FROM THE AMERICAN PEOPLE

DEVELOPING THE FEE FOR WHEELING ELECTRICITY IN PAKISTAN

SUSTAINABLE ENERGY FOR PAKISTAN (SEP) PROJECT

Submission Date: November 3, 2020

Contract No.: AID-OAA-I-13-00028
Task Order: AID-391-TO-16-00005
Activity Start Date and End Date: August 3, 2017 to April 26, 2021

Submitted by: Tetra Tech ES, Inc.
1320 North Courthouse Road, Suite 600
Arlington, VA 22201
Tel. +1-703-387-2100 | Fax +1-703-243-0953

www.tetratech.com

This document was produced for review by the United States Agency for International Development. It was prepared by Tetra Tech ES, Inc. for the Sustainable Energy for Pakistan (SEP) Project.

DEVELOPING THE FEE FOR WHEELING ELECTRICITY IN PAKISTAN

SUSTAINABLE ENERGY FOR PAKISTAN (SEP) PROJECT

Submission Date: November 3, 2020

EXECUTIVE SUMMARY

For many countries, "wheeling" is the first step in introducing open access. Wheeling is the act of disconnecting load (i.e., consumption) from specific locational generation. Put differently, wheeling is, in large part, the act of disconnecting load from decisions made on their behalf in the past about transmission, distribution and generation. *Wheeling allows load to search for, and potentially find generators that offer lower prices or other characteristics that better serve their needs than was purchased for the load in the past.*

To guide our analysis, we developed and adhered to the following set of high-level guiding principles:

a. Counterparties to a wheeling transaction should not be able to avoid or escape appropriate costs.

b. The wheeling fee should be consistent with the principles of short and long-term economic efficiency. And, in particular, should send the correct investment signals. To this end, the assignment of the fee for a service should be based on either cost causation or on who benefits.

c. The incentives produced by the wheeling fee should promote reliable operation of the electricity network and economic efficiency.

d. The methodology by which the wheeling fee is calculated or updated should be known in advance.

e. The wheeling fee should be understandable, transparent and replicable.

For the nine issues raised by the participants at the LUMS Conference from December 9-11, 2019, a summary of our recommendations is as follows:

1. STRANDED ASSETS.

In conclusion, with respect to the wheeling fee, the solution to the issue of cost recovery for the stranded assets in Pakistan is straightforward.

- All load should pay these costs. It makes no difference whether a specific load is a party to a wheeling contract or not, all load benefits from the provision of capacity. Moreover, since the generation backing the wheeling transaction may not be able to perform, the load from a wheeling transaction is just as dependent on capacity as any other load on the system.

- The basis for the cost allocation should be usage and, in particular, peak demand. Capacity is procured to ensure the reliability of the system at all times rather than just certain periods. Finally, while cost recovery should be DISCO-based to reflect the fact that each DISCO has different capacity needs, we did not have complete cost data at the level of the DISCO and, as a result our, calculated fees for stranded assets are at an aggregated level.

- It follows then that we recommend that all loads - both wheeling and non-wheeling - pay a per usage charge that reflects their peak demand.

  Finally, we recommend that the cost for stranded assets be recovered through an appropriately adjusted per kWh fee based on average monthly peak demand for all consumption. Alternatively, the costs for Stranded Assets may be charged as a fixed fee based on the peak demand of each DISCO in PKR per kW per month.

2. TREATMENT OF LOSSES

Our recommendations with respect to how the costs for transmission and distribution network losses should be included in the wheeling fee, are as follows:

- Wheeling transactions should only be charged for technical, and not commercial losses.
- To the greatest extent possible, commercial losses should be minimized prior to implementing an electricity market.
- The recovery of the costs associated with technical losses should be based solely on usage.
- The actual fee for losses should be apportioned to all load, including load being served via a wheeling contract.
- The fee should be based on the total cost of losses and the amount of energy purchased by the total retail sector.

3. USE OF SYSTEM CHARGES

The services provided by NTDC, including the provision of an electricity transmission network and the reliable despatch of electricity, are both necessary for wheeling to take place. The only exception is the case where the power flows, associated with a specific wheeling transaction, never use the high voltage network. Even in that case, there is a strong argument that the specific transaction is beneficiary of a reliable system with access to ancillary services. Given that, and under the current regulatory rate setting paradigm, we agree with the recommendations arrived at by the participants at the LUMs conference. Specifically, we recommend:

- The UoSC should be treated uniformly across all power that flows in the system.
- When power is wheeled across two or more DISCOs, the NTDC system is utilized and, therefore, the UoSC should be a part of the wheeling charge.
- When the wheeler of power and the associated Bulk Power Consumers (BPC) are located in the same DISCO system, the BPC is still utilizing the NTDC system for system stability and reliability. Furthermore, even though the actual flow of power to the BPC is not using the transmission network, the network is providing reliability services that are different from system operation.
- Additionally, when calculating the demand charge, we recommend the use of a coincidental peak rather than a non-coincidental peak.

4. HYBRID BPC'S

Our recommendations with respect to Hybrid BPC's and the wheeling fee are as follows:

- The existence of Hybrid BPC contracts necessarily means that the system operator (or some entity) must have excess electricity available in real time – through some mechanism – to provide for the shortfall between the Hybrid BPC contracted for amount and the actual amount consumed
- Wheeling contracts should be required to schedule their anticipated generation and load with the system operator.
- A "usage" fee based on the cost of operating reserves necessary to reliably and efficiently operate the system. The usage fee will be determined by the cost of operating reserves

divided by the total load. This per kW monthly fee is then aplied equally to all type of loads including hybrid in each DISCO.

- A cost for the consumption of energy that is consumed by the hybrid BPC customer that is not covered by the hybrid BPC contract. This cost applies to all purchases made by the Hybrid BPC customers beyond their contracted amounts of energy.

5. BANKED ENERGY

With respect to Banked Energy and the methodology for calculating the Wheeling Fee, our recommendations are as follows:

- Review the rules pertaining to self-scheduling to ensure that there are no artificial incentives that encourage specific types of transaction that rely on self-scheduling.
- Monitor self-scheduling so that reliability is not compromised by having an excess amount of self-scheduling in constrained regions of the grid.
- Banked energy should be paid on the basis of system marginal cost for the specific operating interval during which there was excess production from self-scheduled resources.
- Banked energy credits must be used within twelve months of when they were accrued.
- Monitor when and how the banked energy credits are used.

6. ECONOMIC DISPATCH

We strongly recommend using Security Constrained Economic Dispatch as the basis for the electricity market in Pakistan for the following primary reasons:

- Nodal pricing reduces the necessary discretion of the system operator.
- Nodal pricing-based markets do not rely on necessarily false assumptions about the state of the transmission network in real time.
- Nodal pricing markets better allow for naturally occurring risk to be more efficiently managed through financial instruments rather than physical capital investment.
- Nodal pricing markets provide location and time-of-use based price signals for generation and load.
- Nodal pricing markets provide an explicit signal for the cost of congestion.
- Nodal pricing markets are economically more efficient and more reliable.
- Nodal pricing markets are far better suited to accommodate intermittent resources.
- Nodal pricing markets are less likely to be manipulated or subject to market power.

7. WHEELING FROM RENEWABLE RESOURCES (FIRM CAPACITY FACTOR)

Given the nature of intermittent renewable reserves, the system operator will be required to carry additional operating reserves in order to maintain reliable operations. As such, we recommend that the Grid Code be reviewed to ensure that the reliability guidelines reflect the added volatility of the generation mix.

Until such time as nodal pricing is implemented, our alternative recommendation is to require intermittent generation to schedule their output one hour in advance of real time operations for every operating interval. Furthermore, these schedules should be continuously reviewed for accuracy.

The deviations between the scheduled and actual amounts will be absorbed by the wider grid, i.e., other generators will be ramped up/down on a reliability and economic basis to absorb the deviations from intermittent resources and the cost will be spread across all users of the system.

8. SPECIALIZED ROLE OF ENTITIES

Given that there is no recommended market design from which to evaluate the specific role of entities, we make the following recommendations:

- All the functions/activities necessary to operate the eventual electricity market should be precisely identified and then defined.
- For each activity, a detailed methodology describing how the function is to be accomplished should be developed and adopted.
- Assuming there is need for a Market Operator based on the adopted market design, after the first two steps have been accomplished all of the functions and responsibilities should be assigned to either the system or market operator.
- Both the System and (potentially) Market Operators are not Principals in the electricity market insofar as neither should purchase or sell electricity. Instead both are service providers to the market and the governance and business objectives of both should reflect this fact.
- The activities of the DISCO's should be focused solely on the physical operation of the low voltage network. Accordingly, they should have no relationship with the electricity market.
- Regardless of the initial market design, legacy decisions regarding the institutional structure should not hamper the efficient evolution of the electricity market.

9. CROSS SUBSIDY SURCHARGE

This section of the report is not yet finalized and is, accordingly, absent from this Draft Report. It will be included in the final report.

Lastly, we note that it was not always possible to obtain the precise data necessary to calculate our recommendations. In such cases we made the best and most appropriate assumptions in regard to the data we were provided.

CONTENTS

1. TREATMENT OF STRANDED ASSETS WITHIN THE WHEELING FEE 10

1.1 BRIEF SUMMARY 10
1.2 BACKGROUND 11
1.3 ISSUE OF STRANDED ASSETS – PAKISTAN VIS-À-VIS INTERNATIONAL CONTEXT 11
1.4 TREATMENT OF STRANDED ASSETS WITHIN THE PROPOSED WHEELING FEE – THE INTERNATIONAL PERSPECTIVE 13
1.5 METHODOLOGY FOR ANALYSING AND CALCULATING THE STRANDED ASSETS COST 16
1.6 CONCLUSION AND RECOMMENDATIONS 18

2. TREATMENT OF NETWORK LOSSES WITHIN THE WHEELING FEE 19

2.1 BRIEF SUMMARY 19
2.2 BACKGROUND 20
2.3 CONCLUSION AND RECOMMENDATIONS 27

3. TREATMENT OF NTDC'S USE OF SYSTEM CHARGE (UOSC) WITHIN THE WHEELING FEE 28

3.1 BRIEF SUMMARY 28
3.2 BACKGROUND 28
3.3 ISSUE OF NTDC'S USE OF SYSTEM CHARGE – PAKISTAN VIS-À-VIS INTERNATIONAL CONTEXT 29
3.4 TREATMENT OF NTDC'S USE OF SYSTEM CHARGE WITHIN THE PROPOSED WHEELING FEE – THE INTERNATIONAL PERSPECTIVE 31
3.5 DETERMINATION OF NTDC UOSC AND RECOMMENDATIONS 34

4. TREATMENT OF HYBRID BPC'S IN THE WHEELING FEE 37

4.1 BRIEF SUMMARY 37
4.2 INTRODUCTION 37
4.3 ISSUE AND TREATMENT OF HYBRID BULK POWER CONSUMER'S – PAKISTAN VIS-À-VIS INTERNATIONAL CONTEXT 38
4.4 CONCLUSION AND RECOMMENDATIONS 39

5. TREATMENT OF BANKED ENERGY IN THE WHEELING FEE 41

5.1 BRIEF SUMMARY 41
5.2 BACKGROUND 41
5.3 ISSUE OF BANKED ENERGY – PAKISTAN VIS-À-VIS INTERNATIONAL CONTEXT 42
5.4 CONCLUSION AND RECOMMENDATIONS 44

6. ECONOMIC DESPATCH AND THE WHEELING OF ELECTRICITY 45

6.1 BRIEF SUMMARY 45
6.2 BACKGROUND 46
6.3 INTRODUCTION 46

6.4 FUNDAMENTAL ISSUES WITH ECONOMIC DISPATCH 48
6.5 PRICE VIS-À-VIS ECONOMIC DISPATCH 51
6.6 PAKISTAN VIS-À-VIS INTERNATIONAL CONTEXT 55
6.7 CONCLUSION AND RECOMMENDATIONS 57

7. THE TREATMENT OF RENEWABLE RESOURCES IN THE WHEELING FEE 58

7.1 BRIEF SUMMARY 58
7.2 TREATMENT OF ALLOCATING FIRM CAPACITY FOR WHEELING TRANSACTIONS FROM RENEWABLE RESOURCES – PAKISTAN VIS-À-VIS INTERNATIONAL CONTEXT 59
7.3 CONCLUSION AND RECOMMENDATIONS 60

8. THE ROLE OF SPECIFIED ENTITIES IN AN ELECTRICITY MARKET 61

8.1 BRIEF SUMMARY 61
8.2 BACKGROUND 62
8.3 ANALYSIS ON ROLE OF ENTITIES IN AN ELECTRICITY MARKET 62
8.4 CONCLUSION AND RECOMMENDATIONS 64

ANNEX A: INITIAL COMMENTS TO NEPRA'S METHODOLOGY FOR ESTABLISHING THE CROSS SUBSIDY SURCHARGE (CSS) 65

TABLES:

Table 1: Statistics from PJM Electricity Market 26

Table 2: Comparison of Transmission and Distribution Rates for four regulated Transmission and Distribution wires businesses in Texas (September 1, 2018 to February 28, 2019) 32

Table 3: System Operation Fee in Alberta (Canada) for 2019 and 2020 32

Table 4: Regulated rates/charges for the users of Amprion's system....................34

Table 5: Appropriate the Cost recovery for the NTDC UoSC Equally Between the Demand and Usage Charges 36

FIGURES:

Figure 1: Difference between Technical Versus Non-technical Losses....................20

Figure 2: Average Fee per Megawatt hour (left hand column) for System Operations for Each RTO/ISO (2018)....................31

Figure 3: Simple Electrical System – Single Transmission Constraint 49

Figure 4: Simple Electrical System – Single Transmission Constraint 50

Figure 5: Simple Electrical System – Single Transmission Constraint, 300MW of System Load....................50

Figure 6: Simple Electrical System – Single Transmission Constraint, 600MW of System Load....................51

Figure 7: Simple Electrical System – Single Transmission Constraint, 270 MW of System Load – Optimal Dispatch 52

Figure 8: Simple Electrical System – Single Transmission Constraint, 360MW of System Load, Optimal Dispatch 53

Figure 9: Simple Electrical System – Single Transmission Constraint, 360 MW of System Load, Power Flows Under Optimal Dispatch 54

1. TREATMENT OF STRANDED ASSETS WITHIN THE WHEELING FEE

1.1 BRIEF SUMMARY

Ultimately the success of all wheeling regimes regardless of the country or jurisdiction - as well as the success of open access itself - will require the separation of generation from transmission and distribution activities and unbundling, i.e., disaggregation, of the services provided by both the transmission and distribution wire businesses as well. The services provided by the wire businesses and the costs of providing those services must be defined and made transparent because not all parties to a wheeling transaction will need or want to purchase all of the services from the wire businesses. Nevertheless, wheeling transactions will use the services of the wire business and should be charged accordingly, i.e., there needs to be an appropriately determined "wheeling fee" that reflects the services the wheeling transaction is "purchasing" through the wire businesses. *This paper addresses the cost recovery mechanism of one such "service" - the recovery of "stranded assets."*

It is important to both characterize and understand the nature of "stranded assets" within the context of Pakistan. First, with respect to their characterization as "assets", these are not "assets" in the traditional meaning of the term within the electricity or energy regulatory paradigm. Within the traditional regulatory paradigm, an asset almost always refers to something that is physically tangible, i.e., a power plant, transmission or distribution line(s), transmission substation, etc. This is not the case with "stranded assets" in Pakistan. Rather the "asset" is in fact a type of a power purchase agreement for generation capacity. As such, there is no physical aspect to the "asset" rather it is purely financial. They are "stranded" not because they are no longer useful but rather because they are out-of-the-money, i.e., the terms and conditions of the contracts (including the prices) do not reflect the current conditions within the electricity sector in Pakistan.

As with any "stranded asset", we note that the real issue is neither the correct determination of the fee or the efficient assignment of the costs of these contracts, but rather the burden, inefficiency (both in the near term as well as in the long term), and fairness of the costs themselves. These are valid and significant questions that are, however, largely beyond the reach of NEPRA, as well as the issue at hand, i.e., the determination of the wheeling fee. In other words, the solution for the problem of out-of-the-money contracts will not be found in either the determination of the wheeling fee or in any other rate-setting (tariff) exercise.

Based on the nature of costs, i.e., stranded assets, and the aforementioned guiding principles we make the following recommendations:

Ignoring that an argument can be made that generation and transmission be responsible for these costs as they are the entities that cause the need for capacity, all load should pay these costs. It makes no difference whether a specific load is a party to a wheeling contract or not, all load benefits from the provision of capacity. Moreover, since the generation backing the wheeling transaction may not be able to perform, the load from a wheeling transaction is just as dependent on capacity as any other load on the system.

The basis for the cost allocation should be usage and in particular peak demand. Capacity needs to be procured to ensure the reliability of the system at all times rather than just certain periods.

It's also recommended that all loads - both wheeling and non-wheeling - pay a per usage charge that reflects their peak demand. This fee should not be based on average or total consumption because that methodology is not aligned with economic efficiency. To understand why, consider two different systems with exactly the same total or average usage. However, one system has a constant level of electricity consumption of 25,000 MW per hour around the clock while the other has a peak demand of 28,000 MW for two hours and 24,727 MW the other twenty-two hours.

Both would have the same average consumption, but the latter would require more reserve capacity, i.e., higher capacity costs to meet the peak demand.

Finally, it's recommended that the cost for stranded assets be recovered through an appropriately adjusted per kw per month fee based on peak demand on all consumption.

1.2 BACKGROUND

The LUMS conference rightfully identified the significance of the costs associated with stranded assets:

Issue 5: Stranded assets cost

5.1 Participants agreed that the current wheeling regulations do not address the stranded assets cost. There is a need to evaluate the quantum of stranded assets cost. The Regulator shall analyse that quantum and the period by which these costs are to be recovered. It was suggested that the Regulator may conduct an independent study in this regard.

5.2 Furthermore, following mechanisms were proposed for the recovery of stranded assets cost that are employed in various global markets:

a. Stranded assets cost is charged to those consumers that are leaving the market for wheeling;

b. All categories of Bulk Power Consumers (BPCs) are charged with stranded assets cost;

c. The regulated consumers (other than those leaving the market for wheeling) are charged with stranded assets cost;

d. All categories of consumers are charged with stranded assets cost; or

e. The government provides subsidy to offset the stranded assets cost.

5.3 Participants agreed that the Regulator may propose the appropriate methodology from above and amend the wheeling regulations accordingly. However, it was argued that stranded assets cost results from inadequate policies/regulatory regime as well as differences in demand forecast and supply planning. It would be a barrier if the Wheeler of power is levied part or all of its competitor's cost. Thus, it should be a policy decision keeping in view the competitiveness of wheeling.

5.4 In comparative analysis, it was pointed out that if the stranded assets cost is charged to all consumers (including the consumers leaving the market), the overall impact will be minimal and may be a good choice to consider.

5.5 Participants agreed that the policies/regulatory regimes as well as NTDC planning process of supply and demand shall be adjusted henceforth to cover the potential wheeling applications in advance and NTDC shall consider the wheelers of power as committed projects in the IGCEP. [1]

1.3 ISSUE OF STRANDED ASSETS – PAKISTAN VIS-À-VIS INTERNATIONAL CONTEXT

It is important to both characterize and understand the nature of "stranded assets" within the context of Pakistan. First, with respect to their characterization as "assets", these are not "assets" in the

[1] LUMS Energy Institute, Summary of discussions from a stakeholder consultation at the LUMS Energy Institute Conference: Resolving regulatory and market impediments to the wheeling of electric power in Pakistan, December 9-11, 2019, page 3.

traditional meaning of the term within the electricity or energy regulatory paradigm. Within the traditional regulatory paradigm, an asset almost always refers to something that is physically tangible, i.e., a power plant, transmission or distribution line(s), transmission substation, etc. This is not the case with "stranded assets" in Pakistan. Rather the "asset" is in fact a type of a Power Purchase Agreement for generation capacity. As such, there is no physical aspect to the "asset" rather it is purely financial. They are "stranded" not because they are no longer useful but rather because they are out-of-the-money, i.e., the terms and conditions of the contracts (including the prices) do not reflect the current conditions within the electricity sector in Pakistan.

This is important with respect to the issue at hand because the traditional regulatory tools and processes for defining and dealing with so-called stranded assets are inapplicable in the Pakistani situation, i.e., increase the rate of allowable depreciation, decrease the useful life of the asset when setting rates, partial or full denial of requests for rate increases, etc. While NEPRA is involved at the very beginning through the approval and licensing process, the eventual PPA contracts lie completely outside of their purview as well as the regulatory process for electricity "assets". Thus, while the costs of the contracts are recovered through regulated rates set by NEPRA, the specific costs for the PPA's are neither reviewable nor capable of being amended by NEPRA unlike a tangible electricity asset.

The significance of this bears repeating. In particular:

- From the perspective of the regulation of the Pakistani electricity sector, these contracts are not assets,
- Once the contracts were signed, they were "outside" of the regulatory process that NEPRA uses in a determination, i.e., the costs are not considered when determining the tariff rate.
- The contracts are "out-of-the-money". In essence, NEPRA's only real role, once they have issued an approval, is to develop the cost recovery mechanism – the level/amount of the costs are exogenous to NEPRA and the entire regulatory process.

As with any "stranded asset", we note that the real issue is neither the correct determination of the fee or the efficient assignment of the costs of these contracts, but rather the burden, inefficiency (both in the near term as well as in the long term), and fairness of the costs themselves. These are valid and significant questions that are, however, largely beyond the reach of NEPRA, as well as the issue at hand, i.e., the determination of the wheeling fee. In other words, the solution for the problem of out-of-the-money contracts will not be found in either the determination of the wheeling fee or in any other rate-setting (tariff) exercise.

It is entirely possible for the Government of Pakistan to evaluate the benefits and costs of the contracts going forward and decide, assuming there are termination clauses in the contracts, whether or not to terminate or renegotiate the contracts. Indeed, there are several examples where governments have taken this step.

Ghana, faced with a similar situation as Pakistan, terminated eleven out of the thirty Power Purchasing Agreements (PPAs) signed by the Electricity Company of Ghana (ECG) following a recommendation by a review committee led by the Energy Commission, which was constituted by the Ministry of Energy to review all the PPAs signed by ECG. The Minister for Energy, John-Peter said, "Pursuant to the review exercise, Government stands to make significant savings from the deferment and/or termination of the reviewed PPAs. The estimated cost of the termination is US$402.39 million, compared to an average annual capacity cost of US$586 million each year."[2]

Similarly, in 2018, the *Bulgarian government* began the process of terminating a 15-year PPA with two privately owned coal plants that produced approximately 20% of the total output for the country. Bulgaria coordinated their decision to terminate the PPA with the European Commission's

---

[2] https://www.ghanaweb.com/GhanaHomePage/business/Gov-t-terminates-11-Power-Purchasing-Agreements-signed-under-Mahama-686747. Accessed 5 September 2020.

Competition Directorate who had already initiated proceedings based on a complaint from the Bulgarian Energy and Water Regulatory Commission (EWRC) charging that the PPA constituted incompatible state aid. The termination of the PPA was not expected to incur any penalties since the initial investment had been repaid in full.[3]

In a more complicated but relevant situation to Pakistan, in *Alberta (Canada),* the PPAs with four generators (Sundance A, Sundance B, Sundance C and Sheerness) were terminated by having the Balancing Pool replace the original buyers of the power and become the contractual buyer of the power from the four generators. The Balancing Pool is a specific feature of the re-structured Alberta electricity sector and has the following legislatively determined responsibilities: (1) Act as a buyer for the PPAs that were not sold in public auction or that were subsequently terminated by third party buyers and manage those assets in a commercial manner; (2) Sell the energy and capacity associated with the PPAs into Alberta's wholesale electricity market; (3) Act as a risk backstop in relation to extraordinary events, such as force majeure, affecting the PPAs; (4) Allocate (or collect) any forecasted cash surplus (or deficit) to (from) electricity consumers in Alberta in annual amounts over the life of the Balancing Pool; (5) Hold the Hydro Power Purchase Arrangement ("Hydro PPA") and manage the associated stream of receipts or payments; (6) Participate in regulatory and dispute resolution processes.[4]

In each of the three previous examples, PPAs were terminated by their respective governments because the net benefits of continuing the contractual relationship were less than the costs of terminating the contracts.

1.4 TREATMENT OF STRANDED ASSETS WITHIN THE PROPOSED WHEELING FEE – THE INTERNATIONAL PERSPECTIVE

Turning to how the stranded assets should be incorporated within the wheeling fee, it is beneficial to have a set of high-level Guiding Principles to inform the recommendations:

- Counterparties to a wheeling transaction should not be able to avoid or escape appropriate costs.
- The wheeling fee should be consistent with the principles of short- and long-term economic efficiency. And, in particular, should send the correct investment signals.
- The incentives produced by the wheeling fee should promote reliable operation of the electricity network and economic efficiency.
- The methodology by which the wheeling fee is calculated or updated should be known in advance.
- The wheeling fee should be understandable, transparent and replicable.

At the LUMS Conference, the participants identified five possible methodologies for recovery of the costs associated with stranded assets (see 5.2 above). The first three are based roughly on the notion that one specific type or class of consumer should pay for the stranded assets. The provision of generation capacity benefits all consumers, perhaps disproportionately, but all consumers, nonetheless. As such, assigning all of the costs to a subset of load rather than to all load is not economically efficient.

We assume that the "Stranded Assets" are synonymous with "capacity" that is no longer on average economic. That is, we assume the contracts that are now defined as being "stranded" refer to generation capacity that would no longer pass the common regulatory metric of being "used and

3 https://www.cms-lawnow.com/ealerts/2019/09/bulgarian-government-to-negotiate-with-contour-global-over-ppa-termination-for-maritza-east-3. Accessed 6 September 2020.

4 http://www.balancingpool.ca/. Accessed 6 September 2020.

useful."[5] Despite the uneconomic characteristics of this capacity, there may be periods when the system operator can and will decide to use the available capacity. Thus, while the capacity is uneconomic and is, therefore, "stranded" it is available for the system operator to use.

Thus, we are left with the final two possibilities; either all consumers pay the stranded assets, or the government provides a subsidy to offset the cost associated with the out-of-the-money contracts.

The latter is the best option from the perspective of economic efficiency but faces a number of obstacles. A government subsidy to offset the costs effectively separates current consumption and future investment decisions from choices made in the past under different circumstances and conditions. While the decisions were prudent at the time, the circumstances and conditions have changed and a government subsidy to separate the legacy decisions from current decisions regarding how much electricity to use as well as investment decisions about where to locate, what type of plant to build, etc. will improve economic efficiency.

However, as already mentioned, a government subsidy has a number of obstacles that may be difficult and time consuming to overcome. This leaves the fourth mechanism – all categories of consumers are charged with the costs associated with stranded assets. While this is a second-best option, it is consistent with the nature of the service being provided by the PPAs. There is no question that capacity is a needed service in order to reliably operate an electricity system. Moreover, since all load benefits from having a reliable system it follows that all load should share in the costs of procuring capacity.[6]

Building on this particular mechanism, i.e., that all consumers should be charged with the costs of stranded assets, it is recommended that all similar loads in a given zone (i.e., a specific DISCO) should pay the same per unit fee for Stranded Assets. This applies to load that is being served through a wheeling transaction as well as load that is not receiving wheeled power. Thus, with respect to the recovery of the costs associated with Stranded Assets there is no difference whatsoever between a load that is receiving power from a different DISCO as compared to a load that is receiving power generated locally, i.e., within the same DISCO as the load is located.

Moreover, since these are costs associated with the local provision of capacity, the specific location of the load and the distance that power has to travel to serve the load is irrelevant.

Thus, the per unit fee should be DISCO specific and similar for all similarly situated consumers. The fact that some consumers may be procuring their electricity via a wheeling transaction is irrelevant.

We turn now to the basis for the cost allocation. There are several options available, and the two most likely are either a flat fee unrelated to consumption or a variable fee that is related to usage. Again, since the service being provided by the PPAs is predominantly capacity, it is economically efficient to link usage to the cost allocation, i.e., higher usage creates the need for more capacity. Therefore, it is recommended that the basis for the cost allocation of stranded assets be megawatts (i.e., usage) and more specifically megawatt hours.

---

5 The term "used and useful" is commonly used in regulatory proceeding as metric for determining what assets should be included in the rate base.

6 The issue of who should pay for capacity has a rich history. On the one hand, as has been stated here, all load benefits from having available capacity. However, the need for capacity, i.e., the entity who causes the need to procure capacity, arises because generating and transmission facilities may fail causing the need to deploy capacity into energy. Moreover, the amount of capacity required depends on the nature of both the generating and transmission facilities. A system with a single 1000MW generator will require far more capacity than a system with 1000 1MW generators. It is interesting that the participants at the LUMs Conference did not entertain the idea that the costs of stranded assets should be apportioned to the generators and transmission/distribution providers.

Lastly, capacity procurement is linked to peak levels of demand, i.e., there should be enough capacity procured so that the system can be operated reliably under peak load conditions.[7] Thus the basis for the cost allocation should be directly related to peak demand.

In summary, it is recommended that: (1) all load, regardless of whether or not the load is party to a wheeling transaction, pay the costs associated with stranded assets, (2) all similarly placed load should pay the same per unit fee, and (3) the fee should be based on peak usage.

This methodology is consistent with how the costs of stranded assets are dealt with in other jurisdictions. To the extent that a regulated entity is allowed to recover the costs associated with stranded assets, the regulator will establish a temporary or interim charge type – sometimes called a "transition charge" – that is used to recover the costs. The transition charge has a limited life and only specifically defined costs are allowed to be recovered through this mechanism. There are other mechanisms such as securitization which is a financing mechanism through which an independent enterprise is established to (1) issue bonds; (2) sell the bonds to investors; (3) use the proceeds from the bond sales to buy out the utilities' stranded assets (which removes the stranded assets from the utilities' rate bases); and (4) place charges on consumers' electric bills for a limited amount of time to re-pay the bond investors.

In *Texas*, legislation was passed that allowed for the creation of a "Competition Transition Charge" that established the right to recover stranded costs: "An electric utility is allowed to recover all of its net, verifiable stranded costs incurred in purchasing power and providing electric generation service. Recovery of retail stranded costs by an electric utility shall be from all existing or future retail customers, including the facilities, premises, and loads of those retail customers, within the utility's geographical certificated service area as it existed on May 1, 1999. A retail customer may not avoid stranded cost recovery charges by switching to on-site generation. In multiple certificated areas, a retail customer may not avoid stranded cost recovery charges by switching to another electric utility, electric cooperative, or municipally owned utility after May 1, 1999. Recovery of stranded cost from wholesale customers. Nothing in this section shall alter the rights of utilities to recover wholesale stranded costs from wholesale customers." The transition charge is based on the amount of kWh used.

Similarly, in *the Philippines,* stranded costs associated with IPP contracts were recovered by a (capped) per kWh charge. The "Universal Charge" is levied on grid-connected end users and (i) covers the cost of stranded NPC PPA contracts for both energy and capacity that were entered into with independent power producers during the 1990s. The process of dealing with the stranded assets "began in 2001 with the passage of the Electric Power Industry Reform Act (Republic Act No. 9136, or "EPIRA"). This process involved: (i) the privatization of all Napocor generation and transmission assets, (ii) state absorption of Napocor's stranded debt (roughly 200 billion pesos), (iii) a congressional investigation and review of all current IPP contracts and (iv) the mandated unbundling of rates…responsibility for implementing the findings was given to the Power Sector Assets and Liabilities Management Corporation ("PSALM"). PSALM is a state-owned corporation tasked with privatizing Napocor's assets in the EPIRA regime, and is staffed by electricity sector experts, and former private sector bankers and lawyers. PSALM was mandated in the EPIRA law to implement the findings of the IAC Review and to "diligently seek to reduce stranded costs, if any. At the same time, PSALM is responsible for privatizing Napocor's assets, and to "optimize the value and sale prices" of Napocor's assets in that process.[8]

---

7 Recent events in California where renewable generation under-performed relative to the forecast by more than 4000MW causing the system operator to implement rolling blackouts, is an example of what happens when not enough capacity is procured.

8 Woodhouse, Eric. "The IPP Experience in the Philippines". Program on Energy and Sustainable Development, Stanford University. September 2005. P 27

1.5 METHODOLOGY FOR ANALYSING AND CALCULATING THE STRANDED ASSETS COST

Our methodology to produce the costs of stranded assets was as follows:

1. We use the actual peak load for July and August 2020 and the forecasted peak load for September 2020 to June 2023, i.e., three years.

2. We use the actual and forecast generation capability from July 2020 to June 2023.

3. The difference between the peak load and generation capability is the power balance.

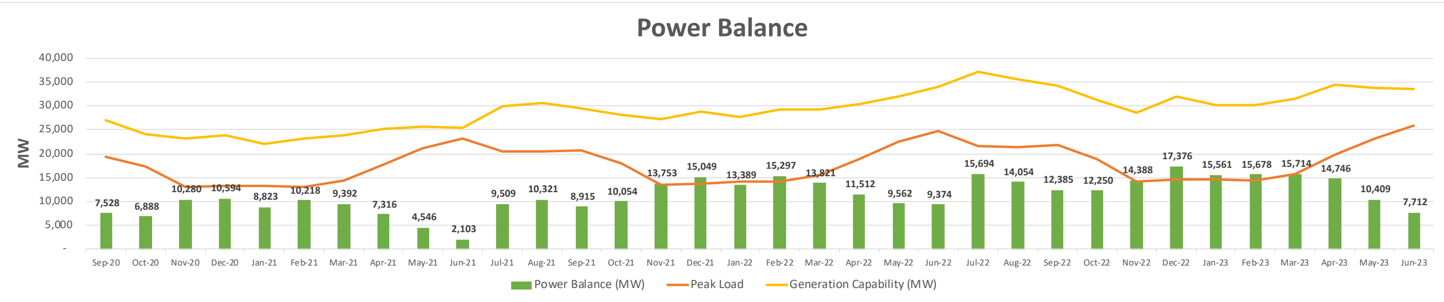
4. We then use the Total Capacity Payment.

5. The $Stranded\ Asset\ Cost = \left[\frac{Power\ Balance}{Generation\ Capability}\right] * Total\ Capacity\ Payment$

6. The first term to the right of the equal sign expresses excess capacity as a fraction of total generation.

7. Finally, we convert this kW/Month.

8. The graph below shows (1) generation capability (the yellow line), (2) peak load (the orange line) and (3) the power balance (the green bars) for September 2020 to June 2023.



Similarly, the table below provides our estimates for the fee applicable for Stranded Assets.

MONTH	FEE FOR STRANDED ASSETS (millions of PKR/kW/Month)
July-20	1,248.7
August-20	1,144.8
September-20	1,030.0
October-20	1,182.6
November-20	2,437.8
December-20	2,395.1
January-21	2,413.7
February-21	2,676.6
March-21	2,194.1
April-21	1,293.0
May-21	668.1
June-21	284.5
July-21	1,278.6
August-21	1,402.0
September-21	1,402.3
October-21	1,968.7
November-21	3,735.9
December-21	3,778.4
January-22	3,397.3
February-22	3,786.6
March-22	3,140.3
April-22	2,066.1
May-22	1,382.6
June-22	1,157.7
July-22	2,346.6
August-22	2,225.1
September-22	2,027.7
October-22	2,553.9
November-22	4,340.0
December-22	4,600.6
January-23	4,276.4
February-23	4,357.5
March-23	3,801.7
April-23	2,631.9
May-23	1,611.8
June-23	1,083.4

We note that the expected fee ranges from PKR284.5 (June 2021) to PKR4600.6 (December 2022) with a mean of PKR2314.5 and a standard deviation of PKR1178.62 over the three-year time period. To reduce this volatility NEPRA could base the fee on the anticipated average annual costs of PKR1580.8 for 2020-21, PKR2374.7 for 2021-22, and PKR2988.0 for 2022-23. Relative to the monthly charge mechanism based on actual expenditures, using the average value would require NEPRA to have some "true-up" mechanism to address any differences between the expected and actual costs.

1.6 CONCLUSION AND RECOMMENDATIONS

The issue of so-called stranded assets is ubiquitous within the electricity industry in all countries. While 20-25 years ago it was the stranding of assets due to the implementation of open access and competition, it is currently an issue because of the imperatives of de-carbonization. While over time it is desirable that private equity increasingly absorb the costs of decision making, to the extent that is not possible - as is the case in Pakistan - then we should find cost recovery mechanisms which are grounded in one of two concepts - either aligning the costs with those who created the need for them or aligning the costs with those who benefit. Our analysis and recommendations are based on the latter approach.

In conclusion, with respect to the wheeling fee, the solution to the issue of cost recovery for the stranded assets in Pakistan is straightforward.

a. Ignoring that an argument can be made that generation and transmission be responsible for these costs as they are the entities that cause the need for capacity, all load should pay these costs. It makes no difference whether a specific load is a party to a wheeling contract or not, all load benefits from the provision of capacity. Moreover, since the generation backing the wheeling transaction may not be able to perform, the load from a wheeling transaction is just as dependent on capacity as any other load on the system.

b. The basis for the cost allocation should be usage and in particular peak demand. Capacity is procured to ensure the reliability of the system at all times rather than just certain periods. Finally, while cost recovery should be DISCO based to reflect the fact that each DISCO has different capacity needs, we did not have complete cost data at the level of the DISCO and, as a result our, calculated fees for stranded assets are at an aggregated level.

c. It follows then that we recommend that all loads - both wheeling and non-wheeling - pay a per usage charge that reflects their peak demand. This fee should not be based on average or total consumption because that methodology is not aligned with economic efficiency. To understand why, consider two different systems with exactly the same total or average usage. However, one system has a constant level of electricity consumption of 25, 000MW per hour around the clock while the other has a peak demand of 28,000 for two hours and 24,727MW the remaining twenty-two hours. Both would have the same average consumption, but the latter would require more reserve capacity, i.e., higher capacity costs to meet the peak demand.

d. Finally, we recommend that the cost for stranded assets be recovered through an appropriately adjusted per kW per month fee based on peak demand for all consumption.

2. TREATMENT OF NETWORK LOSSES WITHIN THE WHEELING FEE

2.1 BRIEF SUMMARY

During normal operation of an electricity network, some electricity that is injected onto the grid "disappears" through conductors, transformers and lines. From a cost causation perspective, these "losses", which are called "technical losses", are almost entirely the result of usage[9] and location. Technical losses refer to energy converted to heat in power lines and transformers, resulting from the laws of physics. In addition to technical losses, there are non-technical losses, or "commercial losses" as they are called in Pakistan, that refer to energy delivered and consumed, but for some reason, not recorded by a meter. The measurement of technical losses is a relatively straightforward process. In contrast, an accurate measurement of non-technical losses is difficult if not impossible to determine.

There are four primary questions related to losses:

1. how is the energy required to offset the losses procured;
2. what is the price for the energy used to offset losses;
3. who should pay for the energy; and
4. how are the costs allocated amongst those who pay?

A wheeling transaction is fundamentally different from other types of transactions in that the electricity comes from a pre-identified source and flows to a pre-identified sink. This transaction will result in technical losses due to the actual power flows. However, the wheeling transaction is neither responsible for nor contributes to any non-technical losses. As a result, from the perspective of economic efficiency, we recommend, like the participants at the LUMs conference, that wheeling transactions not be charged for commercial losses.

Technical losses arise from usage and the magnitude of the losses is positively related to the distance between generation and consumption. Thus, from a theoretical perspective, cost recovery should be related to usage and distance/location. While the theoretical solution to the issue is simple, in actual practice, things may be a bit more complex. Measuring the actual amount of technical losses should be as easy as taking the difference between the energy injected into the grid and the energy that exits the grid. Thus, we recommend that the cost of technical losses be allocated to all load on the basis of the amount of their usage, i.e., an energy charge, expressed in Rs/kWh.

[9] Including potentially time of use.

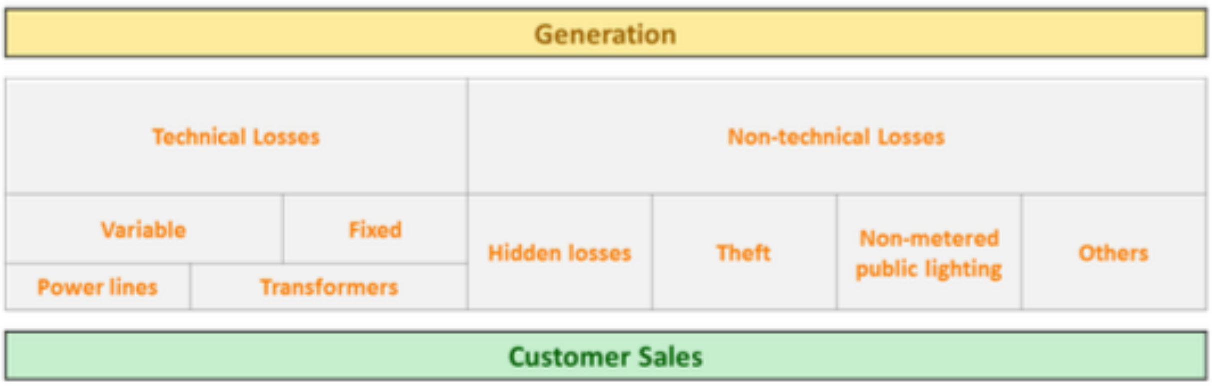
2.2 BACKGROUND

For many countries, "wheeling" is the first step in introducing open access. Ultimately, the success of all wheeling regimes regardless of the country or jurisdiction - as well as the success of open access itself - will require the separation of generation from transmission and distribution activities, and unbundling, i.e., disaggregation, of the services provided by both the transmission and distribution wire businesses as well.

During normal operation of an electricity network, some electricity that is injected onto the grid "disappears" through conductors, transformers and lines. From a cost causation perspective, these "losses", which are called "technical losses", are almost entirely the result of usage[10] and location.

Technical losses refer to energy converted to heat in power lines and transformers, resulting from the laws of physics. In addition to technical losses, there are non-technical losses, or "commercial losses" as they are called in Pakistan, that refer to energy delivered and consumed, but for some reasons, not recorded by a meter. The measurement of technical losses, which are the result of using transformers and lines to transmit energy and by how far the electricity has to travel, is a relatively straightforward process. In contrast, an accurate measurement of non-technical losses is difficult if not impossible to determine. The following figure provides an outline for understanding the differences between the two types of losses:[11]

Figure 1: Difference between Technical Versus Non-technical Losses



We note that while technical losses will always be present in any operational electricity system, this is not the case with non-technical losses, i.e., there is no reason for non-technical losses to exist. To the extent that non-technical losses exist, it is one measure of the inefficiency of the system.

There are four primary questions related to losses, (1) how is the energy required to offset the losses procured, (2) what is the price for the energy used to offset losses, (3) who should pay for the energy, and (4) how are the costs allocated amongst those who pay?

With respect to the allocation of the costs to wheeling transactions associated with network losses, the participants at the LUMs conference arrived at the following conclusions:

[10] Including potentially time of use.

[11] Council of European Regulators (CEER), "CEER Report on Power Losses". 18 October 2017. P. 10.

Issue 2: Network Losses

2.1 The current wheeling regulations do not address the technical losses in the transmission and/or distribution system when wheeling power.

2.2 Participants agreed that the fair cost of technical losses should be included in the wheeling charge and accordingly allocated to the Wheeler of power and/or BPC as per the bilateral wheeling contract.

2.3 Different methods for allocation of these technical losses were discussed and CPPA-G provided the past trend and future directions this allocation method may take. Some methods of allocation based on direct and indirect distances are now obsolete whereas some more advanced methods like nodal pricing may require a high level of automation with sub-hour input of actual load flow scenarios.

2.4 There was general consensus that the postage stamp method for determining the cost of technical losses is equitable and has been adopted in many global markets, till the time nodal pricing mechanism is established or any other advanced loss calculation techniques are implemented. However, some participants argued that alternative method of load flow analysis (as per existing practice) should continue on a case to case basis, till the time CTBCM[12] is implemented. There was another recommendation of implementing zonal pricing mechanism for accounting technical losses in the wheeling charges, till the time nodal pricing mechanism is established.

2.5 It was pointed out that the Regulator shall decide the methodology of determining the cost of technical losses in the wheeling charge and the output of the consultative session is an objective assessment of different options.

2.6 Participants acknowledged that due consideration should be given to those wheelers of power that reduce overall system technical losses. It was discussed that the instantaneous power flow with changing load conditions and future addition of generation may alter the *technical losses* in both positive and negative direction. Static methods, like instantaneous power flow analysis, are not real depiction of individual contribution to the overall loss of the network.

2.7 Participants agreed that administrative and commercial losses of DISCOs shall not be part of the wheeling charge.[13]

The answers/solutions to the four questions raised above depends primarily on the type of electricity market. In general, there are two fundamentally different models for an electricity market. Put simply, there are two distinctly different electricity market models because there are two fundamentally different ways of looking at the problem of implementing open access; the focus can either be on the objective of open access, *i.e.,* a competitive electricity market, or, alternatively the focus can be on open access to the transmission grid itself, and then allowing the competitive electricity market to develop around that structure. The choice of which to focus on is fundamental and affects all related decisions.

The choice of which electricity market model to adopt rests on the answer to a single question: How accurately can the actual conditions on the electricity network be known in advance? Equivalently, can the conditions on the electricity network be predicted accurately enough in advance so as to

12 We note in passing that the design and operation of a Competitive Trading Bilateral Contract Market is almost always inconsistent with a market based on nodal pricing.

13 LUMS Energy Institute, Summary of discussions from a stakeholder consultation at the LUMS Energy Institute Conference: Resolving regulatory and market impediments to the wheeling of electric power in Pakistan, December 9-11, 2020, Pages 1-2.

minimize the actions the system operator must take to maintain reliable operations? The greater the accuracy the less discretion and action needed by the system operator and vice versa.

New Zealand, because of the expected significance of congestion, operating reserves, losses, voltage and other characteristics – all of which reduce the accuracy of predicting the state of the grid in the future – designed and implemented an electricity market based on nodal pricing, i.e., locational marginal pricing.[14] This model has subsequently been replicated in the United States, Central and South America, Mexico, the Philippines, Singapore, and parts of Canada. The foundational structure of this market model is the optimization of actual real time power flows (most often energy and ancillary services are co-optimized) by the System Operator using Security Constrained Economic Dispatch (SCED). The use of SCED results in nodal prices, i.e., a price for electricity at every electrical bus/node. In this model the system and market operators are the same entity and in almost all cases are entirely separated from the wires businesses.[15] In the United States, this function is called either an Independent System Operator (ISO) or Regional Transmission Operator (RTO).

Within the paradigm of the nodal pricing market design, most jurisdictions choose to have the SCED algorithm optimize, i.e., minimize the production cost, of energy, congestion and technical losses. That is, for every operating interval in real time, the SCED algorithm will find the lowest production cost for the sum of energy, congestion and technical losses. In this market design, the nodal price or, equivalently, the location marginal price (LMP), is defined as:

$$LMP_t^n = MCE_t^n + MCC_t^n + MLC_t^n$$

Where: LMP_t^n = the locational marginal price at node *n* for time *t*,

MCE_t^n = the marginal cost of energy at node *n* for time *t*,

MCC_t^n = the marginal cost of congestion and node *n* for time *t* and

MLC_t^n = marginal loss component at node *n* for time *t*.

In this way, all four of the aforementioned primary questions related to losses are answered: losses are procured via the dispatch process in real time, the price is the marginal loss component of the locational marginal price, and load pays on the basis of how much energy they consume. The effect of distance is incorporated via the dispatch algorithm, i.e., the SCED.

For an example, consider the following table from the Midcontinent ISO (MISO) for September 25, 2020 at 8:10 Eastern Standard Time[16]:

[14] "The MISO market design establishes a price for losses – like energy and congestion – at the margin, i.e., the nodal price is determined by the marginal or incremental cost of losses, energy and congestion."

[15] New Zealand being a notable exception.

[16] https://www.misoenergy.org/markets-and-operations/real-time--market-data/real-time-displays/ accessed on September 25, 2020. This represents only a snapshot of the available node.

LMP Consolidated Table

25-Sep-2020 - Interval 08:10 EST

Search: Location ● All ○ North ○ Central ○ South

	Hourly Day-Ahead Ex Ante			Hourly Day-Ahead Ex Post			Five Minute Real-Time			Last Hour Estimated		
	HE 8			HE 8			08:10			HE 07		
Location	**LMP**	**MLC**	**MCC**	**LMP**	**MLC**	**MCC**	**LMP**	**MLC**	**MCC**	**LMP**	**MLC**	**MCC**
EAI.ANO1	17.51	-1.12	0.26	17.51	-1.12	0.26	1.97	-1.55	-16.48	16.50	-1.40	0.00
EAI.AECCHYDRO9	17.80	-0.81	0.24	17.80	-0.81	0.24	2.68	-1.22	-16.10	16.80	-1.10	0.00
LAGN.BYCT4	18.09	0.32	-0.60	18.09	0.32	-0.60	19.82	-0.18	0.00	17.75	-0.15	0.00
SME.BATESV_2	17.53	-0.04	-0.80	17.53	-0.04	-0.80	19.63	-0.37	0.00	18.00	0.10	0.00
CLEC.ACA11	17.92	0.17	-0.62	17.92	0.17	-0.62	19.87	-0.13	0.00	17.81	-0.09	0.00
CONS.WPSC_2.AZ	19.74	1.37	0.00	19.74	1.37	0.00	21.53	1.53	0.00	19.40	1.51	0.00
MPW.MPW	18.48	0.37	-0.26	18.48	0.37	-0.26	34.72	0.18	14.54	16.71	0.06	-1.25
EAI.CC.UPS3	16.97	-0.62	-0.78	16.97	-0.62	-0.78	18.86	-1.14	0.00	16.91	-0.99	0.00
INDIANA.HUB	19.08	0.64	0.07	19.08	0.64	0.07	26.86	0.62	6.24	18.52	0.63	0.00
AMMO.MERAMEC4	17.66	-0.71	0.00	17.66	-0.71	0.00	-9.62	-0.76	-28.86	15.83	-0.68	-1.39
CLEC.TPS4	18.45	0.71	-0.63	18.45	0.71	-0.63	20.30	0.30	0.00	18.18	0.29	0.00
WPS.WESTON3	23.35	0.32	4.66	23.36	0.33	4.66	20.72	0.72	0.00	31.31	0.57	12.84
CIN.CC.SUGRCK	18.53	0.24	-0.08	18.53	0.24	-0.08	41.35	0.21	21.14	18.98	0.18	0.90
EES.L_CREEK1	18.15	-0.09	-0.13	18.14	-0.10	-0.13	4.13	-0.67	-15.20	17.04	-0.86	0.00
AMMO.LABADIE3	17.21	-1.16	0.00	17.21	-1.16	0.00	-15.42	-1.13	-34.29	14.28	-0.98	-2.64

The MISO market design establishes a price for losses – like energy and congestion – at the margin, i.e., the nodal price is determined by the marginal or incremental cost of losses, energy and congestion. Both losses and congestion have to be relative to a set location which is called a reference bus. As shown in the table above, the marginal "price" of losses at EAI.AN01 (the node in the first row) is -$1.55 for the interval from 8:10 to 8:15. However, for CONS.WPSC_2_.AZ the marginal price of losses is $1.53. The difference in prices, and in particular the difference between the negative and positive prices at the two nodes, reflects where the nodes are located relative to the reference bus.[17]

Thus, the cost and allocation of technical losses are not directly regulated[18] within the MISO electricity market as they are both determined by the operation of the electricity market.

The alternative model, developed first in the United Kingdom and in Nord Pool, is used primarily in Europe. While not refuting that actual transmission capacity in real time can be affected by the issues identified above, this model assumes that the commercial/financial/reliability effects of those issues are insignificant.

This model is not based on optimized real time dispatch through the use of SCED and, as such, there are no nodal prices or a marginal loss component. Rather the electricity system is divided into "zones" and real time power flows are managed internally by a "Transco". The Transco, or more precisely the Transmission System Operator (TSO), owns the transmission assets and operates the wires business as well as being the system operator.

In contrast to the MISO example, the TSO model is based on the complete separation of electricity and transmission. Who produces the power and at what price is determined by the price mechanism through a properly designed electricity market. A for-profit regulated transmission monopolist, that is both the owner of the transmission assets and the system operator, i.e., the TSO, will then be

17 While the MISO prices and bills losses at the margin, they rebate the difference between marginal and actual losses back to the end user.

18 The MISO Tariff is subject to approval by the Federal Energy Regulatory Commission. The tariff specifies how dispatch will take place but does not define a price for energy, congestion or losses.

responsible for the transportation of the power at regulated rates from the producer to the consumer. This necessarily means that the price/rate of the service provided by the TSO are determined through a regulatory process. As a result, the Tariff Rate is for a bundled product offering, i.e., the rate covers the cost of dispatch, losses, congestion, transmission wires, maintenance, etc.

For either market design, a significant amount of non-technical losses is extremely problematic. As a result, we recommend that, prior to implementing any type of electricity market, non-technical or commercial losses should be minimized.

A wheeling transaction is fundamentally different from other types of transactions in that the electricity comes from a pre-identified source and flows to a pre-identified sink. This transaction will result in technical losses due to the actual power flows. However, the wheeling transaction is neither responsible for nor contributes to any non-technical losses. As a result, from the perspective of economic efficiency, we recommend, like the participants at the LUMs conference, that wheeling transactions not be charged for commercial losses.

From a regulatory perspective, when there are no non-technical losses, then technical losses are just the difference between the electricity that is injected and the electricity withdrawn, i.e., it is the metered difference between the grid injection point and the grid withdrawal point. When non-technical losses are present, then this difference will include both technical and non-technical losses. This creates a problem for rate setting insofar as the rate will combine losses that are necessary and appropriate with those that are unnecessary and should be eliminated.

In the situation where there are significant non-technical losses, since it is difficult/impossible/inefficient to rely on actual metered data, the regulator has two broad choices:

- they can estimate the technical losses from the physical characteristics of the transmission assets and the expected/actual power flows and then apply a factor to include the non-technical losses; or

- they can simply apply a predetermined factor to total production, i.e., a "loss" factor, that is meant to account for both technical and non-technical losses.

In either case the regulator can then determine the cost of the "allowable" amount of losses (both technical and non-technical) that can be recovered by the transmission/distribution utility.

As previously stated, technical losses arise from usage. The magnitude of the losses is positively related to the distance between generation and consumption. Thus, from a theoretical perspective, cost recovery should be related to usage and distance/location. While the theoretical solution to the issue is simple, in actual practice, things may be a bit more complex. Measuring the actual amount of technical losses should be as easy as taking the difference between the energy injected into the grid and the energy that exits the grid.

Of the two cost drivers – usage and location – the former is primarily related to current consumption while the latter is primarily related to investment decisions. Both are important but, given where Pakistan stands with respect to implementing an electricity market, we believe that the additional complexity of including a locational component is very unlikely to affect decision-making. We add, however, that this is one reason why Pakistan should continue the process of implementing an electricity market. As such, we recommend that the recovery of the cost associated for transmission losses be based solely on usage. Furthermore, since all usage contributes to the need to procure energy to mitigate transmission losses, there is no basis for excluding the load associated with wheeling transactions for the charge.

Thus, we recommend that the cost of technical losses be allocated to all load on the basis of the amount of their usage, i.e., an energy charge, expressed in Rs/kWh. Before providing some international examples, we note that actual regulated rates across jurisdictions are largely non-comparable. The general format for determining the revenue requirement is as follows:

$$RR_t = ((RB_t)*R_t) + OC_t + D_t + T_t + F_t$$

Where:

RR = Revenue requirement,

R = Rate of return,

RB = Rate base (Gross Investment – Accumulated Depreciation),

OC = Operating costs,

D = Depreciation expenses,

T = Taxes,

F = Other costs,

t = Test year

In determining the Revenue Requirement – which determines the amount that is to be recovered from users of the services – the regulator must make a number of assumptions regarding the cost of equity, the rate of depreciation, the appropriate debt-to-equity ratio, the tax rate, the rate of inflation, etc. Moreover, the transmission and distribution assets used in different jurisdictions will reflect the local requirements and will not be similar across jurisdictions. Thus, from the perspective of the precise regulatory determined rate, there is little value at comparing the rates between two countries – the situation in Singapore is very different from that in Germany.

ELECTRIC BILL				
METERING INFORMATION				
Meter Number	Previous Reading	Present Reading	Multiplier	Registered
3302ZN92097	9337	9659	1	322 kWh

RATE: Residential		
Generation Charge	322 X 3.4029	1,095.73
Transmission Charge	322 X 0.9605	309.28
System Loss Charge	322 X 0.5493	176.87
Distribution Charge	322 X 1.1628	
METERING CHARGE		
Retail Customer Charge	5 X 1 mo	5.00
Metering System Charge	322 X 0.2435	78.41
Supply Charge	322 X 0.5271	169.73
Lifeline Rate Subsidy	322 X 0.0761	24.50
Interclass Subsidy	322 X -0.7130	-229.59
Power Act Reduction	322 X 0.3000	-96.60
CERA	374.42 X 11.87%	44.44
FRANCHISE TAX		
National	1952.19 X 2%	39.04
UNIVERSAL CHARGES		
Missionary	322 X 0.0168	5.41
Environmental Fund	322 X 0.0025	0.81
OTHER CHARGES		
Feb-Mar Missionary Electrification Charges (MEC)	546 kWh X 0.0168	9.15
TOTAL CURRENT AMOUNT		P 2,006.60

Table-I to the right is a representative bill from the Philippines for a residential customer in 2017. The System Loss Charge "represents recovery of the cost of power lost due to technical and non-technical losses currently pegged at 9.5% for private distribution utilities and 14% for electric cooperatives, including company used power."[19] The regulator determined rate is 0.5493 Philippine Pesos per kWh.

Similarly, in the Table above that provided the decomposition of the nodal prices in MISO for September 25, 2020 at 8:10 Eastern Standard time, the marginal loss component is generally equivalent to the technical losses for that dispatch interval. Thus, a load-weighted average of these values would be equivalent to the annual usage fee in Pakistan. We provide this for PJM – the largest electricity market in the world – in the table below:[20]

[19] www.erc.ph/files/media/812_pub-unbundled accessed September 27, 2020

[20] https://www.monitoringanalytics.com/reports/PJM_State_of_the_Market/2019/2019-som-pjm-sec11.pdf accessed September 29, 2020.

Table 1: Statistics from PJM Electricity Market

	Real-Time LMP	Energy Component	Congestion Component	Loss Component
2008	$71.13	$71.02	$0.06	$0.05
2009	$39.05	$38.97	$0.05	$0.03
2010	$48.35	$48.23	$0.08	$0.04
2011	$45.94	$45.87	$0.05	$0.02
2012	$35.23	$35.18	$0.04	$0.01
2013	$38.66	$38.64	$0.01	$0.02
2014	$53.14	$53.13	($0.02)	$0.02
2015	$36.16	$36.11	$0.04	$0.02
2016	$29.23	$29.18	$0.04	$0.01
2017	$30.99	$30.96	$0.02	$0.01
2018	$38.24	$38.19	$0.04	$0.02
2019	$27.32	$27.28	$0.02	$0.02

These values are per MWh so that the average technical loss per kWh "fee" for 2019 would be US$0.00002. It is important to recognize that this is the average per kWh *price* and not *cost* of the technical loss. This number would be applied to all power purchased in the PJM footprint for 2019.

Similarly, in New Zealand, "consumers face the cost of all losses, whatever the cause. For most consumers, this cost is bundled in the price per kWh they pay their retailer."[21]

As previously stated, our recommended fee for technical losses in Pakistan, is in the form of an energy charge:

$$\textit{Fee For Technical Losses} = \frac{Cost\ of\ technical\ losses\ on\ the\ NTDC\ and\ DISCO\ networks}{Amount\ of\ energy\ purchased}$$

For 2018-19, in kWh, the fee would have been: $\frac{Rs\ 151{,}126{,}651{,}951.74}{93{,}877{,}000{,}000\ kWh} = Rs\ 1.6097$

This rate should be applied to all load (outside of Karachi Electric's region) – including the load being served through a wheeling transaction.

[21] https://www.ea.govt.nz/operations/distribution/losses/ accessed September 27, 2020.

2.3 CONCLUSION AND RECOMMENDATIONS

To conclude, our recommendations with respect to how the costs for transmission and distribution network losses should be included in the wheeling fee, are as follows:

- Wheeling transactions should only be charged for technical, and not commercial losses.
- To the greatest extent possible, commercial losses should be minimized prior to implementing an electricity market.
- The recovery of the costs associated with technical losses should be based solely on usage.
- The actual fee for losses should be apportioned to all load, including load being served via a wheeling contract.
- The fee should be based on the total cost of losses and the amount of energy purchased by the total retail sector.

3. TREATMENT OF NTDC'S USE OF SYSTEM CHARGE (UOSC) WITHIN THE WHEELING FEE

3.1 BRIEF SUMMARY

The services provided by the wire businesses, including the case of Pakistan's system operation, and the costs associated with providing these services must be defined and made transparent because not all parties to a wheeling transaction will need or want to purchase all of the services from the wires businesses. Nevertheless, wheeling transactions will use the services of wire business, as well as the services provided by the system operator, and should be charged accordingly, i.e., there needs to be an appropriately determined "wheeling fee" that reflects the services the wheeling transaction is "purchasing". *This paper addresses the cost recovery mechanism for one such "service" - the recovery of the Use of System Charges (UoSC) incurred by the National Transmission and Despatch Company (NTDC).*

The services provided by NTDC, including the provision of an electricity transmission network and the reliable despatch of electricity, are both necessary for wheeling to take place. The only exception is the case where the power flows, associated with a specific wheeling transaction, never use the high voltage network. Even in that case, there is a strong argument that the specific transaction is the beneficiary of a reliable system with access to ancillary services. Given that, and under the current regulatory rate setting paradigm, we agree with the recommendations arrived at by the participants at the LUMs conference. Specifically, we recommend:

- The UoSC should be treated uniformly across all power that flows in the system.
- When power is wheeled across two or more DISCOs, the NTDC system is utilized and, therefore, the UoSC should be a part of the wheeling charge.
- When the wheeler of power and the associated Bulk Power Consumers (BPC) are located in the same DISCO system, the BPC is still utilizing the NTDC system for system stability and reliability. Furthermore, even though the actual flow of power to the BPC is not using the transmission network, the network is providing reliability services that are different from system operation.
- Additionally, when calculating the demand charge, we recommend the use of a coincidental peak rather than a non-coincidental peak.
- Finally, we recommend the UoSC charge be based on a combination of a demand charge and an energy charge. This reflects the effects of managing electricity flows during peak periods (demand charge) as well as non-peak periods (energy charge). From the perspective of economic efficiency, it is desirable that the rate sends the correct signals when the system is at capacity as well as when conditions are less tight. *Using both demand and energy charges is consistent with international practices and sends improved economic signals.*

3.2 BACKGROUND

With respect to the allocation of *cost of the services* provided by the National Transmission and Despatch Company (NTDC) in regard to wheeling transactions, the participants at the LUMs conference arrived at the following conclusions:

Issue 3: Use of System Charge (UoSC) of NTDC system

3.1 It was agreed that the NTDC system is necessary to provide stability (frequency and voltage control) and reliability to the overall grid system. Accordingly, NTDC undertakes investments in its system. Thus, UoSC should be treated uniformly across all power that flows in the system.

3.2 Participants agreed that, when power is wheeled across two or more Distribution Companies (DISCOs), NTDC system is utilized and, therefore, the UoSC shall be a part of the wheeling charge.

3.3 When the Wheeler of power and its BPC are located in the same DISCO system, there was general consensus that the BPC is still utilizing the NTDC system for system stability and reliability. The actual flow of power to the BPC for the Wheeler of power may or may not be coming directly from the transmission network. Thus, UoSC should still be a part of the wheeling charge. However, some participants argued that there is no additional burden on the DISCO for using the NTDC's system and UoSC should be excluded from the wheeling charges, till the time Competitive Trading Bilateral Contract Market (CTBCM) is implemented.[22]

3.3 ISSUE OF NTDC'S USE OF SYSTEM CHARGE – PAKISTAN VIS-À-VIS INTERNATIONAL CONTEXT

Currently, NTDC operates two core businesses:

- The Wire Business
  - Transmission & Generation Planning
  - Design and Engineering
  - Project Development and Execution
  - Operation & Maintenance of Transmission Assets
- System Operation and Despatch
  - Economic Generation Despatch
  - Power System Operation and Control23

Importantly, while highlighting and defining the two core businesses, the regulatory process has not yet treated them as separate. Instead, the National Electric Power Regulatory Agency (*NEPRA's)* regulatory rate setting process aggregates the costs and derives a single bundled rate for all the services provided. This rate is called the Use of System Charge (UoSC).

The fact that services provided by NTDC have been separately identified but do not have separate rates is both interesting and significant for the reason that the operational activities provided by NTDC are fundamental to the ultimate electricity market design. As such, it is necessary to explain why.

In general, there are *two fundamentally different models for an electricity market.* Put simply, there are two distinctly different electricity market models because there are two fundamentally different ways of looking at the problem of implementing open access. The focus can either be on the objective of open access, *i.e.,* a competitive electricity market. Alternatively, the focus can be on open access to the transmission grid itself and then allowing the competitive electricity market to develop around that structure. The choice of which to focus on is fundamental and affects all related decisions.

---

[22] LUMS Energy Institute, Summary of discussions from a stakeholder consultation at the LUMS Energy Institute Conference: Resolving regulatory and market impediments to the wheeling of electric power in Pakistan, December 9-11, 2019, P. 2.

[23] National Electric Power Regulatory Authority, "Determination of the Authority in the matter of Petition filed by National Transmission & Despatch Company Ltd. (NTDC) for Determination of Transfer/Wheeling Charges for the FY 2017-18 and FY 2018-19 - Case No. NEPRA/TRF-450/NTDC-2018. July 31, 2019. P. 6.

The choice of which electricity market model to adopt rests on the answer to a single question i.e., how accurately can the actual conditions on the electricity network be known in advance? Equivalently, can the conditions on the electricity network be predicted accurately enough in advance so as to minimize the actions the system operator must take to maintain reliable operations? The greater the accuracy, the less discretion and action needed by the system operator, and vice versa.

New Zealand, because of the expected significance of congestion, operating reserves, losses, voltage and other characteristics – all of which reduce the accuracy of predicting the state of the grid in the future – designed and implemented an electricity market based on nodal pricing, i.e., locational marginal pricing.[24] This model has subsequently been replicated in the United States, Central and South America, Mexico, the Philippines, Singapore, and parts of Canada. The foundational structure of this market model is the optimization of actual real time power flows (most often energy and ancillary services are co-optimized) by the System Operator using Security Constrained Economic Despatch (SCED). The use of SCED results in nodal prices, i.e., a price for electricity at every electrical bus/node. In this model, the system and market operators are the same entity and, in almost all cases, are entirely separated from the wires businesses.[25] In the United States, this function is called either an Independent System Operator (ISO) or Regional Transmission Operator (RTO).

The aforementioned separation of the business activities of NTDC, into the wires business and system operation and despatch, is entirely consistent with the path taken by New Zealand and then a host of other countries. Under this structure, system dispatch and the wires business are two separate activities and two completely separate businesses. As a result, the costs of providing system operation are separate from the costs of owning and operating the wires business. There are two regulated rates - one for system operation and one for the wires business. This is not how NTDC is currently regulated. Rather the costs of the two businesses are aggregated and a single rate is determined for the combined businesses.

The alternative model, developed first in the United Kingdom and in Nord Pool, is used extensively in Europe. While not refuting that actual transmission capacity in real time can be affected by the issues identified above, this model assumes that the commercial/financial/reliability effects of those issues are not material.

This model is not based on optimized real time despatch through the use of SCED and, as such, does not rely on nodal prices to guide decisions. Instead, the electricity system is divided into "zones" and real time power flows are managed by a "Transco". The Transco, or more precisely the Transmission System Operator (TSO), owns the transmission assets and operates the wires business as well as being the system operator.

Both the "Transco" and Nodal Pricing approaches agree that the generation and transmission of electricity should be separated. The difference occurs due to the assumption of the former model that real time power flows can be accurately predicted. Who produces the power and at what price can and should be determined by price mechanism through a properly designed electricity market? A for-profit regulated transmission monopolist, that is both the owner of the transmission assets and the system operator, i.e., the TSO, will then be responsible for the transportation of the power at regulated rates from the producer to the consumer.

This model is, therefore, inconsistent with the separation of NTDC according to its core business – the distinction between the wires business and system operation is irrelevant in this model. However, NEPRA's current rate setting process does not make any distinction between NTDC's two core businesses and is, therefore, aligned with the way in which a TSO should be regulated.

---

[24] Nodal pricing creates a price based on voluntary bids and offers that generators will be paid or load will pay for electricity at every electrical bus/node in the system for every dispatch interval. The prices are, in effect, the dispatch instruction from the system operator.

[25] New Zealand being a notable exception.

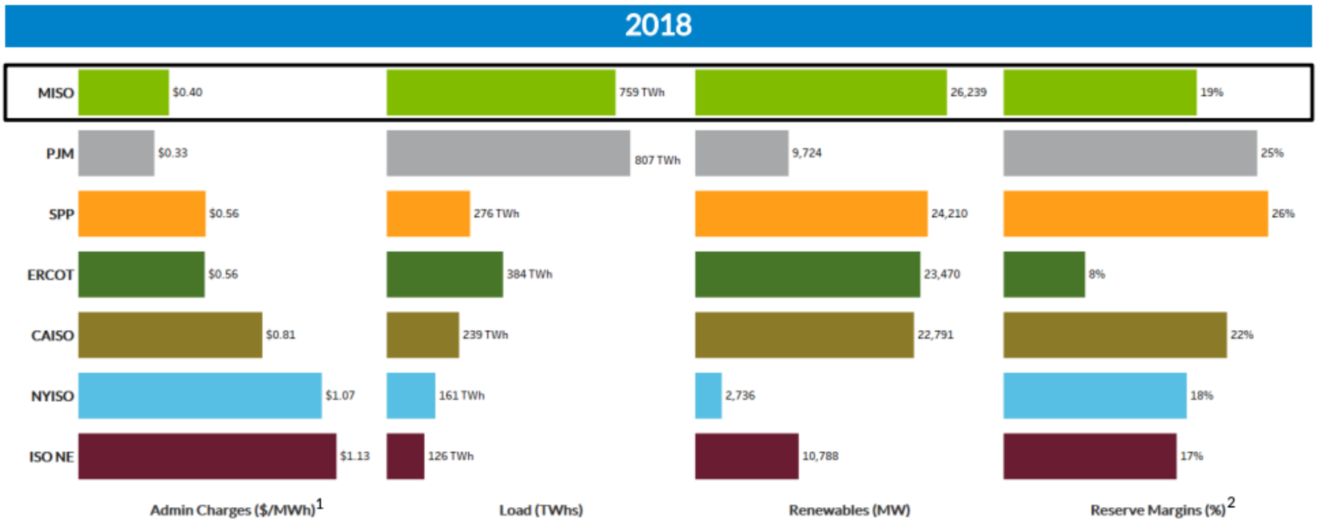
3.4 TREATMENT OF NTDC'S USE OF SYSTEM CHARGE WITHIN THE PROPOSED WHEELING FEE – THE INTERNATIONAL PERSPECTIVE

Currently, there is an inherent inconsistency with how NTDC is being treated within the regulatory/rate setting process. On the one hand, NTDC is being defined as having two core businesses – the wires business and system operation – that not only can be separated but definitely should be separated, if the eventual electricity market design is based on real time optimization through nodal pricing. However, NTDC is currently being regulated as if it is a single integrated transmission asset owner and system operator. This regulatory/rate setting paradigm is not consistent with the stated unbundling of NTDC into a wire business and a system operator.

This paradox directly affects the adopted cost allocation methodology for both the UoSC but also the portion of the UoSC that is recovered through the wheeling fee.

There are seven separate system operators in the *United States* – the Midcontinent ISO ("MISO"), the PJM Interconnection ("PJM"), the Southwest Power Pool ("SPP"), the Electricity Reliability Council of Texas ("ERCOT"), the California ISO ("CAISO"), the New York ISO ("NYISO") and ISO New England ("ISO-NE"). All but ERCOT are regulated similarly by the Federal Energy Regulatory Commission. ERCOT is regulated by the Public Utility Commission of Texas. The figure[26] below shows the average fee per megawatt hour (left hand column) for system operations for each RTO/ISO in 2018. This fee is determined via a standard revenue requirement exercise and is recovered through a charge to load based on energy usage and a demand charge.

Figure 2: Average Fee per Megawatt hour (left hand column) for System Operations for Each RTO/ISO (2018)



For purposes of comparison, **Table 2** compares the Transmission and Distribution Rates for the four regulated Transmission and Distribution wires businesses in Texas for the period September 1, 2018 to February 28, 2019. Thus, a small residential customer within the service territory of CenterPoint using 1250 kWh for a given month would have paid: [$1.62 + $3.85 + ($0.008439 * 1250kWh) + ($0.016489 * 1250kWH)] = $36.63 to the wires business. The charges due to ERCOT and the charges due to CenterPoint for the consumer would have been separate on the consumer's bill. The rates for the wires businesses will vary dramatically across the United States because of a number of factors.

[26] https://cdn.misoenergy.org/2020%20Operating%20and%20Capital%20Budgets406850.pdf accessed on 17 September 2020.

Table 2: Comparison of Transmission and Distribution Rates for four regulated Transmission and Distribution wires businesses in Texas (September 1, 2018 to February 28, 2019)

Generic Transmission and Distribution Rates

Class	CenterPoint - Docket 38339/41072			Oncor - Docket 46957		AEP	TCC	TNC	TNMP - Docket 38480	
	Charges			Charges		Charges	D-33309	D-33310	Charges	
Residential	Customer Charge (per customer)	$1.62		Customer Charge (per customer)	$0.89	Customer Charge (per customer)	$3.19	$2.94	Customer Charge (per customer)	$4.00
	Metering Charge (per customer)	$3.85		Metering Charge (per customer)	$2.60	Metering Charge (per customer)	$3.55	$5.24	Metering Charge (per customer)	$1.25
	Transmission Charge (per kWh)	$0.008439		Transmission Charge (per kWh)	$0.000000	Transmission Charge (per kWh)	$0.005190	$0.005803	Transmission Charge (per kWh)	$0.000000
	Distribution Charge (per kWh)	$0.016489		Distribution Charge (per kWh)	$0.021141	Distribution Charge (per kWh)	$0.013915	$0.019007	Distribution Charge (per kWh)	$0.017347
Secondary	Customer Charge (per customer)	$1.61		Customer Charge (per customer)	$2.03	Customer Charge (per customer)	$3.20	$4.25	Customer Charge (per customer)	$2.50
≤ 10 kW	Metering Charge (per customer)	$4.41		Metering Charge (per customer)	$6.19	Metering Charge (per customer)	$3.68	$7.50	Metering Charge (per customer)	$2.20
(or kVA)										
(TNMP 5 kW)	Transmission Charge (per kWh)	$0.004437		Transmission Charge (per kWh)	$0.000000	Transmission Charge (per kWh)	$0.002512	$0.003148	Transmission Charge (per kWh)	$0.000000
	Distribution Charge (per kWh)	$0.012218		Distribution Charge (per kWh)	$0.022625	Distribution Charge (per kWh)	$0.015489	$0.031948	Distribution Charge (per kWh)	$0.033323
Secondary	Customer Charge (per customer)			Customer Charge (per customer)	$9.18	Customer Charge (per customer)				
> 10 kW	Non-IDR customers	$2.26				Non-IDR customers	$3.26	$4.25	Customer Charge (per customer)	$2.56
(or kVA)	IDR Customers	$65.83		Metering Charge (per customer)	$31.35	IDR Customers	$26.52	$26.00		
(TNMP 5 kW)									Metering Charge (per customer)	$10.74
	Metering Charge (per customer)			Transmission Charge (per kW)	$0.00	Metering Charge (per customer)				
	Non-IDR customers	$18.82				Non-IDR Customers	$15.81	$18.68		
	IDR Customers	$63.07		Distribution Charge (per kW)		IDR Customers	$15.81	$35.00	Transmission Charge (per kW)	
				NCP ≤ 20 kW	$4.775600				Non-IDR customers	$0.0000
	Transmission Charge			NCP > 20 kW, Load Factor 0 - 10%	$6.664054	Transmission Charge			IDR Customers	$0.0000
	Non-IDR Customers (per NCP kVA)	$1.4318		NCP > 20 kW, Load Factor 11 - 15%	$5.901778	Non-IDR Customers (per NCP kW)	$1.286	$1.245		
	IDR Customers (per 4CP kVA)	$2.2387		NCP > 20 kW, Load Factor 16 - 20%	$5.550602	IDR Customers (per 4CP kW)	$1.793	$1.953	Distribution Charge (per NCP kW)	
				NCP > 20 kW, Load Factor 21 - 25%	$5.366679				Non-IDR Customers	$6.0981
	Distribution Charge (per kVA)	$3.059429		NCP > 20 kW, Load Factor ≥ 26%	$4.775600	Distribution Charge (per NCP kW)	$3.314	$3.21	IDR Customers	$5.2808
Primary	Customer Charge (per customer)		Primary	Customer Charge (per customer)	$6.17	Customer Charge (per customer)			Customer Charge (per customer)	$34.50
	Non-IDR customers	$3.58	≤ 10 kW	Metering Charge (per customer)	$18.91	Non-IDR customers	$3.80	$4.25		
	IDR Customers	$76.73		Transmission Charge (per kWh)	$0.000000	IDR Customers	$28.41	$26.00		
				Distribution Charge (per kWh)	$0.003629				Metering Charge (per customer)	$204.98
	Metering Charge (per customer)		Primary	Customer Charge (per customer)	$14.97	Metering Charge (per customer)				
	Non-IDR customers	$181.35	> 10 kW	Metering Charge (per customer)	$41.00	Non-IDR Customers	$154.62	$151.75		
	IDR customers	$138.40	Distrib.			IDR Customers	$154.62	$168.65	Transmission Charge (per kW)	
				Transmission Charge (per kW)	$0.00				Non-IDR customers	$0.0000
	Transmission Charge					Transmission Charge			IDR Customers	$0.0000
	Non-IDR Customers (per NCP kVA)	$1.7033		Distribution Charge (per kW)	$3.944984	Non-IDR Customers (per NCP kW)	$1.628	$1.189		
	IDR Customers (per 4CP kVA)	$2.1546	Primary	Customer Charge (per customer)	$150.37	IDR Customers (per 4CP kW)	$1.925	$1.963	Distribution Charge (per NCP kW)	
			> 10 kW	Metering Charge (per customer)	$255.07				Non-IDR Customers	$4.7102
	Distribution Charge (per kVA)	$2.002820	Substat.	Transmission Charge (per kW)	$0.00	Distribution Charge (per NCP kW)	$2.945	$1.88	IDR Customers	$5.1286
				Distribution Charge (per kW)	$0.53					
Transmission	Customer Charge (per customer)	$154.44		Customer Charge (per customer)	$161.34	Customer Charge (per customer)	$38.84	$24.80	Customer Charge (per customer)	$214.51
	Metering Charge (per customer)	$1,449.82		Metering Charge (per customer)	$263.30	Metering Charge (per customer)	$1,869.15	$850.00	Metering Charge (per customer)	$1,751.67
	Transmission Charge (per 4CP kVA)	$2.1188		Transmission Charge (per 4CP kW)	$0.00	Transmission Charge (per 4CP kW)	$1.718	$1.356	Transmission Charge (per 4CP kVA)	$0.0000
	Distribution Charge (per 4CP kVA)	$0.463296		Distribution Charge (per kW)	$0.2612	Distribution Charge (per NCP kW)	$0.199	$0.0182	Distribution Charge (per 4CP kVA)	$0.0000

As shown in the **Table 3**, the system operation fee in Alberta (Canada) for both 2019 and 2020 was roughly C$0.46 per megawatt hour.[27] This includes the charge for the Market Surveillance Administrator (MSA) and the Alberta Utilities Commission (AUC) for certain mandate activities.

Table 3: System Operation Fee in Alberta (Canada) for 2019 and 2020

AESO Energy Market Trading Charge	2020 Budget	2019 Budget
AESO Component		
General and administrative costs, interest, amortization and other industry costs	29.8¢	34.7¢
Energy market deferral deficit (surplus)	6.6¢	3.0¢
Total	36.4¢	37.7¢
AUC Component	6.2¢	4.8¢
MSA Component	4.0¢	3.4¢
AESO Energy Market Trading Charge	46.6¢	46.0¢

27 https://www.aeso.ca/assets/Uploads/2020-Energy-Market-Trading-Charge.pdf accessed 17 September 2020.

With a mean monthly load of 600kWh the average monthly fee for the use of the transmission and distribution networks for customers in *Alberta* ranged from C$0.035 to C$0.046 and C$0.037 to C$0.143 respectively.[28]

Finally, in *Singapore* the fee for system operation in 2020 is S$0.0046 per kWh, while the fee for the wires businesses is S$0.0544 per kWh.[29]

In contrast to the three previous examples, which were all examples where system operation and the wires business were completely separate, *Europe provides examples where these activities are performed within the same company.*

Amprion GmbH defines their business activities very precisely in their 2019 Annual report:

> "*Business activities of the company Amprion GmbH*, headquartered in Dortmund, is one of four transmission system operators (TSOs) in Germany. In a control area that stretches from Lower Saxony to the Alps, Amprion operates its network at voltage levels of 220 and 380 kilovolts (kV) and is expanding it in accordance with market requirements. The extra-high-voltage grid links the generation units to the main centres of consumption and is a vital component of the transmission network in both Germany and Europe. *Amprion uses its grid to serve customers from industry, distribution system operators, electricity traders and power utilities.*
>
> In addition, Amprion controls and monitors the safe transport of electricity within the EHV grid in its control area. For this purpose, the grid operations managers in Brauweiler/Pulheim ensure that electricity consumption and generation are kept in balance at all times. The system services required (primary control, secondary control and tertiary control (minute reserve)) and the electricity necessary to compensate grid losses are sourced using transparent tender procedures in line with regulations. The company also coordinates the exchange programmes and the subsequent volume balancing, both for the entire transmission network in Germany and for the northern section of the integrated European grid.[30] (emphasis added)"

Thus, Amprion is very similar to the way NTDC currently operates in that the company both owns the transmission assets as well as operates the system. **Table-4** below provides the regulated rates/charges for the users of Amprion's system:[31]

[28] https://ucahelps.alberta.ca/electricity-transmission-and-distribution-charges.aspx#:~:text=In%202019%2C%20monthly%20distribution%20charges,of%20a%20customer's%20total%20bill. Accessed 17 September 2020.

[29] https://www.spgroup.com.sg/what-we-do/billing accessed on 17 September 2020.

[30] https://www.amprion.net/Dokumente/Amprion/Gesch%C3%A4ftsberichte/2019/amprion_annual-report-2019_financial-report.pdf accessed 16 September 2020.

[31] https://www.amprion.net/Dokumente/Strommarkt/Netzkunden/Netzentgelte/Entgelte/Entgelte-Amprion-g%C3%BCltig-ab-01-01-2019-englische-Version.pdf accessed 17 September 2020.

Table 4: Regulated rates/charges for the users of Amprion's system

Amprion-Charges for grid usage since 01.01.2019

1) Annual demand rate system

	Annual utilization time			
	< 2500 h/a		≥ 2500 h/a	
	Demand rate	Energy rate	Demand rate	Energy rate
	€/kWa	ct/kWh	€/kWa	ct/kWh
Tapping point in extra high voltage system	**8.03**	**1.89**	**45.43**	**0.40**
Amprion-individual grid charge part	6.03	1.41	33.86	0.30
Unified grid charge part	2.00	0.48	11.57	0.10
Tapping point in extra high voltage system incl. transformation	**13.25**	**2.09**	**56.25**	**0.37**
Amprion-individual grid charge part	10.16	1.54	40.98	0.31
Unified grid charge part	3.09	0.55	15.27	0.06

Prices plus extra costs *), **)

The data in the above table reveals that cost recovery is based on a combined demand charge and an energy charge and that on a "per unit" basis the rate for higher usage customers is biased more towards the demand rate.

The following table provides the average tariff rate across the European Network of Transmission System Operators (ENTSO) for 2016.[32] The UTT is the Uniform Transmission Tariff and in 2016 the average tariff rate for transmission and system operation across most of Europe was €0.01188 per kWh.

	2016	Δ 2016/2015
Average European UTT	**11.88 €/MWh**	+ 11.91 %
• Due to TSO Costs	8.32 €/MWh	+ 3.70 %
• Due to Non-TSO Costs	3.56 €/MWh	+ 31.14%

3.5 DETERMINATION OF NTDC UOSC AND RECOMMENDATIONS

With respect to the portion of the wheeling fee related to UoSC, Pakistan has two possibilities:

- Unbundle the two businesses operated by NTDC, i.e., the wires business and system operation, and perform a proper cost-of-service regulatory exercise. This would entail separating the assets and costs of each business, determining the revenue requirement and then assigning the costs to the users of those particular businesses.

32 https://eepublicdownloads.azureedge.net/clean-documents/mc-documents/ENTSO-E_Transmission%20Tariffs%20Overview_Synthesis2016_UPDATED_Final.pdf accessed 16 September 2020.

- Alternatively, NEPRA can continue to treat NTDC as a bundled business and determine the revenue requirement for this aggregated entity and assign the associated costs in an appropriate manner, i.e., on a usage, peak and fixed cost basis.

Eventually, NEPRA should move to a true cost-of-service regulatory regime regardless of what electricity market design they choose. This would allow NTDC to perform the two services internally even if they are functionally[33] *unbundled.*

The services provided by NTDC, including the provision of an electricity transmission network and the reliable despatch of electricity, are both necessary for wheeling to take place. The only exception is the case where the power flows, associated with a specific wheeling transaction, never use the high voltage network. Even in that case, there is a strong argument that the specific transaction is beneficiary of a reliable system with access to ancillary services. Given that, and under the current regulatory rate setting paradigm, we agree with the recommendations arrived at by the participants at the LUMs conference. Specifically, we recommend:

- The UoSC should be treated uniformly across all power that flows in the system.
- When power is wheeled across two or more DISCOs, the NTDC system is utilized and, therefore, the UoSC should be a part of the wheeling charge.
- When the wheeler of power and the associated Bulk Power Consumers (BPC) are located in the same DISCO system, the BPC is still utilizing the NTDC system for system stability and reliability. Furthermore, even though the actual flow of power to the BPC is not using the transmission network, the network is providing reliability services that are different from system operation.
- Additionally, when calculating the demand charge, we recommend the use of a coincidental peak rather than a non-coincidental peak.[34]

Currently NTDC provides a bundled service that includes: (1) system operation; (2) access to the high voltage transmission system; and (3) system planning. These are three very different businesses from the perspective of CAPEX and OPEX. The wires business is a capital-intensive business, while the latter two are relatively low on capital requirements. Looking at the associated costs for these three activities, the salient point is that regardless of the cost allocation mechanism, all of the costs must be recovered. Given this fact, the specific cost allocation mechanism is of secondary importance. However, within that context, it is important for the individual cost allocation mechanisms for each of the three activities to provide the best economic incentives. There are three types of cost recovery mechanisms: (1) fixed fees that are independent of either usage or peak demand; (2) fees based on usage; and (3) fees based on peak demand. Given the current way that NTDC is being regulated, i.e., as an integrated firm, we recommend a mix of the latter mechanisms. This is because we lack the granularity to understand which, if any, of the costs are unrelated to usage.

Currently NEPRA allocates the entire revenue requirement for NTDC[35] through a demand charge.[36] Insofar as this cost allocation mechanism puts all of NTDC's costs on peak periods, it provides a

---

[33] While it is common to legally/commercially separate the wires business from system operation to avoid any possibility of "self-dealing", this is not a strict requirement. For example, Transpower New Zealand is both the high voltage wires business and the system operator, albeit with functional separation between the two.

[34] Coincidental Peak is a better indicator of when the system is being used to its maximum capacity. The Coincidental Peak is the total system Peak occurring at one time. It is not the arithmetic sum of different peaks occurring at different times.

[35] The costs of NTDC are recovered through a demand charge while the costs associated with losses are recovered through a variable usage charge.

[36] National Electric Power Regulatory Authority, "Decision of the Authority in the Matter of Motion for Leave for Review Filed by National Transmission and Dispatch Company Limited With Respect to the Determination of the Authority Dated July 31, 2019 - Case NO. NEPRA/TRF-450/NTDC-2018. March 31, 2020. p. 10.

focused economic signal regarding the effect of peak usage. However, both a demand charge based on either the non-coincident peak or the coincident peak does not result in efficient economic signals:

> But traditional demand charges are not the answer. These are typically based on the customer's peak usage period, even if it does not coincide with the system's peak. Many argue this "non-coincident demand charge" sends the wrong price signal to reduce system costs. Instead, policymakers have proposed demand charges for the customer's peak usage during the system's peak. But, in economic terms, this "coincident demand charge" still fails to match cost causation as well as dynamic pricing. If the customer's peak use is not precisely at the system peak, the coincident demand charge fails to address the actual level of system stress caused by the customer's usage.[37]

From a purely economic perspective, the ultimate goal is to introduce dynamic pricing, i.e., pricing electricity according to the actual supply and demand conditions on the grid during a specified interval. This is another example of where the choice in regard to the ultimate market design is important. A market design based on nodal pricing will include the effect of scarce transmission capacity in the nodal price.[38] In this market design the nodal price increases when transmission capacity is constrained. As a result, the regulation of the wires businesses under this market model is very straightforward. There is no need to attempt to mirror the scarcity of the transmission system in regulated rates. Under the "Transco" model the issue of signaling when the transmission system is under stress exists. This is why the transmission rates in ERCOT look so different from those of Amprion. The former is a nodal pricing market regime while the latter is not.

Given the issues mentioned above regarding a demand charge, we recommend that NEPRA move to a cost recovery mechanism for NTDC that is based on a combination of a demand charge and an energy charge.

In the **Table 5**, which is based on information contained in the recent Determination for NTDC (March 31, 2020, see footnote 16), we provide the recommended cost recovery for the UoSC. The methodology assumes that 90% of NTDC costs are assignable to the wires business and that there is out of merit dispatch, i.e., transmission constraints, during 75% of the operating intervals. This allocation methodology should be applied to all loads.[39]

Table 5: Appropriate the Cost recovery for the NTDC UoSC Equally Between the Demand and Usage Charges

Year	Revenue Requirement (in milions of Rs)	Amount recovered from demand charge	Amount recovered from energy charge	Demand Charge Rs/kW/M	Energy Charge Rs/kWh	User Inputs	
2018-19	Rs 50,087.00	Rs 33,808.73	Rs 16,278.28	Rs 118.91	Rs 0.13	Percent allocated to peak (demand charge)	67.5%
						Percent allocated to usage (energy charge)	32.5%
						Share of Costs for Wires Business vs System Operation and Planning	90.0%
						Percentage of Hours of Scarce Transmission Capacity	75.0%
						Average MDI in megawatts	23694
						Total electricty production in GWh	121000

---

37 https://www.utilitydive.com/news/could-rate-design-help-californias-struggle-with-flat-demand/513234/#:~:text=Another%20option%20are%20demand%20charges.&text=But%20traditional%20demand%20charges%20are,signal%20to%20reduce%20system%20costs. accessed September 20, 2020.

38 The nodal price on the constrained side of scarce transmission capacity will increase.

39 The values for the revenue requirement as well as the Average MDI (Maximum Demand Indicator) are taken from the aforementioned Determination for NTDC (see fn 2) on page 76. The value for the Total Electricity Production for 2018-19 was taken from NEPRA's 2019 State of Industry Report (page 20). See: https://nepra.org.pk/publications/State%20of%20Industry%20Reports.php

4. TREATMENT OF HYBRID BPC'S IN THE WHEELING FEE

4.1 BRIEF SUMMARY

The basic underlying contractual structure of a so-called "Hybrid BPC" is that the quantity supplied under the contract is intentionally less than the total quantity consumed for specific operating intervals that may or may not be regularly occurring. For example, a contract that relies on solar generation will obviously not be able to provide power during non-daylight operating intervals. Similarly, contracts relying on wind generation, will not be able to supply power when there is no wind. More generally we note that the technical characteristics of the underlying generating technologies are not the only limiting characteristics. Financial considerations, reliability, input fuel availability, load capacity factors, etc., may all provide significant reasons why both counterparties, i.e., load and generation may want to contract for level below the total expected consumption.

This "hybrid" relationship necessarily means the load, by not contracting for the full amount of their expected demand, must purchase an "option" from the rest of the market for every operating interval. While this type of contracting may be in its infancy in Pakistan, it has become the norm in other parts of the world. This is particularly true in regions where there is a spot market for electricity. This contracting paradigm gives the hybrid BPC the right, i.e., the option, but not the obligation, to purchase the energy they need at a price determined by NEPRA.

The structure of the Hybrid BPC contract represents a serious threat to the current operating procedures of the Pakistani electricity sector. The fundamental premise of the current structure is that load and generation will be balanced prior to real time. While not explicitly stated, this premise influences both the reliability standards (as presented in the Grid Code) and system dispatch. In contrast the supply and demand in the Hybrid BPC relationship is *unbalanced by design*. Put differently, the Hybrid BPC relies on acquiring electricity from a real time spot market. This is problematic for the current situation in Pakistan because the system is not based upon a spot market. Consider the point that the system operator procures reserves for reserves and contingencies and not for load that does not have a contract that covers their entire expected load.

Our general recommendations are:

- The existence of Hybrid BPC contracts necessarily means that the system operator (or some entity) must have excess electricity available in real time – through some mechanism – to provide for the shortfall between the Hybrid BPC contracted for amount and the actual amount consumed.
- Wheeling contracts should be required to schedule their anticipated generation and load with the system operator.
- The wheeling fee for the Hybrid BPC's should be based on a charge for operating reserves and a separate energy charge when actual energy consumption exceeds the contract amount.

4.2 INTRODUCTION

At the Conference held at the LUMs Energy Institute on December 9-11, 2019, the participants recognized and discussed the issues created by a "Hybrid BPC" with respect to wheeling and, in particular, the wheeling fee. The summary of their discussion is as follows:

Issue 6: Hybrid BPC

6.1 It was argued that hybrid BPCs are allowed in various global markets where BPCs wheel part of their load while the rest is procured from the grid at regulated rates. However, some participants stated that it is against the global practice in wholesale markets.

6.2 The current volumetric tariff structure does not adequately address the impact of capacity and leaves room for the BPC to contract baseload requirement from a Wheeler of power and take services of variable load from the DISCOs. This creates a room for arbitrage and needs to be addressed. Furthermore, such BPC's avoid their fair share of capacity costs which results in putting additional burden on remaining regulated consumers.

6.3 Participants agreed that any withdrawal of energy from the grid system by the BPC, other than that provided by its Wheeler of power, shall be settled at the marginal price determined by the Market Operator. It would otherwise be impractical for the Regulator to determine tariffs on case to case basis.

6.4 Participants highlighted that captive power plants also have the same effect on the grid and they should be treated in the same manner. Their tariff structure shall be reviewed and modified to the regime that existed before 2001 (charging of fixed cost on MDI /sanctioned load). The Regulator shall, therefore, provide a level playing field with fair price signals for Wheelers of power and captive generation. However, it was argued that if a consumer in his application had already decided to use the grid for his net demand, exclusive of captive generation, then it is not fair to charge that consumer additional system costs.

4.3 ISSUE AND TREATMENT OF HYBRID BULK POWER CONSUMER'S – PAKISTAN VIS-À-VIS INTERNATIONAL CONTEXT

The basic underlying contractual structure of a so-called "Hybrid BPC" is that the quantity supplied under the contract is intentionally less than the total quantity consumed for specific operating intervals that may or may not be regularly occurring. For example, a contract that relies on solar generation will obviously not be able to provide power during non-daylight operating intervals. Similarly, contracts relying on wind generation, will not be able to supply power when there is no wind. More generally we note that the technical characteristics of the underlying generating technologies are not the only limiting characteristics. Financial considerations, reliability, input fuel availability, load capacity factors, etc., may all provide significant reasons why both counterparties, i.e., load and generation may want to contract for level below the total expected consumption.

This "hybrid" relationship necessarily means the load, by not contracting for the full amount of their expected demand, must purchase an "option" from the rest of the market for every operating interval. While this type of contracting may be in its infancy in Pakistan, it has become the norm in other parts of the world. This is particularly true in regions where there is a spot market for electricity. This contracting paradigm gives the hybrid BPC the right, i.e., the option, but not the obligation, to purchase the energy they need at a price determined by NEPRA.

An alternative way to characterize the relationship between the hybrid BPC and all other non-hybrid BPC consumers, i.e., the rest of the market, is that the difference between the contracted amount and the actual usage of a hybrid BPC represents a "spot" transaction. For example, suppose the hybrid BPC has a contract for 2MWh but during a specific interval actually consumes, i.e., purchases, 2.1MWh. The additional 0.1MWh above the contracted amount constitutes a spot purchase of physical energy. The fundamental and immediate problem this raise is that Pakistan does not currently have a physical spot electricity market. Any solution to the issues raised by hybrid BPC's will necessarily involve creating a pseudo spot market of some kind.

For our purposes, the hybrid BPC makes use of two explicit aspects of either an implied or explicit spot market – operating reserves and energy. The system operator must make sure they have enough "operating reserves" (in Pakistan - the reserves necessary to cover primarily regulation and contingencies) to provide the hybrid BPC with the additional power they may need. In this way, the Hybrid BPC is no different than any other load on the system and they should pay for their fair share of the costs of operating reserves. That is, the consumer side of the Hybrid BPC is identical to all other load insofar as it may require that available generation capacity be transformed into energy at any instant in time, i.e., the load that is not the subject of a bilateral contract. In order for them to be able to do this, the system must first have the necessary capacity that can be transformed into energy to fill their "short", i.e., the system must have the capacity that is required to meet the load of the Hybrid BPC customer. Should the Hybrid BPC actually take power above and beyond what they have contracted for, they should, in addition to the appropriate capacity charge, pay the energy cost of that energy procured.[40] Thus, with respect to the wheeling fee, there are two separate components – the cost of reserves, i.e., capacity, and the cost of energy if it is needed. The former charge is determinable from the cost of the reserves that were procured. The latter is, in effect, the spot price of electricity and under the current structure in Pakistan must be determined by NEPRA. The most economically efficient "spot" price would be one that varies by time and location, i.e., the nodal price. However, it is not clear whether NTDC currently has the tools required to calculate nodal prices. In the absence of the nodal price, we recommend that NEPRA use the system marginal cost, also called the system lambda, as the surrogate for the nodal price for energy purchases.

There is a further issue that arises from hybrid BPC's regarding the operation of the electricity system. Because these consumers have a contract for some, but perhaps not all, of their consumption, they can be expected to "lean" on the regulated market during some operating intervals. In order to operate the system reliably and efficiently, the System Operator needs to know how much power the BPC will be receiving under the contract. If the System Operator is "blind" to the contracted amount, they will necessarily over procure reserves. Therefore, one of the parties to the hybrid BPC contract should be required to provide the expected amount of power that will be provided by the System Operator. This will allow the System Operator to procure the correct amount of reserves for the operating intervals.

Thus, our general recommendations are:

- Wheeling contracts should be required to schedule their anticipated generation and load with the system operator.
- The wheeling fee for the Hybrid BPC's should be based on a charge for operating reserves and a separate energy charge when actual energy consumption exceeds the contract amount.

The model for the treatment of the wheeling fee for Hybrid BPC contracts is consistent with international experience in every jurisdiction that has a physical spot market.

- In the electricity market administered by ERCOT in Texas, the average cost of operating reserves in 2018 was US$0.00197 per kWh.[41]
- In Alberta, Canada in the electricity market administered by the Alberta Electric System Operator, the cost of reserves in 2019 was CAN$0.0243 per kWh.[42]
- In Australia, the Benchmark Reserve Capacity Price for 2019 was AUS$0.0144.[43]

We note that operating requirements vary significantly by region/market as do the resulting costs.

4.4 CONCLUSION AND RECOMMENDATIONS

The structure of the Hybrid BPC contract represents a serious threat to the current operating procedures of the Pakistani electricity sector. The fundamental premise of the current structure is that load and generation will be balanced prior to real time. While not explicitly stated, this premise influences both the reliability standards (as presented in the Grid Code) and system dispatch. In

[40] We note that perhaps the most economically efficient mechanism would be to divide the cost of the operating charge into two components – one based on peak demand and another based on general usage. The former component would reflect the need to procure more operating reserves during peak periods while the latter would reflect the general "reliability" service provided by operating reserves. However, given the current state of dispatch, ancillary service procurement, and pricing in Pakistan, we doubt the added granularity would affect decision-making by the load.

[41] https://www.potomaceconomics.com/wp-content/uploads/2019/06/2018-State-of-the-Market-Report.pdf p. 5 (Executive Summary) accessed October 14, 2020

[42] AESO, "AESO 2019 Annual Market Statistics", p. 28.

[43] https://aemo.com.au/en/energy-systems/electricity/wholesale-electricity-market-wem/wa-reserve-capacity-mechanism/benchmark-reserve-capacity-price accessed October 14, 2020.

contrast the supply and demand in the Hybrid BPC relationship is *unbalanced by design*. Put differently, the Hybrid BPC relies on acquiring electricity from a real time spot market. This is problematic for the current situation in Pakistan because the system is not based upon a spot market. Consider the point that the system operator procures reserves for reserves and contingencies and not for load that does not have a contract that covers their entire expected load.

As a result, our specific recommendations for the wheeling fee are as follows:

a. The existence of Hybrid BPC contracts necessarily means that the system operator (or some entity) must have excess electricity available in real time – through some mechanism – to provide for the shortfall between the Hybrid BPC contracted for amount and the actual amount consumed.

b. Wheeling contracts should be required to schedule their anticipated generation and load with the system operator.

c. A "usage" fee based on the cost of operating reserves necessary to reliably and efficiently operate the system. The usage fee will be determined by the cost of operating reserves divided by the total load:

$$Operating\ Reserve\ Fee = \frac{Total\ annual\ cost\ of\ "operating\ reserves"\ for\ all\ DISCOs}{\text{Aggregated}\ monthly\ peak\ demand\ for\ all\ \ DISCOs\ in\ kW}$$

$$= per\ kW\ per\ month$$

This per kW monthly fee is then aplied equally to all type of loads including hybrid in each DISCO.

d. A cost for the consumption of energy that is consumed by the hybrid BPC customer that is not covered by the hybrid BPC contract.

$$Energy\ Charge = (Total\ Energy\ Consumed\ per\ Operating\ Interval\ \ - Qty\ Covered\ By\ Contract) * \text{Marginal}\ cost\ per\ kWh$$

This cost applies to all purchases made by the Hybrid BPC customers beyond their contracted amounts of energy.

5. TREATMENT OF BANKED ENERGY IN THE WHEELING FEE

5.1 BRIEF SUMMARY

The concept of "banked energy" has been primarily associated with renewable generation due to the so-called "must run" characteristics of intermittent resources. That is, intermittent generation like wind and solar are dependent on variable fuel sources and not specifically price. As a result, there will be times when there is production from these units that is not consumed by the owner or the contract counterparty.[44]

The focus of this short paper is on "banked energy" as it applies to wheeling in Pakistan. While the concept is similar to what has been applied to intermittent resources the context is different. The primary way in which wheeling transactions are operationalized on a day-to-day basis is through the use of self-scheduling. In essence, what this means is that the generator produces whatever level of output they need to produce in order to satisfy the wheeling contract.

Accordingly, there may be specific operating intervals in which the generation output of the wheeling transaction either exceeds or is less than the amount of electricity consumed by load side of the wheeling transaction. When this happens the excess energy, i.e., the electricity that is produced but not consumed by the wheeling parties, is purchased by the DISCOs at rates that are set by NEPRA.

Banking should be allowed only with appropriate attention paid to (1) the price received for the banked energy as well as (2) the length of time that the energy can be banked. With respect to banked energy, our specific recommendations are as follows:

- Review the rules pertaining to self-scheduling to ensure that there are no artificial incentives that encourage specific types of transaction that rely on self-scheduling.
- Monitor self-scheduling so that reliability is not compromised by having an excess amount of self-scheduling in constrained regions of the grid.
- Banked energy should be paid on the basis of system marginal cost for the specific operating interval during which there was excess production from self-scheduled resources.
- Banked energy credits must be used within twelve months of when they were accrued.
- Monitor when and how the banked energy credits are used.

5.2 BACKGROUND

At the Conference held at the LUMs Energy Institute on December 9-11, 2019, the participants recognized and discussed the issues created by "banked energy" with respect to wheeling and, in particular, the wheeling fee. The summary of their discussion is as follows:

[44] See for example, https://mercomindia.com/mnre-directs-rollover-banked-energy-lockdown/ which explains that the Ministry of New and Renewable Resources (MNRE) in India has recommended that three states allow a rollover of banked energy, "The Ministry observed that due to the dip in demand for power, the generated and banked units in previous months could not be utilized by the consumers. According to the Ministry, the lapse of such banked units or purchase at the average pooled purchase cost (as is typically done) would severely affect the profitability of both the developers and consumers of solar PV rooftop projects and open access renewable energy generating stations."

Issue 9: Banked Energy

9.1 Participants noted that injections and actual withdrawals of energy cannot be equal at all times and results in imbalances which need to be settled.

9.2 Participants agreed that marginal pricing mechanism is more efficient in settling imbalances. Participants were apprised that the System Operator has started committing the power plants through state-of-the-art unit commitment tool i.e. NCP on trial basis. Moreover, the System Operator is sharing this information with the Regulator on daily basis. The Market Operator is also in the process of implementing IT systems to automate the settlement process based on marginal pricing, which will be implemented in near future.

9.3 However, it was argued that till the implementation of marginal pricing mechanism, the existing practice of banking, at rates approved by the Regulator, may continue. The DISCOs concerns may be addressed by placing suitable caps on maximum units and period for banked energy. It was also suggested that a band or range of acceptable deviations from contracted amount may be set and restrict the frequency and magnitude of such deviations within a month. Withdrawals within this band should be balanced out by requiring similar injections during the month without any additional costs. Beyond these bands, reasonable charges can be applied on the BPCs

5.3 ISSUE OF BANKED ENERGY – PAKISTAN VIS-À-VIS INTERNATIONAL CONTEXT

The concept of "banked energy" has been primarily associated with renewable generation due to the so-called "must run" characteristics of intermittent resources. That is, intermittent generation like wind and solar are dependent on variable fuel sources and not specifically price. As a result, there will be times when there is production from these units that is not consumed by the owner or the contract counterparty.[45]

In these cases, the energy is "banked" in that the excess energy is bought by the overall system and the parties to the transaction are credited with this sale. This revenue can then be used at a later date to offset future charges.

The primary way in which wheeling transactions are operationalized on a day-to-day basis is through the use of self-scheduling. In essence, what this means is that the generator produces whatever level of output they need to produce in order to satisfy the wheeling contract.

Accordingly, there may be specific operating intervals in which the generation output of the wheeling transaction either exceeds or is less than the amount of electricity consumed by load side of the wheeling transaction. When this happens the excess energy, i.e., the electricity that is produced but not consumed by the wheeling parties, is purchased by the DISCOs at rates that are set by NEPRA.

Given the quantity of power associated with wheeling transactions in Pakistan, the amount of these purchases, i.e., the amount of banked energy, should be insignificant.

That being said, the real underlying issue associated with banked energy is not the financial significance but rather the concept of "self-scheduling". In an earlier paper on economic dispatch, we noted that there are a few unique and significant characteristics of electricity as a commodity. One is the need

[45] See for example, https://mercomindia.com/mnre-directs-rollover-banked-energy-lockdown/ which explains that the Ministry of New and Renewable Resources (MNRE) in India has recommended that three states allow a rollover of banked energy, "The Ministry observed that due to the dip in demand for power, the generated and banked units in previous months could not be utilized by the consumers. According to the Ministry, the lapse of such banked units or purchase at the average pooled purchase cost (as is typically done) would severely affect the profitability of both the developers and consumers of solar PV rooftop projects and open access renewable energy generating stations."

to keep physical supply and demand in tight balance which necessitate the need for an "air traffic controller" or dispatcher. Another such characteristic is that electricity relies on network production

Network production means not only that the actions of all connected parties have to be coordinated (i.e. through the dispatch process) but also that the actions of every party potentially has an effect on the potential actions of every other connected entity, i.e., potential externalities arise from every decision in regard to connected assets.

Self-scheduling, i.e., the act of producing electricity regardless of what is happening on the grid has the potential for any specific operating interval to both increase costs to other users of the grid as well as threaten reliable operation of the grid.

In 2005, when the Midcontinent Independent System Operator (MISO) began using Security Constrained Economic Dispatch to balance the system, wind generation was such a small percentage of overall generation that the rules allowed wind to autonomously "self-schedule." That is, wind was allowed to run regardless of what was happening on the rest of grid or with other generators. Wind generation was fully and completely accommodated by other "dispatchable" generation. Thus, other generation was ramped up/down based on what the wind generation output was.

As should have been expected, this type of self-scheduling created a partial incentive for both an increase in wind generation but also an increase in the specific locations of new wind generation.[46] Over time with more wind generation, located in specific areas, the ability of Security Constrained Economic Dispatch to solve the "dispatch problem", i.e., the instantaneous balancing of physical supply and demand became compromised. Since a self-scheduled generator, by definition, does not respond to economic incentives, the higher the amount of self-scheduled generation the more difficult it is for the dispatcher to change the output levels of non-self-scheduled generation to maintain reliable operation.

The problem escalated so quickly that within 10 years of implementing economic dispatch the MISO changed the rules such that every new wind generator had to be capable of being ramped up/down, i.e., be dispatchable. Wind generation, like any other generator under the MISO rules, is still allowed to self-schedule but new wind turbines must be capable of ramping up or down.

Depending upon how Pakistan defines the rules pertaining to self-scheduling, a similar situation could arise. The self-scheduling of wheeling transactions that are held harmless for the costs may provide an artificial incentive for parties to enter into these types of transactions.

Other than the potential incentive effects arising from self-scheduling, banking should be allowed with appropriate attention paid to (1) the price received for the banked energy as well as (2) the length of time that the energy can be banked.

Regarding the price received under Security Constrained Economic Dispatch, it would be the nodal price and would vary by both the time and location. Absent nodal pricing, it should be the system marginal price, or system lambda, which can be calculated by NTDC/NPCC. The system lambda is the cost of serving the next increment of load, i.e., the marginal cost of electricity.

Regarding the length of time that energy can be banked, we believe that it should be for no more than a rolling twelve-month period. Furthermore, if the amount of banked energy increases significantly, we suggest re-visiting this recommendation and possibly shortening the allowed period for holding banked energy.

Pakistan wants to avoid situations whereby parties can acquire a large banked energy "position" that can then be used in significant quantities during certain periods and may, accordingly, negatively affect

[46] In 2005, wind capacity in the MISO was 908 MW. By 2010 wind capacity had grown to 8,179MW and by 2019 it was 20,452. See: https://cdn.misoenergy.org/2020%20Wind%20&%20Solar%20Capacity%20Credit%20Report408144.pdf, p. 9.

dispatch and reliability in the future. The expectation should be that banked energy is held in small amounts and for relatively short periods of time.

5.4 CONCLUSION AND RECOMMENDATIONS

Our specific recommendations are as follows:

- Review the rules pertaining to self-scheduling to ensure that there are no artificial incentives that encourage specific types of transaction that rely on self-scheduling.
- Monitor self-scheduling so that reliability is not compromised by having an excess amount of self-scheduling in constrained regions of the grid.
- Banked energy should be paid on the basis of system marginal cost for the specific operating interval during which there was excess production from self-scheduled resources.
- Banked energy credits must be used within twelve months of when they were accrued.
- Monitor when and how the banked energy credits are used.

6. ECONOMIC DESPATCH AND THE WHEELING OF ELECTRICITY

6.1 BRIEF SUMMARY

Electricity is a *real time* commodity. That is, within very strict bounds, the instantaneous physical supply of electricity must equal the instantaneous physical demand for electricity at every moment in time. Until such time as large-scale storage is commercially viable, electricity consumed at every instant in time must have been generated at the same time, i.e., electricity cannot be produced yesterday for consumption today at any viable scale. The simultaneity of supply and demand in electricity is one of a few characteristics that makes electricity – as a commodity – truly unique and different from all other commodities.

Electricity markets that are based on the actual real time capacity of the grid, i.e., what is actually physically feasible at every instant in time, do not necessarily anticipate that any *ex-ante* market will solve the dispatch problem, i.e., the instantaneous balancing of supply and demand. As a result, these markets rely almost exclusively on a physical spot market. The underlying tool of the physical spot market is a constrained optimization algorithm called Security Constrained Economic Dispatch (SCED). The SCED uses current information about the grid to reliably and efficiently allocate (scarce) transmission capacity as quickly as every 5 minutes. In these markets, the Day Ahead market is just another, albeit near term forward market and the results have no special importance in terms of reliably and efficiently matching supply and demand in real time. Accordingly, the focus of this market structure is on the dispatch process and the forward or bilateral electricity market. In theory, the dispatch prices and the associated nodal prices are expected to provide the foundation for a robust financial futures market and bilateral contract market.

The physical characteristics of electricity necessarily mean that the capacity of the transmission system cannot be perfectly defined and "rationed" prior to real time. Moreover, the capacity of the system is not simply a function of the physical infrastructure for generation, and transmission facilities and load, but also a result of the decisions made by the dispatcher in matching supply and demand. This is not an opinion rather it is a fact derived from the physics of electricity and it has implications for the design and operation of any electricity market.

As a result, in real time – when electricity generation and load are being balanced through the dispatch/balancing function – the "commodity' is being rationed/allocated is transmission capacity. This fact is true regardless of the choice of market design or institutional structure. Whenever the transmission grid is constrained, there is not enough transmission capacity to meet the *ex-ante* plans of the participants. Thus, given current technology, the "commodity" being rationed in real time *is not electricity*, but rather the available capacity on the transmission network. The primary question to be addressed is the process through which this scarce resource will be allocated, i.e., who will and who will not receive transmission capacity and at what price. Any deliberation regarding the dispatch function is simultaneously a discussion on how to implement open access as well as a discussion about what electricity market design should be chosen. The three issues – dispatch, open access, and electricity market design – are fundamentally interdependent.

We strongly recommend using Security Constrained Economic Dispatch as the basis for the electricity market in Pakistan for the following primary reasons:

- Nodal pricing reduces the necessary discretion of the system operator.
- Nodal pricing-based markets do not necessarily rely on false assumptions about the state of the transmission network in real time.
- Nodal pricing markets better allow for naturally occurring risk to be more efficiently managed through financial instruments rather than physical capital investment.

- Nodal pricing markets provide location and time-of-use based price signals for generation and load.
- Nodal pricing markets provide an explicit signal for the cost of congestion.
- Nodal pricing markets are economically more efficient and more reliable.
- Nodal pricing markets are far better suited to accommodate intermittent resources.
- Nodal pricing markets are less likely to be manipulated or subject to market power.

6.2 BACKGROUND

At the conference held at the LUMs Energy Institute on December 9-11, 2019 the attendees noted the following with respect to the "issue" of economic dispatch:

Issue 8: Economic Dispatch

8.1 It was discussed that Economic Dispatch may allow benefits to both the Wheeler of power and the system to reduce costs at both ends.

8.2 If the Wheeler of power is dispatched when the marginal price is higher than its own variable cost, and its BPC does not draw power from the system, the Wheeler of power recovers additional revenue by getting payment according to the marginal price. The system benefits by avoiding the dispatch of an expensive generator.

8.3 If the Wheeler of power is not dispatched when the marginal price is lower than its own variable cost, it will recover the contracted revenue from its BPC. And, it will procure from the market at a cost below its own variable cost to serve the demand of its BPC which will result in additional revenue.

8.4 It was argued that settlement of imbalances through marginal pricing has tax implications such that there are different tax implications for electricity generation and sale, and electricity purchase and sale (trading). The Regulator shall review the implications and propose a way forward accordingly.

8.5 In addition to above, there are other ancillary services such as start-up and ramp-up/ ramp-down costs which should be recovered by the Wheeler of power in the economic dispatch system. The Regulator shall determine the costs of those ancillary services which can be practically determined.

8.6 Moreover, participants agreed that dispatch under system constraints has no bearing on the marginal price. The System Operator may from time to time dispatch a generator to ensure system stability and reliability – the variable cost of that generator dispatched under technical constraint is not included in the determination of marginal price.

6.3 INTRODUCTION

We note first that the term "wheeling" is only relevant under a specific institutional structure and has no meaning with respect to the actual transmission of electricity. The laws of physics and not the type of transaction or underlying contract determine how electricity flows. Thus, the issue at hand – "economic dispatch" – is entirely unrelated to wheeling. Rather "wheeling" and "economic dispatch" are both a function of how Pakistan chooses to fully implement open access.

Keeping physical supply and demand in equilibrium, i.e., keeping the system in balance, requires dispatching generation and increasingly dispatchable load so as to constantly keep the flow of electricity that is entering the system equal to the load, i.e., the demand. Thus, the dispatch function is the process of reliably maintaining the power balance such that the physical supply and demand of electricity are kept within a very narrow band at all times. In this way the dispatch function cannot – nor should not – be separated from the wider electricity market design.

That this definition of the "dispatch function" does not include "economic" or "least cost" is significant. The question of whether or not the objective of the dispatch function will be to reliably maintain the power *at least cost* gets to the core of how open access to the transmission system will be implemented. And in particular, how transmission constraints will be addressed, while maintaining the power balance.

As a result, transmission capacity, specifically short-term transmission capacity, is *the* fundamental issue in implementing open access. With respect to electricity market design, the central issue from which all other decisions flow in regard to the final design of the market is:

> How will transmission capacity be allocated in real time when the system is either constrained or cannot otherwise match any/all combinations of power flows from the supply source to the demand sink that participants may wish to transact?

Therefore, any deliberation regarding the dispatch function is simultaneously a discussion about how to implement open access as well as a discussion about what electricity market design should be chosen. The three issues – dispatch, open access, and electricity market design – are fundamentally interdependent.

This is the central question – the question at the core of implementing open access and, therefore, of every electricity market ever designed and implemented. In effect, the question of how best to implement open access directly relates to how best to dispatch the system. The answer to this question will directly determine the ultimate market design. There are two ways this question can be answered.

- At any moment in time the electricity transmission system is either constrained or not.[47] This is not a matter of opinion. Rather it is a simple, provable observable fact. Either the grid can facilitate every possible transaction that buyers and sellers have entered into for that moment or it cannot. By facilitate we mean allow the physical power flow from the source of the transaction i.e., the generator, to the sink of the transaction i.e., the load. Again, given the state of the network, this result is an incontrovertible fact.

In physical reality there are only two states of the world – either the grid is constrained or it isn't. The distinction is vitally important for the following reason:

- If the transmission system is unconstrained, i.e., if the transmission system can facilitate any/all possible physical transactions between producers and consumers in real time, then the transmission system is irrelevant to the buying and selling of electricity.

- If, on the other hand, the transmission system is constrained, i.e., the transmission system cannot facilitate any/all possible physical transaction between producers and consumers in real time, then, by definition, some transactions cannot take place and electricity, as a commodity cannot and should not be separated from transmission as a service.

These two points define what is at stake with respect to implementing open access and the ultimate electricity market design (i.e., dispatch) and, hence, why the decision is fundamental. If the transmission

[47] For purposes of simplicity, in the current context the term "constrained" means that the transmission system is, for whatever reason, unable to facilitate the power flows behind any/all transactions that producers and consumers of electricity wish to enter into. We acknowledge that this definition is technically not correct because there are instances when the grid may not be "constrained" yet any/all transactions are not feasible, i.e., contingencies, etc. Nevertheless, the purpose of using the simplified definition of the term is to highlight the question of whether or not the grid can facilitate any/all transactions.

system *can* facilitate any/all possible transactions between producers and consumers, then the electricity market, i.e., electricity as the commodity, can be separated from the complexity of dispatch and the transmission system, i.e., transmission as a service.[48] The following example illustrates why this is true:

Assume the transmission system is such that any/all possible (*ex-ante*) transactions between all producers and all consumers can be accommodated. In this scenario, producers and consumers can enter into (necessarily) *ex-ante* bilateral contracts or forward positions (i.e., from the day ahead market). In real time, these bilateral contracts and forward positions will then be "dispatched" and the system operator will manage so-called "balancing energy," i.e., volume risk, in real time by incrementing or decrementing generation to ensure the power balance stays in equilibrium.

To re-iterate the basis for this "vision", the transmission grid can accommodate any/all of the *ex-ante* bilateral contracts and/or forward positions. To "fine tune" the results and reduce the involvement of the system operator in supplying "balancing energy", the market design usually incorporates an "intra day market" whereby the (electricity) market participants can re-organize their *ex-ante* positions to better reflect current conditions.

Thus, for example if a generator and a load have a contract for 3MW in place prior to real time but due to unforeseen cold weather the actual load was higher, say 3.5MW then the load could use the Intra Day Market to purchase the additional 0.5 MW of energy. In theory the operation of the transmission system never constrains the transaction.

If however, the transmission system *cannot* accommodate any/all possible transactions between producers and consumers, then the electricity market and the operation of the transmission system cannot be separated or at least should not be separated if reliability and efficiency are desired objectives of the market and any attempt to do so will risk both reliability and efficiency.

Continuing with the previous example, suppose that in any given period one or more of the physical contracts or physical forward positions that market participants have entered into before real time cannot be accommodated. In this situation, the system operator (or dispatcher) will have to re-dispatch the entire system, because of the laws of physics.

6.4 FUNDAMENTAL ISSUES WITH ECONOMIC DISPATCH

The fundamental problem is that the dispatcher has not been provided with the information necessary to re-dispatch the system while simultaneously maintaining both reliability and efficiency. In short, there is no spot market for electricity whereby generation and transmission capacity can be allocated. There is only a day ahead market and an intra-day market. Neither of which provide sufficient guidance to the dispatcher in real time.[49] The so-named *market-and-re-dispatching process* is structurally incompatible with achieving efficiency in real time. Not only is there no possible methodology by which this paradigm can simultaneously achieve reliability and efficiency, the paradigm is incapable of producing efficient outcomes under any circumstance.

---

[48] This was called the "market separation hypothesis" in the initial electricity market designed and implemented in California.

[49] A situation that caused Polskie Sieci Elektroenergetyczne (PSE) the system operator for the Polish electricity system to remark in a recent press release (7 December 2018), "European market design does not facilitate optimization of generation and grid to the benefit of consumers. TSO has to ensure the secure supply of energy. When demand for transmission exceeds physical capabilities of power lines, TSOs intervene to prevent the risk of emergency disconnections, or even blackouts, by means of re-dispatching, i.e. increasing the generation that sits downstream of congestion, and decreasing it by the same amount upstream. It results in additional payments to generators. The problem with this market-and-re-dispatching sequence is that it treats power generation and grid resources separately, instead of optimizing them together. As a consequence, the energy is not priced at individual points of the system, and the price does not include the real costs of generation, transmission and distribution. Furthermore, the higher RES penetration in the system, the more energy volume is needed for re-dispatching" in a recent press release.

In fact, the statistical likelihood of any given actual transmission system being able to physically and reliably accommodate any/all possible source/sink transactions in real time is zero. Thus, the fundamental decision is not in regard to whether the system is constrained or not but rather whether or not the costs associated with the inefficient operation of the system in real time outweigh the benefits of supposedly simplifying the electricity market.

In practice, electricity markets that rely on the premise of an unconstrained grid in real time also rely heavily on a (physical) Day Ahead Market to determine how the grid will be used in real time. Therefore, the results of the Day Ahead market produce a (binding) physical schedule of generation and load which is assumed to be close to what will take place in real time. The thought is that the capacity of the grid and the plans of generators and load in real time should not change significantly from what was expected 24-hours prior to a given operating day.[50] Given the results of the Day Ahead market (i.e., the schedule), any deviations from what the capacity of the grid was expected to be 24-hours prior and what actually occurs are to be managed by the system operator - hence the common use of the phrase "market-and-re-dispatch process" to describe the structure of this type of market. Accordingly, the focus of this market structure is on the market and not the re-dispatch process. In theory, if not in actual practice, the "market" is assumed to deliver a feasible "dispatch." To the extent it is not feasible – for whatever reason – the dispatcher will take actions, which are necessarily outside of the electricity market, to maintain the reliability of the system.

In contrast, electricity markets that are based on the actual real time capacity of the grid, i.e., what is actually physically feasible at every instant in time, do not necessarily anticipate the *ex-ante* market will solve the dispatch problem. As a result, these markets rely almost exclusively on a physical spot market. The underlying tool of the physical spot market is a constrained optimization algorithm called Security Constrained Economic Dispatch (SCED).[51] The SCED uses current information about the grid to reliably and efficiently allocate (scarce) transmission capacity as quickly as every 5 minutes. In these markets, the Day Ahead market is just another forward market and the results have no special importance in terms of reliably and efficiently matching supply and demand in real time. Accordingly, the focus of this market structure is on the dispatch process and the forward or bilateral electricity market. In theory, the dispatch prices and the associated nodal prices are expected to provide the foundation of a robust financial futures market and bilateral contract market.

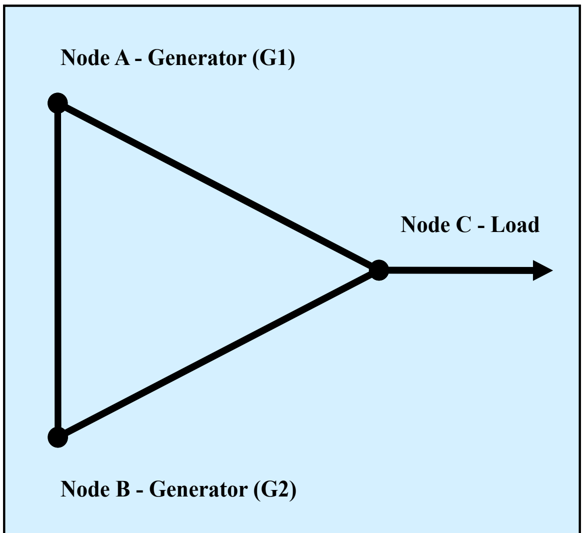


Figure 3: Simple Electrical System – Single Transmission Constraint

The physical characteristics of electricity necessarily mean that the capacity of the transmission system cannot be perfectly defined and "rationed" prior to real time. Moreover, the capacity of the system is not simply a function of the physical infrastructure of generation, and transmission facilities and load, but also a result of the decisions made by the dispatcher in matching supply and demand. This is not an opinion rather it is a fact derived from the physics of electricity and it has implications for the design and operation of any electricity market.

To understand the implications of this fact, we develop and make use of a simplified 3-node model of an electricity system, i.e., the simplest possible interconnected system; three nodes connecting two generators and one load. The system is

[50] While we use 24-hours as the benchmark in this paper, in actuality market close can be anywhere from 24- to 1-hour before real time.

[51] Note that this is specifically not called "economic dispatch" but rather Security Constrained Economic Dispatch, meaning that the dispatch is reliable, i.e., "security constrained", as well as least cost, i.e., "economic".

shown in **Figure 3**. The generators are located at Nodes A and B (G1 and G2 respectively) and the load is located at Node C. To keep things as simple as possible we assume the transmission lines, AB, BC, and AC are identical in both length and size and there are no losses when electricity is flowing In this abstract system, with infinite and lossless transmission capacity, load can be served either from G1 or G2.

This simplistic model doesn't allow for much understanding or analysis of the problems facing the system operator, i.e., the dispatcher, in real time. In this hypothetical world with infinite generation and transmission capacity the job of system operator in dispatching or balancing supply (generation) and demand (load) is easy and straightforward; simply choose the cheapest generator necessary to meet the load.

However, if we relax just a single assumption regarding the transmission system, the simple model is very useful in demonstrating why the determination of capacity *a priori* is problematic.

Specifically, we will assume there is a line limit on the transmission line between A and C. A line limit is an example of a transmission constraint. In reality there are many other transmission constraints that the system operator must take into consideration when matching actual/physical supply and demand in real time.

In **Figure 4**, we relax the simplistic assumption of unlimited transmission capacity by introducing a thermal constraint on AC. Specifically, the line now has a 200 MW limit, i.e., it cannot transmit more than 200 MW of power from Node A to Node C or vice versa. At this point we have made no assumptions about the capacity or the cost of either generator.

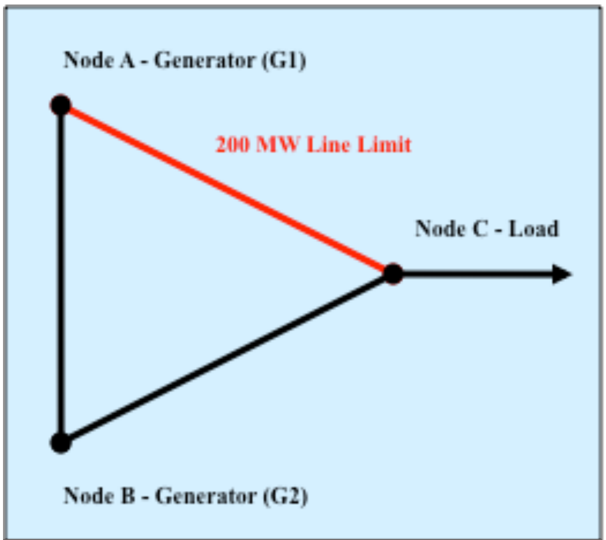


Figure 4: Simple Electrical System – Single Transmission Constraint

We can now use the simple model to derive some important conclusions regarding the calculation of the transmission capacity available on the system.

First, if we assume that load at Node C is 300 MW then either G1 or G2 is capable of supplying the load. **Figure 5** shows the situation when G1 produces 300 MW. In this case, 200MW from G1 will flow along AC to the load at Node C, while 100 MW will flow along AB and then BC to the load at Node C. If G2 produced the entire amount, then 200 MW would flow along BC, with the other 100 MW flowing from BA to BC. The 1/3 "relationship" between the transmission lines is a direct product of our assumption that the lines are of equal length and Kirchhoff's Law.[52] Based on this example it would appear that the available

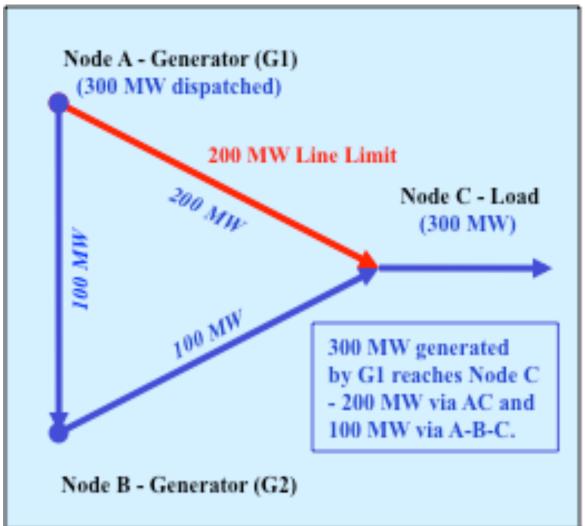


Figure 5: Simple Electrical System – Single Transmission Constraint, 300MW of System Load

[52] Gustav Kirchhoff's Current Law is one of the fundamental laws used for circuit analysis. His current law states that for a parallel path the total current entering a circuit's junction is exactly equal to the total current leaving the same junction. This is because it has no other place to go, as no charge is lost. Kirchhoff's current law is $\sum_{k=1}^{n} I_k = 0$, where n is the total number of branches with currents flowing towards or away from a node.

transmission capacity of the 3-node system is 300 MW. Assuming G1 is producing 300 MW then if either G1 or G2 tries to produce an additional MW of electricity some of it will flow along AC and the 200MW line limit will be exceeded.

However, suppose instead that load at Node C was not 300 MW but 600 MW. Note first that G1 cannot supply the entire load. If G1 produces 600MW, 2/3 or 400MW will flow along AC which will be a violation of the line limit and AC will overheat. G2 on the other hand, can supply the entire load. As shown in **Figure 5**, if G2 produces 600 MW, 400 MW will flow along BC and the remaining 200 MW will follow BA to AC without violating the 200 MW line limit on AC. Thus, it now appears that the available transmission capacity of the 3-node model is 600MW, i.e., 600MW of electricity can be generated and put on the transmission system

For our purposes, **Figures 5** and **6** demonstrate several very important and fundamental points, in regard to, the design of an electricity market:

> *Most importantly the available transmission capacity of any electricity network depends on the actual level of load, location of generation and the configuration of the network all of which will not be known precisely until real time, i.e., when power is actually flowing.*

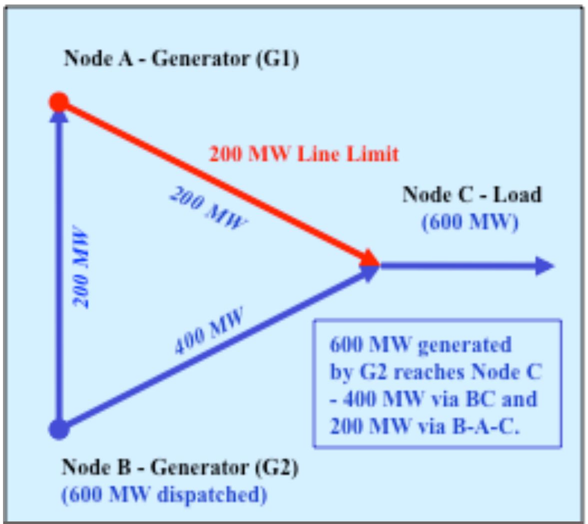


Figure 6: Simple Electrical System – Single Transmission Constraint, 600MW of System Load

Electricity generation, consumption and transmission/distribution wires operate like a single integrated machine, i.e., an integrated system. If supply and demand are not kept within close balance of each other at all times, the frequency of the system will rise or fall and the individual components (e.g. generation plants) of the "machine" will most likely be damaged. While storage would eliminate the need for current generation to always equal current load, the large-scale application of electricity storage is not yet economically feasible. As a result, the reliable operation of any electricity system requires that current generation must be equal to load at every moment in time.

6.5 PRICE VIS-À-VIS ECONOMIC DISPATCH

For most commodities and services, the market mechanism works to coordinate the activities of participants by creating a price that incentivizes behavior on the part of its participants that leads to a balance between supply and demand. For example, when the quantity supplied exceeds the quantity demanded at a given price, the price will fall which induces suppliers to produce less and for consumers to purchase more and the "problem" of excess supply is eliminated. This simple textbook explanation glosses over the fact that, in the real world it takes time for the price mechanism to "work", i.e., it takes time for prices to change and for participants to react to new prices.

Unfortunately, when it comes to electricity, we do not have the luxury of ignoring time. In the time it takes the price mechanism to work, lives could be lost, expensive machines could ruin, and the system could either go black or burn down. The integrated electrical system operates at a far faster pace than can the market. Although physical demand and supply on an electrical system must be balanced within a narrow band at every instant in time, the price mechanism cannot coordinate buyers and sellers that quickly.

The time frame within which we cannot expect the market to work to allocate resources is called real time or the "dispatch period/interval" and in some markets is as short as five minutes or as long as sixty minutes.

In essence, the "dispatch period" is the time frame within which the market cannot be used to solve supply/demand imbalances. Rather within this time frame the system operator has the authority to direct/order the participants to act in a way that is consistent with reliable operation of the integrated system.

The longer the dispatch period the more authority and autonomy the system operator requires in order to operate the system reliably. For example, compare the very different requirements of 5- and 60-minute dispatch periods. Under the former a price – reflecting supply and demand conditions – is produced every 5 minutes. Both generation and load are expected to respond to that price, and in a majority of scenarios, this response is expected to eliminate supply/demand imbalances. In this case the system operator only needs the tools necessary to get him/her to the next 5-minute interval. Suppose for example the 5-minute load forecast was below the actual level of load. Within the 5-minute interval no new price is being produced to signal generators that more output is needed. Instead the system operator must instruct a generator to increase their production. However, at the start of the next 5-minute interval, prices will rise to reflect the inaccurate load forecast and generation will increase their production in response to this price increase. In contrast, compare what happens under market design based on a 60-minute dispatch interval. In this case the system operator has to wait a much longer time for the price mechanism to work and so will most likely have to order much larger quantities of generation to mitigate the inaccurate load forecast.

We know that non-discriminatory open access requires that all connected parties will be able to acquire the transmission capacity they need in a non-discriminatory manner.

In real time - when electricity generation and load are being balanced through the dispatch/balancing function – the "commodity' that is being rationed/allocated is transmission capacity. This fact is true regardless of the choice of market design or institutional structure. Whenever the transmission grid is constrained, there is not enough transmission capacity to meet the *ex-ante* plans of the participants. Thus, given current technology, the "commodity" being rationed in real time *is not electricity*, but rather the available capacity on the transmission network. ***The primary question to be addressed is the process through which this scarce resource will be allocated, i.e., who will and who will not receive transmission capacity and at what price.***

While it is common to use the term "re-dispatch" to refer to situations where, as a result of transmission constraints, the non-constrained dispatch solution must be "re-dispatched", there is no such thing in actual practice – there is only dispatch, i.e., the system is not physically dispatched and then simultaneously "re-dispatched."

We have shown that the available capacity depends upon system load, which generators are used to meet the load, and the location of the load and generation *vis-a-vis* the topology of the transmission system.

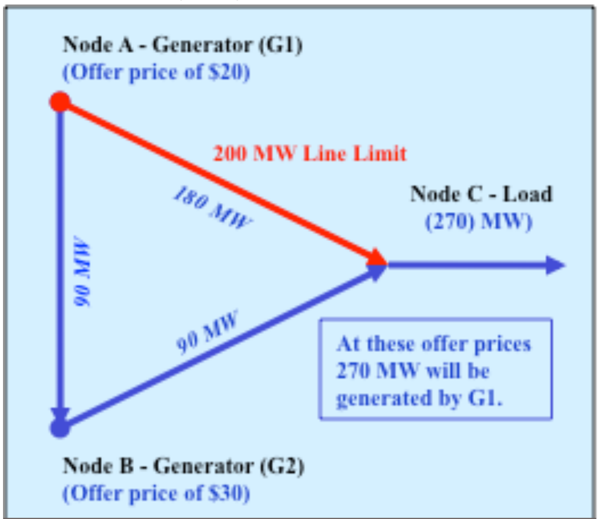


Figure 7: Simple Electrical System – Single Transmission Constraint, 270 MW of System Load – Optimal Dispatch

Now let's assume that the short run marginal cost (SRMC) of G1 is $20 and that of G2 is $30 and both have unlimited generating capacity. As shown in **Figure 7**, if the load at C is 270 MW then the optimal generation dispatch would be for G1 to produce all of the output (total cost would be 270 MW * $20 = $5,400).

Figure 7 provides the detail of how the power would actually flow from G1 to Node C. In this case, the *system-wide price* would be $20. That is, everybody on the system would either pay, in the case of load, or be paid, in the case of G1, $20. G2 would receive no revenue since they did not produce any electricity

But what happens if load rises from 270MW to 360MW as in **Figure 8**. Imagine for the moment the thermal transmission constraint of 200 MW on AC does not exist then having G1 produce all 360MW would be the least cost option. The total cost – without any transmission constraint – would be 360 MW * $20 = $7,200.

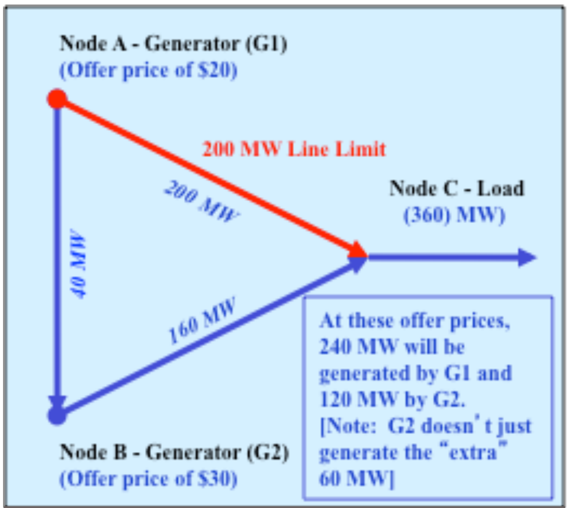


Figure 8: Simple Electrical System – Single Transmission Constraint, 360MW of System Load, Optimal Dispatch

But the existence of the line limit on AC means that the G1 cannot provide all 360 MW. Furthermore, since we saw in Figure 3 the most that G1 can produce is 300 MW, it might *seem* optimal to have G1 produce 300 MW, while G2 produces the remaining 60MW. Total cost under this scenario would be: (300 MW * $20) + (60 MW * $30) = $7,800. However, if G1 produces 300MW then there is no available capacity on AC. Thus, when G2 produces 60MW and 1/3 of that output flows on AC, the line limit will have been exceeded (200W from G1 and 20MW from G2). *So, while this may be a least cost solution, it is not a feasible solution*.

Rather than dispatch according to the simple Economic Merit Order, the system operator is going to have to change the least cost level of output because it does not respect the transmission constraints in the system. Given the offer prices and the system, the least cost solution will be when G1 produces 240MW and G2 produces 120MW. The total cost will be (240MW * $20) + (120MW * $30) = $8,400, which is $600 more than the "unconstrained" solution. **Figure 8**. shows the final solution.

We will walk through the process incrementally to show what takes place. Assume the system dispatcher started with the assumption that the cheapest generator (G1) was going to produce all of the power necessary to serve the load, i.e. 360MW. But he/she knows that will cause AC to have 240MW of power flowing across it, when the constraint is 200MW. So, the operator has to relieve 40MW of flow across AC while keeping the lights on at C. Suppose it was possible for the operator to sequentially solve this problem, *i.e.*, they start with the lowest cost solution, regardless of constraints, and then find a solution that is least cost while not violating the constraints, *i.e.*, they determine the optimal dispatch solution.

For every MW reduction in output by G1, the operator in effect, "buys" 2/3MW of space on AC, which then allows him to "buy" 2MW of output from G2. In this way the optimal dispatch will occur when G1 and G2 are producing 240MW and 120MW respectively. This solution is "optimal" because it minimizes the cost of meeting the demand without violating the constraint. No other solution will achieve this result. Suppose for example, the dispatcher chose instead to use 238MW from G1 and 122MW from G2, then total production costs for this solution would be $20 higher at $8,420. Alternatively, suppose they chose G1 and G2 to produce 241MW and 119MW respectively. This would lower production costs to $8,390 but would cause the flow on AC to be 200.33 and would violate the line limit. **Figure 9** provides a disaggregated view of the power flows from both G1 and G2 under the optimal solution.

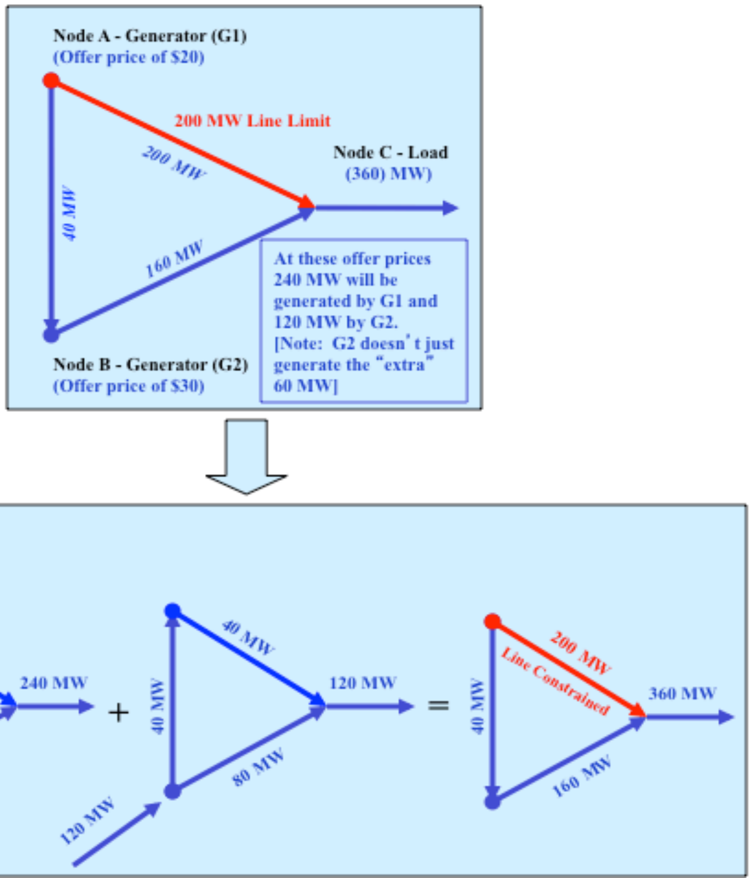


Figure 9: Simple Electrical System – Single Transmission Constraint, 360 MW of System Load, Power Flows Under Optimal Dispatch

We can use this example to determine the nodal prices, or equivalently the locational marginal prices, at each node.

The nodal price reflects what actually took place including the steps the dispatcher needed to take in order to match supply and demand while recognizing the constraints.

Suppose that load at C was 361MW rather than 360MW. We know how much G1 and G2 produce when load is 360MW, but how would the dispatcher acquire one more MW. He/she would reduce G1 by 1MW which saves $20 and then increase G2 by 2MW (1MW to make up for the reduction in G1's output and 1MW to meet the added load) which would cost $60, i.e. 2MW * $30. Thus, the marginal price at Node C, when load is 360 is: (-1 * $20) + (2 * $30) = $40. We can do the same analysis at the other two nodes and determine that prices are $20 at A and $30 at B. Thus, the price that reflects the actual dispatch is not a simple single, system-wide price but rather three prices, one for each node:

- Price at Node A = $20,
- Price at Node B = $30,
- Price at Node C = $40.

It is useful to reflect on these prices and in particular the information that is being conveyed. For consumers at Node C, an extra MW of demand will require the dispatcher to reduce output from G1 by 1 unit and increase G2 by 2 units. Given the assumed marginal costs of G1 and G2 the cost is $40. Under nodal pricing, generators will receive the price at their node times, the amount of output they produced or $8,400 in revenue (G1 will receive $20 * 240MW = $4,800 and G2 will receive $30 * 120MW = $3,600) and the load will pay $14,400 (360MW * $40/MW).

The difference between the amount the generators received, i.e., $8,400, and what the load paid, i.e., $14,400, is $6,000, and is the cost of the 200 MW line limit on AC. That is the economic cost of congestion for this interval given the offer prices, the load at C and the line limit is $6,000.[53]

6.6 PAKISTAN VIS-À-VIS INTERNATIONAL CONTEXT

From an international perspective nodal pricing – or Security Constrained Economic Dispatch – forms the foundation for the following electricity markets:

- Alberta *(AESO),*
- California *(CAISO),*
- Central America *(CDMER),*
- Mexico *(MEM),*
- Midwestern United States *(MISO),*
- New England *(ISO-NE),*
- New Zealand *(NZEM),*
- New York *(NYISO),*
- Philippines *(WESM),*
- PJM (Mid- Atlantic and Central Regions of the United States) *(PJM),*
- Singapore,
- Southwest Power Pool (Southwest and Mid Continent Region of the United States) *(SPP),*
- Texas (ERCOT).

While the focus of this paper has been on Security Constrained Economic Dispatch, there is an alternative market design - one that does not rely on or use SCED. In these markets, which necessarily rely upon the assumption that the capacity of the transmission grid can be accurately known in advance, the system operator manages transmission constraints in a non-optimal manner. Examples of these markets are:

- Belgium (Elia System Operator SA),
- Czech Republic (*ČEPS a.s.),*
- Denmark (*Energinet*),
- Finland (*Fingrid*),
- France (*RTE*),
- Germany (TransnetBW GmbH, Tennet TSO GmbH, Amprion GmbH, and 50Hertz Transmission GmbH),
- Italy (*Terna*),

[53] There are several available mechanisms for returning this "surplus", but that discussion is beyond the scope of this paper.

- Netherlands (*TenneT TSO B.V.*),
- Norway (*Statnett*),
- Sweden (Svenska Kraftnät)
- Switzerland (*SwissGrid*),
- United Kingdom (*National Grid*).

The following table provides a summary of the key characteristics of nodal and non-nodal electricity markets:

Market Characteristic	Non-Nodal Pricing	Nodal Pricing
Self-dispatch	Yes	Yes
Economic Dispatch	No	Yes
Separate Market and System Operators	Yes	No
Co-Optimization of Energy and Ancillary Services	No	Yes
Day Ahead Market	Yes	Yes
Market Based Congestion Management	No	Yes
Pricing	Zonal	Locational/Nodal
Forward, futures and bilateral contracts	Yes	Yes
"Re-dispatch"	Yes	Not required

We stated earlier that both wheeling and economic dispatch are a function of the choice that Pakistan makes regarding how they want to implement open access. There are two choices. Either Pakistan chooses to accept the assumption that the transmission system is unconstrained, or it does not. The former choice – that the transmission system is unconstrained – necessarily means that dispatch can be based on a simple generator merit order. However, when the transmission system is constrained, then generation cannot be dispatched according to the generation stack. To maintain the power balance the system operator will have to use out-of-merit generation and this is problematic for an open access regime predicated on the notion that there is no reason to use out-of-merit order generation.[54] Alternatively the latter choice regarding open access, is based on the fact that the transmission system will be unpredictably and regularly constrained. As such, this approach fundamentally, as compared to an *ad hoc* basis, incorporates out-of-merit order dispatch through the use of Security Constrained Economic Dispatch and some form of nodal pricing. In this model wheeling, as a method of transacting has no meaning and is eliminated. Hence any discussion on "economic dispatch" necessarily eliminates any discussion on "wheeling." Economic dispatch supports bilateral contracting of which a wheeling contract is an example, but there is no need to have what is currently called a "wheeling contract".

[54] We have used the extreme case of no transmission constraints to implicitly include both few transmission constraints as well as predictable (in terms of locational and time period) transmission constraints.

6.7 CONCLUSION AND RECOMMENDATIONS

We strongly recommend using Security Constrained Economic Dispatch as the basis for the electricity market in Pakistan for the following primary reasons:

- Nodal pricing reduces the necessary discretion of the system operator.
- Nodal pricing-based markets do not rely on necessarily false assumptions about the state of the transmission network in real time.
- Nodal pricing markets better allow for naturally occurring risk to be more efficiently managed through financial instruments rather than physical capital investment.
- Nodal pricing markets provide location and time-of-use based price signals for generation and load.
- Nodal pricing markets provide an explicit signal for the cost of congestion.
- Nodal pricing markets are economically more efficient and more reliable.
- Nodal pricing markets are far better suited to accommodate intermittent resources.
- Nodal pricing markets are less likely to be manipulated or subject to market power.

7. THE TREATMENT OF RENEWABLE RESOURCES IN THE WHEELING FEE

7.1 BRIEF SUMMARY

The discussion held at the LUMs Conference regarding this issue reflects a paradigm with respect to the operation of the grid, contracting for electricity and the nature of an electricity market that is no longer relevant. And, in fact, the demise of this paradigm has, in part, been hastened by increased intermittent generation.

In particular, the concept of "firm capacity" has no operational meaning with respect to intermittent resources. By that we mean, that the system operator cannot rely on a "scientific study" in real time to operate the system reliably. It is one thing to determine "firm capacity" for purposes of contracting and quite another to reliably operate a system that has intermittent connected generation resources. To the extent there is a difference between the firm capacity from the probabilistic "scientific" study and what is actually happening in real time, the system operator will be required to make up or dispose of the difference.

Given the nature of intermittent renewable reserves, the system operator will be required to carry additional operating reserves in order to maintain reliable operations. As such, we recommend that the Grid Code be reviewed to ensure that the reliability guidelines reflect the added volatility of the generation mix.

Furthermore, until such time as nodal pricing is implemented, our alternative recommendation to that of establishing a "firm capacity" is to require intermittent generation to schedule their output one hour in advance of real time operations for every operating interval. These schedules should be continuously reviewed for accuracy. The deviations between the scheduled and actual amounts will be absorbed by the wider grid, i.e., other generators will be ramped up/down on a reliability and economic basis to absorb the deviations from intermittent resources and the cost will be spread across all users of the system.

At the Conference held at the LUMs Energy Institute on December 9-11, 2019, the participants recognized and discussed the need to address the specific entities that will be involved in operating the eventual electricity market. The summary of their discussion is as follows:

Issue 7: Wheeling from renewables / firm capacity factor

7.1 Participants agreed that the current wheeling regulations do not take into account the concept of firm capacity of different technologies especially renewables.

7.2 It was discussed that a methodology to determine the firm capacity allocation to renewables is required to allow them to adequately and justifiably participate in the power system that takes into account the intermittent nature of resource, contribution during periods of peak loads etc.

7.3 There are many methods adopted globally to determine firm capacity of renewables. For instance, a scientific study is required to look at contribution of renewable technologies during n-sampled (say 100) periods of peak load to allocate a firm capacity factor to renewable technologies with variable resources. The wheeler of renewable power can therefore choose to:

a. contract not more than the allocated firm capacity factor with its BPC and any imbalances are settled through the marginal pricing mechanism, including the effect of variability in load.

b. include multiple technologies offering firm capacity in contracting with BPC to cover BPC's peak demand.

The Regulator may approve the appropriate methodology to calculate firm capacity factors for generation technologies with variable resources.

7.2 TREATMENT OF ALLOCATING FIRM CAPACITY FOR WHEELING TRANSACTIONS FROM RENEWABLE RESOURCES – PAKISTAN VIS-À-VIS INTERNATIONAL CONTEXT

Unfortunately, the discussion reflects a paradigm with respect to the operation of the grid, contracting for electricity and the nature of an electricity market that is no longer relevant. And, in fact, the demise of this paradigm has, in part, been hastened by increased intermittent generation.

In particular, the concept of "firm capacity" has no operational meaning with respect to intermittent resources. By that we mean, that the system operator cannot rely on a "scientific study" in real time to operate the system reliably. It is one thing to determine "firm capacity" for purposes of contracting and quite another to reliably operate a system that has intermittent connected generation resources. To the extent there is a difference between the firm capacity from the probabilistic "scientific" study and what is actually happening in real time, the system operator will be required to make up or dispose of the difference.

This necessarily means there will be negative and positive externalities built into the system that will have potentially significant effects on the wider pool of generation and load.

The core of the issue – how best to reliably and efficiently integrate intermittent resources – relates directly to the choice of electricity market design. To that end, Polskie Sieci Elektroenergetyczne (PSE) the Polish Transmission Operator, offers the following observations and advice:

> "In the opinion of PSE, the Polish transmission system operator (TSO), current European market design does not facilitate optimization of generation and grid, resulting in increase of costs of energy for consumers. It also contributes to higher CO2 emissions. PSE states that introducing Locational Marginal Pricing market (nodal pricing) would be a more adequate solution for the future. PSE plans to pilot a similar solution on the Polish market within the next few years. PSE's position on electricity market re-design was presented by Professor Leszek Jesień, Director for International Cooperation in PSE at POLITICO event "Giving the EU's Electricity Market a Facelift" organized in Brussels.
>
> Today's European energy market is divided into bidding zones and is based on "copper plate" assumption that the physical capacity of electric power transmission is unlimited. This model proved effective at the initial stage of energy market liberalisation in Europe and in the early period of development of renewable energy sources; yet, today, it is by no means flawless.
>
> European market design does not facilitate optimization of generation and grid to the benefit of consumers. TSO has to ensure the secure supply of energy. When demand for transmission exceeds physical capabilities of power lines, TSOs intervene to prevent the risk of emergency disconnections, or even blackouts, by means of re-dispatching, i.e. increasing the generation that sits downstream of congestion, and decreasing it by the same amount upstream. It results in additional payments to generators. The problem with this market-and-re-dispatching sequence is that it treats power generation and grid resources separately, instead of optimizing them together. As a consequence, the energy is not priced at individual points of the system, and the price does not include the real costs of generation, transmission and distribution. *Furthermore, the higher RES penetration in the system, the more energy volume is needed for re-dispatching.*
>
> According to PSE, these flaws in European market design can be overcome with fundamental changes. The LMP model, already operating in the United States, allows generation units to be continuously optimized in a way that minimizes energy supply costs, given the available transmission grid. Current and expected system conditions are fed into the market engine, so that generation set points are updated and implemented every five minutes. Harmony between market and system operations allows for trading up to very real-time, without endangering security of

supply. This model also translates into savings on electricity generation; CO_2 emissions are likely to decrease regardless of the generation mix.[55]

The presumptive electricity market model underlying the discussion at the LUMs conference forces NTDC, NEPRA or some other entity to forecast generation from intermittent resources rather than to simply allow these resources access to a physical spot market for energy where generation and transmission capacity are allocated in real time. Furthermore, the forecast capacity will almost never be correct and, as highlighted by PSE, this will force the system operator to intervene almost constantly to maintain reliability.

7.3 CONCLUSION AND RECOMMENDATIONS

While we accept that establishing a fictious level of "firm capacity" based on a probabilistic study/analysis could be used as an interim step we caution against doing so for several reasons. First, the concept of firm capacity will be enshrined in wheeling contracts that will potentially be 20 or more years in length. Second, once those contracts are entered into, changing the electricity market rules will affect the commercial terms and will, in some sense, amount to a regulatory taking. Third, deriving a hypothetical number for firm capacity does not aid the system operator in their task. Rather it necessarily makes their job more difficult.

Given the nature of intermittent renewable reserves, the system operator will be required to carry additional operating reserves in order to maintain reliable operations. As such, we recommend that the Grid Code be reviewed to ensure that the reliability guidelines reflect the added volatility of the generation mix.

Until such time as nodal pricing is implemented, our alternative recommendation is to require intermittent generation to schedule their output one hour in advance of real time operations for every operating interval. Furthermore, these schedules should be continuously reviewed for accuracy.

The deviations between the scheduled and actual amounts will be absorbed by the wider grid, i.e., other generators will be ramped up/down on a reliability and economic basis to absorb the deviations from intermittent resources and the cost will be spread across all users of the system.

---

55 https://www.pse.pl/web/pse-eng/-/pse-calls-for-electricity-market-design-facelift. See also https://vimeo.com/303528930 for the video that accompanied the press release. Also see https://www.pse.pl/documents/31287/20965583/PSE_16102018_The_good_the_bad_the_ugly.pdf for the written report on the problems arising from a non-nodal pricing based market.

8. THE ROLE OF SPECIFIED ENTITIES IN AN ELECTRICITY MARKET

8.1 BRIEF SUMMARY

We believe it is useful to discuss the actual functions that are required to operate an electricity market before discussing who should carry out those functions. Furthermore, while mindful of existing legislation regarding specific activities we also believe that it is helpful to at least recognize the principle that "form follows function", i.e., that the ultimate individual components of the institutional structure of the industry should be closely related to the fundamental function or purpose of the activity itself.

While the actual functions to be performed in operating an electricity market are largely the same regardless of the specific market design. Where market design significantly matters is in prescribing specifically *how* the functions will be carried out. For example, whether or not the dispatch function will be based on Security Constrained Economic Dispatch or not, i.e., the type of market design, has fundamental and significant implications for the scope and scale of the function.

The separation of electricity market activities into those performed by a "Market Operator" and those performed by a "System Operator" largely presupposes a specific type of electricity market design. In particular, it suggests that the System Operator will not operate a Spot Market based on nodal prices. While this dichotomy is consistent with the electricity market design used primarily in Europe, it is not consistent (or at least not necessary) for a market based on Security Constrained Economic Dispatch. In the latter type of markets, the system and market operators are one in the same because system operation is based upon a unification of the market and system operation.

Addressing this issue precisely is, therefore, dependent on the type of market design. Hence it is difficult for us to address the concerns and solutions raised by the LUMs participants because we do not know what the design of the Pakistani electricity market will be.

Given that there is no recommended market design from which to evaluate the specific role of entities, we make the following recommendations:

- All of the functions/activities necessary to operate the eventual electricity market should be precisely identified and then defined.
- For each activity, a detailed methodology describing how the function is to be accomplished should be developed and adopted.
- Assuming there is need for a Market Operator based on the adopted market design, after the first two steps have been accomplished all of the functions and responsibilities should be assigned to either the system or market operator.
- Both the System and (potentially) Market Operators are not Principals in the electricity market insofar as neither should purchase or sell electricity. Instead both are service providers to the market and the governance and business objectives of both should reflect this fact.
- The activities of the DISCO's should be focused solely on the physical operation of the low voltage network. Accordingly, they should have no relationship with the electricity market.
- Regardless of the initial market design, legacy decisions regarding the institutional structure should not hamper the efficient evolution of the electricity market.

8.2 BACKGROUND

At the Conference held at the LUMs Energy Institute on December 9-11, 2019, the participants recognized and discussed the need to address the specific entities that will be involved in operating the eventual electricity market. The summary of their discussion is as follows:

Issue 1: Specialized role of entities:

1.1 Participants discussed the need for specialized role of entities for stable and reliable system operations and a hassle-free, speedy process for settlement.

1.2 Participants agreed that the System Operator should be the sole entity to perform the centralized dispatch of all generators, including the Wheelers of power, at 132kV and above voltage level, to balance demand and supply, frequency control and voltage control. The existing wheeling regulations should be amended and provide clarity accordingly.

1.3 Participants agreed that the Market Operator should perform the settlement function as mandated through the NEPRA Act, MO Authorization (granted by NEPRA) and Market Operator Rules. The DISCOs do not have the legal mandate for settlement under the respective Distribution Licenses, nor is it a global practice. It is also prudent for Market Operator to be the central entity when the information for settlement needs to be gathered from different entities for BPCs in multiple DISCO systems. Therefore, the existing wheeling regulations should be amended to provide clarity on assigning this settlement role to Market Operator only.

1.4 It was also discussed that the Market Operator must demonstrate its readiness to perform the settlement function resulting from trade as envisaged under the legal and regulatory framework.

1.5 Participants discussed that the role of DISCOs includes retail business and distribution network operator and should be continued as such under the wheeling regulations. Moreover, DISCOs may through their power distribution control centers perform coordination of dispatch with the wheelers of power at 11kV and below voltage level.

1.6 In future, it was further proposed that standardized/model contracts can accelerate the process of wheeling. The Regulator and/or Market Operator may develop such standardized/model contracts in consultation with stakeholders.

8.3 ANALYSIS ON ROLE OF ENTITIES IN AN ELECTRICITY MARKET

While we do not necessarily disagree with the conclusions reached by the participants in their discussion, we believe it is useful to discuss the actual functions that are required to operate an electricity market before discussing who should carry out those functions. Furthermore, while mindful of existing legislation regarding specific activities we believe that it is helpful to at least recognize the principle that "form follows function", i.e., that the ultimate individual components of the institutional structure of the industry should be closely related to the fundamental function or purpose of the activity itself.

We further note that the actual functions to be performed in operating an electricity market are largely the same regardless of the specific market design. Where market design significantly matters is in prescribing specifically *how* the functions will be carried out. For example, whether or not the dispatch function will be based on Security Constrained Economic Dispatch or not, i.e., the type of market design, has fundamental and significant implications for the scope and scale of the function.

We further note that once the institutional structure, i.e., the creation of specific electricity market institutions and the assignment of the functions to be carried out by those entities, has been defined it has not only long-lasting, perhaps permanent, effects on the operation and evolution of the market, but also is very difficult to change. As a result, it is potentially unfortunate that Pakistan has already legislatively enshrined certain parts of the institutional structure prior to making decisions about the final market design.

For example, the separation of electricity market activities into those performed by a "Market Operator" and those performed by a "System Operator" largely presupposes a specific type of electricity market design. In particular, it suggests that the System Operator will not operate a Spot Market based on nodal prices. While this dichotomy is consistent with the electricity market design used primarily in Europe, it is not consistent (or at least not necessary) for a market based on Security Constrained Economic Dispatch. In the latter type of markets, the system and market operators are one-and-the same because system operation is based upon a unification of the market and system operation.

Addressing this issue precisely is, therefore, dependent on the type of market design. Hence it is difficult for us to address the concerns and solutions raised by the LUMs participants because we do not know what the design of the Pakistani electricity market will be.

To date, it appears to us that Pakistan expects to implement a "TSO-type" market initially and then at some (currently undefined) point in the future move to an ISO-type market. The two market designs are fundamentally different and there is no marginal "incremental" step, once implemented, to move from one to the other, The former market design does not rely on a physical spot and does not integrate dispatch with the electricity market. Rather, the TSO, or Transmission System Operator, manages all the transmission constraints outside, or separate, from the electricity market. In contrast, the ISO-type market integrates the electricity market with the management of all transmission constraints and the fundamental tool to do this is voluntary bid-based Security Constrained Economic Dispatch.

In general, we can characterize the functions that must be performed to operate an electricity market into three distinct time periods. We define the operating interval as the time interval during which generation and interruptible load must follow the commands of the system operator. This is more commonly called "real time". There is then the time period prior to real time and the time period after real time.

Prior to real time, there are a number of activities that must be done[56] - pre-commitment of slow-start generation, reliability contingency analysis, the submission of generation schedules, unit commitment, etc. These "scheduling activities" are typically, but need not be, the responsibility of the system operator. Similarly, after the operating interval, there are a number of functions that must be performed – price validation (if using nodal pricing), final approval of the nodal prices, settlement, invoicing, credit analysis, etc. Depending on the market design these ex post activities are either the responsibility of the system operator or the "market" operator. During real time, the primary activity performed by the dispatcher or system operator is to maintain the power balance within the prescribed reliability parameters, this involves potentially committing or de-committing fast start units, developing and communicating dispatch instruction (a set point) for every generator, (including ramping units up or down).

The activities of the local distribution companies – the DISCOs – should be confined to the low voltage network, i.e., maintenance of the lines, managing outages, etc. Thus, the DISCOs should not participate in the "energy" market. That is, with one possible exception, the DISCOs should have no relationship with the operation of the electricity market. The one exception is related to how interruptible demand is integrated into the electricity market.

[56] We ignore transmission planning and focus directly on the functions required for real time operations.

8.4 CONCLUSION AND RECOMMENDATIONS

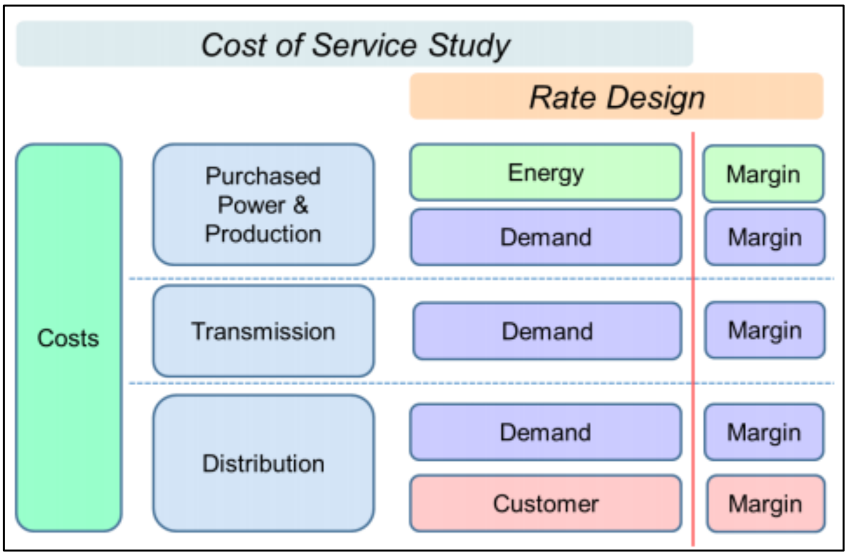
Given that there is no recommended market design from which to evaluate the specific role of entities, we make the following recommendations:

- All of the functions/activities necessary to operate the eventual electricity market should be precisely identified and then defined.
- For each activity, a detailed methodology describing how the function is to be accomplished should be developed and adopted.
- Assuming there is need for a Market Operator based on the adopted market design, after the first two steps have been accomplished all of the functions and responsibilities should be assigned to either the system or market operator.
- Both the System and (potentially) Market Operators are not Principals in the electricity market insofar as neither should purchase or sell electricity. Instead both are service providers to the market and the governance and business objectives of both should reflect this fact.
- The activities of the DISCO's should be focused solely on the physical operation of the low voltage network. Accordingly, they should have no relationship with the electricity market.
- Regardless of the initial market design, legacy decisions regarding the institutional structure should not hamper the efficient evolution of the electricity market.

ANNEX A: INITIAL COMMENTS TO NEPRA'S METHODOLOGY FOR ESTABLISHING THE CROSS SUBSIDY SURCHARGE (CSS)

We have been asked to provide initial comments to NEPRA's methodology for establishing the Cross Subsidy Surcharge (CSS)

The methodology used by NEPRA to derive the CSS is based on a "Cost of Service Study" or CoSS for each DISCO. In essence, a CoSS is a mechanism for apportioning the total costs of a DISCO to each of the various customer classes. In Pakistan the customer classes are defined by their connected voltage level, i.e., 132 kV, 11kV, and 220 & 400 V. By altering the rates, NEPRA can allocate the costs to each of the three customer classes. The CoSS allocates the total costs of providing electricity to each of these three customer classes and then takes the rates for each customer class and shows how much each customer class contributes to the overall utility margins.



The generic CoSS process is shown in the diagram on the right. What NEPRA apparently does is to set the DISCO rates so as to achieve the desired level of cross subsidization. Any customer class with a margin that is less than the overall margin is being subsidized by other customer classes, i.e., those with a margin higher than the overall margin for the company.

As a rate setting device, the CoSS methodology is commonly used and is a standard regulatory mechanism. It is important to understand that the tool is more commonly used when the rates have already been set and the objective of the exercise is to *show the cross subsidies that exist within the existing rate structure*. In contrast, NEPRA appears to be using the methodology to determine the level of the cross subsidy. In other words, in most cases the rate is outside, or exogenous, to the CoSS exercise, which is not the case with how NEPRA is using it. NEPRA is using the tool to set the rates based on the levels of cross subsidization. Mathematically there is nothing wrong with doing this but it needs to be understood that a CoSS is not capable of answering the question of "what amount should the subsidy be?" Given the information that we have been provided, we have no ability to evaluate the efficacy of the actual level of the subsidy that the CoSS is being used to allocate.