance, and vice versa. This is what Howard Illian achieves by his "regression", equations (12) through (19). In other words we seek to find the term, label it $10\overline{\boldsymbol{b}_{ij,h}}$, that, when multiplied by the frequency deviation $\overline{\Delta F}_h$, will give us something very close to U_{ij} which is the energy part of entity j's unscheduled energy in BA i, or is the energy part of BA i's unscheduled energy where $j = 0$. [Notation refinement of Howard's equations (22) and (23), with i indexing a BA, j indexing an entity inside that BA, and the convention that entity $j = 0$ is the BA itself or $ij = i$ when $j = 0$:

$$\sum_{j>0} U_{ij} = I_i \quad ; \qquad U_{i0} \equiv I_i \quad ; \qquad \sum_i \sum_{j>0} U_{ij} = \sum_i I_i = 0 \qquad (22)$$

$$\sum_{j>0} \overline{\boldsymbol{b}_{ij,h}} = \overline{\boldsymbol{b}_{i,h}} \quad ; \qquad \overline{\boldsymbol{b}_{i0,h}} \equiv \overline{\boldsymbol{b}_{i,h}} \quad ; \qquad \sum_i \sum_{j>0} \overline{\boldsymbol{b}_{ij,h}} = \sum_i \overline{\boldsymbol{b}_{i,h}} = 0 \qquad (23)\text{ .]}$$

You find this $10\overline{\boldsymbol{b}_{ij,h}}$ by solving for the $10\overline{\boldsymbol{b}_{ij,h}}$ that will give you the lowest value of the sum of the square of the difference between $10\overline{\boldsymbol{b}_{ij,h}\Delta F}_h$ and U_{ij} over all the measured values of $\overline{\Delta F}_h$ and U_{ij}. [It's "the square" of the difference that you minimize in order that what you're minimizing is positive.] The metric for the "unscheduled" part (vs the "energy" part) of i's unscheduled energy is therefore:

$$10\overline{\boldsymbol{b}_{ij,h}} = \frac{\frac{1}{n}\sum_n U_{ij} \times \overline{\Delta F}_h}{\frac{1}{n}\sum_n \overline{\Delta F}_h^2} = \frac{\sum_n U_{ij} \times \overline{\Delta F}_h}{\sum_n \overline{\Delta F}_h^2}, \qquad (11)$$

which is Howard Illian's equation (11), where n is the assessment period in hours. He labels this expression as FCC_h, to express it as a deficit to be paid for or bought when the expression (11) is positive, and a surplus, or "contribution" to frequency control, to be compensated for or sold when it is negative. This metric is readily computed on a spreadsheet.

5. *Applying a fixed unit price to the Illian* FCC_h *metric would violate the "covariance" principle embodied in* $CPS1$, *give a perverse incentive for control, and motivate the concern for tolerance bands. An appropriate price formula can be chosen to mimic the market and normalize the Illian* FCC_h *metric back to the "covariance" principle and properly incent control.* Rather than the covariance essence $I_i \times \overline{\Delta F}_h$ of Howard's $CPS1_{60}$ equation (9), the metric $10\overline{\boldsymbol{b}_{i,h}}$ [(equation (11)] is an estimator for $10\boldsymbol{b}_{i,h} = \frac{I_i}{\overline{\Delta F}_h}$ which is the geometric opposite of the "*covariance" principle*. When control is good (i.e. small $\overline{\Delta F_h} \approx 0$) the penalties and rewards $\overline{\boldsymbol{b}_{i,h}}$ are huge. When control is bad (i.e. big $\overline{\Delta F_h}$) the penalties and rewards $\overline{\boldsymbol{b}_{i,h}}$ are tiny. This stands in direct contravention of the *"covariance" principle* embodied in $CPS1$ and given above as Criterion 3. So, we need to mimic what a market price would do and apply a metric for unit price to bring the Illian FCC metric back to the *"covariance" principle.* Applying the price

$$p_{10\bar{b}} = k \times \frac{1}{n}\sum_n \overline{\Delta F}_h^2$$, where k is some fixed constant set by NERC, to get

$$FCC_{\bar{p}} = -10\overline{\boldsymbol{b}_{ij,h}} \times p_{10\bar{b}} = -k \times \frac{\sum_n \left(U_{ij} \times \overline{\Delta F}_h\right)}{\sum_n \overline{\Delta F}_h^2} \times \frac{1}{n}\sum_n \overline{\Delta F}_h^2 = -k \times \frac{1}{n}\sum_n \left(U_{ij} \times \overline{\Delta F}_h\right)$$

would do the trick. So would applying a price

$p_{10\underline{b}} = k \times \overline{\Delta F}_h^2$ to each nth hourly component of $\overline{\boldsymbol{b}_{ij,h}}$ to get

$$FCC_{\underline{p}} = -\frac{\sum_n \left(U_{ij} \times \overline{\Delta F}_h \times p_{10\underline{b}}\right)}{\sum_n \overline{\Delta F}_h^2} = -k \times \frac{\sum_n \left(U_{ij} \times \overline{\Delta F}_h^3\right)}{\sum_n \overline{\Delta F}_h^2}.$$

This gives us an *augmented* FCC_p metric that brings us back to the $I_i \times \overline{\Delta F}_h$ or $U_{ij} \times \overline{\Delta F}_h$ covariance world of $CPS1$. Furthermore, price p moves in the same direction as frequency deviation as we would expect in a market for volatility-driven options products like frequency response. That would make the *augmented* FCC_p metric more attractive to both marketers and control rooms than the FCC_h metric. In particular it would alleviate the concern for tolerance bands expressed in reaction to the sensed perverse incentive effect of the FCC_h metric, and in an attempt to protect against astronomical levies when F_h is near the center of the control band. In order to get FCC_p back to $CPS1$ covariance $U_{ij} \times \overline{\Delta F}_h$, p needs to increase quadratically (in $\overline{\Delta F}_h^2$) rather than linearly (in $\overline{\Delta F}_h$). Once $CPS1$ is implemented and there becomes a market for trading $\overline{b_i}$, the *augmented* FCC_p metric could revert to the plain FCC_h metric provided that p is the market price and assuming that p does indeed vary exponentially with $\overline{\Delta F}_h$ as expected under enforcement of $CPS1$.

6. *Shared frequency control sacrifices "equity"*. It would seem unfair for FCC to assess i on the basis of shares of $\overline{\Delta F}_h$ contributed by others. To correct that inequity you would need to multiply FCC by $\left|\frac{I_i}{I^+}\right|$, where I^+ is the sum of all the inadvertents in the same direction as frequency error, and you need the absolute value to preserve the polarity of FCC. This would "normalize" FCC from overcollection/overcompensation for others' contributions while preserving some dependence on $\overline{\Delta F}_h$, but just that amount of $\overline{\Delta F}_h$ that i is responsible for. However this detracts from the concept/amount of shared frequency control that underlies $CPS1$. To see this, suppose that all I_i s are small such that each i is responsible for a tiny amount of some huge $\overline{\Delta F}_h$. $\overline{\Delta F}_h$ is, in the interest of equity, doing little to sensitize each i's control to overall system conditions. On the other hand, if there is only one $i*$ responsible for all of some huge $\overline{\Delta F}_h$, $\overline{\Delta F}_h$ is sensitizing his control much more than it may be collectively sensitizing the control of a whole bunch of little contributors $i = 1,\ldots,n$ to the same big $\overline{\Delta F}_h$. Proof (for all $I_i \geq 0$):

$$I_{i*}\overline{\Delta F}_h =_{set} (I_1 + \cdots + I_n)\overline{\Delta F}_h = (I_1 + \cdots + I_n)\sum_{i=1}^{n}\left|\frac{I_i}{I^+}\right|\overline{\Delta F}_h > \left(I_1\left|\frac{I_1}{I^+}\right| + \cdots + I_n\left|\frac{I_n}{I^+}\right|\right)\overline{\Delta F}_h$$

, since $\sum_{i=1}^{n}\left|\frac{I_i}{I^+}\right| = 1$. In other words, shared control suffers as equity and/or the number of contributors of bad inadvertent increases. In fact, shared control collapses to control of only one small component's bad inadvertent (and an ever smaller component as the number of components increases), versus control of the entire system's bad inadvertent, in case the bad components are contributing equal bad inadvertents. Suppose $i = 1,2$ and $\left|\frac{I_i}{I^+}\right| = ½$ for all $i = 1,2$.

$$\left(I_1\left|\frac{I_1}{I^+}\right| + I_2\left|\frac{I_2}{I^+}\right|\right)\overline{\Delta F}_h = (½I_1 + ½I_2)\overline{\Delta F}_h = I_1\overline{\Delta F}_h = I_2\overline{\Delta F}_h = I_n\overline{\Delta F}_h$$

and I_n decreases as n increases.

7. *No double counting*. There is no double counting of unscheduled energy by applying $\overline{b_{ij,h}}$ both to the entities j and to the BAs i that contain those entities. Metric $\overline{b_{i,h}} = \sum_{j>0}\overline{b_{ij,h}}$ (equations 23) is just a pass-through of net unscheduled energy of the entities j in BA i with neighboring BAs i^*. Instead of entities j in BA i settling compensation directly with entities j^* in a neighboring BA i^*, these entities

j settle with their BA i which in turn settles with the neighboring BA i^* which in turn settles with its constituent entities j^*.

8. *Frequency support cannot be provided by entities from what is already being supplied*. When a BA i wants to reduce its unscheduled deficit $\overline{b_i}$, rather than just pay for it to a BA i^* with unscheduled surplus $\overline{b_{i*}} = -\overline{b_i}$ and collect from j in BA i with deficit $\overline{b_{ij}} = \overline{b_i}$, it can buy from BA i^* or entity j' in BA i with surplus $\overline{b_{ij'}} > 0$, frequency response which is defined by accounting. When frequency response is delivered under a specific delivery contract, it is classified as being (1) *scheduled by the seller* rather than being unscheduled by the seller, and it is classified as being (2) *unscheduled by the buyer*. If the frequency response called on by the buyer is (3) *expressly produced* for the buyer, the seller j''s unscheduled position $\overline{b_{ij'}}$ remains the same but the buyer's unscheduled position is improved by $-\overline{b_{ij}}$. If the frequency response called by the buyer is (4) a *reassignment of response* that would still be produced without the contract, and so would be classified as unscheduled, the seller's unscheduled position $\overline{b_{ij'}}$ is changed by $-\overline{b_{ij}}$ toward the direction of frequency error because the frequency support has become scheduled but no longer changes unscheduled position $-\overline{b_{ij}}$ toward the direction opposite the frequency error. In that case a reduction $\overline{b_i} < 0$ in BA i^*'s unscheduled surplus $\overline{b_{i*}}$ offsets the improvement $-\overline{b_i}$ in BA i's unscheduled deficit and frequency is not improved. When it's entity j' in BA i that BA i is buying from, BA i's unscheduled deficit $\overline{b_i}$, and therefore frequency, is not improved because BA i is just acting as a paying agent to compensate surplus entity j' whose unscheduled position $\overline{b_{ij'}}$ has worsened and then contract with deficit entity j to deliver the frequency response and reduce his unscheduled deficit by $-\overline{b_{ij}}$. The choice is the seller's how he wants to produce for the contract and, so, whether he wants to produce more to earn extra revenue, or just wants to be paid for reducing some of his existing unscheduled surplus $\overline{b_{ij'}}$ by scheduling it or be paid for increasing his existing unscheduled deficit $\overline{b_{ij'}}$ by scheduling what he doesn't already have. However, BA i needs to know that the seller j' is not just changing his $\overline{b_{ij'}}$ in a bad direction and just supplying some of his actual $\overline{b_{ij'}}$, and needs not to discover that the frequency support purchased has no impact on frequency nor therefore on BA i's $\overline{b_i}$, and this needs to be written into ancillary service option contracts and enforced. (Of course BA j' has to pay for his worsened $\overline{b_{ij'}}$ at settlement time.) BA i would not need to be concerned about this when buying from another BA i^* who is bearing the risk from his seller $j'*$.

9. *Metric $10\overline{b_{ij}}$ is a special case of metric $10\overline{\Delta b_i}$ applied under $CPS1$. A least-squares estimate $\overline{b_i}$ of BA i's variable bias $\overline{B_i}$. $FCC_{\bar{p}}$ is a monetization of $CPS1$ with bias $\overline{B_i} =_{set} 0$; so, $FCC_{\bar{p}}$ can have no tolerance band. $FCC_{\bar{p}}$ can facilitate and re-decentralize $CPS1$ compliance. $\overline{b_i}$ serves as the currency used by BAs i to eventually trade their $CPS1$ scores.* $10\overline{\Delta b_{ij}}$ is derived by using ACE_i instead of energy imbalance U_{ij} in $10\overline{b_{ij}}$

$$10\overline{\Delta \boldsymbol{b}_{ij}} = \frac{\frac{1}{n}\sum_n \left(T_{ij} - 10\overline{B_{ij}\,\Delta F}\right)\overline{\Delta F}}{\frac{1}{n}\sum_n \left(\overline{\Delta F}^2\right)} \qquad (11')$$

$$= \frac{\sum_n \left(T_i - 10\overline{B_i\,\Delta F}\right)\overline{\Delta F}}{\sum_n \left(\overline{\Delta F}^2\right)} = \frac{\sum_n ACE_i\,\overline{\Delta F}}{\sum_n \left(\overline{\Delta F}^2\right)}$$ where $j = 0$ and intertie error T_i (in Mw) replaces inadvertent I_i (in MwH). This means that, instead of

$$\underset{b_{i,h}}{Min} \sum_n \left(U_i - 10\overline{\boldsymbol{b}_{i,h}\,\Delta F}_{\ h}\right)^2$$ whose solution $10\overline{\boldsymbol{b}_{i,h}}$ solves

$$AVG\left(U_i\,\overline{\Delta F}_{\ h}\right) = 10\overline{\boldsymbol{b}_{i,h}}\,AVG\left(\overline{\Delta F}_{\ h}^{\ 2}\right)$$, we are now solving

$$\underset{b_i}{Min} \sum_n \left(T_i - 10\overline{B_i}\,\overline{\Delta F} - 10\overline{\Delta \boldsymbol{b}_i}\,\overline{\Delta F}\right)^2$$ whose solution $10\overline{\Delta \boldsymbol{b}_i}$ solves

$$AVG\left[\left(T_i - 10\overline{B_i}\,\overline{\Delta F}\right)\overline{\Delta F}\right] = 10\overline{\Delta \boldsymbol{b}_i}\,AVG\left(\overline{\Delta F}^2\right)$$

$$AVG\left(T_i\,\overline{\Delta F}\right) = 10\,AVG\left(\overline{B_i} \times \overline{\Delta F}^2\right) + 10\overline{\Delta \boldsymbol{b}_i}\,AVG\left(\overline{\Delta F}^2\right)$$

$$AVG\left(T_i\,\overline{\Delta F}\right) = 10\overline{\boldsymbol{b}_i}\,AVG\left(\overline{\Delta F}^2\right)$$ wherein

$\overline{\boldsymbol{b}_{ij}}$ has been constructed as $\overline{\boldsymbol{b}_{ij}} = \overline{b_{ij}} + \overline{\Delta \boldsymbol{b}_{ij}}$, where

$$\overline{b_{ij}} =_{df} \frac{\frac{1}{n}\sum_n \overline{B_{ij}} \times \overline{\Delta F}^2}{\frac{1}{n}\sum_n \overline{\Delta F}^2} = \frac{\sum_n \overline{B_{ij}} \times \overline{\Delta F}^2}{\sum_n \overline{\Delta F}^2}$$

is the portion of $\overline{\boldsymbol{b}_{ij}}$ required to meet frequency-support obligation expressed by $\overline{B_{ij}}$ (which is bias). In the case of entities, i.e. $j \neq 0$, the two metrics would be the same, $\overline{\boldsymbol{b}_{ij}} = \overline{\Delta \boldsymbol{b}_{ij}}$, because $\overline{\boldsymbol{b}_{ij}} = \overline{b_{ij}} + \overline{\Delta \boldsymbol{b}_{ij}}$ and, by the definition above of $\overline{b_{ij}}$, $\overline{b_{ij}} = 0$ when $j \neq 0$ because $\overline{B_{ij}} = 0$ when $j \neq 0$.

In other words, $10\overline{\boldsymbol{b}_{ij}}$ is implemented in the general form $10\overline{\Delta \boldsymbol{b}_{ij}}$ which applies to BAs i under $CPS1$. $CPS1$ then reduces to a simple form of equation that relates $\boldsymbol{b}$ or $\Delta \boldsymbol{b}$ to b as we see by substituting $\overline{b_i}$ for $\overline{B_{i,\,h}}$ in Howard's equation (9)

$$CPS1_{60} : AVG\left(\frac{I_i \times \overline{\Delta F}_{\ h}}{-10\overline{\overline{B}_{i,\ h}}}\right) \leq \boldsymbol{e}_h^2 - AVG\left(\overline{\Delta F}_{\ h}^{\ 2}\right)$$

$$AVG\left(\frac{T_i \times \overline{\Delta F}}{-10\overline{\overline{b}_i}}\right) + AVG\left(\overline{\Delta F}^2\right) \leq \boldsymbol{e}^2$$

$$\frac{\left(\dfrac{AVG(T_i \times \overline{\Delta F})}{AVG(\overline{\Delta F}^2)}\right)}{-10\overline{\overline{b}_i}} AVG\left(\overline{\Delta F}^2\right) + AVG\left(\overline{\Delta F}^2\right) \leq \boldsymbol{e}^2$$

$\frac{10\overline{\boldsymbol{b}_i}}{-10\overline{\overline{b_i}}} AVG\left(\overline{\Delta F}^2\right) + \frac{-\overline{b_i}}{-\overline{\overline{b_i}}} AVG\left(\overline{\Delta F}^2\right) \leq \boldsymbol{e}^2$, by equation (11) for metric $\overline{\boldsymbol{b}_i}$ and so

$\frac{\overline{\boldsymbol{b}_i} - \overline{b_i}}{-\overline{\overline{b_i}}} AVG\left(\overline{\Delta F}^2\right) \leq \boldsymbol{e}^2$. This equation for $CPS1$ becomes

$$\frac{\overline{\Delta \boldsymbol{b}_i}}{-\overline{\overline{b_i}}} AVG\left(\overline{\Delta F}^2\right) \leq \boldsymbol{e}^2.$$

This makes FCC basically a monetization of $CPS1$ with bias $\overline{B_i} =_{set} 0$, since $CPS1$ then becomes

$$\left(\overline{\boldsymbol{b}_i} - \overline{b_i}\right) AVG\left(\overline{\Delta F}^2\right) \leq -\overline{b_i}\boldsymbol{e}^2$$

$$\overline{\boldsymbol{b}_i} AVG\left(\overline{\Delta F}^2\right) \leq 0$$

$$\frac{AVG(T_i \times \overline{\Delta F})}{AVG(\overline{\Delta F}^2)} AVG\left(\overline{\Delta F}^2\right) \leq 0$$

$$AVG(T_i \times \overline{\Delta F}) \leq 0,$$

and $$-k \times AVG(I_i \times \overline{\Delta F}_h) = FCC_{\overline{p}}$$

through the presumed effect (in Comment 5) of market pricing of $\overline{\boldsymbol{b}_i}$ while $CPS1$ is actually operating in the background with $\overline{B_i} \neq 0$ and within tolerance band $-\overline{b_i}\boldsymbol{e}^2 > 0$. This proves that a tolerance band $\pm \boldsymbol{e} \neq 0$ cannot be assigned to $FCC_{\overline{p}}$ as defined because $CPS1$ tolerances for individual BAs i must necessarily be weighted $-\overline{b_i}\boldsymbol{e}^2$ by the BA i's negative bias $-\overline{b_i}$, and $-\overline{b_i}\boldsymbol{e}^2 = 0$ when the bias $\overline{B_i} =_{set} 0$. $FCC_{\overline{p}}$ provides a means of decentralizing frequency control to entities j to alleviate BA i's job of $CPS1$ compliance, and to re-decentralize frequency control when $CPS1$ compliance is centralized into fewer and bigger BAs i. $\overline{\boldsymbol{b}_i}$ is the currency BAs i would use to eventually trade their $CPS1$ scores.

10. *Without price mediation, BAs i with bad inadvertent will improve their $CPS1_i$ scores by trading their primary response $\overline{b_i}$ (≤ 0) rather than other kinds of $\overline{\boldsymbol{b}_i}$. Then primary response $\overline{b_i}$ is more valuable than other kinds of $\overline{\boldsymbol{b}_i}$.* Rather than reduce $\overline{\boldsymbol{b}_i}$ by x when $0 < x < \overline{\boldsymbol{b}_i}$, BA i prefers to improve $CPS1_i$ performance by increasing primary response (bias) by x because increasing the denominator $-\overline{b_i}$ has a bigger impact on decreasing the term $\frac{\overline{\boldsymbol{b}_i} - \overline{b_i}}{-\overline{b_i}}$ than decreasing the numerator $\overline{\boldsymbol{b}_i}$ does. When you increase the denominator $-\overline{b_i}$ you do not change the numerator $\overline{\boldsymbol{b}_i} - \overline{b_i}$ because it's only by purchasing some other i's primary response that you are buying primary response obligation from him; otherwise he experiences no change in $\frac{\overline{\boldsymbol{b}_i} - \overline{b_i}}{-\overline{b_i}}$ and there's no point to selling obligation alone. In other words, increasing your usage/obligation $-\overline{b_i}$ of primary response already means an offsetting decrease in your $\overline{\boldsymbol{b}_i}$. Since $\overline{b_i}$ is preferred it would be priced higher and therefore allow for more lower cost secondary response to be used. To improve $CPS1_i$ performance when the interconnection I is outside its $CPS1$ target band $\boldsymbol{e}^2 < \overline{\Delta F}^2$ and to get $\frac{\overline{\boldsymbol{b}_i} - \overline{b_i}}{-\overline{b_i}} < 1$ sufficiently, BA i then prefers to decrease the numerator $\overline{\boldsymbol{b}_i}$ by $x \geq \overline{\boldsymbol{b}_i} > 0$ or $x > 0 \geq \overline{\boldsymbol{b}_i}$ because decreasing the numerator $\overline{\boldsymbol{b}_i}$ by x has a bigger impact on decreasing the term $\frac{\overline{\boldsymbol{b}_i} - \overline{b_i}}{-\overline{b_i}}$ than increasing primary response (bias) does.

This is a mathematical proof of Howard Illian's observation (provided, for correction x, $0 < x < \overline{\boldsymbol{b}_i}$) that ultimately the quality (rapidity) of the support should affect its cost, ranging from most expensive and immediate frequency response, to AGC, to operating reserves, to slower-to-deploy load following. [If $\boldsymbol{e}^2 < \overline{\Delta F}^2$, and $x \geq \overline{\boldsymbol{b}_i} > 0$ or $x > 0 \geq \overline{\boldsymbol{b}_i}$, this order of cost is reversed because restoring frequency to get the interconnection I back within target range $\overline{\Delta F}^2 \leq \boldsymbol{e}^2$ takes priority over stabilizing frequency.] It is efficient for frequency response to be replaced by AGC to be replaced by operating reserves, to be replaced by load following, as these ramp up with respectively longer lags, for as long as the support is needed. The cheapest support is, of course, regular scheduled energy. For a given $\overline{\boldsymbol{b}_{ij,h}}$ it's in the support provider's interest to supply the cheapest form of support as long as it is scheduled as support under a contract (such as an option) which it would be in the buyer's/exerciser's interest to replace with a regular energy contract as soon as possible.

Here is the proof:

$$\frac{\left(\overline{\boldsymbol{b}_i} - x\right) - \overline{b_i}}{-\overline{b_i}} > \frac{\overline{\boldsymbol{b}_i} - \overline{b_i}}{-\overline{b_i} + x}$$

$$\left(\overline{\boldsymbol{b}_i} - x - \overline{b_i}\right)\left(-\overline{b_i} + x\right) > \overline{\boldsymbol{b}_i}\left(-\overline{b_i}\right) + \overline{b_i}^2$$

$$\overline{\boldsymbol{b}_i}\left(-\overline{b_i}\right) - x\left(-\overline{b_i}\right) + \overline{b_i}^2 + \overline{\boldsymbol{b}_i}x - x^2 + x\left(-\overline{b_i}\right) > \overline{\boldsymbol{b}_i}\left(-\overline{b_i}\right) + \overline{b_i}^2$$

$$\overline{\boldsymbol{b}_i}x - x^2 > 0$$

$$0 < x < \overline{\boldsymbol{b}_i}.$$

11. *No "tolerance band" need be stipulated for* $\overline{\boldsymbol{b}_{ij,h}}$ because we're just paying/charging for individual service, not evaluating/incenting system performance relative to some target. The notion of a tolerance band is a performance-evaluation concept in violation of the neutrality principle (criterion 2) above and comes from the $\boldsymbol{e}^2$ tolerance band in the $CPS1$ world of frequency-support obligation $\overline{B_{i,h}}$. If we exit the $CPS1$ world of frequency-support obligation $\overline{B_i}$, we get (from Comment 5) unbounded

$$FCC_{\bar{p}} = -10\overline{\boldsymbol{b}_{ij,\,h}} \times p_{10\bar{b}} = -k \times AVG\left(U_{ij} \times \overline{\Delta F}_{\ h}\right).$$

This lack of a tolerance band for the metric $\overline{\boldsymbol{b}_{ij,h}}$ can be explained verbally as follows. The metric $\overline{\boldsymbol{b}_{ij,h}}$ allocates responsibility for or against frequency error $\overline{\Delta F}_{\ h}$ to the extent that $\sum_i \sum_{j>0} \overline{\boldsymbol{b}_{ij,h}} = 0$ since $\sum_i \sum_{j>0} \overline{\boldsymbol{b}_{ij,h} \Delta F}_{\ h} = \sum_i \sum_{j>0} U_{ij} = 0$ by equations (22) above and $\overline{\Delta F}_{\ h} \neq 0$ by hypothesis. But outside of economic motivation to increase $FCC_{\bar{p}}$, $\overline{\boldsymbol{b}_{ij,h}}$ by itself does nothing to stabilize frequency F or to return frequency F to, or keep frequency error $\overline{\Delta F_h}$ within, any target range. It certainly doesn't set or enforce the 60 Hz core frequency. When $j=1$'s unscheduled energy U_{11} in BA $i=1$ induces a frequency change $\overline{\Delta F_h}$, that frequency change, by changing the power level on the system, induces primary frequency response $\sum_i \sum_{j\_} \overline{B_{ij\_,h} \Delta F}_{\ h}$ in the form of offsetting aggregate unscheduled energy U_{ij} in the opposite direction distributed among $j's > 1$ in BA $i=1$ and $j's \neq 0$ in the other BAs $i > 1$. Those entities j (call them $j'_\_$) providing enough response to prevent frequency from deteriorating further would be compensated (through assessment of the resultant $\overline{\boldsymbol{b}_{ij,h}}$) by the entity $j=1$ in BA $i=1$ causing the frequency error. In this framework there is no explicit mechanism, besides avoidance of higher cost, incenting this response, nor the reversal of frequency error $\overline{\Delta F_h}$ provided reversal is cheaper than containment, no explicit mechanism incenting the actual targeting of frequency F_h , certainly specifying the 60 Hz center of the target range.

There can only be an *implicit* mechanism outside of (in other words hidden "inside") the metric $\overline{\boldsymbol{b}_{ij,h}}$ to incent targeting of frequency error $\overline{\Delta F}_{\ h}$ to a band $\boldsymbol{e}$ around F = 60 Hz. That mechanism resides in the coordinated agency of the BAs i acting on an agenda independently of their constituent entities j and independently of what the BA i's own $\overline{\boldsymbol{b}_{i,h}}$ happens to be, in order first to stabilize frequency deviation $\overline{\Delta F}_{\ h}$, then to reverse it and to actually restore frequency F to the target range unstated in $\overline{\boldsymbol{b}_{ij,h}}$. In that capacity the BAs i act as a broker, rather than a principal, assessing the entities j whose inadvertent $10\sum_i \sum_j \overline{\boldsymbol{b}_{ij,h} \Delta F}_{\ h} = -10\sum_i \sum_{j\_} \overline{B_{ij\_,h} \Delta F}_{\ h}$ triggered the frequency deviation $\overline{\Delta F}_{\ h}$, and then paying entities $j\_$ to incur sufficient offsetting stabilizing inadvertent and eventually entities j' to incur sufficient additional correcting inadvertent, or secondary response,
$10\sum_i \sum_{j'} \overline{\boldsymbol{b}_{ij',h} \Delta F_h} = 10\sum_i \sum_{j\_} \overline{B_{ij\_,h} \Delta F}_{\ h}$, in the direction opposite to the original deviation. This secondary, correcting inadvertent induces offsetting, stabilizing inadvertent $-10\sum_i \sum_{j'_\_} \overline{B_{ij'_\_,h} \Delta F}_{\ h}$ that

cancels out the stabilizing (primary) inadvertent $-10\sum_i \sum_{j\_} \overline{B_{ij\_,h} \Delta F}_h$ offsetting the original inadvertent $-10\sum_i \sum_j \overline{B_{ij,h} \Delta F}_h$ that triggered the frequency deviation $\overline{\Delta F}_h$. So the final result of frequency-correcting intervention by the BAs i is to replace the offsetting inadvertent or primary response $10\sum_i \sum_{j\_} \overline{b_{ij\_,h} \Delta F_h}$, that only stabilized frequency deviation $\overline{\Delta F_h}$, by a frequency-correcting inadvertent or secondary response $10\sum_i \sum_{j'} \overline{b_{ij',h} \Delta F_h}$ that triggers frequency correction $-\overline{\Delta F_h}$ to keep F near enough to 60 Hz. $CPS1$ *puts the onus of stabilization on all the BAs* i *for any single BA* i's *error, but puts the onus of reversal on the offending BA* i. $\sum_i \sum_{j'} \overline{b_{ij',h}}$ represents the portion of the $\sum_i \sum_{j>0} \overline{b_{ij,h}}$ responsible for frequency correction, versus the portion $\sum_i \sum_j \overline{b_{ij,h}}$ responsible for the frequency deviation, and both portions sum near enough to zero when added together to satisfy to $CPS1$. (In the frequency-correcting portion) the correction-triggering $\overline{b_{ij',h}}$ and the offsetting stabilizing primary response $\overline{b_{ij\_',h}}$ sum to zero, and (in the frequency-deviation portion) the triggering $\overline{b_{ij,h}}$ and the offsetting stabilizing primary response $\overline{b_{ij\_,h}}$ sum to zero. The offsetting stabilizing primary responses $\overline{b_{ij\_,h}}$ and $\overline{b_{ij\_',h}}$ sufficiently cancel each other, leaving the correcting secondary responses $\overline{b_{ij'',h}}$ to sufficiently reverse the triggering $\overline{b_{ij',h}}$ to keep frequency within allowed $CPS1$ range.

In other words, in the world of frequency-control performance targeting-and-evaluation of and by BAs not explicitly captured in the metric $\overline{b_{ij,h}}$, but compatible with and implicit or *presumed* in $\overline{b_{ij,h}}$, there is a mechanism or market *tiering* under which entities j charged $\overline{b_{ij,h}}$ for triggering frequency deviation $\overline{\Delta F_h}$ ultimately wind up paying not the offsetting entities $j\_$, but instead the corresponding BA i. The BA i is also charged $\overline{b_{i,h}}$ (for triggering the frequency deviation) by all the BAs i providing stabilizing response. Under its mandate to control frequency, the offending BA i either preempts the $\overline{b_{i,h}}$ charge by in turn paying $\overline{b_{ij',h}}$ to entities j' to correct frequency deviation or is charged $\overline{b_{i,h}}$ and in turn is paid for $-\overline{b_{i,h}}$ triggering frequency correction $-\overline{\Delta F_h}$ and uses that reward to pay the entities j' for producing the $\sum_{j'} \overline{b_{ij',h}} = -\overline{b_{i,h}}$. The offending BA i has the choice of paying another BA i^*, say, or directly paying j's inside BA i, to correct the frequency deviation. And a BA i^* receives the funds paid by offending BA i and in turn uses the funds to pay entities j' inside BA i^* for correcting the frequency deviation. Transfers and assessments actually get done on a net, time-averaged basis; so, on a momentary basis, these procedures are more a matter of interpretation than mechanics. Under such an implicit mechanism as just described or under an explicit one like $CPS1$, in addition to the assessment metric $\overline{b_{ij,h}}$ for entities j and, indirectly BAs i, BAs i are performing (under assessment and an explicit tolerance targeted by NERC in the case of $CPS1$) a frequency support/correction/restoration function as a default agent for the entities j' who actually provide that support but might otherwise be economically incented to provide stabilization or primary frequency response, but insufficient support or secondary response, under assessment metric $\overline{b_{ij,h}}$ alone.

That frequency-correction mechanism for BAs i is the place for a target or tolerance band e for frequency deviation $\overline{\Delta F_h}$. If you sum $CPS1_i$ over all the BAs i (with $j=0$) you wind up with the equation $AVG\left(\sum_i \frac{\overline{\Delta b_{i,h}} \times \overline{\Delta F}_h^{\,2}}{\overline{b_{i,h}}}\right) \le e^2$ which is the target band of frequency deviation. (You use the square only so that deviations all add up, rather than cancel each other out when they are in opposite directions!) In other words, once $CPS1$ (including ACE) is applied to BAs i (i.e. $j=0$), frequency is *explicitly* being maintained within a target band $60hz \pm e$. $CPS1$ (including ACE) provides the incentive to BAs i for sufficient reversal $-\overline{\Delta F}_h$ of frequency deviation $\overline{\Delta F}_h$, to actually restore frequency F to the target range determined by 60 Hz & e. $CPS1$ and ACE do this by assessing a BA i's performance in terms not of its contribution to the frequency deviation $\overline{\Delta F}_h$ (i.e. in terms not of how much $\overline{b_i}$ its intertie error T_i contributed to the deviation $\overline{\Delta F}$ or contributed to keeping the frequency at F wherever F happens to have deviated to), but in terms of by how much $10\overline{\Delta b_i \Delta F}$ i's intertie error fell short of or exceeded both

(1) i's fair contribution $10\overline{B_i \Delta F}$ ($\overline{B_i} < 0$) to stabilizing frequency deviation $\overline{\Delta F}$ and

(2) in reversing i's own previous contribution $10\overline{\Delta b_i \Delta F}$ to frequency error $\overline{\Delta F}$ (in terms not of by how much i's triggering or offsetting intertie error $T_i = 10\overline{b_i \Delta F}$ exceeded or fell short of 0 (of how much $T_i < 0$ or $T_i > 0$), but of by how much $10\overline{\Delta b_i \Delta F}$ i's intertie error $T_i = 10\overline{b_i \Delta F}$

(1) failed or succeeded in meeting i's frequency support obligation $10\overline{B_i \Delta F}$ (of how much $T_i < \left|10\overline{B_i \Delta F}\right|$ or $T_i > \left|10\overline{B_i \Delta F}\right|$) and

(2) offset BA i's own previous contribution $10\overline{b_i \Delta F}$ to frequency error.

$CPS1$ indirectly determines what may be called the "frequency-deviation stabilizing-and-correcting"-obligation, or primary and secondary response obligations, through a performance target band e (for $\overline{\Delta F}$) that is set in $CPS1$ and reveals the primary response obligation $\overline{b_i}$ that is otherwise hidden or implicit in $\overline{b_i} = \overline{b_i} + \overline{\Delta b_i}$, with $\overline{b_i} =_{df} \frac{\sum_n \overline{B_i} \times \overline{\Delta F}^2}{\sum_n \overline{\Delta F}^2}$. With $j=0$, $\overline{\Delta b_{ij}}$ measures and compensates BA i's performance for (or relative to) $\overline{B_i}$ (while $CPS1$ "evaluates" it by setting frequency at 60 Hz and target band e) and, since $\overline{B_i} = 0$ for individual entities j, simply compensates or charges for the service level or demands of entities j, including the entities $j\_$ who are deployed by the BAs i to provide the frequency support $\overline{B_i}$. The metric $\overline{b_{ij}}$ includes the frequency support for BA i being provided by entities $j\_$ in BA i, but without isolating them out and measuring them for purposes of performance evaluation of the BAs i. In other words, the metric $\overline{b_{ij}}$ is neutral relative to performance evaluation but is readily amenable to a performance evaluation mechanism by the simple flip of measuring $\overline{\Delta b_{ij}}$ within the

$CPS1$ framework. The metric $\overline{\Delta b_{ij}}$ allocates responsibility to all the BAs i for primary response to stabilize frequency and, under $CPS1$, to the BA i at fault for the frequency deviation to correct it: $\sum_i \overline{\Delta b_i} = \sum_i \overline{b_i} - \sum_i \overline{\bar{b}_i} = -\sum_i \overline{\bar{b}_i} = -b$, by equations (23) above, with $\bar{b} =_{df} \dfrac{\sum_n \overline{B} \times \overline{\Delta F}^2}{\sum_n \overline{\Delta F}^2}$.

$CPS1$ "stacks the deck" against the BAs i to get them to stabilize frequency, and against the BAs i specifically at fault for hurting frequency to get them to correct their individual fault. To contain, then avoid, assessment BAs i execute their agency function by paying entities $j\_$, j' to provide frequency support with funds collected from entities j contributing to frequency deviation. If aggregate $CPS1$ is such that $AVG\left(\sum_i \dfrac{\overline{\Delta b_i} \times \overline{\Delta F}^2}{\overline{\bar{b}_i}}\right)$ is within the error band e^2, BAs i within the band don't get assessed and they even get rewarded for their shortfall within the band by selling additional room for $\overline{\Delta b_i}$ to BAs i with excess $\overline{\Delta b_i}$ that puts them outside the band who want to avoid being assessed a higher cost for that excess and who in turn assess entities j contributing to frequency deviation in order to get the funds to buy the other BAs' shortfall within the band. By setting a targeted tolerance band e and by setting frequency stabilization obligations B_i in the metric $\overline{\Delta b_i}$ $CPS1$ enforces BA i s' compliance with obligations for frequency support. In the case of metric $\overline{b_i}$, $\sum_i \overline{b_i} = 0$ automatically, and this says simply that the system is at a given disturbance level $\overline{\Delta F}$, and therefore the raw $\overline{b_i}$ metric doesn't explicitly incent system performance relative to any *particular* frequency.

12. *Unit price(s) for* $\overline{b_{ij}}$ *must ultimately derive from the action of an institution like NERC. Two-tiered market for* $\overline{b_{ij}}$ *:* $CPS1$ *market for frequency support to determine prices for settling the entire market of* $\overline{b_{ij}}$. Outside BA i's $CPS1$ obligation to support frequency, overall supply and demand of $\overline{b_{ij}}$ s is always balanced between triggering unscheduled energy and offsetting unscheduled energy; so, $CPS1$ obligation would drive, and therefore determine, price. As unit price rises, paying entities j are incented to minimize $\overline{b_{ij,h}}$ and charging entities j' are incented to maximize $\overline{b_{ij',h}}$, while other entities $j\_$ or $j'\_$ are incented to provide offsetting-inadvertent or frequency-support. Since we demonstrated in comments 8 & 9 that secondary support to correct frequency should be somewhat cheaper than primary support used just to stabilize frequency, paying entities j can have an economic incentive to correct frequency and not just maintain it. But under $\overline{b_{ij,h}}$ alone it's still optional to provide secondary frequency support to replace primary offset. So, we can't be sure that, at whatever level the j minimize unscheduled energy in the direction of frequency deviation, sufficient support will be provided in the opposite direction to maintain frequency at $60Hz$ nor do BAs i need to have $60Hz$ in mind outside $CPS1$. In other words, as long as there is no requirement in the form of a penalty for lack of frequency support, such as that levied on BAs i under $CPS1$, there can't be any pricing mechanism for $\overline{b_{ij}}$ alone. Without a penalty mechanism like $CPS1$ establishing the necessity and range for frequency support, there is no mechanism for pricing $\overline{b_{ij}}$, for guaranteeing the sufficiency of frequency support.

Frequency support is a public good (or "social choice") like clean air and, so, can be maintained only by common agreement which sets a standard enforced by a simple institution needed to drive market-pricing of frequency support. So, under the $\overline{b_{ij}}$ metric, NERC first needs to apply a unit price formula that appropriately incents supply of frequency support the more it's needed to keep frequency from deteriorating from the $60Hz$ target--a process in which the notion of a tolerance band is inappropriate since entities aren't directly assessed for meeting the target or not. Under the $CPS1$ standard, the targeted $60Hz$ frequency, and $\pm e$ tolerance band reflecting the frequency deviation history of the system, determine the average level of $\overline{\Delta b_i}$ at which BA i begins being assessed by NERC positively within the band, or negatively outside the band. ($e = 0$ is the strictest band, actually a point, from which BAs i get assessed only negatively and between whom there can thus be no "trading" of $\overline{b_{ij}}$ to incent good frequency restoration performance.) That assessment will determine the level of demand for $\overline{b_i}$ for support and therefore the unit price of $\overline{b_i}$ s to *assure* frequency support. A sufficient amount of frequency support will emerge provided the penalty price set by NERC for a bad $CPS1$ score is high enough. In that case a market for pricing $\overline{b_i}$ will emerge in which BAs i with bad $CPS1$ scores will buy b_{i*} from BAs $i*$ with good $CPS1$ scores rather than pay the NERC penalty. A supply/demand equilibrium should occur at a unit price for b as in the market for pollution rights and this $CPS1$ market for primary and secondary support will determine the price at which $\overline{b_i}$'s are settled. Otherwise, if NERC sets the penalty too low, NERC faces the challenge of finding the frequency support.

Absent applicable markets, and subject to FERC approval, NERC would set a formula for unit price of the "unscheduled" part that appropriately motivates/compensates accurate scheduling by varying the price in sympathy with frequency deviation. Since this is an artificial mechanism, it becomes urgent to establish a market mechanism to determine that price. (2) Moreover, since no price, set or market, by itself can assure that the frequency target is maintained, the metric alone would not assure good control without a frequency performance standard like $CPS1$ applied to BAs implicitly (as now is the case) or explicitly; so, it becomes just as urgent to marketize and implement high enough penalties for $CPS1$ to get an explicit performance standard. (3) Furthermore, use of the market clearing mechanism needs to be incented through enforcement of the $CPS1$ performance standard.

Once there is a market mechanism in place for setting the price of $\overline{b_{ij}}$, based on an intertie error performance standard in place for BAs, NERC would withdraw the unit price formula and apply instead an unreasonably high penalty to incent clearing in the $CPS1$ market for support and a price at which $\overline{b_{ij,h}}$'s are settled:
(1) As the DOE did very successfully to drive the pollution rights market, NERC would set an unreasonably high penalty on BAs for intertie error performance that would incent BAs i with bad error $\overline{b_i}$ to avoid that penalty by (a) buying on the $CPS1$ market good error $\overline{b_{i*}} = -\overline{b_i}$ in the form of support from BAs $i*$ with good error, or (b) buying good error $\overline{b_{ij'}} = -\overline{b_i}$ from BA i's own entities j' on BA i's own (ancillary services) options market for secondary frequency support, and turning around and selling that $\overline{b_{ij'}}$ at that cost to the deficit entities j in the BA i's own jurisdictional market for trading the unscheduled part between the surplus entities j' and the deficit entities j or on the ancillary services (options) market for frequency support. Entities could trade the frequency component $\overline{b_{ij}}$ of their scheduling error directly between themselves or settle the $\overline{b_{ij}}$ and avoid having to settle with or to buy from BA i or on the ancillary services (options) market, or a deficit entity j could buy directly on the ancillary services (options) market to avoid having to buy from his BA i.
(2) Entities j settle at closing time through the settlement agent (say, NERC) either among themselves or with their BA i, based on the the prices paid by BAs for any $-\overline{b_i}$ required to keep them $CPS1$ compliant.

The need for a $CPS1$ penalty big enough to drive the market for scheduling error and ancillary services (options) embodies the so-called "dictatorial outcome" or cost required of such mechanisms by social choice theory. Any penalties must stay with NERC and cannot flow back to the supply side of the market; if they do, the mechanism fails.

Two-tiered market but no double charges. In $CPS1$, having entities j pay their $\overline{b_{ij}}$ and funnel inter-BA deficits & surpluses through the BAs can't be done without $CPS1$ assessment of the BAs. The BAs i *pay again* if any paid-for $\overline{b_{ij}}$ isn't stabilized and exceeds the $CPS1$ limit. So, under $CPS1$ not only would entities j need to pay for their $\overline{b_{ij}}$ under FCC_p, but BAs i need to stabilize and reduce it and charge those added costs and pay those added benefits through to the entities j. That (prospect of) intervention by the BAs i doesn't compound charges but sets the price for the $\overline{b_{ij}}$ not already settled. It preempts entities j from incurring even higher charges for a less reliable level of support, for example for too much primary instead of secondary support. It's further stabilization and reduction in $\overline{b_i}$ (beyond what entities j have already done by themselves) that come from the purchase of some of another BA $i*$'s $\overline{b_{i*}} < 0$ as frequency support. When BA i buys secondary reduction $-\overline{b_i}$ it eliminates $\overline{b_i}$ it would otherwise have to pay for, plus BA i avoids $CPS1$ penalty. Accordingly a $CPS1$ penalty provides a needed economic incentive to buy stabilization and reduction of $\overline{b}_i$ when merely paying for $\overline{b}_i$ isn't incentive enough. When you contract for primary support in the FCC_p market you don't reduce your $\overline{b}_{ij}$; you are just paying for it. The producer $j_{\_}$ of primary support makes his $\overline{b}_{ij_{\_}} < 0$ even more negative as an offset to some $-\overline{b}_{ij_{\_}}$ that occurs anyway. When you buy $\overline{b}_i$ primary support, on the $CPS1$ market from a producing BA $i*$ you don't reduce your $\overline{\Delta b}_i > 0$ but you do reduce your $\overline{b}_i > 0$.

(i) NERC determines the system/frequency requirement. Imposes a penalty to drive the $CPS1$ market.
(ii) $CPS1_i$ determines the amount of BA i's $\overline{b}_i$ requirement. Drives the FCC_p settlement price.
(iii) FCC_p determines calculation and settlement of entities js' $\overline{b}_{ij}$.

Producing unscheduled energy versus producing frequency support (under contract):

Frequency support	Producing without (option) contract	Producing under (option) Contract
Secondary	Affects producer j''s $\overline{b}_{ij'}$	Not affects producer j''s $\overline{b}_{ij'}$ because "scheduled"; so, settled
	Entitlement to payment for reversing some system $\overline{b}$	Improves buyer j's $\overline{b}_{ij}$ because part reversed & settled with j'
Primary	Affects producer $j_{\_}$'s $\overline{b}_{ij_{\_}}$	
	Entitlement to payment for part of some entity $j_{\_}$'s $\overline{b}_{ij_{\_}}$	Not affects buyer j's $\overline{b}_{ij}$; rather j prepays part of $\overline{b}_{ij}$ to $j_{\_}$

Differential pricing of primary and secondary response can lighten the mediating role of the BA.
When entity j pays $\overline{b}_{ij}$ to an entity j' with $\overline{b}_{ij'} = -\overline{b}_{ij}$ or buys $-\overline{b}_{ij_{\_}}$ primary from entity $j_{\_}$, he fills his $\overline{b}_{ij}$ settlement requirement.
When BA i buys $-\overline{b}_{ij}$ secondary, or $\overline{b}_{i*}$ from another BA $i*$, BA i reduces his $\overline{b}_i$ settlement requirement.
If there is a market in which primary and secondary response $-\overline{b}_{ij} < 0$ are priced differently, entities j (not subject to $CPS1$) should prefer reducing rather than filling their $\overline{b}_{ij}$ settlement requirement insofar as secondary support should be cheaper than primary. Unless primary and secondary response $\overline{b}_{ij}$ were separately priced products, entities j would be indifferent between filling or reducing their $\overline{b}_{ij}$ settlement requirement. That indifference is otherwise broken by the action of the BA i (whose job, once there is differentiated market pricing of the two products, becomes more a job of oversight and monitoring $CPS1$ compliance than actual intervention to support frequency through substituting secondary for primary support. BA i is ultimately the agent of $CPS1$ assessment by NERC until such a time as entities j were individually assigned that responsibility.) BAs i procure added response for stabilizing and reducing their $\overline{b}_i$ settlement requirement, and thereby set the $\overline{b}_i$ settlement price..

4 market operations

<table>
<tr><td colspan="6">Entities j may be indifferent, but BAs i are not indifferent, between</td></tr>
<tr><td colspan="3">Paying out your $\overline{b}_{ij}$ to another i^*, j', $j\_$</td><td colspan="3">Paying another i^*, j', $j\_$ to reduce your $\overline{b}_{ij}$</td></tr>
<tr><td rowspan="2">By settling $-\overline{b}_{ij}$ of his $\overline{b}_{i*}$, $\overline{b}_{ij\_}$ or $\overline{b}_{ij'}$</td><td colspan="2">For increasing his $\overline{b}_{ij\_}$ (or $\overline{b}_{i*}$) primary by $-\overline{b}_{ij}$</td><td rowspan="2">By buying $-\overline{b}_{ij}$ of his $\overline{b}_{i*}$, $\overline{b}_{ij\_}$ or $\overline{b}_{ij'}$</td><td>For increasing his $\overline{b}_{ij'}$ (or $\overline{b}_{i*}$) secondary by $-\overline{b}_{ij}$</td><td>By increasing his scheduled</td></tr>
<tr><td>No option contract</td><td>Option contract</td><td>No option contract</td><td>Option contract</td></tr>
<tr><td colspan="2">Operation 1:</td><td>Operation 2:</td><td>Operation 3:</td><td>Operation 1:</td><td>Operation 4:</td></tr>
<tr><td colspan="2">Settlement payout of $\overline{b}_{ij}$</td><td>Buy primary support $\overline{b}_{ij}$</td><td>Trading $\overline{b}_{ij}$</td><td>Settlement payout of $\overline{b}_{ij}$</td><td>Buy secondary support $\overline{b}_{ij}$</td></tr>
</table>

Two-tiered market mechanism for "tickets"

FCC: Market for paying/redeeming & extinguishing tickets

$CPS1$: Market for extinguishing tickets unextinguished in FCC

Rule: By assessment time (monthly)

(1) all tickets must be paid/redeemed and

(2) enough tickets must be extinguished for $CPS1$: $\overline{\Delta F}^2 \leq e^2$.

Extinction => Payment

Payment ≠> Extinction

# of tickets outstanding since last assessment varies according to the cumulative average frequency state F of the system

Market actions

0. Tickets are reissued at the beginning of each assessment period, the # outstanding is adjusted at measurement interval n, and they cumulate or reverse cumulate payment obligation or receipt as intervals n cumulate.

   -Good tickets are awarded to j's for cumulative $\overline{b}_{ij'} < 0$, entitling the holder to a cash receipt at the prevailing price by assessment time, from some holder of an equal value of bad tickets.

   -Bad tickets are assigned to j s for cumulative $\overline{b}_{ij} > 0$, obligating the holder to a cash payment at the prevailing price by assessment time, to some holder of an equal value of good tickets

   -Good tickets are denominated according to whether $\overline{b}_{ij}$ is

   primary response or

   not primary response (called "secondary" response here for brevity despite including coincidental reversing $\overline{b}_{ij}$ that is not strictly response).

1. Good (bad) ticket holders make (receive) payment, and their tickets may be extinguished, by selling good tickets to holders of bad tickets at a prevailing/negotiated price.

   Purchase of good secondary response tickets by holders of bad tickets extinguishes the matched good and bad tickets.

   Purchase of good primary response tickets by holders of bad tickets settles but does not extinguish the matched good and bad tickets.

2. A holder can pay and extinguish a bad ticket, or extinguish a paid but unextinguished bad ticket for the same amount, by exercising an option on secondary response. If the option writer j' does not produce, j' is issued a bad ticket to reflect the settlement of a good ticket with the option exerciser j.
3. A holder can pay but not extinguish a bad ticket by exercising an option on primary support. If the option writer j' does not produce, j' is issued a bad ticket to reflect the settlement of a good ticket with the option exerciser j.
4 A holder can pay and extinguish a bad ticket, or extinguish a paid but unextinguished bad ticket for the same amount, by buying entity j''s secondary $b_{ij'} < 0$.

5.. A holder can pay and extinguish a bad ticket by buying entity $j\_$'s primary $b_{ij\_} < 0$.

By trading b_i under options on primary response, a BA i sets/adjusts his primary response obligation $\overline{b_i}$ in the $CPS1$ market.

Discussion

Entities j are indifferent in the choice within 1, and indifferent between 2 & 3, or between 4 & 5 if there is no price differentiation between primary and secondary support $-b_{ij} < 0$.

BAs i are not indifferent over those choices thanks to the action of $CPS1$. BAs i are incented to extinguish, and therefore reduce, bad tickets to a level allowed by $CPS1$, and maintain their bias share of unextinguished tickets for $\overline{b_i}$ of primary response, through their market interaction with other BAs i and ultimately with the constituent entities j. Since BAs i pass this extra cost on to the entities j holding the unextinguished tickets, those entities have an added economic incentive to have extinguished the tickets themselves.

The $CPS1$ market

$CPS1$ determines the second-tier market for trading $\overline{b_i}$ to extinguish unextinguished bad tickets and in the process set/adjust the BA i's bias amount $\overline{b_i}$ of unextinguished tickets.

BA i extinguishes, & thereby reduces, the # of non-extinguished bad tickets it holds in the name of its constituent entities j, by resorting to

action 2 , or

action 4 with BAs $i*$.

BA i adjusts bias share $\overline{b_i}$ of primary response by resorting to

action 3, or

action 5 for $\overline{b_{i*}}$ from BAs $i*$.

[In Comment 9 it was shown that BAs i would increase their $CPS1$ score by opting first to increase their bias share of non-extinguishing good tickets because the favorable impact on $CPS1$ of doing this is more immediate (especially when interconnection frequency is well within the $CPS1$ tolerance limit) than reducing the # of non-extinguished bad tickets by buying extinguishing good tickets which must be cheaper and ultimately make more economic sense. The optimal economic amount of primary response for BA i to produce should be his natural bias share $\overline{b_i}$.]

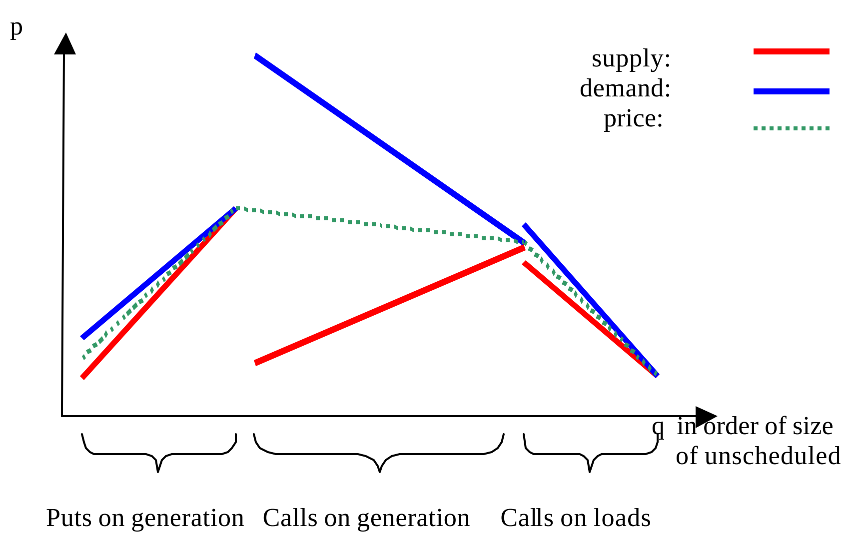
13. *Options market for ancillary services. $CPS1$ assessment required to make this market work*. In electricity, physical volatility occurs in the form of unscheduled energy and can be subject to economic mediation. The option price a generator would require (buyer would pay--buyer being a load who underschedules or a generator who overschedules) just for the generator to be standing by ready to generate

during some period would be lower (higher) the more frequently/likely the generator is to be called or the higher (lower) the exercise price. Frequent exercise makes the option more attractive to both buyer and seller, provided the buyer actually does exercise it often enough--the seller's moral hazard problem. The lower the exercise price, the likelier a call option is to be exercised. Since attracting the seller with a high exercise price lowers the frequency of exercise, discouraging both buyer and seller, frequency of exercise, rather than exercise price, is a likelier driver of this call option price which must cover a generator's opportunity cost of standing by and not offering in the energy market.

At the extremes, where the generator is expected to be called (1) all the time or (2) hardly at all, it might be more attractive to the buyer to make the following two equivalent purchases of options that it expects to exercise infrequently: in the case of (1), overschedule load or underschedule generation and buy a put option from a generator not to generate and, in the case of (2), buy a call option from a load to decrease load. The option price a seller would require, and a buyer would pay, just for the seller to be ready to reduce operation during some period would be higher the more frequently/likely the buyer is to exercise, or the higher the exercise price. Frequent exercise makes the option more attractive to the buyer and less attractive to the seller, all the less to the seller if the buyer exercises the option more than expected--the seller's moral hazard problem. The higher the exercise price, the likelier a put is to be exercised. Since attracting the seller with a low exercise price lowers the frequency of exercise, discouraging the buyer, exercise price is as likely a driver of these option prices as is frequency of exercise. But in this case the option price is trying to cover, not the opportunity cost of standing by and not offering in the energy market, but the opportunity cost of not producing at a specific moment.

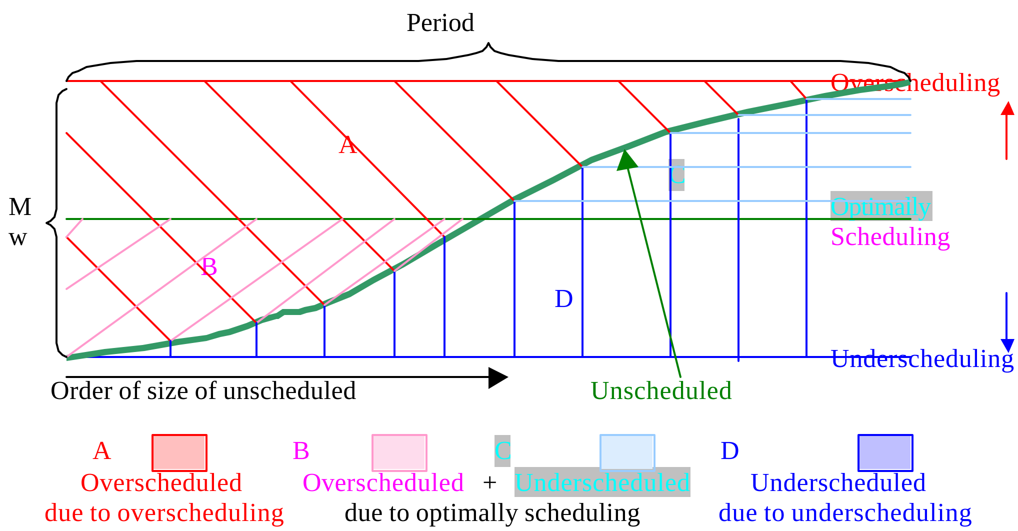
The frequently exercising option buyers would be the earliest and strongest demanders, willing to pay a price higher than infrequent exercisers whom a generator would want to charge a higher price to but who wouldn't be willing to pay more than they would have to pay for their unscheduled part. At the extreme of infrequent exercise one of two things is possible. (1) A generator could sell multiple infrequently exercised options at a lower price and incur the moral hazard of frequent exercise. Like insurance companies, generators could pool those options with other generators' to evenly spread the frequency of calls among themselves. (2) Loads would tend to serve the demand for infrequently exercised options and face the moral hazard of frequent exercise. While generators prefer the options they sell to be frequently exercised, loads prefer the options that they sell to curtail their consumption to be infrequently exercised. Loads' standby capacity cost is limited to the cost of carrying inventory to cover lost production. They could pool the risk of frequent exercise among themselves or to generators.

Possible portfolio of options
Price should arbitrage buyer's avoided cost *b* with supplier's opportunity cost



An equilibrium may be reached when entities schedule somewhere up the merit order of the amount of their intrahourly variability, buy call options on the remaining overscheduled generation or underscheduled load, including call options from loads for the least frequent amounts, and buy some amount of put options on underscheduled generation or overscheduled load.

In a complete options market for frequency response
the optimal scheduling point minimizes the cost of unscheduled B + C



$$MinValue(B+C) \leq Min[Value(A), Value(D)]$$

If unscheduled energy were distributed diagonally in order of amount, optimizing may reduce the demand for options, reduce the range of moral hazard, and cut the expected frequency of exercise. It may also encourage generators to write bi-directional options: to increase or reduce generation.

The price of the option per frequency share of the energy exercised should be close to a unit market price of the unscheduled part it actually corrected. The exercise price would be the closest thing to an intrahourly energy price but would compete with the hourly price used to price the energy component of unscheduled energy not supplied as ancillary service. It would be meaningless to attempt to use the exercise price to indicate the instantaneous value of intrahourly energy since the "instantaneousness", and the difference between the exercise price and the expected hourly energy price, are being captured in the optionality or price of the option itself that reflects $\overline{b_{ij}}$ on an averaged basis.

14. *Megawatthours U_{ij} not more intuitive than megawatts $\overline{b_{ij,h}}$ per hertz.* There may be a reluctance to consider using the construct $\overline{b_{ij,h}}$ (measurable in terms very operational and material of unscheduled energy U_{ij} and frequency deviation $\overline{\Delta F_h}$) as the metric for assessment of the "unscheduled" aspect of inadvertent or energy imbalance performance because it lacks the direct "feel" of "megawatthours" U_{ij} , which is itself but a construct of measured megawatts and measured hours.