

TESTIMONY

June 25, 2013

Cost Containment in California's Cap-and- Trade System

Dallas Burtraw

Prepared for the California Air Resources Board

1616 P St. NW
Washington, DC 20036
202-328-5000 www.rff.org



Cost Containment Workshop California Air Resources Board

Presentation of Dallas Burtraw
Darius Gaskins Senior Fellow
Resources for the Future
June 25, 2013

My name is Dallas Burtraw and I am the Darius Gaskins Senior Fellow at Resources for the Future (RFF). RFF is a 61 year old independent, non-advocacy research institution in Washington, DC. RFF does not take positions on issues; the opinions I express today are my own.

I want to speak to you about cost management within California's cap-and-trade program, and specifically the possibility that the compliance reserve would be exhausted causing prices to rise above the top tier imagined in the program currently. In summary, I make one key observation and identify four possible policy options.

First, there is a small but nonzero probability that the compliance reserve will be exhausted, but if that were to happen it would not happen immediately. It would unfold over several years. This gives regulators time to consider responses in a timely way, although the steps that would be taken should be identified in advance. Four options could be considered:

1. Exercise of the governor's discretion to relax compliance obligations. This authority is granted in AB 32. The Air Resources Board should support the governor by anticipating this possibility and develop criteria for administrative intervention and the steps that would be taken in that eventuality.
2. The compliance reserve can be repopulated through the acquisition of allowances from other existing cap-and-trade programs. Difference in stringencies of the programs can be overcome such that multiple tons of reductions might be equivalent to one allowance entered into the compliance reserve, providing emission reduction benefits. The institutional details for how this would occur should be anticipated in advance and proof of concept demonstrated, but enactment could be postponed.
3. Compliance entities could be given the opportunity to settle accounts and restart their three-year compliance clocks, thereby introducing overlapping compliance periods among the regulated entities. This would enable limited borrowing within the market.
4. The development of plans for beyond 2020 would provide a signal to innovators and investors to take measures that could be expected to lower the marginal abatement cost.

* * * * *

The use of incentive-based approaches to environmental regulation such as emissions trading allows environmental goals to be achieved at less overall costs than would occur through prescriptive regulation. However, the market price will vary and it may exceed the range viewed acceptable in designing the program. The purpose of cost containment is to prevent unanticipated swings in the price that are so severe that they undermine the environmental objectives, political acceptability or regulatory success of the program. If prices are much lower than anticipated, they will remove the incentive for innovation and investment to reduce emissions. If prices are much higher than anticipated they will remove political support for the program by inflicting high costs on consumers and industry, and they may put California business at an unacceptable disadvantage in competitive interstate and international markets.

It is important to note that cost containment is truly a two-edged problem. Experience indicates that at the outset of existing market-based programs the regulated entities expressed anxiety that prices might be higher than expected. Ex post, in virtually every program the outcome has been that prices are less than expected. Unfortunate outcomes can result in either case. The California program has anticipated this through a symmetric approach to cost management, with a reserve price in the auction providing a floor and numerous features providing a buffer against high prices. Those features include trading, banking, offsets and the compliance reserve which makes a large sum of allowances collected from under the cumulative cap through 2020 available in any individual year at three tiers of prices. The concern that motivates this workshop is articulated in Air Resources Board Resolution 12-51—that these buffers may be exhausted and prices could rise above the top tier of the compliance reserve.

The focus of my remarks are on the performance and design of the compliance reserve and the possibility that prices may be greater than expected and may exceed the top price tier in the reserve. Although the size of the reserve is large, the fact that it is ultimately limited is the key feature of concern. This design is distinguished from what has been sometimes referred to as a safety valve or hard price ceiling, which would provide an unlimited supply of allowances at a ceiling price. The hard price ceiling (safety valve) could be implemented either through the availability of an unlimited supply of allowances at the ceiling price or through an alternative compliance payment (fee) set at that level that could be paid to offset emissions.

In 2007 the Market Advisory Committee made general recommendations to outline the design of a cap-and-trade program. The committee’s report identified a symmetric safety valve as one desirable option, including a price floor, which was adopted in the cap-and-trade program. The committee recommended against a safety valve providing an unlimited supply of allowances at the ceiling price. The market design ultimately turned away from a safety valve because of

concern that it would erode the environmental objective of the program by potentially introducing allowances that would bust the emissions goals under AB 32.

Instead, the program adopted a soft price ceiling providing a limited supply of allowances in the compliance reserve drawn forward from future year emissions targets under the cap. Modeling exercises (Fell et al. 2011) show that this approach captures the majority of expected cost savings that would be achieved if the supply was unlimited. This is because in most realizations of an uncertain outcome a small extra supply is sufficient to prevent prices from exceeding the desired ceiling. At the same time, this has the advantage that it limits the amount of emissions in excess of an annual allocation that can occur, yielding lower expected emissions and providing assurance about environmental integrity. The fact that the reserve is populated with allowances from future years under the cumulative emissions reduction goal through 2020 makes environmental integrity a certainty. Hence a limited supply of allowances delivers most of the benefits to industry and retains environmental support because it can strengthen overall political support for the program.

Under the assumption that there is a nonzero probability that the compliance reserve could be exhausted and prices could rise above the top tier, the central question in deciding how to mitigate against this (addressing Board Resolution 12-51) is whether to implement a design in which this could never happen in contrast to one in which it would be very unlikely to happen.

My first observation is that if the top tier of prices in the compliance reserve were reached and sustained for an extended period of time, one might expect it would trigger a program review. This would occur whether the price was supported by an unlimited supply of additional allowances or not. Hence, the effort to enact a permanent market design that is robust against all eventualities may not be achievable. It might be as effective to enact a design that made the probability of exhausting the compliance reserve very small.

In this vein, the primary functionality of cost management in general and the compliance allowance reserve in particular is to provide buoyancy to the program in the case that it encounters short-run disruptions that cause an imbalance in supply and demand. This short-run phenomenon would likely be corrected over months. In the meantime the compliance reserve would ensure that the price in the trading program would not sky-rocket. On the other hand, if the changes in the market were structural and the demand for extra allowances was to continue for the long term, an unlimited supply of allowances would likely lead to the same outcome as a limited supply—that is, a program review. For this reason, efforts to reduce the probability of triggering a review by bolstering the compliance reserve with a limited additional supply could be equally effective as an unlimited supply of reserve allowances.

What are the chances in the California program that the quantity in the compliance allowance reserve will be exhausted and prices rise above the top tier? Bailey et al. (2013) solve a dynamic stochastic model to argue that the chance is small, but nonetheless possible. One

reason is that the supply schedule of additional emissions reductions is steep, so that if events precipitated the need for additional emissions reductions, the price could rise substantially. This is consistent with the results of our model at Resources for the Future. There are situations that seem plausible that could lead to a supply-demand imbalance, such as a series of years with low hydro availability. This could lead to a transitory imbalance between supply and demand for allowances. The loss of low-emitting electricity generation, as we have observed recently, or the failure of the complementary programs such as the renewable portfolio standard or the low carbon fuel standard to achieve their intended results could lead to long-term imbalance. If events were to occur that increased the demand for additional emissions reductions beyond what is already envisioned in the program planning, those additional reductions might be hard to come by.

In cases like this, the price level might reach the trigger in the compliance reserve, but the drawdown of the reserve would not be instantaneous. It is important to note the difference in time between when the price reaches the reserve and when it might exhaust the reserve. The analysis in Bailey et al. is a cumulative analysis of demand and supply through 2020. However, the reserve is large relative to the emissions budget in any given year. In an economic model, the solution is achieved instantly and inter-annual behavior is captured in a single outcome. This seems to imply that investors would recognize the price path and the possibility that the reserve could be exhausted and purchase the entire reserve in anticipatory behavior. In practice, this would require substantial capital and it would require depositing allowances for use in a future year in the holding account, potentially violating holding limits. The firm also would have to consider the risk associated with prospect of administrative measures or federal legislation which could expand supply or lower demand and undermine the value of that investment. Consequently, in the eventuality that the reserve was ultimately exhausted, regulators would see signals of the possibility years in advance.

In sum, the chances of breaching the top tier of the compliance reserve are small, but if it were to happen it would not happen suddenly. Instead, a series of events would unfold over time that would send regulators a signal that the structural economic changes were occurring.

In planning to mitigate the possibility of prices rising above the top tier of the compliance reserve, as directed in the board's resolution (12-51), the Air Resources Board should consider whether a policy response should be identified and enacted in advance (now) or on a real-time basis if and when they are necessary. In general, there is a disadvantage to leaving the response up to administrative discretion at a future date. The time of a market disruption may not provide a political environment to make the best decisions. In addition, the uncertainty that would be associated with anticipated administrative decisions may affect decisions in the market. Identifying decision rules and criteria that would be used in deciding when to enact possible changes in the program would help mitigate uncertainty for market participants. Moreover, there is no way to decidedly preclude administrative involvement in the program. The goal of program design is to mitigate the need for such an intervention. There are steps that can be taken ex ante

to indicate to the market how decisions would be made and what steps would be taken if the price ceiling were reached. This involves modifications to the program design now and the identification of criteria and additional actions that would be evaluated in the future if necessary.

There are four measures that could be taken now to make the program design more robust to potential price fluctuations.

1. Governor's discretion

The governor is given discretion by Part 7 of AB 32 to adjust applicable deadlines for individual regulations in extraordinary circumstances. Although discretion remains with the governor, the Air Resources Board can provide guidance with respect to criteria that would constitute extraordinary circumstances. It could also develop a blueprint for how those measures might be implemented, for example, by executing one or more of the measures identified in the staff document.

One might fear that developing a contingency plan would make its execution more likely to happen, but the opposite can be equally as likely. Under changes in political circumstances, the executive might pursue actions to deconstruct AB 32. Clear criteria laid out in advance might provide clarity and accountability for future administrative decisions that might be considered and help ensure that changes were not made for arbitrary or short-term political reasons.

2. Expansion of the (potential) supply of allowances in the compliance reserve

The primary opposition to a hard price ceiling (safety valve) is the erosion of environmental integrity. However, climate change is a global problem and emissions reductions anywhere help to mitigate that threat. Emissions reductions outside California might provide a way to populate the compliance reserve while preserving environmental integrity. Measures taken outside California would not yield ancillary reductions in conventional pollution in California, with the exception of geographically near investments such as along the northern Mexican border. They also would not yield employment or investment in the state. However, they can help to build momentum for global action through learning and coordination across jurisdictions leading to cooperation and the incremental alignment of climate policies.

One possibility is to purchase allowances from other existing cap-and-trade programs. A central challenge in doing so would be the difference in stringency. Allowances trading at a market price that is much less than that in California might introduce moral hazard because California entities would look to purchase low-cost allowances from outside the state, and if those programs were less stringent they might be viewed as “hot air”—allowances that do not map into meaningful emissions reductions.

The different levels of stringency provide an opportunity, in fact, for California to incentivize the reduction of multiple tons at the cost of a single allowance if the allowance price

is equal to the top tier of the allowance reserve. This could be accomplished through the aggregation of allowances from another program so that the compliance reserve in the California reserve would be re-populated at a ratio of multiple tons in the other program per allowance. For example, a purchase of five allowances and the reduction of five tons from a program with a price of \$10 could be made available at \$50 through the reserve.

There are several ancillary benefits from supplementing the compliance reserve with allowances from other programs. First, it would help support the price in those programs. Second, it builds bridges to those programs. Third, the emissions reductions would be effectively guaranteed by the monitoring and verification procedures already in place in those programs. Fourth, environmental performance could benefit by aggregating multiple allowances from other programs with a low price so as to establish equivalency with the price of an allowance sold through the compliance reserve.

The possible use of allowances from other programs to populate the compliance reserve would not constitute linking, but there are nonetheless complicated institutional questions about how it would occur, for example, what body would aggregate emissions allowances from other programs and deposit them in the compliance reserve. The steps that would be needed to implement this option could be thought out now and proof of concept could be demonstrated through the purchase of a small lot of allowances from another program that would be contributed to the allowance reserve. It is conceivable that a quasi-governmental organization could serve this function. There may be interest at international organizations such as the World Bank to support the development of an institution that could provide this function across markets. After proof of concept is established, this option could then be put on ice until it might eventually be needed again. It is useful to remember that a substantial depletion of the compliance reserve will not happen immediately, but will unfold over months and years, providing time for re-population of the reserve to be implemented.

Finally, there is the possibility of populating the reserve with emissions reductions from outside other programs, essentially with sectoral offsets or other new types of offsets. This would be different from adding additional offsets within the market and subject to existing protocols and to the limits on offset use for compliance. As a supplement to the compliance reserve, offsets could be aggregated to yield more than one ton per allowances. California might also look for the opportunity to support new institutions encouraging efforts in the developing world that ultimately could help those efforts move toward engagement with international carbon markets. Again, there may be interest at the World Bank in helping to catalyze institutions that could fulfill this ambition. Engagement with climate-related activities in the developing world could leverage California's contribution to mitigation of carbon emissions on the international stage and this might be done in a way that would reinforce the cost containment provisions of the California market design.

3. Overlapping compliance periods

Another option is to draw additional allowances from future annual allocations to replenish the compliance reserve if the price ceiling is reached. This would have a limited impact on the allowance price in the near term because the market would recognize the cumulative supply of allowances is unchanged. However, it could provide relief in the face of a short-run and transitory disruption in the market.

To a similar end, a simple and straightforward way to add resiliency to cost management within the market would be to institute overlapping compliance periods to stagger the settlement dates in the market. This market feature would provide additional buoyancy in the short run if the market were to experience a temporary disruption due to exogenous events such as weather or disruption in fuel markets.

The most likely event that could trigger a price spike in the California market, in my opinion, is a confluence of events near the end of a compliance period. The three-year period gives regulated entities an opportunity to plan their compliance activities in response to expected prices, but if there is a disruption in the market at the end of the compliance period there would be little time to adjust. Overlapping compliance periods would enable limited borrowing among firms to overcome a short-run imbalance in demand and supply.

One consideration is how to assign different entities to different periods, especially now that the program is underway. It may be inviting controversy to force such an assignment, but entities may voluntarily settle their compliance obligations early in order to restart a three-year compliance clock.

Another consideration is the transition of the program in 2020. If the trading program were to be discontinued, all entities would have to come into compliance at the terminal date. However, if the program were to continue perhaps with adjustments in the annual targets, the overlapping periods might provide a smooth way to implement that result. In effect, overlapping periods would complement banking as a way to link compliance periods.

4. Planning for beyond 2020

Allowance banking provides a link in market prices over time. The short-run price is expected to reflect the discounted value of marginal abatement costs that will be necessary in the future. The determination of plans for beyond 2020 will provide a signal to regulated entities about their future obligations. On the one hand, this could provide an incentive to bank allowances now if the program is expected to be more stringent in the future, and this would provide upward pressure on allowance prices. On the other hand, it could provide the incentive for innovation and investment that will bring down abatement costs.

I believe it is unlikely that compliance entities will rely heavily on banking in any circumstance, even if the holding limits were not in place. One reason is the faint recurring whisper of congressional action at the federal level. If federal action were to happen it would be likely to push down the price of allowances in the California market. The marginal abatement cost for regulated entities would be equal to the sum of prices under federal and state systems. Given that there will be moderate banking a declaration of intent for beyond 2020 would be unlikely to introduce an upward pressure on allowance prices in the near term. However, it could provide clarity in the market that helps smooth short-run disruptions in allowance prices and advances the planning process in general.

* * * * *

In summary, there is a small possibility that allowance prices will reach the trigger price of the compliance reserve. If it does so, it will take a substantial amount of time before the compliance reserve is exhausted and prices rise above that level. In that time frame regulators will have the chance to implement a program review or take actions to mitigate the price increase. It would be preferable that such actions and criteria for taking action are set out in advance.

One action that is always available is discretionary intervention by the governor. The Air Resources Board could provide guidance with respect to criteria that would constitute extraordinary circumstances. It could also develop a blueprint for how those measures might be implemented, for example, by executing one or more of the measures identified in the staff document.

A second action would be to expand the supply of allowances in the compliance reserve. The most progressive way to do that would be to draw allowances from other cap-and-trade programs. The institutional vehicle for doing this would need development.

A third action would be to institute overlapping compliance periods. This would guard against an unfortunate confluence of events that might cause a short-term spike in allowance prices.

Fourth, regulators should begin to signal their intent for beyond 2020. The influence of extending the program on allowance prices is ambiguous, but given various reasons that banking is expected to be limited under the program, the most important effect may be to provide a signal for innovators and investors to take action that is expected to help reduce compliance costs.