Investigation of the Effects of Emission Market Design on the Market-Based Compliance Mechanism of the California Cap on Greenhouse Gas Emissions

University of Virginia – PEAR Project Team

February 12, 2013

University of Virginia

PEAR
Power & Energy Analytic Resources

Acknowledgements

We would like to express our appreciation for the support and guidance of Michael Gibbs and his colleagues at the California Air Resources Board.

We gratefully thank those who sponsored the study:

- *Pacific Gas & Electric Company*
- *Southern California Edison Company*
- *Chevron*
- *NRG Energy*
- *Los Angeles Department of Water & Power*
- *Sacramento Municipal Utility District*
- *Northern California Power Agency*
- *Southern California Public Power Authority*

And we acknowledge the gracious assistance of those who volunteered their time and insights to participate in the simulations, as well as our research assistants Sijia Yang, Elsa Schultze, and Annie Rorem.

UVA PROJECT TEAM

Project Manager	Jan Mazurek, Ph.D., *ICF International*
Deputy Project Manager	Michelle Osborn, *ICF International*
University of Virginia Team	Charles Holt, Ph.D., *Chair, Economics Department*
	William Shobe, Ph.D., *Professor of Public Policy*
PEAR Market Experts Team	Thaddeus J. Huetteman, *President, PEAR Inc.*
	Andy Van Horn, Ph.D., *Van Horn Consulting*
	Kedin Kilgore, *CO_2 Capital LLC.*
	John Melby, *The Melby Group*
Application Development	Doug Simpson, *Adasphere*
	Nancy Brown, *Adasphere*

Investigation of the Effects of Emission Market Design on the Market-Based Compliance Mechanism of the California Cap on Greenhouse Gas Emissions

TABLE OF CONTENTS

Investigation of the Effects of Emission Market Design on the Market-Based Compliance Mechanism of the California Cap on Greenhouse Gas Emissions

TABLE OF CONTENTS *(cont.)*

APPENDICES

Investigation of the Effects of Emission Market Design on the Market-Based Compliance Mechanism of the California Cap on Greenhouse Gas Emissions

FIGURES AND TABLES

I. Executive Summary

- In January 2013 the California Air Resources Board (ARB) launched the world's first economy-wide greenhouse cap and trade system. The state's experience with electricity deregulation **convinced ARB of the need to add unique design elements aimed at preventing**:
  - **buyers from gaining dominant market positions** that could give rise to undue market influence, for which ARB specified a ceiling on the maximum number of allowances one company could maintain ("holding limits"); and
  - **excessive price volatility for California carbon allowances** that would disadvantage the state's consumers and businesses, leading ARB to set aside a reserve of allowances to be sold at pre-established prices ("allowance price containment reserve, or APCR"). In addition, some economists have urged ARB to keep allowance prices from rising above the third tier of the APCR, which would effectively impose a "hard cap" on allowance prices.

- The **holding limits and APCR provisions have been criticized for potentially creating their own problems,** which could hinder the development of the California carbon market. In addition, concern has been expressed over the viability of these features during randomly occurring periods of especially high demand for emission allowances (e.g., years of low hydroelectric power availability.)
  - In our analysis, we studied the impact of ARB's program design on market behavior using experiments with student subjects in carefully controlled laboratory settings and simulations with experienced professionals in a richer but less controlled market setting.
  - These complementary methods generated mutually reinforcing results but also provide added information about what factors might be driving the results.

- In both our experimental sessions and market simulations, **we were able to detect patterns of differences attributable to the imposition of holding limits and the APCR on four key indicators of market performance**, including 1) price discovery, 2) efficiency, 3) volatility, and most critically, 4) liquidity, which is regarded by many economists as the key factor in defeating market manipulation.

- In summary, we found that **holding limits can have a significant negative impact across the board on each of the four market performance indicators.** Tight holding limits (those that constrain normal operations of entities with large obligations) contribute to reduced liquidity, higher price volatility, less effective price discovery, and lower efficiency.
  - Much of this effect occurs through reduced banking of allowances in holding accounts, which is critical given that the ability to bank reflects sources' flexibility to manage risk and to mitigate the effects of increasing scarcity over time.
  - The resulting **reduced banking also translates into delayed reductions in greenhouse gases**.

- Our results suggest that an APCR serves a useful function for emission markets. APCR did change subject behavior by lowering their risk of large price spikes. We found that subjects resorted to **the APCR as structured by ARB more frequently and used it in different ways than we expected, both as a price insurance mechanism** viewed by participants as a "seller of last resort" [1] and as a source of "borrowed" allowances used to hedge against future scarcity.

- In experimental sessions, subjects used (primarily) lower price tiers to smooth[2] anticipated long-term scarcity and short-term risk of price spikes. The patterns of reserve use observed indicate that allowances available for direct purchase may have a value apart from what would be gained from selling additional allowances into the auction alone.
- In the professional simulations, we saw participants buy allowances to meet unfulfilled compliance obligations when close to a compliance deadline, even when market prices were well below the threshold tier prices.

- We conclude that **the unexpected uses of the APCR in our experiments and simulations justifies further research into different forms the APCR could take, including its possible combination with other tools** such as in-auction release of reserves or a hard price cap.

- There are still unresolved issues regarding the impact of California cap and trade market design elements. While our work focused almost exclusively on the electricity sector, **the prospect in 2015 of inclusion of vehicle fuels under the cap represents a huge shift in allowance demand for the sector**.
  - From 2013–2015, our refiner test portfolio goes from being responsible for its direct carbon emissions to the carbon content of gasoline (a multiple of nearly 7.0 times),
  - while its holding limits and limited exemption are based on its reported direct emissions.

[1] This is somewhat analogous to a central bank, (e.g., the Federal Reserve), which by standing ready to provide funds to financially-stressed institutions functions as a "lender of last resort."
[2] Buying in anticipation of future scarcity should allow participants to avoid being exposed to the large increases in scarcity during the "low hydro" periods, and so should tend to "smooth" some of the high price peaks.

II. Introduction

In January 2013, the California Air Resources Board (ARB) launched the world's first economy-wide greenhouse gas cap-and-trade system. Our study investigated the performance of two design features of the market-based compliance mechanism, which is part of the California cap on greenhouse gas emissions specified under the California Global Warming Solutions Act of 2006. In recognition of California's earlier experience with energy markets, and the experience of the European Union Emissions Trading Scheme (EUETS), ARB added two unique design features designed to minimize the prospects of excessive price movements and market manipulation: 1) holding limits, which restrict the number of allowances that any one of its 350 capped entities[3] may hold, and 2) a three-tiered allowance price containment reserve (APCR) to contain costs in the event of volatility of sustained high allowance prices.

California imports about one quarter of its power. Thus, the state is especially vulnerable to unexpected fluctuations in electricity supply, which will place upward demand pressure on allowances. The Board noted that while the reserve provided a measure of price certainty, its size was not unlimited, and that under extreme conditions it could be exhausted.

The lessons of decades of trading emissions in various markets taught us that while markets excel in providing flexibility and aggregating information, market design details matter for market performance (Carmona, Fehr, Hinz, & Porchet, 2010; Ellerman, Convery, & de Perthuis, 2010). The holding limits and APCR provisions have been criticized for potentially creating their own impediments to a robust California carbon market, and we endeavored to evaluate the validity of those claims.

We used a combination of laboratory experiments and more complex simulations (both with live subjects) to measure the effects that these design features have on the likelihood of manipulation, price volatility, the costs of compliance, incentives for early emission reduction, and the quality of price discovery. Specifically, we set out to investigate three key questions.

1. Are the holding limit provisions likely to prevent market manipulation or otherwise improve market performance? And what are the ancillary effects of these limits, if any?
2. Does the APCR help reduce price volatility or otherwise improve market outcomes?
3. How robust are our conclusions about these design elements to randomly occurring periods of especially high demand for emission allowances (as, for example, in a year with low availability of hydropower to the California market)?

Since any comprehensive program to reduce greenhouse gas (GHG) emissions must necessarily be quite complex, an investigation into one component of that larger policy has to balance the need for abstraction with the need to model "real world" conditions to allow conclusions about the California regulation. We supplemented our necessarily abstract laboratory investigations with market simulations designed to mimic more closely many details of the California energy market.

[3] Ranging from electric utilities (both private companies and public entities) to industrial companies (including refineries, cement, glass, other manufacturing, etc.) comprising approximately 350 entities starting in 2013 and expanding to include distributors of vehicle fuels and natural gas in 2015, culminating in more than 600 capped entities.

It was our intention, in the design of this research program, to provide careful and neutral analysis of elements of the current market design so that the market instrument can be as useful as possible in achieving the goals of the California cap on GHG emissions.

III. Objective: Understanding the Use of Markets in Emission Reduction Programs

It has now been amply demonstrated that emission markets can lower the costs associated with achieving a given reduction in levels of pollution (Ellerman, Joskow, Schmalensee, Bailey, & Monteiro, 2000; Stavins, 1998). These cost savings are not an economic abstraction. They represent real resources that, if not used for compliance, can be used to achieve other valuable social ends. Understanding how markets achieve this and what might go wrong helps us craft market institutions that have the best chance of achieving the greatest savings. Since Assembly Bill (AB) 32 sets ambitious targets relative to what other such programs have attempted to date, it is especially important that the emission market component of that program achieve the greatest measure of those potential savings. Toward that end, we briefly review the economic properties of these markets that will be critical to this investigation.

Economic theory strongly suggests that, *within the constraint imposed by a given environmental quality goal*, costs are minimized when flexibility is maximized: flexibility across space, across time, and across the identity of the emitter. It is easy to show that if the costs of controlling the next unit of emissions at one source are different from that cost at another source, then you can save money by allowing the sources to trade their emission reduction responsibilities. This is true across time, as well, but with the added complication that time itself changes the value of owning something, so we have to use discounting to put the present and future values on equal footing.

Like other cap and trade programs, the California regulation provides for a gradually declining limit or cap on emissions, in the California case declining 13 percent between 2013 and 2020. As the cap declines and fewer allowances enter the market, the increased scarcity of allowances puts upward pressure on prices. Traders in the market anticipate the future scarcity and may find it profitable to protect themselves against future increases by making additional cuts in current emissions and banking the unused allowances for future use. Such intertemporal trading helps smooth price increases and actually shifts greenhouse gas reductions forward in time, generating reductions earlier than would have occurred otherwise.

Sources of emissions face considerable uncertainty about the value of their actions to reduce emissions. Some of this uncertainty concerns actual present circumstances, such as the costs of emission reductions faced by other sources covered, and some concerns how things will play out in the future as the cap declines and the demand for allowances adjusts to external circumstances. Starting in 2013, the California program covers more than 350 entities (expanding to 600 in 2015) encompassing all sectors of the economy and many different types of organizations. It also envisions reductions that go beyond the experience in other GHG control programs. It makes sense that sources will take account of this uncertainty in their decisions about investment in emission reductions. They will expend resources to manage their exposure to risk and to gather information that reduces the magnitude of their uncertainties. Providing covered sources with cost-effective means of managing and reducing their risk can generate substantial savings on compliance costs.

A related, and sometimes under-appreciated, value of emission markets is their ability to gather the best information about the value of emission reductions and summarize that information in the market prices of current and future allowances. The price discovery function of markets supports the efficient allocation of resources dedicated to reducing emissions and reduces risk by providing signals about current and future values of allowances.

That emission markets are subject to substantial price risk has been recognized for some time. Some of this risk arises from the very design of the market institution itself. This has led to a large body of research about how emission market design might be changed to lower this price risk. Since this extra price risk arises because of the fixed nature of the supply of allowances no matter what the price, much attention has been paid to how to add flexibility so that the supply of allowances can adjust in a way that will not generate added environmental harm. The risk of higher-than-anticipated prices led the ARB to add the APCR to the emission market design, while the risk of oversupply leading to extremely low prices and curtailing emission reduction opportunities led the ARB to institute a slowly escalating price floor. The APCR's three-tier, post-auction sale of allowances has not been used in previous emission markets. How it contributes to controlling price risk will be one focus of this report.

We will show in our results how different elements of market design help or hurt the ability of participants to voluntarily allocate emission reductions to the cheapest reducers. We will also show that different design elements can change how well emission markets reduce the costs of risk and change the ability of the market to accurately summarize information about present and future circumstances. And we will examine the effectiveness of the design features intended to reduce unnecessary price variability.

III.1 Measures of market performance

Markets are not without their problems. A market participant may take advantage of periods when few allowances are available for trade, to manipulate price in a way that earns that trader profits at the expense of other participants in the market. A market participant who owns a large enough share of allowances may be able to use this dominant position to artificially raise the price and earn profits purely on the basis of the large share of allowances owned. Both of these forms of market manipulation operate by moving the actual price away from the true competitive market price. This unfairly imposes costs on other market participants and interferes with the price discovery and cost-saving function of the market. Market manipulation, if successful, imposes real costs on society.

This is not merely a theoretical concern. The California energy market suffered a very damaging episode of manipulation in 2001. And there are a number of documented cases in a variety of asset markets where periods of low liquidity, that is low asset availability in the market, have allowed price manipulation.

While the details of each instance of market manipulation, including the case of the California energy market, are different, studies of market manipulation have reached a strong consensus that improving the liquidity of a market is the best defense against manipulation (Pirrong, 2009; Hart, 1977; Jarrow, 1992; Putnins, 2012; Comerton-Forde and Putnins, 2009). Liquidity is a measure of how many units of the market asset (in our case allowances) are available to come into the market as the price rises. The greater this amount relative to the size of the market, the greater is market liquidity. It not only makes profiting from price manipulation more difficult and more risky, but also improves the overall competitiveness of the market, which can be expected to enhance price discovery. In light of this definition, a dominant position in a market may be seen as a particular form of low liquidity; there are few allowances for sale except by the owner in the dominant position.

i. *Liquidity:*

Since holding limits were added to the California allowance market to reduce the likelihood of market manipulation and because low market liquidity is generally viewed as making a market more vulnerable to manipulation, market liquidity is one important measure of market performance. Market manipulation generally depends on making a purchase or a sale that would likely not be profitable were it not for the effect on market prices. If there are lots of allowances available for sale, then manipulating price is at once more difficult and more risky. In our experiments, liquidity depends only on the number of allowances in holding accounts and on the number of allowances consigned to the auction. The auction itself is an important source of liquidity, but the base auction amount is identical across treatments.[4] We used the number of allowances in holding accounts as a key measure of liquidity and also looked at the relationship between the number of allowances in holding accounts and the number in compliance accounts.

The holding limit provisions of the California allowance market are unique among emission trading programs and have been the subject of considerable discussion (Linklaters, 2011). The purpose of the holding limit rule is to prevent one party from obtaining a dominant position in the market and to prevent market manipulation. This provision is accompanied by a rule limiting the share of allowances that may be purchased in any single auction. One of the key purposes of this study is to examine the effect of the holding limit provision on the liquidity and overall performance of the allowance market. The results of our investigation can help shed light on the issue of market manipulation and on the use of the holding limit to avoid it. If it is true that, contrary to expectations, holding limits may result in increased risk of certain kinds of manipulation, then there should be an effort to determine whether holding limits are a necessary and effective tool for preventing market dominance.

ii. *Price discovery:*

It is a well-established result that the price of emission allowances should be close to the "marginal cost of abatement." If we could line up all of the opportunities to reduce emissions from the cheapest to the most expensive, then the very most expensive unit of reduction that would be needed to meet our chosen environmental standard is the one we call the marginal unit of abatement. The cost of controlling this last unit should determine the market price of allowances. The reason for this is simple. Given the choice of the target level of emissions, we want to abate as cheaply as possible, so we want to use all of the abatement possibilities that are cheaper than that marginal unit. If there were a price on emissions equal to that amount, this is what sources would *choose* to do. And, if you have a market for trading allowances, sources would trade allowances in such a way that the cheaper to clean up sources would reduce their emissions and sell their allowances to the more costly to clean up sources. The market price that would result from this

[4] Some commentators on the California market have noted that one straightforward way to increase liquidity would be to increase the frequency of auctions. The Regional Greenhouse Gas Initiative (RGGI) program settled on quarterly auctions largely due to the uncertainty over the administrative burden that the auctions would represent. As experience with allowance auctions has grown, the rationale for low frequency auctions has faded. Increasing auction frequency seems to provide a low cost way of greatly enhancing market liquidity.

trading would be the cost of the last unit controlled. Any other price would leave the sources with a profitable trade, and result in higher than necessary costs of compliance. It follows that we should use the proximity of the market price to the marginal cost of control as one standard of market performance.

In our experimental sessions, since we constructed all of the costs of production and, consequently, the value of reducing emissions from each source in each period, it was possible to identify that marginal reduction unit and to calculate the exact marginal cost of reducing emissions. Given all of the uncertainties faced by subjects in the lab, we would not expect them to get the market price exactly right, but we can assert that, if one market design element results in market prices that are systematically further away from the marginal abatement cost, then that design element harms the market's price discovery function. We use the sum of the squared difference of market price from marginal abatement cost as a measure of market performance.[5]

iii. *Efficiency:*

Efficiency is closely related to price discovery. As the price moves away from the marginal abatement cost benchmark, businesses are receiving incorrect signals about which units of abatement are worthwhile and which ones are not. This results in cases where some production units worth running do not run and some units not worth running do run. In both cases, society loses valuable resources either as costs that are higher than they should be or benefits that are lower than they should be. We measure efficiency as the net value to society of a given unit of production; that is, the price of the output minus the cost of operation minus the cost of GHG emissions.[6] Efficiency is easily the most widely accepted measure of market performance among economists.

iv. *Price volatility:*

Uncertainty adds to the cost of doing business. Added volatility in allowance prices, other things equal, tends to increase the cost of complying with emission reduction requirements. It has been well-understood for quite some time that emission markets are subject to high price volatility due to the fixed supply of the asset being traded (Weitzman, 1974). As a result, all cap and trade programs, including California's, have had design features such as APCR, banking and use of offsets, features intended to control excess price volatility. Price volatility, then, should be one of our metrics of market performance. We measure price volatility as the standard deviation of auction closing prices in each session.

IV. Complementary Methods: Applying the Lessons of Experimental Economics to Market Simulations

Experimental economics can complement conventional mathematical economic models. Simulation makes it possible to create a more detailed representation of market responses to complex policies than allowed by conventional tools. As described in Appendix 1, simulation exercises relax some of the constraints in economic models, relying instead on the actual behavior of individual subjects playing simulation roles to ground the results. A frequent outcome of experiments and simulation

[5] Another possible measure of difference is the mean absolute deviation. This measure leads to the same conclusions in our experimental tests.

[6] The social cost of GHG emissions is assumed to be the dynamically efficient price, which is the marginal cost of reductions, in these sessions $14.75.

exercises is unanticipated behavior and consequences. These results may yield insights unattainable from mathematical models alone.

To provide some assurance of the robustness of the experimental results to the complexities of the actual California market, we have supplemented the student-based experiments in the controlled laboratory environment with a set of "professional" simulations where the allowance auction and price containment sale institutions from the lab are linked to our adaptation of a standard emission market simulation program, ET-Sim (an emission trading simulation tool developed by U.S. EPA for training purposes).

Our simulation participants were experienced in energy markets, including some in the California energy markets specifically. We believe that the concurrence between the results of the laboratory and the professional simulations should provide some confidence in the external validity of the more abstract, but carefully controlled, laboratory sessions.

IV.1 Structure of Basic Auction Experiments

Our laboratory sessions were implemented with the University of Virginia's auction software (Veconlab) and were designed to mimic key features of the California allowance market, maintaining details that would likely affect subject responses to variables of interest: holding limits, price containment reserve, and demand spikes. Each session comprised a sequence of 12 periods (or rounds) of "permit" markets[7], where the permits (allowances) were required for production of an "output" good that is sold for a price that is announced at the start of each round.

The number of allowances allocated either to the base auction quantity or to grandfathered (free) allowances declined by 5 each period from 69 in the first period to 14 in the last. This schedule was published in advance and available for reference by the subjects throughout the session. The allowances could be banked into future periods in either a holding account or a compliance account, with the compliance account allowances no longer available for any use other than to meet compliance obligations for that subject. "Compliance periods" were defined such that true-up occurred every third round, at which time any deficits in allowances relative to requirements were tripled.[8] Allowances had no redemption value after the end of the last period.[9]

Each session had either a tight holding limit of 7 for low emitters and 14 for high emitters, or just greater than the auction purchase limit for each subject; or a loose holding limit of 13 and 26 for low and high emitters respectively, or twice the auction purchase limit. The tight holding limit, at just greater than the purchase limit, was still greater than any routine contemporary compliance needs. Loose holding limits allowed subjects to hold more than twice as many permits as needed to cover maximum possible production of 5 during any period. The holding limit is the treatment variable for

[7] The language that we used in the laboratory sessions was neutral as to the purpose of the market exercise referring to "permits" rather than "allowances," and no mention was made of emission markets.
[8] In the last period of a session, the penalty was calculated as three times the closing price of the auction in that period and was subtracted from earnings.
[9] This may result in some differences between what we observed in our experiments and what we would have observed in sessions with many more rounds, due to permits losing their value at the end of the 12th period. There is no easy way around this endpoint dilemma, but it may not be as significant as it first seems since the same thing is true in all treatments. So for this effect to be important, it would have to have a substantially stronger effect in one treatment than the other. There is nothing about our results that indicates such an asymmetric effect.

the first set of sessions. Since we were examining the effect of tight holding limits, the treatment with the loose holding limits was our baseline treatment, and the tight holding limit case was the alternative treatment.

An APCR of 36 allowances is established at the start of each session. The reserve is divided into three tiers of 12 each. Allowances in these tiers will be sold at $23, $28, and $33 respectively. Rules for the APCR sale match those of the California APCR. Unsold allowances are carried forward at the same price tier, and a sale will occur in each period so long as there are any allowances left in any tier.

Each session had 12 subjects, who were students at the University of Virginia. Subjects were paid a $6 show-up fee plus their earnings during the sessions, with an average total of around US$40. All subjects had 5 "capacity units," each of which could produce 1 unit of output per period. Subjects were assigned one of four different roles, which stayed the same throughout the session. Half were "low users," requiring one allowance to operate each capacity unit, and half were "high users" requiring two allowances per capacity unit.[10] All subjects received a free allocation that declined after each compliance period, reflecting the gradual reduction in free allocations in the actual program. Free allocations were zero in the last three rounds. High users received an allocation twice as large as that for low users. Half of the subjects in each of these groups were required to consign all of their free allocations of allowances to auction and half were not required to consign any but had the option to consign.

The costs of operating capacity units were randomly assigned each period and were uniformly distributed in the range of $4–20 per unit for high emitters and $14–22 for low emitters. The maximum value of a permit for a given capacity unit is equal to the output price minus the cost of operating the unit, all divided by the number of allowances required to operate the unit.

For all sessions, there was a two-thirds probability of a $35 output price and a one-third probability of a $60 output price. The higher output price induces a high demand for allowances and mimics periods of high emissions with accompanying high allowance demand, such as might occur in a drought year with low availability of hydroelectric power. During "normal" periods, low users were required to run 3 of their 5 units while high users had no "must-run" requirement. During "low hydro" periods, the must-run requirement for low users rose to 4 and high users had a must-run requirement of 1 unit.

Periods comprise a sequence of information disclosures and decisions in which the subjects are asked to:

1. Observe the price of the output for that period, any must-run requirements, and one's own current holding and compliance account totals.
2. Choose how many allowances to consign to the auction in addition to any required amounts.
3. Choose how many allowances to move from holding account to compliance account.
4. Submit bids in auction for any number up to the current purchase limit and view auction results.
5. Post offers to buy and sell in the post-auction spot market and view own spot market results.

[10] One can think of this as reflecting the difference between, say, natural gas and coal-fired generators.

6. Post offers to buy in price containment reserve sale (if APCR not exhausted) and view results of APCR sale.
7. Decide how many units to run above any must-run requirement.
8. Observe earnings, account totals, cash balances, and any deficits relative to compliance requirements.

The auction in each period had a uniform-price, sealed-bid format. The number of allowances for sale included the base auction quantity, which declined each period, plus any allowances consigned to the auction. The number of bids that a subject could submit was determined by the purchase limit or the holding limit. A bid that would satisfy the purchase limit for a subject was not allowed if winning that bid would result in a violation of that subject's holding limit. Bidders with winning bids paid only the closing price of the auction for any accepted bids.[11]

The spot market was a single-round, limit-order call market where each subject could post buy and sell orders. A buy order consisted of a quantity offered and a maximum price to be paid, and a sell order was a quantity along with a minimum acceptable price. The bids and asks were ordered to determine the market clearing price. All bid and ask orders were filled at that spot market price, and bidders were informed of any changes in their cash and permit holding accounts.[12]

We tested hypotheses about the four key measures of market performance described in Section III.1. These hypothesis tests are listed in Table 1. All four tests focus on the issue of whether, in our experimental setting, tight holding limits can be said to reduce the effectiveness of the emission market. For each hypothesis, we use the Wilcoxon rank sum test (also known as the Mann-Whitney test) for differences in the key measures of market outcomes: liquidity, efficiency, price discovery, and price volatility. In each case we present the significance (or p-value) of the Wilcoxon rank sum statistic. The significance can be interpreted as the likelihood, given the data from the experiments, that the two treatments were not different. So, a p-value of 0.05 means that we can be 95% sure that the differences between the treatments are not just due to random variation, but rather represent different outcomes.

Table IV-1: Hypotheses Tested

	Null (H_0)	*Alternative (H_1)*
Price discovery	*Tight holding limits do not increase the difference between the observed price and the efficient market price*	*Tight holding limits increase the difference between the efficient price and the observed price.*
Efficiency	*Tight holding limits do not reduce efficiency*	*Efficiency is lower under tight holding limits*
Price volatility	*Tight holding limits do not increase price volatility*	*Price volatility is higher under tight holding limits*
Liquidity	*Holding limits do not reduce market liquidity*	*Liquidity is lower under tight holding limits*

[11] Half of the sessions used the first rejected bid to set the price, the rule in the RGGI auction, and half used the lowest accepted bid, which is the rule in the California auction. No differences in outcomes between these closing rules were observed.
[12] Ties are determined by a random device.

IV.2 Structure of Professional Market Simulations

The goal of the professional simulations was to build upon the insights gained through the laboratory sessions into the incentives inherent in the structure of the California program and determine how they apply in a more complex environment. To that end, we endeavored to broaden the alternatives from which companies may choose to comply, as well as make the affected parties more closely resemble actual covered entities.

We also relaxed the restrictive standards of economic experiments in order to better represent the California program design features. So we looked for patterns of differences in behavior, either between different roles in the same circumstances (by evaluating cross-sectional data) or between actual choices and predicted optimum results (through comparison of model-based projections). Where we identified apparent differences between expected or projected behavior in a particular role, we conducted a post-simulation survey in order to characterize the motivation of the participants.

For a more realistic decision set, the professional simulation design used experienced subjects addressing a structural representation of the California market. The objective was to provide a realistic context for participants to choose among alternative means of compliance with carbon reduction requirements, weighing their costs and risks as reflected in behaviorally-determined prices of environmental instruments. The participants were asked to assume the general role of portfolio manager, taking responsibility for a representative entity's compliance under the California carbon cap and trade program mandated by AB 32.

The configuration of the simulation rounds depended upon 1) the number of parties available, and 2) the number of years included in the simulation. For 12-person simulations (which were two of the seven simulations performed) we set up two-person teams, where one player adopted the role of compliance planner and their partner was designated to be the team's emissions trader. In six-person simulations, each participant adopted both roles.

For all seven of the professional simulations, the specified time interval was Compliance Period 1 (CP-1), which for calendar purposes was assumed to be auctions conducted in calendar 2013–2014. While the program includes quarterly auctions, for purposes of simplification we specified three auctions per year, with compliance true-up scheduled for the end of the sixth auction round. The actual program schedule set by ARB beyond CP-1 includes CP-2 (2015–2017) and CP-3 (2018–2020.)

The professional simulation integrates the transaction functions already employed in the Veconlab application with the compliance planning tool ET-Sim, offering additional compliance options (e.g., offsets, energy efficiency project development) as well as financial analysis capabilities. For the electric sector focus with this cycle in ET-Sim, the participants evaluated one of six "representative" California-based electric generation portfolios based on data from 2010.

- Two investor-owned utilities: one low-emitting, one high-emitting.
- Two publicly-owned utilities: one low-emitting, one high-emitting.
- Two merchant generators: both low-emitting.

Each portfolio was further subdivided into three parts: 1) "marginal" generation, for which participants buy California carbon allowances in the auctions and spot market, 2) "committed" generation, for which participants were responsible for allowance account management for the remainder of their fossil portfolio, and 3) non-emitting generation, which is unaffected. In the integrated environment, participants managed the "marginal" component of representative portfolios, and adopted compliance obligations appropriate for their role type, including allocation specifics and ARB holding limits.

"Marginal" generation for the block of all six portfolios was identified by estimation of the carbon allowance shortfall in the last period if the generation mix is preserved, i.e., a generator might be forced to reduce output or otherwise change operating practices if it is among the units competing for a declining block of allowances. Thus, we used our modeling of Western power markets under specified baseline conditions to determine the compliance obligation on affected sources in the electric generation sector.

For example, in the baseline case, the reported 11% decline in California carbon emissions from 2008 to 2010 just exceeds the 10% cut in emissions from 2008 levels that was specified by ARB to determine the 2012 allocations baseline, so a reasonable first approximation of emissions reductions required from generators in CP-1 is simply the decline in cap over the period (a 2% decrement each year, or 3% cumulative by 2014). We used that 3% decline in carbon as the definition of "excess demand" and use the shortfall to define a block of generation at risk, and evenly distribute the total carbon "generation at risk" among all portfolios at their 2010 generation mix. We asked participants to bid for allowances to cover these marginal units in their portfolio, or to develop an alternative compliance plan.

In contrasting the steps in the professional simulation cycle with those of the laboratory sessions, the development of a compliance strategy (in ET-Sim) was the first step, followed by a series of activities in Veconlab that are identical to the laboratory sessions (see steps 1–6 in section IV.1 above), followed by another series of steps that required the user to:

1. Make unit commitment decision; and
2. Purchase replacement wholesale power; and/or
3. Buy offsets; and/or
4. Invest in energy efficiency projects.

The intent of creating the professional simulation was to adapt the lessons from the experiments while adding these program elements, which were excluded from the laboratory sessions and were thought to be critical to formation of a robust California carbon market. These elements included the ability to make offsets purchases, invest in energy efficiency projects, import power or fuel from out of state, and for the longer term simulations, invest in lower-emitting plant and equipment (e.g., renewable energy projects). At a minimum, adding to the potential range of choices available to participants in our professional sessions increased their "degrees of freedom," which likely influenced the simulation results.

i. Offsets markets:

In an efficient carbon market, offsets with identical compliance value to carbon allowances will trade at or near the same market price as allowances. Contributing factors to a higher-risk offset market in California include usage restrictions (8% of a participant's total compliance obligation) and the ARB's requirement that would make the buyer liable for invalid offsets. In order to price offsets in ET-Sim, we identified three general categories of offsets with different types of market risk.

- Guaranteed delivery of offsets, priced at 85% of the current carbon dioxide (CO_2) price
  - Lowest risk (primarily counterparty risk) and thus lowest discount to allowance price.
- Unit contingent, priced at 70% of the current CO_2 price
  - Medium risk category with volumetric and counterparty performance risk.
- Pre-compliance spot, priced at 55% of the current CO_2 price.
  - Highest risk (regulatory risk), highest cost (economic, contractual, administrative), and offered at greatest relative discount to allowance pricing.

At settlement, volumes delivered, net pricing, and regulatory performance issues were factored in and deliveries and payments likely differed from expectations.

ii. Wholesale power:

In ET-Sim, participants were also offered choices between power purchase options that could be selected to replace an existing higher-emitting generator in their portfolio. Thus, the portfolio manager's compliance strategy must weigh the 1) cost-effectiveness of reducing their emissions compliance obligation, and 2) benefit of locking in the current CO_2 allowance price against their expected prices. For CP-1 simulations, two wholesale power options were offered.

1. *An existing natural gas-fired steam generator* at an implied market heat rate of 10,250 Btu/kWh and an emissions rate of 0.541 per metric ton CO_2/MWh, and
2. *Unspecified gas-fired power purchases* at a nominal heat-rate of 8,260 Btu/kWh and an emissions rate of 0.436 per metric ton CO_2/MWh (an ARB-assigned default emission rate for power delivered from unspecified sources).

The price of wholesale power in ET-Sim is updated after three auction rounds using the closing price from the last auction.

iii. Energy efficiency – electric sector:

The professional simulation allows market participants simulating the portfolios of electric distribution utilities to reduce electric load by either:

1. Replacing more expensive energy purchases with less costly energy efficiency projects; this benefit only occurs if energy/load reduction costs are less than replacement energy costs (EE $/MWh < RE $/MWh),
2. Reducing operating costs by lowering the amount of offsets or allowances needed; this benefit only occurs if it is less costly to install an EE project than it is to operate the marginal generation that the project would replace (EE $/MWh < [Production Costs + cost of allowances and/or offsets necessary to cover the emissions from the marginal generation])

With the addition of a price on carbon, utilities with higher-carbon intensity are likely to see the value of energy efficiency increase more rapidly than will lower-carbon intensive utilities. For greater detail on the structure of ET-Sim, see *Appendix 2*.

V. Results – Holding Limits

The California cap and trade regulation establishes limits on the number of allowances that may be held at any time by any market participant. This holding limit provision was established to prevent one entity from obtaining a dominant market position and using that position to manipulate the market. The regulation also establishes a limit on the number of allowances that can be purchased in a given auction, and provides for the establishment of compliance accounts from which allowances may only be withdrawn to cover actual emissions but may not be sold directly or consigned to the auction for sale. Protecting confidence in the market for both the public and participants was a key concern in settling on these particular market design features. The prospect that one firm may obtain a position sufficiently strong to profit from controlling the supply of allowances is obviously not a purely theoretical concern to California, which experienced just this problem with Enron in 2000, and given that there will be a mix of participants in the emission market from the quite small to the very large.

The use of holding limits for accomplishing this end was one of the recommendations in an analysis of market manipulation provided to the ARB in its deliberations on market design.[13] Holding limits have been applied in other contexts, primarily in derivatives markets. ARB followed the holding limits precedent from the Commodity Futures Trading Commission (CFTC) rule[14] designed to address systematic risk in commodity futures markets. Purchase limits have also been used in other markets, such as the market for U.S. treasury bonds and the Regional Greenhouse Gas Initiative (RGGI) allowance market. The compliance account provision has, to our knowledge, not been used before.

Because holding limits restrict the range of options by market participants, they will, necessarily, also restrict the opportunities that participants have to undertake transactions that lower their compliance costs. It is one of the key functions of the research reported here to assess these costs and to evaluate any evidence from our experiments on how holding limits might change the likelihood of market manipulation.

Our laboratory and simulation results both indicate that, as presently applied in the California cap and trade regulation, holding limits may have a substantial effect on market performance and may actually have the unintended effect of increasing the probability of market manipulation. Aside from the costs of reducing compliance flexibility and limiting opportunities for risk management, tight holding limits reduce market liquidity, which analysts generally consider to be the first line of defense against those attempting to manipulate a market for profit. We found that holding limits contributed to higher price variability, less effective price discovery, lower efficiency, and ultimately reduced banking, which translates into delayed reductions in greenhouse gases.

In endeavoring to address the possibility of undue market influence exercised by large players, ARB focused on the size of positions held. To that end, the rules set an absolute maximum on the total number of allowances that may be held by any company or its affiliates. The holding limits apply to all allowances combined held in any of three types of accounts maintained by the company (its

[13] See the market analysis prepared for the Western Climate Initiative Markets Committee: *"Report on Holding Limits,"* J.H. Harris, University of Delaware, May 6, 2010. http://www.westernclimateinitiative.org/news-and-updates/108-markets-committee-invites-comments-on-holdings-limits-report

[14] The CFTC position limits rule has been stayed by court order.

General Holding, Compliance, and Exchange Accounts). Purchases that would cause holding limits to be exceeded will not be approved by the Executive Officer. Penalties may be assessed when the holding limit is exceeded.

In recognition of the structure of the market and the differential value of allowances by vintage, ARB defined two sets of holding limits, for 1) allowances eligible for compliance use in current years, and 2) allowances for use in future years. Sticking close to the precedent established in the CFTC position limit rule, ARB set the holding limit for allowances eligible for compliance use in the current year as:

*Holding Limit = 0.1*Base + 0.025*(Annual Allowance Budget – Base)*
where:

"Base" =	25 million metric tons of CO_2e.
"Annual Allowance Budget" =	number of allowances issued for the current budget year.

The resulting holding limit for current allowances by year is in *Figure V-1* below. The maximum number of California greenhouse gas allowances that may be held at any time is thus independent of the size of a covered entity's allowance requirements. At nearly 6 million metric tons for the first two years of the program, the holding limit is greater than the annual emissions of the vast majority of affected companies, a first pass indication that they should have little difficulty maintaining reserves of allowances needed to meet their obligations irrespective of market uncertainties. However, for those companies whose annual compliance obligation does exceed the holding limit (and some exceed it by a significant margin), there is question regarding their ability to hold sufficient allowances to comply.

To retain the link between a firm's ability to satisfy holding limits requirements and meet its compliance obligations, ARB provides a "limited exemption" from holding limits for allowances that are acquired by a covered entity and placed into its compliance account (rather than its general holding account). While the limited exemption undoubtedly provides latitude for larger firms to acquire and bank additional allowances, some commenters have expressed concern that market liquidity will be reduced as a result, because once placed into compliance accounts allowances cannot be withdrawn again and made available for trading.

Furthermore, to prevent dominant positions from emerging out of any single auction, ARB has also set a ceiling on the total number of allowances that any one bidder may acquire at auction, varying by type of buyer. For auctions of current-period allowances, electric distribution utilities may purchase no more than 40 percent of allowances sold at any one auction, non-utility covered entities may purchase no more than 15 percent, and other entities are limited to 4 percent of allowances sold at auction.

Figure V-1: Cap and Trade Holding Limits without the Limited Exemption

Year	Holding Limit (million metric tons)
2013	5.945
2014	5.868
2015	11.738
2016	11.435
2017	11.135
2018	10.833
2019	10.533
2020	10.230

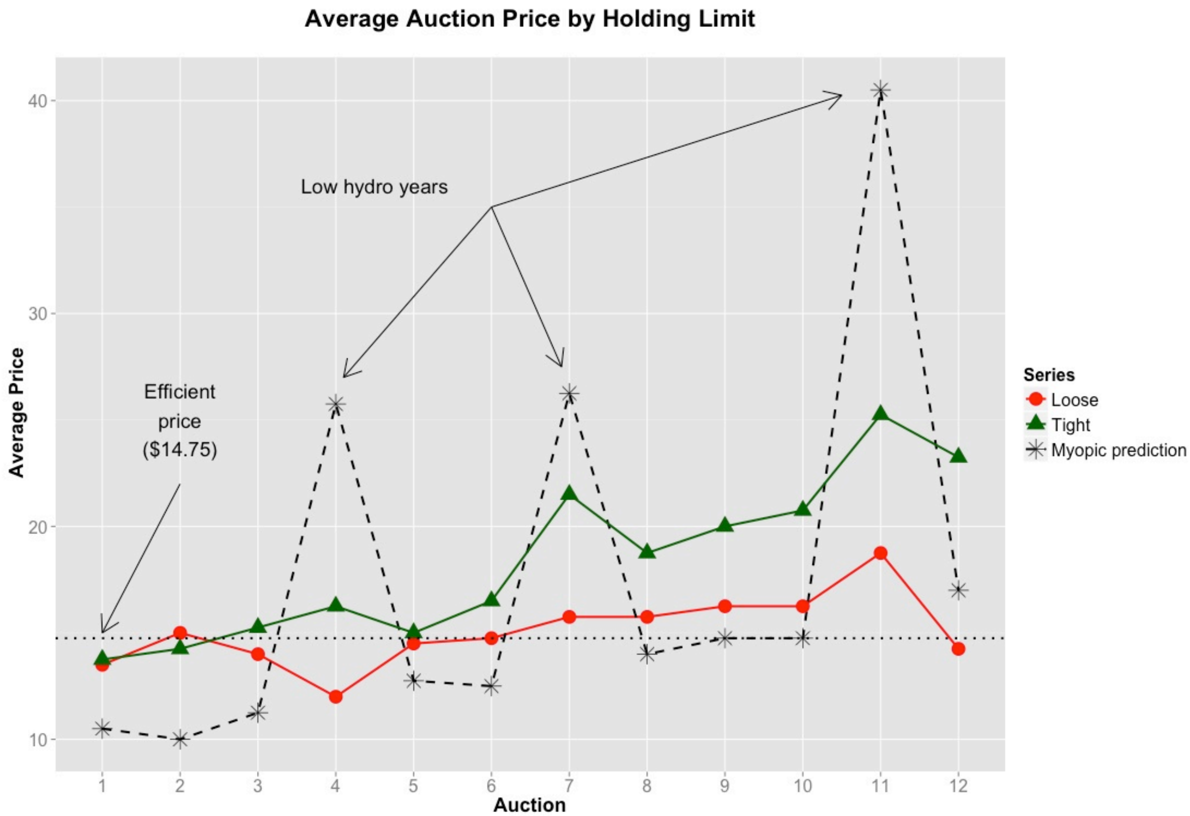
According to our static analysis, while leaving small to medium-sized entities unaffected, ARB's holding limits force a few of the largest players to place 50–70% of their allowances into compliance accounts. In 2011, in the electricity sector there were only three companies in this position: Pacific Gas & Electric, Southern California Edison, and Los Angeles Department of Water and Power. While few in number, taken together these companies represent a significant share of the market, accounting for 74% of total allowances allocated to the electricity sector.

Finally, in addition to ARB holding and auction purchase regulations, investor-owned utilities (IOUs) are also subject to the annual allowance purchase limits specified by the California Public Utilities Commission (CPUC) in their most recent Long Term Procurement plan. The Commission's limits determine the maximum quantity of compliance instruments (including both allowances and offsets) that regulated entities would be allowed to purchase in the current year for all vintages of allowances combined.

V.1 Student Laboratory Sessions – Effect of Holding Limits

Definition of holding limits in the laboratory sessions: As indicated earlier, each laboratory session had either a tight holding limit or a loose holding limit, with tight holding limits being just greater than the auction purchase limit for each subject and loose holding limits being twice the auction purchase limit. The holding limit is the treatment variable for this set of sessions. Since we were examining the effect of tight holding limits, the treatment with the loose holding limits was our baseline treatment, and the tight holding limit case was the alternative treatment.

Figure V-2: Average Auction Price by Holding Limit



Detailed discussion of results – holding limits – laboratory sessions: Figure V-2 effectively summarizes the results from the laboratory sessions. The dotted line and the dashed line represent two extreme theoretical predictions. The dashed line gives the predicted market-clearing price in each period of a session under the assumption that subjects only take into account the current period cap, costs and output prices. It represents the "perfectly myopic" competitive equilibrium where subjects act as if only the current period matters and do not account for any benefits they would receive from saving allowances from one period to the next.

The dotted horizontal line ($14.75) shows the dynamically efficient price, which is the value of the marginal control cost combining all 12 periods in the session and represents the "perfect foresight" case, where subjects correctly anticipate all future scarcity and completely balance their allowance scarcity across all periods. If the market price exactly matched this price each period, this would maximize the value to society of the trading program because it would equate marginal emission reduction costs across both space and time, and would minimize the cost of price risk due to periodic shifts in demand.

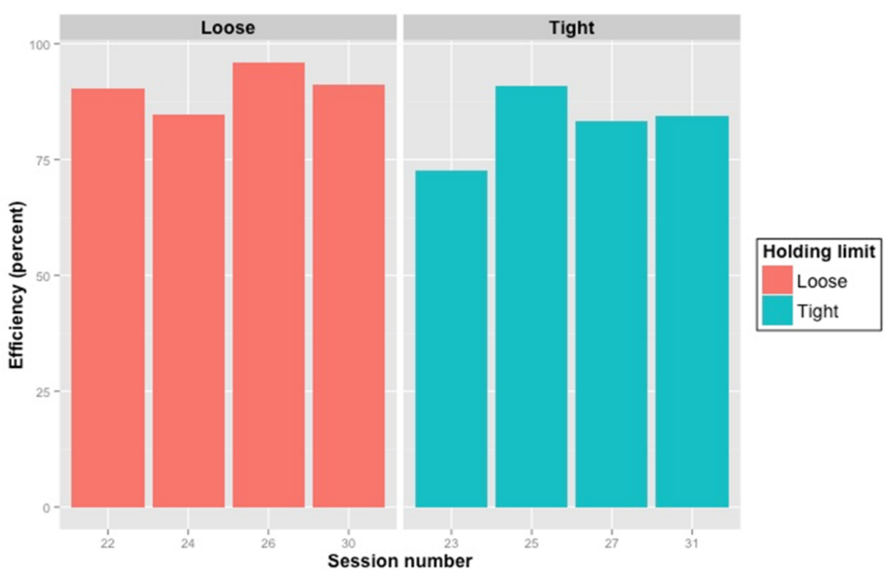
Market participants who anticipate future scarcity can profit by accumulating allowances early and saving them for later periods, when they will have greater value as the cap declines. This would cause the observed price in the lab to be smoother and flatter than the predicted myopic clearing price. The more the market induces traders to smooth and flatten the price path away from the myopic price path toward the efficient (flat) price path, the closer the market comes to the least-cost pattern of emissions.

Smoothing the demand for allowances in this way has the salutary effect of generating earlier reductions in emissions than would occur otherwise. Anticipation of future allowance scarcity raises the true current cost of GHG emissions in the present. Incentives for early investment in new low GHG technologies rise, in turn.

Hypothesis test 1: Price Discovery – We can reject the null hypothesis that the mean squared difference between the actual price and the dynamically efficient price is the same in sessions with tight holding limits as in sessions with loose holding limits (p = 0.045).

We conclude from this that tight holding limits are likely to reduce the effectiveness of price discovery.

Figure V-3: Average Session Efficiency: Tight vs. Loose Holding Limits



Turning back to *Figure V-2,* the remaining two lines represent the average across sessions of auction prices for each auction period separated by whether the holding limit was loose or tight. The difference between these two lines dramatically illustrates our first result, that tight holding limits tend to drive the auction price further from the dynamically efficient price path, reducing the ability of traders to prepare for high demand periods, and limiting trader ability to prepare for long-term future scarcity by making additional reductions in the present and banking the unused allowances for use in later years. Price discovery is less effective with tight holding limits than with loose holding limits. We use the mean-squared difference between the actual prices and the dynamically efficient price to measure the effectiveness of price discovery in each session.

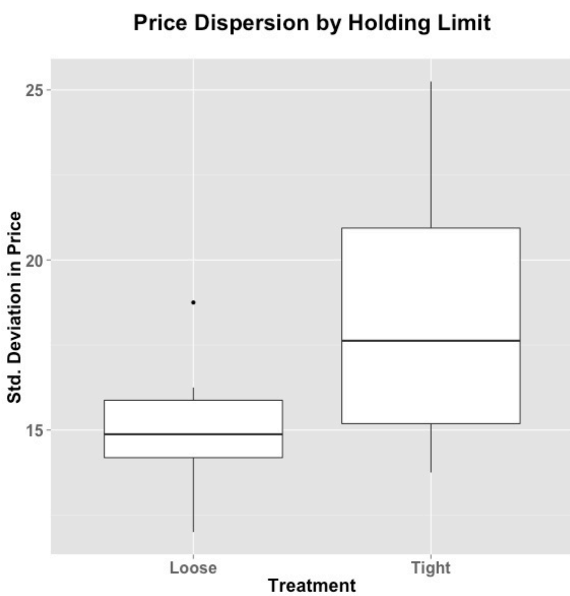
The difficulty that lab subjects have in shifting their stock of allowances across time periods is reflected in lower efficiency for tight holding limits relative to loose holding limits. *Figure V-3* shows session efficiencies for the eight sessions divided by the holding limit treatment. The sessions in the left panel (Loose) tend to be higher than in the right panel (Tight). This difference is statistically significant.

Hypothesis test 2: Efficiency – We can reject the null hypothesis that session efficiency is higher or equal under tight holding limits than under loose holding limits (p = 0.043) .

We conclude from this that efficiency is likely to be lower under tight holding limits.

Given the greater difficulty that subjects have smoothing costs, it should not come as a surprise that prices are also more volatile. *Figure V-4* indicates that the standard deviation in price is higher in tight holding limit sessions than in loose holding limit sessions, averaging 4.2 and 2.1 respectively. Once again, we reject the null hypothesis that the variances are equal across treatments in favor of the alternative hypothesis that variation is greater in tight holding limit sessions.

Figure V-4: Price Volatility: Tight vs. Loose Holding Limits



These losses of efficiency, price discovery, and price stability might be a price worth paying for a substantial reduction in the likelihood of the allowance market falling prey to damaging bouts of market manipulation. We now turn to address this central issue directly by measuring the effect of tight holding limits on market liquidity. Our focus will be primarily on the size of holding accounts and, to some extent, on the relative sizes of holding and compliance accounts. In our lab sessions, the main differences in liquidity arise from differences in holding accounts. Consignment to the auction can contribute to liquidity as well, but we did not observe statistically significant differences in the levels of consignment between holding limit treatments.

Hypothesis test 3: Price volatility – We can reject the null hypothesis that price volatility is lower or equal under tight holding limits (p = 0.041).

We conclude from this that prices are more volatile under tight holding limits.

Figure V-5 shows density plots for the size of holding accounts with the dashed (green) line representing tight holding limit sessions and the solid (red) line representing loose holding limit sessions. These distributions are quite different, showing substantially higher numbers of allowances available for trade in the loose sessions than in the tight sessions. *Figure V-6* provides another view of this difference. This is a scatterplot of the number of allowances in holding accounts by the total number of banked allowances, so it is a way of viewing holding accounts relative to non-tradable compliance accounts. As the size of the total bank rises, we see the number of allowances in holding accounts rises as well. There is a striking difference in the shares of allowances in holding accounts between the two treatments. The lines represent the line of best fit through the points for each treatment and the shading shows the 95% confidence interval for that line.

The holding accounts are distinctly larger under loose holding limits. Indeed, total banking never reaches 100 allowances under tight holding limits. Even for very small banks, where subjects are unlikely to be constrained by the holding limit, the holding accounts are smaller under tight holding limits and the share of allowances in holding accounts is higher. The differences apparent in this graph reflected a statistically significant difference between the experimental treatments.

Hypothesis test 4: Liquidity – We can reject the null hypothesis that allowance banking is greater or equal under tight holding limits ($p = 0.021$).

We conclude from this that market liquidity is likely to be lower under tight holding limits.

Figure V-5: Market Liquidity: Holding Account Volume – Tight vs. Loose Holding Limits

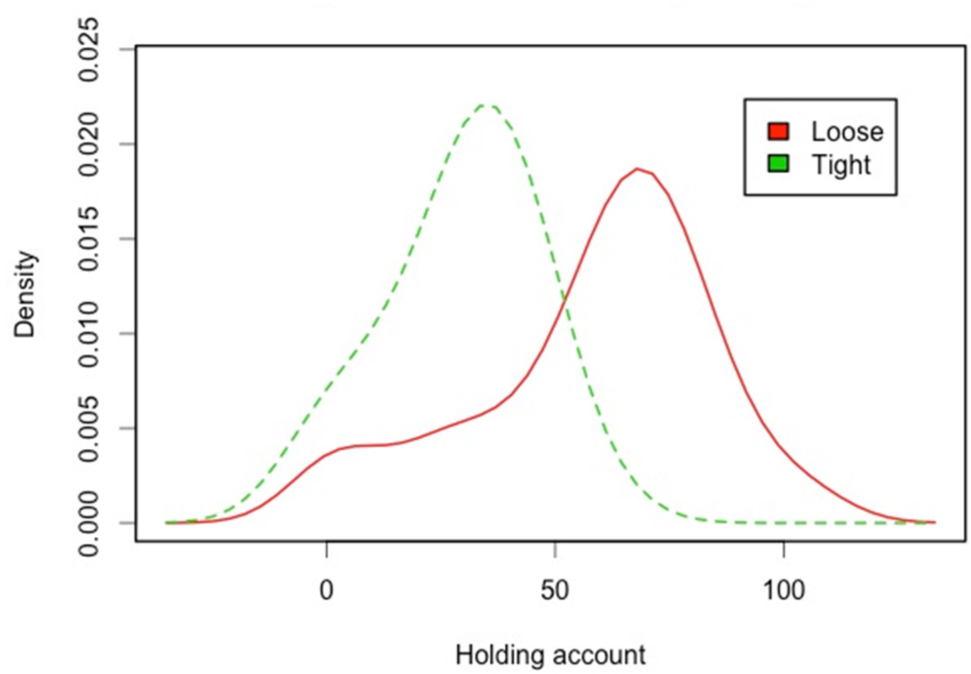
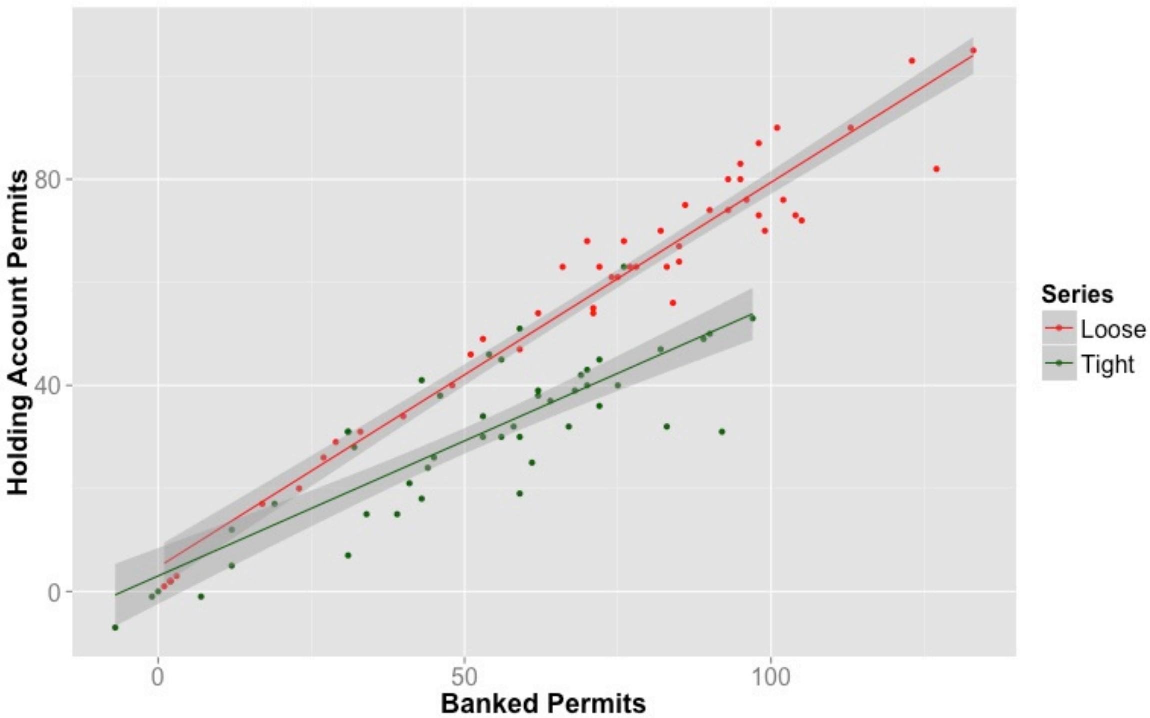


Figure V-6: Market Liquidity: Holding Account and Compliance Accounts – Tight vs. Loose Holding Limits



These results give strong weight to the argument that tight holding limits may impose considerable costs on the economy by interfering with the valuable function of the allowance market for allocating emission reduction responsibilities between sources and across time. Firms in the market face higher risk and receive less accurate signals concerning where resources should be invested in emission reduction activities. It is especially noteworthy that tight holding limits tend to operate directly counter to one of the key defenses against market manipulation. We conclude that tight holding limits make the allowance market *more* vulnerable to some of the commonly discussed strategies for market manipulation. What we do not yet know is whether any protections provided by limiting the share of allowances owned by one or a few large providers outweighs the costs and risks imposed by the holding limits themselves. In the least, it does seem clear that the holding limits should be not tighter than necessary for this function and that minimizing the damage from protections against manipulation by dominant participants should be a continuing priority for the ARB.

V.2 Professional Simulations – Effect of Holding Limits

The advantage of proceeding from experiments run with students to sessions staffed with professionals is that one can first identify the key elements of the cap and trade program's incentive structure in the simplified architecture of the lab and then see if the results hold up under the greater complexity of our simulations. Of the seven professional simulations conducted, five had no holding limits, while two were faced with ARB holding limits. Initial analysis of holding limits indicated five of six portfolios represented in the study would be compelled to transfer a substantial majority of their allowances into their compliance account. Based on this analysis, we conclude that allowances held by large entities will be predominantly held in compliance accounts, which means lower liquidity in the California carbon allowance market.

For the professional simulations, we addressed the following series of questions regarding the impact of holding limits.

- Who is most affected by holding limits? Are its effects more pronounced on particular classes or sectors?
- How do holding limits contribute to reduced banking, i.e., simply lower share of allowances in holding accounts?
- Is the impact of holding limits uniform over time, or are there critical decision points? I.e., does the timing of contributions to compliance accounts matter?

Definition of holding limits – professional simulations: Entities with annual emissions near or above the ARB holding limits shown in *Figure V-1* must pay close attention to managing their positions in order to avoid exceeding the current allowance holding limit including their limited exemption. In addition, advance allowance holding limits apply to future allowance vintages. Moreover, for investor-owned utilities, which are rate-regulated entities, the potential overlap of the California Public Utilities Commission purchase limits may also affect purchase strategies. Under these circumstances, managing the transfer of allowances into the compliance account while building an allowance bank and keeping some current allowances available for trading will become a complex exercise, particularly for companies with annual emissions above the holding limits.

To accommodate the structure of the professional simulation, position limits must be applied at two levels: 1) in the transactions rounds (in the Veconlab segments), where allowances are purchased to cover "marginal" generation, and 2) in the compliance planning activities (in ET-Sim) where allowances purchased to cover marginal units are combined with allowances accrued on behalf of "committed" capacity to complete the compliance obligation. Since marginal generation represents a small share of total emissions in the initial compliance periods for the transaction segments, holding limits are applied to their auction purchases in Veconlab just as they are in the laboratory sessions.[15]

Holding limits are calculated for six portfolios meant to be representative of affected sources under the program. For the electric generation sector cases, the portfolios included four utilities (two investor-owned and two publicly-owned) and two merchant generators. For the compliance-planning segment of the professional simulation, portfolio managers engage in allowance account management in a three-step process.

First, on an annual basis, they may transfer a portion or all of their purchased allowances into their compliance accounts (all transfers into ET-Sim are regarded as transfers to compliance accounts). The purchased allowance transfer partially fulfills the annual requirement that participants transfer and hold at least 30 percent of the preceding year's emissions in compliance accounts. In addition, the closing price of the most recent allowance auction is used to a) determine a transfer price for the committed capacity allowances transfers made in that period, and re-set values of b) potential purchases of offsets and c) wholesale power purchase options in ET-Sim.

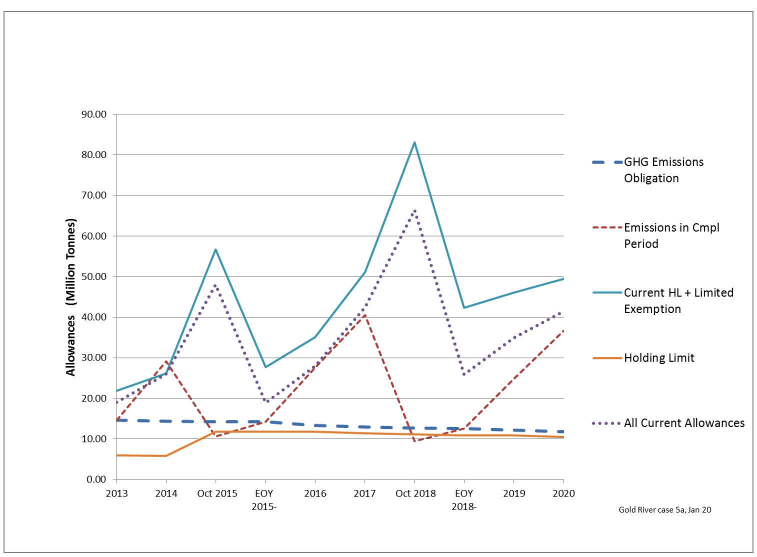
[15] Holding limits are defined in relationship to subject purchase limits, e.g., a tight limit for a low-emitting source is 1 unit over their maximum requirements for the period.

Second, each year portfolio managers must make up any difference between their total purchased allowances and allowances needed to cover 30 percent of last year's emissions, which is the minimum required for transfer each year to the compliance account. The difference is made up by transferring committed capacity allowances into ET-Sim, again valued at the closing price of the most recent auction. Current vintage allowances[16] for committed capacity are issued to participants in even increments into holding accounts overseen by the session administrator, who ensures that the holding limits are not exceeded. At the time of transfer to holding accounts, allowances needed to cover emissions from committed capacity are priced at the value of the last auction completed by the time of issuance.

Third, portfolio managers must bring their allowance accounts into compliance when notified by the session administrator that the next transfers of "committed capacity" allowances into holding accounts would result in a holding limit violation. This excess is projected before the simulation for each portfolio, based on the initial emissions targets and projected market conditions. The information relevant to each portfolio is shared with the portfolio managers/auction participants at the beginning of each simulation in a private information package dealing with plant-specific emissions and costs for each portfolio, and outlines the compliance account transfers needed to ensure they receive sufficient allowances while complying with holding limits.

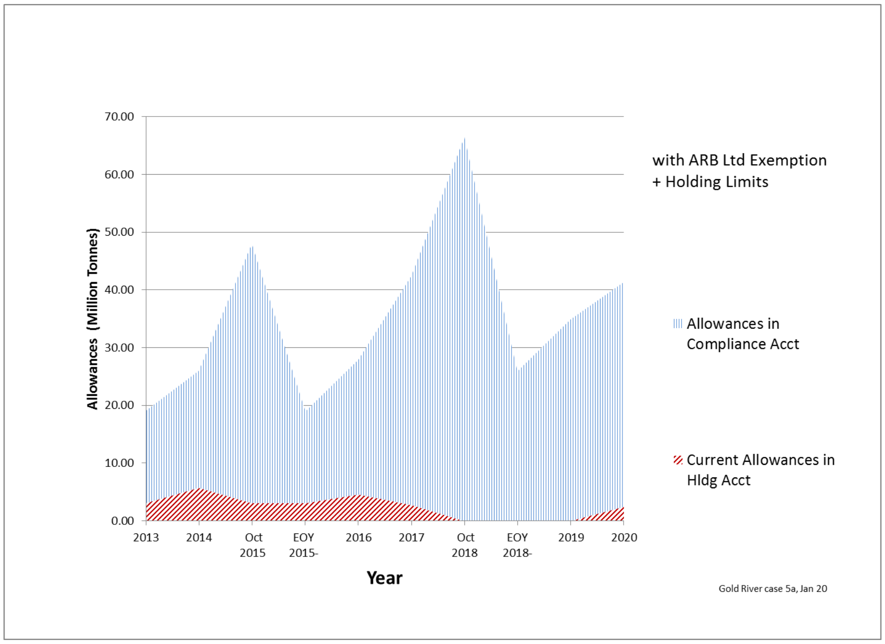
The session administrator signals the individual participants when their limit of allowances for the holding account has been reached by the accruals for committed capacity. The session administrator also informs each auction participant about the level of transfers that will be necessary in subsequent periods to put that participant under the holding limits, inclusive of their limited exemptions. Once deposited into compliance accounts, allowances may no longer be traded; thus professional simulation participants have an incentive to defer transfers to compliance accounts to the maximum extent possible.

Figure V-7. CA Cap and Trade Allowance Account Profile for Representative Large Portfolio
(Gold River 2013 to 2020)



[16] In the scenarios for this report, no advance vintage allowances were offered, either in auction or spot market sales.

Figure V-8. Compliance vs. Holding Accounts for Representative Large Portfolio
(Gold River 2013 to 2020)



In order to provide guidance to professional simulation participants on the timing of allowance transfers needed to cover their "committed capacity" compliance obligations consistent with applicable holding limits, we analyzed patterns of allowance purchase and distribution assuming a base case projected emissions level and expected carbon allowance price path.[17] Our analysis seeks to find the cost-minimizing level of allowance purchases through an iterative process and calculates required transfers over the specified compliance period. As an example, here are the results for a representative large high emitting portfolio (Gold River) presumed to have an emissions obligation of nearly 15 million tons per year in 2013. Gold River's hypothetical situation indicates the difficulties facing similar firms in California, including large electric distribution utilities and integrated oil companies, in implementing their allowance purchase, trading, and account management strategies.

In the analysis to support the allowance account management by portfolio managers, the objective is to minimize the cost of each participant's allowance position through 2020, given current allowance price expectations and each portfolio's total projected CO_2 emissions. For the Gold River portfolio, the emission allowance obligation is presumed to decline in concert with the program cap (i.e., from 14.7 million metric tons in 2013 to 11.9 million metric tons in 2020, blue dashed line in *Figure V-7*). The ARB holding account limit (see again *Figure V*-1) begins at 5.95 million tons in 2013, and then nearly doubles in 2015, with the inclusion of CP-2 sources under the AB 32 cap and trade program (orange line in *Figure VI-7*.)

[17] This is designed to incorporate inputs from the Project Team and projections from the PEAR Competitive Generation Cost Model for WECC power market regions.

Iterating to approximate a least-cost solution, we calculate the applicable current limit (current holding limit plus limited exemption, blue line *Figure V-7*) to be 21.8 million allowances in 2013, rising to 26.05 in 2014, as the result of allowance transfers to the compliance account. It should be noted that the limited exemption and, hence, the current limit rises and falls as allowances are added to or surrendered from the compliance account.[18] Under this pattern of transfers, current allowances held (purple dotted line *Figure V-7)* almost reach the current limit in 2014, but otherwise fall below it throughout the remainder of these compliance periods. These allowance purchases and transfers are projected to cover the emissions forecast shown in the figure by building a reasonable allowance bank and satisfying all holding and purchase limits along the way. Thus, this analysis demonstrates one feasible allowance account management strategy for Gold River's presumed 2020 emissions target, using an assumed allowance price trajectory.

However, the commitment of allowances to compliance accounts to take advantage of the limited exemption while building a sufficient allowance bank for compliance does significantly reduce the supply of current allowances in Gold River's holding account. These reductions are most likely to occur near the end of each compliance period and during the first year of the next compliance period. *Figure V-8* demonstrates that because current allowances must be transferred to the compliance account to maintain the limited exemption, the current allowances available for trading from the holding account will be greatly diminished. Starting from about 3.1 million allowances in 2013 (out of about 19 million total) allowances in the holding account peak at 5.7 million in 2014 (out of 26 million total, or 22 percent), which means that the lowest compliance account share of total allowances for this portfolio is about 78 percent. Again, since allowances cannot be removed from the compliance account (only surrendered to cover emissions,) these current allowances will not be available for trading, reducing allowance market liquidity.

[18] The current limit also increases and decreases when the maximum limited exemption increases each October 1 and decreases on December 31, 2015 and 2018.

Figure V-9. 2014 Compliance vs. Holding Accounts for Six Portfolios: No ARB Limits
(Year 2014 Allowance Levels)

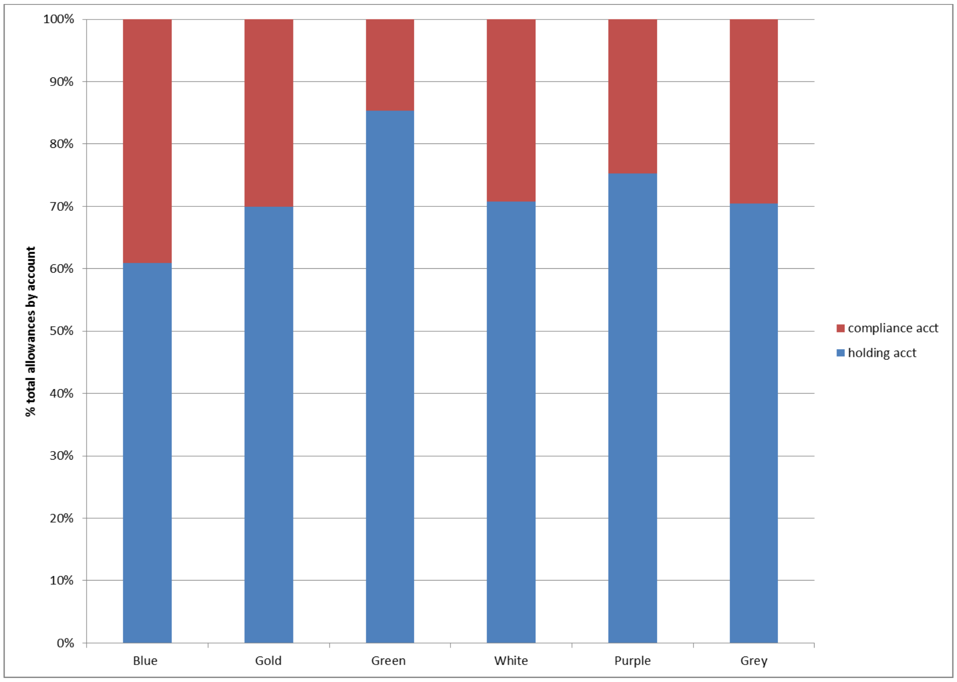
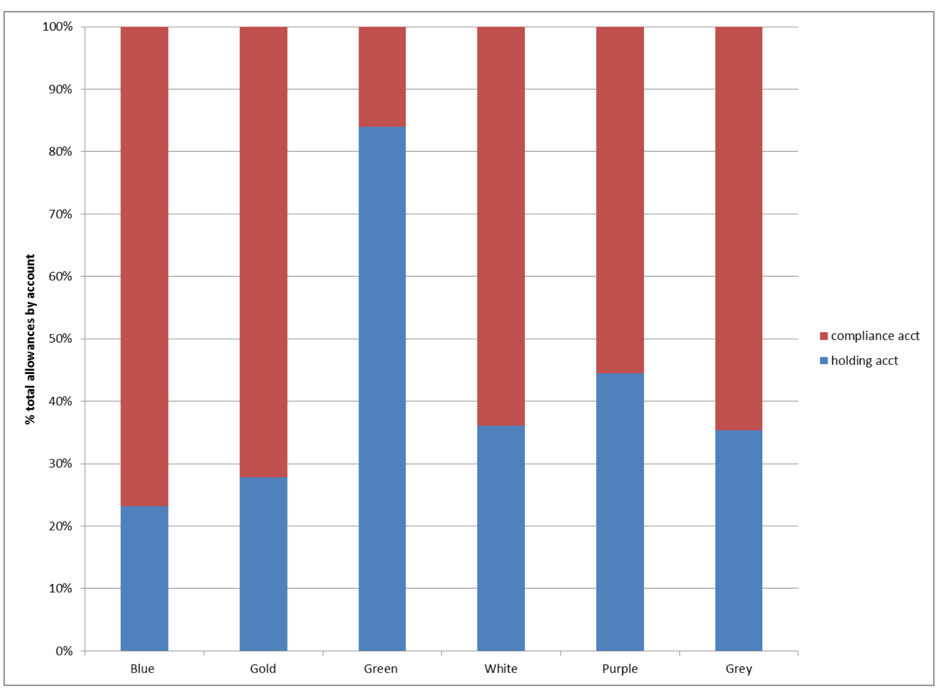


Figure V-10. 2014 Compliance vs. Holding Accounts for Six Portfolios: With ARB Limits
(Year 2014 Allowance Levels)



Detailed results discussion – professional simulation – holding limits: For the five simulations without holding limits, participants were informed of the annual obligation for covered sources to deposit one-third of their preceding year's emissions in compliance accounts each year. For CP-1, this provision means that by October 2014, one-third of 2013 emissions should be transferred. *Figure V-9* indicates the split between compliance and holding accounts for the six managed portfolios in 2014 following fulfillment of this requirement.

For the two simulation cases with ARB limits, participants were provided with a private portfolio briefing package including a description of the timing of allowance transfers to compliance accounts needed to maintain compliance with holding limits, which outlines what they will need to do during the course of the simulation in order to receive the allowances needed to comply. Five of the six portfolios have annual emissions greater than 5.95 million tons per year (the holding limit without limited exemptions), and we noted that once they would have completed their "committed" capacity transfers in 2014, the five portfolios would have deposited over half of their allowances, with two of the five depositing over 75%; see *Figure V-10* for detail by portfolio. More to the point, a comparison of *Figure V-9* with *Figure V-10* indicates the significant reduction in allowances maintained in holding accounts based on required transfers, which translates into a significant reduction in allowances available to trade.

VI. Results – Allowance Price Containment Reserve

Given concerns regarding the uncertainty the APCR sale might impose on California covered sources, we considered a range of question on its potential impact.

- Does the presence of the price containment reserve contribute to a reduction in allowance banking?
- Is price containment reserve associated with reduced price variability? I.e., not lower price levels, but lower price spikes?
- Are there plausible circumstances under which the price containment reserve would be exhausted?
- Can the containment reserve be said to be effective only for transitory vs. protracted events?

In summary, we found that the California APCR as structured by ARB was effective, forming a kind of price insurance mechanism that was apparently viewed by participants both as a reliable supplier of allowances, as well as a "seller of last resort." [19] Sessions with APCR experienced sales during stress periods (as any good insurance policy might) and attendant benefits in the form of lower prices. In the professional simulations, we saw participants buy allowances to meet unfulfilled compliance obligations with true-up at hand, even when market prices were well below the threshold tier prices.

VI.1 Student Laboratory Sessions – Effect of APCR

Economists have known since 1974 that emission cap and trade programs can result in unwanted price volatility (Weitzman, 1974). The underlying source of this volatility is the fixed nature of the supply of allowances; that is, the supply curve for allowances is vertical, there is no supply response to increased price. The increasing use of emission trading programs and their use in GHG regulation

[19] This is somewhat analogous to a central bank, (e.g., the Federal Reserve) which by standing ready to provide funds to financially-stressed institutions functions as a "lender of last resort."

in particular have kindled a substantial academic interest in the subject, even as actual trading programs are building price control mechanisms into their own programs. Offsets have long been seen as one element of a price containment strategy and have been authorized but not much used, at least not enough to have any real effect on price spikes. The California cap and trade regulation continues this pattern, encompassing both offsets and a novel approach to the use of an allowance reserve.

Generally speaking, a price containment reserve is a block of allowances that only becomes available for use in the market once some high price trigger-point is reached.[20] Once the triggering event occurs, some allowances will enter the market. The RGGI, for example, currently is considering a reserve that would become part of the supply of allowances available at auction whenever the auction would otherwise close above a trigger price. The allowances available at that price point would be released to accept additional bids until the supply of allowances available at that price is exhausted or the auction closes, whichever comes first.[21]

Definition of APCR – laboratory sessions: California's approach establishes a post-auction sale where potential buyers may bid for allowances at any of three price points. The reserve allowances are divided into thirds and allocated to these "price shelves." Once an allowance is sold from a shelf, that allowance is not replaced. As long as there are plenty of allowances on a given shelf, then a bidder can make a bid at a higher shelf and expect their order to be filled from the next lower price shelf. But when a shelf gets close to empty, bidders have incentive to bid for as many as are available at that price and not to bid at a higher price because bids at the lower price are filled in proportion to the total number of bids at that price. And, bids at the higher price will likely be filled at that price and not drop to the next lower level.

Because of the different release mechanism and the somewhat complicated incentive structure in the California reserve, it is not clear how participants in the California market will respond to the presence of the reserve or what would be its likely effect on price. The California APCR design was included in our main experimental sessions, and this provides us with an opportunity to make some observations about how market participants use the reserve and how the presence of the reserve changes other outcomes in the market.

What one expects from the presence of the reserve depends on views about the function of the reserve and its likely interaction with the normal market. While these reserves are sometimes viewed as a protection against a cap that is tighter than expected, or, to say it differently, that the cost of reducing emissions is higher than expected, the size of the reserves currently envisioned are not large enough to provide much of a cushion against longer term scarcity. It is easy to envision reasonable cases with large errors in forecasting allowance demand, as in the case of RGGI, the demand for allowances might fall well below the forecast, or it might go the other way, in which case the reserve does not provide much protection.

---

[20] There is considerable debate over the accounting issue of whether a reserve should be seen as a part of the stock of allowances under the cap or outside of the cap. This may be less of a difference than is commonly assumed chiefly because, if the reserve comes from within the cap, prices will be higher and the reserve is more likely to be used, and if the allowances come from outside the cap, prices will be lower and this reserve is less likely to be used.

[21] This approach has been dubbed "the Eastern consensus" PCR mechanism because it is being considered for use in the RGGI market.

If a reserve cannot be relied on to manage long-term patterns of high allowance demand, then it may serve much better as protection against transient spikes in demand, just as a price floor (also called the reserve price in the California regulations) is protection against transient drops in demand. Seen from this point of view, a reserve serves as a short-run relaxation in the otherwise fixed supply of allowances. Just as a spike in demand for normal market goods encourages an increased supply to the market, a spike in demand for allowances expands the available supply and can blunt further increases.

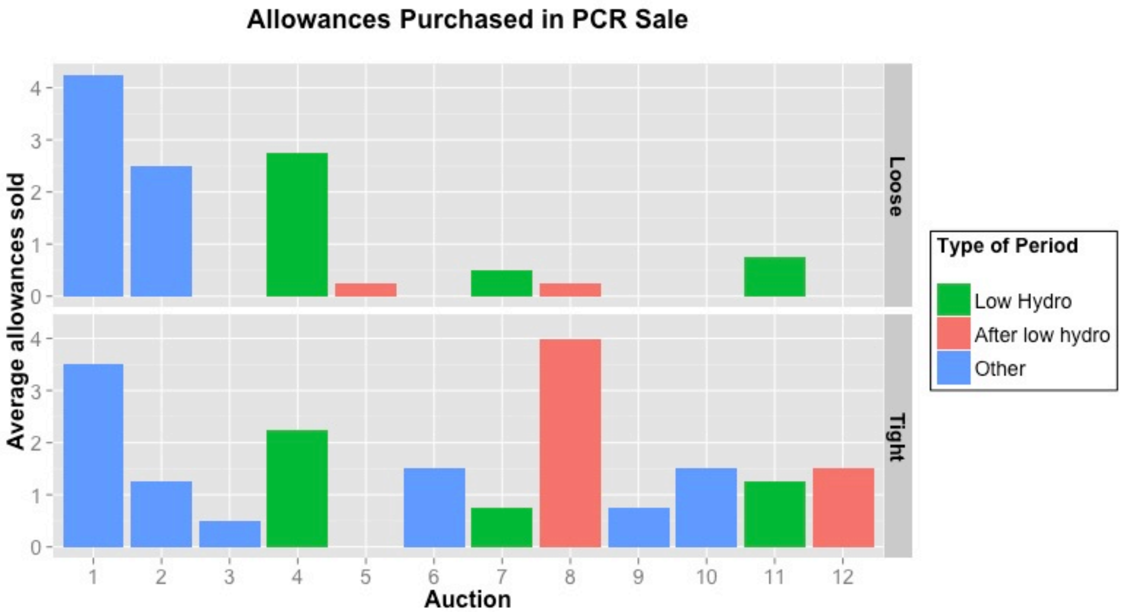
In the case of a potential price bubble or a possible spike in price due to an attempt at price manipulation, the presence of the expanded supply at the available triggers may be enough to prevent the problem from occurring at all. Panic buying is not necessary when a store with fixed price goods is open down the street. A potential manipulator would need to buy right through the expanded supply, at considerable risk, to profit from the manipulative strategy.

We do not know whether one form of price containment mechanism is better than another at accomplishing this short-run control of price spikes. More experience and more research are needed to discover any potential differences.

That said, in this series of experiments and simulations we did see patterns of reserve use that point to unexpected uses of the California-style APCR and to some important differences between the post-auction sale and the in-auction release of reserve allowances. There may be advantages to having an opportunity for individuals to choose when to purchase allowances from the reserve as opposed to only purchasing them when some overall market trigger-point is passed. The difference is that the allowances that can be purchased by individual choice (even at a high price) provide a kind of insurance to the individual source that is not provided by a general relaxation of the fixed market supply curve. The obvious risk in this behavior is that the reserve may be drawn down to a point where it will not serve to reduce price spikes due to transient increases in demand.

Detailed results discussion – APCR – laboratory sessions: The purchase of allowances from the reserve in our laboratory sessions is shown in *Figure VI-1*. The top panel shows purchases under a loose holding limit and the bottom panel a tight holding limit. One would expect to see relatively large purchases during the high demand periods that we label as "low hydro." These high demand periods, 4,7, and 11, do show modest purchases, as expected.

Figure VI-1. Purchases from APCR in Laboratory Sessions



Surprisingly, a significant number of allowances were purchased from the $23 tier of the reserve in the first two periods, when the auction closing price was very close to the floor of $10. This appears to be speculative buying in anticipation of substantial future price increases. These buyers are correctly anticipating that prices will eventually rise above the $23 trigger price. More buying of this sort seems to be taking place in period 10 of the tight holding limit sessions.

The orange bars, labeled "After low hydro" represent periods just following high demand periods when prices were well below the trigger prices. There appears to be some tendency for participants to buy from the reserve to refill individual compliance accounts after high demand periods.

A final observation to make about purchases from the APCR during the lab sessions is the substantial difference between the loose and tight holding limit cases. This clearly shows that there is an interaction between the holding limit requirement and the likelihood that firms will buy allowances out of the reserve. If this incentive is strong enough, the reserve could become depleted, as occurred in one of our tight holding limit sessions when the entire reserve was used up before the end of the session. Participants in tight holding limit sessions tend to draw more on the reserve in later periods because they are unable to keep the price from rising due to their limited ability to bank allowances.

Some of the early buying of permits was simply buying in anticipation of future scarcity. The post-auction sale provides a way for participants to "borrow" from the future allowance supply by buying from the APCR in anticipation of future scarcity. Because our experiments were not designed with this in mind, we cannot say how the presence of this borrowing facility changed bidding and banking behavior, but we can guess that, especially for risk-averse bidders, the presence of the reserve served as a form of insurance that allowed them to bid somewhat more aggressively based on production values rather than on the possibility of falling short and paying penalties in the future.

This insurance function of the APCR may improve efficiency. This should be a fruitful area for future research.

We would expect this property of the APCR to be true of the actual California market as well as in the lab. For emitters, the consequences for non-compliance go beyond the legal penalties. Firms and other institutions receiving penalties face substantial costs of bad publicity in addition to the statutory penalties. It is reasonable to expect that this prospect would cause many market participants to be quite risk-averse toward a compliance penalty. Having the APCR available as a last protection against non-compliance may allow participants to bid closer to their actual values for allowances due to their smaller risk premium.

Finally, there are many possibilities for the design of a post-auction sale or an in-auction release to relax the fixed supply curve and make available additional allowances during high price periods. Some combination of the two may perform even better. Or, a post-auction sale, such as the California APCR, could have a top step that establishes a hard ceiling on price, a ceiling that might only be available should the rest of the reserve be depleted. A number of interesting opportunities present themselves and are worth consideration.

VI-2. Professional Simulation – Effect of APCR

Definition of APCR in professional simulations: The sole change in the professional simulation in the representation of the allowance price containment reserve is that the tiered pricing was adjusted upward to match the pricing in the regulation. Thus for 2013 the tiers were priced at $40, $45, and $50 in the Veconlab application. The shift in focus here clearly was to have the participants weighing the likelihood of the $40 allowance price threshold being reached over the course of the simulation given the scenario conditions as described to them—which seemed a close approximation to what real-world California sources would have to do.

Based on our own Western power market price projections for the baseline and high demand (low hydro) scenario conditions, we expected that the bids would be received for the APCR sales during a low hydro year, or more likely, after recurring low hydro years. In all other respects, we preserved the same representation of the APCR, including the volume of allowances in each tier and fulfillment processes. The sale was also offered to the participants at the end of each auction round, as in the laboratory sessions.

Of the seven professional simulations conducted, the APCR was accessed in one simulation at the end of the session. To our surprise, that simulation was run using baseline rather than high demand (low hydro) scenario conditions.[22] The buyer chose to buy reserve allowances rather than be subject to compliance penalties, just the kind of "seller of last resort" role that would appear to be a logical extension of (and fully consistent with) the "price insurance" role that emerged in the laboratory sessions.

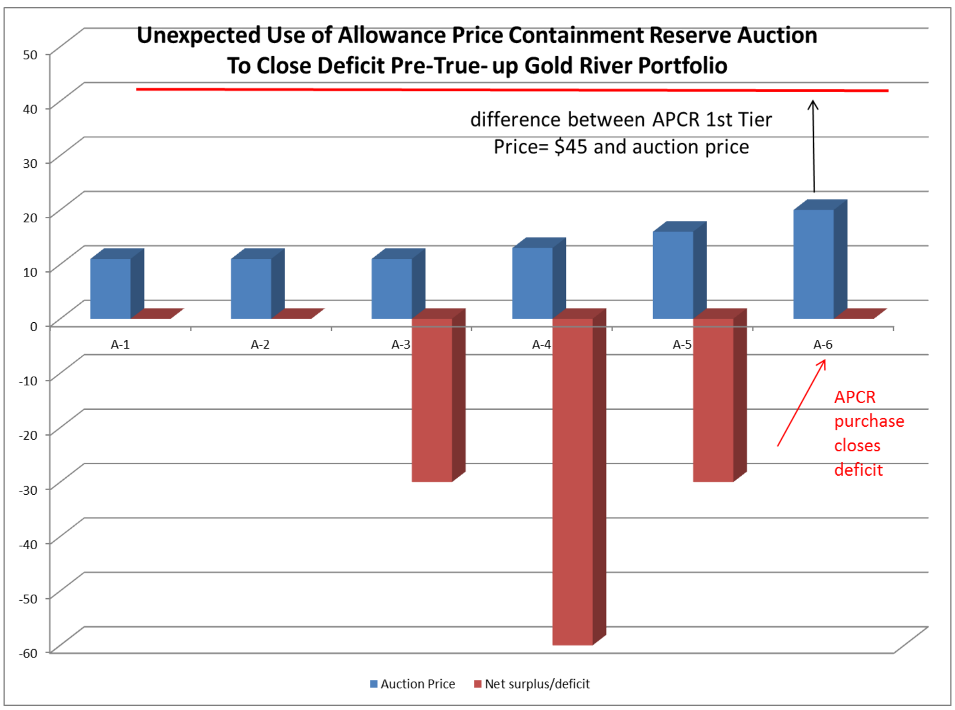
Detailed results discussion – APCR – professional simulations: The buyer in the APCR sale was assigned the role of a high-emitting investor-owned utility (managing the Gold River portfolio). *Figure VII-2* depicts their position as it evolved over the course of a six-round simulation to the compliance

[22] A parallel discussion of the similarly unexpected behavior of the participants in the low hydro case—not accessing the APCR—is included in *Section VII.2.*

deadline at the end of Round 6, including 1) the participant's allowance balance, (maroon bars), and 2) the auction clearing price (blue bar) for each round. As is not uncommon for the high-emitting roles, in the initial rounds their typical valuations (if they bid close to their values) are such that they are able to acquire sufficient allowances to operate, but that position can shift quickly, particularly as the cap tightens.

After remaining in balance for two rounds, with the auction clearing nearing the minimum price (at $11 per metric ton), by the end of Round 4 the Gold River portfolio went into a deficit position down -60,000 allowances (or enough to shut down two units for one round). In a post-simulation survey, the portfolio manager explained, "I was trying to get into compliance and I believed that prices would continue to rise. As the auctions progressed the trend became more clear, my objective was to take every available allowance." When he was unable to do so, he concluded: "With no auction remaining (before true-up) I was forced to use the allowances from the reserve." Thus, unable to get the allowances he wanted for the closing rounds of the auction (where the final closing price was $18 per metric ton CO_2), the Gold River portfolio manager elected to buy allowances at $40 per metric ton.

Figure VI-2. Use of the APCR to Close Allowance Deficit



VII. Results – High Allowance Demand Periods

As noted earlier, our model of allowance scarcity in both our experiments and simulations had two components. First, there was a gradual tightening of the cap in each period. Second, there were randomly occurring periods of particularly high allowance demand. From the individual participant's point of view, there was a third dimension of increasing scarcity, the decline in free allocations, with steps down after each compliance period. This did not tighten the overall cap.

Both sources of increasing scarcity give participants incentive to anticipate the future scarcity. They can reduce their own costs by purchasing more allowances than needed in the early periods to use or sell in later periods. This buying in anticipation of future scarcity should allow participants to avoid being exposed to the large increases in scarcity during the "low hydro" periods, and so should tend to "smooth" some of the high price peaks. It should also "flatten" the growth in the actual, observed price path relative to the growth that would occur in the absence of opportunities to bank allowances in anticipation of future scarcity. The smoothing and flattening that results from market activities are important for emission markets to achieve cost-effective outcomes, and market design features can have a large influence on the ability of market participants to effectively flatten the scarcity profile and smooth risk.

Even though the declining cap makes allowances more valuable each period, the optimal price for allowances would be constant at $14.75. Subjects would cut back early and bank allowances for future, higher scarcity periods, just as we have observed in actual emission markets such as the U.S. SO_2 emission trading program. These transfers would be profitable until the prices were the same in each period. Figure V-2 shows clearly that subjects in the experiment do this time-shifting of allowances, using early emission reductions to reduce future costs. It is also clear that subjects in the loose holding limit case are better able to shift emissions out into the future. Leaving off the last period, prices would have risen 305% with no smoothing. Tight holding limit treatment subjects held this increase down to 77%, even with their reduced ability to bank. Loose holding limit treatment subjects held price increases down to a mere 25%, less than a tenth of the no-smoothing case.

VII.1 Student Laboratory Sessions – Effect of High Allowance Demand

In *Figure V-2*, the low hydro years, 4,7, and 11, are clearly marked by high predicted price spikes. These are the prices that would clear the market if no intertemporal smoothing were possible. Predicted prices rise 130%, 110%, and 175% respectively. Market participants, although not knowing exactly when the high demand periods would arise, were prepared for the sharp rises in price and used their banks to expand production beyond what would have been possible if only that period's allowances had been available.

Table VII-1 shows the percent changes in price over the previous period. In period 4, which was the first high demand period, the price actually dropped in the loose holding limit treatment due to enthusiastic selling into the tight market by participants who had accumulated banks for this very eventuality. Generally speaking, in later periods, these sellers became more circumspect about dumping large quantities into the market. If the expected pattern of smoothing spikes is clear, so is the difference between the loose and tight holding limit treatments. The subjects in loose holding limit treatments were much better able to reduce their exposure to risk by banking allowances. This "self-insurance" banking lowers price risk for the whole market, reducing price volatility.

Table VII-1: Percentage Price Changes in High Demand Periods

Period	Percent price change from previous period		
	Loose	Tight	Predicted
4	-14.3	6.6	128.9
7	6.8	30.3	110.0
11	15.4	21.7	174.6

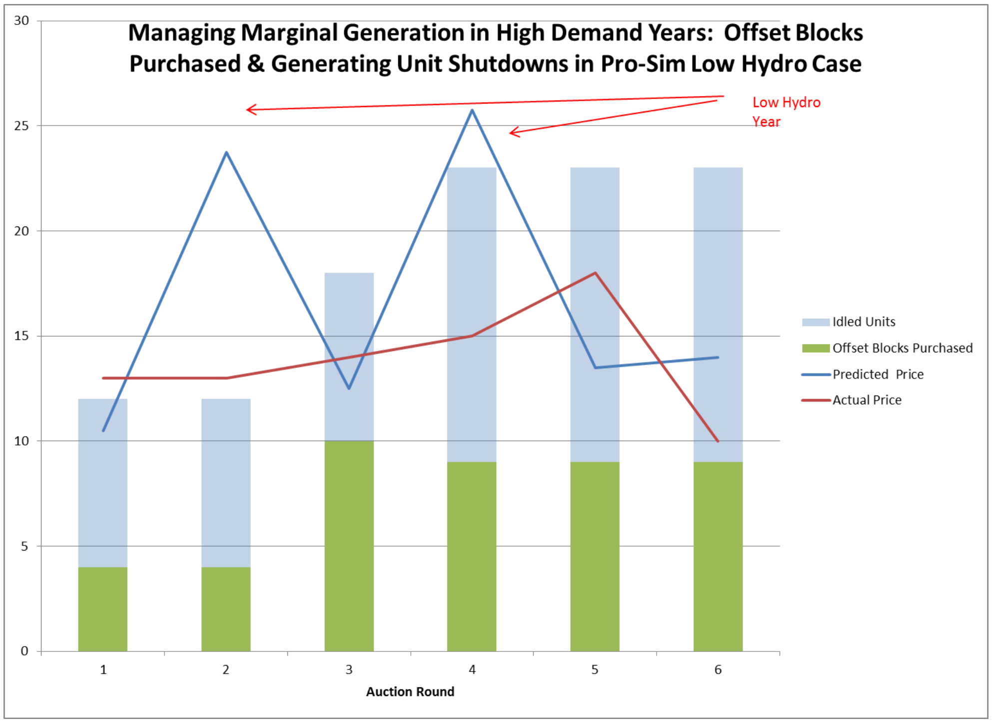
Professional Simulation – Effect of High Allowance Demand

In the professional simulations, we examined whether the presence of the wider range options for meeting compliance obligations available in the simulation helped participants reach better outcomes. To test the difference, we ran the majority of simulation cases under baseline conditions, i.e., with normal hydro availability. After testing the baseline conditions, we managed to complete one full low hydro professional simulation. A focal point in the professional simulations is how the recurrence of low hydro conditions similar to 2001 (and other constraints) may interact with holding limit rules and the APCR. To that end we developed an alternative set of optimal compliance projections, which we used to help define the marginal generation portfolio estimates.

We began the simulation with our optimum compliance plan targets for a low hydro scenario, assuming that each of the portfolio managers would select their least-cost compliance alternative and proceed through the auction, building a bank of allowances in anticipation of the "high demand" period and drawing down their balances accordingly. *Figure VII-2* demonstrates how the two high-emitting unit portfolio managers took an alternative course: rather than focusing on a single, least-cost option, both managers choose to split their compliance strategies between offsets purchases (green bar) and shutting down generating units to be replaced with wholesale power purchases (blue bar). The "split-decision" compliance strategy is in contrast to a projected optimal choice of wholesale power alone, which, although more expensive than the first block of offsets purchases, is less costly overall as offsets escalate in price through repeated purchases.

Even after one low hydro year and a further ratchet downward in the cap, both managers remain committed to their split compliance strategy decision (see Round 3 in *Figure VII-2*). Unlike their counterpart discussed in the preceding section, neither participant needed to cover their position with purchases from the APCR. These findings reinforce the general conclusion from the laboratory sessions that the California markets still offer significant flexibility to manage risk, and the options afforded under the professional simulation only amplify that freedom.

Figure VII-1. Impact of Low Hydro Years: Offsets and Unit Shutdowns



VIII. Conclusions and Potential Relevance for Policy

We applied the methods of experimental economics to study how two central design features of California cap and trade regulation—holding limits and the APCR—could affect trading behavior, potentially signaling changes in actual market performance. Our approach included both laboratory sessions with students and market simulations with professionals. One advantage of experimental economics over conventional modeling is that experimenters can look for differences in behavior in the presence or absence of that design feature, and, with a sufficient number of repetitions, can draw statistically valid conclusions about its effect on behavior.

While the statistical significance (or "internal validity") of the test may be uncontroversial, the applicability to the actual policy regime (its "external validity") depends on how well the laboratory setting captures the essential elements of the policy as applied to the California market. To help assess how well we had done in this regard, we looked for concurrence between the results of the student laboratory sessions and the professional market simulations.

And concurrence is in fact what we found. First, on holding limits, results from both laboratory sessions and professional simulations indicate that tight holding limits may have a substantial effect on market performance and may actually have the unintended effect of increasing the probability of market manipulation. Aside from the costs of reducing compliance flexibility and limiting opportunities for risk management, tight holding limits reduce market liquidity, which analysts generally consider to be the first line of defense against those attempting to manipulate a market for profit. We also found that tight holding limits contributed to higher price variability, less effective price discovery, lower efficiency, and reduced banking, which translates into delayed reductions in greenhouse gases.

Next, both the laboratory sessions and simulations yielded broadly similar results in the case of the three-tiered APCR. Our review of the experimental results suggest that the APCR did change subjects' behavior by lowering their risk of large price spikes and also allowed subjects to "borrow" allowances as a hedge against anticipated scarcity and higher prices due to the declining cap. The professional simulations reinforced these somewhat surprising results, with subjects purchasing allowances from the APCR to meet unfulfilled compliance obligations at the end of compliance periods. Professional subjects purchased APCR allowances even when market prices were well below the threshold APCR tier prices. The early buying in anticipation of the induced scarcity due to the declining cap does raise the issue of the long-term adequacy of the reserve for protection against price spikes. If this incentive is strong enough, it is possible that the reserve could become depleted.

Our results suggest that an APCR serves a useful function for emission markets. *ARB may wish to consider studying the performance of other APCR designs, including possible hybrids that include some combination of post auction sale, in auction sale, and hard cap.*

Since any comprehensive program to reduce GHG emissions must necessarily be quite complex, an investigation into one component of that larger policy has to balance the need for abstraction with the need to model "real world" conditions to allow conclusions about the California regulation. Due to the fact that the scope of the program is expanding,[23] as well as the fact that our laboratory and market simulations cannot exactly mirror the California regulation, we must exercise caution in drawing policy lessons. Nevertheless, our findings suggest some important guides.

The regulation's holding limits primarily reflect concern about traders' positions relative to market size and accordingly, whether one or a few participants might achieve a position of market dominance, especially since:

- although the market has more than 350 covered entities in 2013—rising to 600 in 2015—it is still small relative to the originally anticipated regional WCI, or national, markets; and
- the state lacks a financial regulator comparable to the Securities and Exchange (SEC) or Commodity Futures Trading Commissions (CFTC).

Accordingly, ARB chose to follow the holding limits precedent from the CFTC rule designed to address systematic risk in commodity futures markets. But as the board itself recognized, derivative markets are different from emission markets and as a result are likely to function differently. To ensure that the state's largest covered entities have sufficient allowances to meet compliance, ARB added the limited exemption feature, which we found causes large emitters to transfer substantial shares of their allowances to compliance accounts. Thus, in practice, holding limits will reduce the ability of covered entities to trade actively, and therefore will tend to reduce market liquidity.

We are not in a position to know whether any protections provided by limiting the share of allowances owned by one or a few large providers outweighs the costs and risks imposed by the holding limits themselves. At the very least, it does seem clear that the holding limits should not be tighter than necessary. *If the emissions market is to achieve its potential in lowering the costs of GHG reductions,*

[23] Linkage with the province of Quebec is a pending matter before ARB.

then the damage to market function from protections against manipulation by dominant participants should be a continuing priority for the ARB.

IX. Future Research

The prospect in 2015 of inclusion of vehicle fuels under the cap represents a huge shift in allowance demand for the sector, and we have taken some initial steps to model the problem: in some test laboratory sessions we tried the notion of a subject changing from low- to high-emitter during the course of a single session.

We also set up an initial fuel supplier portfolio in the professional simulation. From 2013–2015, our refiner test portfolio goes from being responsible for its direct CO_2 emissions (at 0.052 metric tons per barrel) to the CO_2 content of its refined products (for gasoline, 0.38 metric tons per barrel), a multiple of nearly 7.0 times. The compliance problem is represented by the limited number of options with which integrated oil companies have to comply (energy efficiency, biofuels, production shifts). Further complications arise from having holding limits and limited exemption for refiners based on their reported direct emissions. While we have sought industry comment on our portfolio structure, project timelines did not afford the time to simulate the refiner position.

Good experiments lead to new questions, and suggest implications for future studies. Do holding limits prevent formation of dominant market positions? The issue was not tested in our analysis and has not been demonstrated in studies to date. Future research should investigate the likelihood of market dominance without holding limits.

Additional issues include:

- Hard price caps and hybrid price containment strategies
- Liquidity, manipulation, and price bubbles
- Post-auction information disclosure
- Policy/litigation-induced changes in offset supply availability and in the low carbon fuel standard
- Uncertainty over future program requirements.

X. References

Bodsky, R., D. Donato, K. James, and D. Porter. 2012. *Experimental Evidence on the Properties of California's Cap and Trade Price Containment Reserve*. Economic Science Institute. Chapman University. http://www.chapman.edu/research-and-institutions/economic-science-institute/academics-research/working-papers.aspx

Burtraw, D., C. Holt, E. Myers, J. Goeree, K. Palmer, and W. Shobe, 2011. Price Discovery in Emissions Permit Auctions in R. Mark Isaac, Douglas A. Norton (ed.) *Experiments on Energy, the Environment, and Sustainability* (Research in Experimental Economics, Volume 14), Emerald Group Publishing Limited, pp.11-36.

Burtraw, D., C. Holt, E. Myers, J. Goeree, and K. Palmer. 2010. An Experimental Analysis of Auctions versus Grandfathering to Assign Pollution Permits, with *Journal of the European Economic Association*, Vol. 8, No. 2-3: 514-525. April/May.

Burtraw, D., C. Holt, J. Goeree, K. Palmer, and E. Myers. 2009. Collusion in Auctions for Emission Permits: An Experimental Analysis, *Journal of Policy Analysis and Management*, Vol. 28, No. 4: 672-691.

California Air Resources Board (CARB). 2012. Teleconference with University of Virginia project team to discuss CARB interpretation of allowance price containment reserve and holding limits. August 28.

CARB. 2012. Addendum to the Final Statement of Reasons for Rulemaking: 2012 Amendments to the California Cap on Greenhouse Gas Emissions and Market-Based Compliance Mechanisms. June 28.

CARB. 2012. Copy of Purchase Limit Examples, 2014, 2016. Internal Memo on File with Authors.

CARB. 2010. Electric Power Entities: Retail Providers and Marketers Aggregated Electricity Transactions and Greenhouse Gas Emissions Reported to the California Air Resources Board.

CARB. 2010. Greenhouse Gas Facility Emissions Reported to the California Air Resources Board.

CARB. n/d. Final Regulation Order. Subchapter 10 Climate Change, Article 5, Sections 95800 to 96023, Title 17, California Code of Regulations, to read as follows: Article 5: California Cap on Greenhouse Gas Emissions and Market-Based Compliance Mechanisms.

CARB. n/d. Subchapter 10 Climate Change, Article 5, Sections 95800 to 96023, Title 17, California Code of Regulations to read as Article 5: Amendments to the California Cap on Greenhouse Gas Emissions and Market-Based Compliance Mechanisms to Allow for the Use of Compliance Instruments Issued by Linked Jurisdictions.

CARB. n/d. Final Regulation Order. Subchapter 10 Climate Change, Article 5, title 17, California Code of Regulations, sections 95802, 95812, 95814, 95830, 95831, 95832, 95833,95834, 95856, 95870, 95892, 95910, 95911, 95912, 95913, 95914, 95920, and 95921, are amended to read as follows Article 5: California Cap on Greenhouse Gas Emissions and Market-Based Compliance Mechanisms.

Carmona, R., Fehr, M., Hinz, J., & Porchet, A. 2010. Market Design for Emission Trading Schemes. *Siam Review*, 52(3), 403-452.

California Public Utilities Commission. 2012. Decision on System Track I and Rules Track III of the Long-Term Procurement Plan Proceeding and Approving Settlement: Order Instituting Rulemaking to Integrate and Refine Procurement Policies and Consider Long-Term Procurement Plans. Decision 12-04-046. April 19.

Chesney, M., L. Taschini, and M. Wang. 2011. Experimental Comparison between Markets on Dynamic Permit Trading and Investment in Irreversible Abatement with and without Non-Regulated Companies. Grantham Research Institute on Climate Change and the Environment Working Paper No. 41 Centre for Climate Change Economics and Policy Working Paper No. 51. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1719026

Comerton-Forde, C. and T.J. Putnins. 2009. Pricing Accuracy, Liquidity and Trader Behavior with Closing Price Manipulation. University of Sydney, NSW. Discussion paper.

Ellerman, A. D., Convery, F. J., and de Perthuis, C. 2010. *Pricing Carbon: The European Union emissions Trading Scheme*. Cambridge: Cambridge University Press.

Ellerman, A.D., Joskow, P.L., Schmalensee, R., Bailey, E.M., and Monteiro, P.K. 2000. *Markets for Clean Air: the U.S. Acid Rain Program*. Cambridge and New York: Cambridge University Press.

Environmental Defense Fund (EDF). 2010. Modeling the Effectiveness of a Strategic Allowance Reserve in a Cap and Trade Program in California. December.

Grüll, G., and L. Taschini. 2011. Cap-and-Trade Properties under Different Hybrid Scheme Designs. *Journal of Environmental Economics and Management* 61:107–118.

Hart, O.D. 1977. On the Profitability of Speculation. *The Quarterly Journal of Economics*, Vol. 91, No. 4: 579–597.

Holt, C., B. Shobe, D. Burtraw, J. Goeree, and K. Palmer. 2007. Auction Design for Selling CO_2 Emission Allowances Under the Regional Greenhouse Gas Initiative. Research sponsored by the New York State Energy Research and Development Agency. October.

Jarrow, R.A. 1992. Market Manipulation, Bubbles, Corners, and Short Squeezes. *The Journal of Financial and Quantitative Analysis*, Vol. 27, No. 3: 311–336.

Ledgerwood, S. D. and P. R. Carpenter. 2012. A Framework for the Analysis of Market Manipulation. Forthcoming in the *Review of Law and Economics*.

Linklaters. 2011. California Agency Unanimously Adopts Cap-and-Trade Regulations: A Review of Key Program Design Elements and Outstanding Issues. *Linklaters' Environment and Climate Change Practice*.

Murray, B., R.G. Newell, and W. Pizer. 2008. Balancing Cost and Emissions Certainty: An Allowance Reserve for Cap-and-Trade. Resources for the Future Discussion Paper 08-24. July.

Pirrong, C. 2009. Market Oversight for Cap-and-Trade: Efficiently Regulating the Carbon Derivatives Market. Brookings Institution, Policy Brief 09-04.

Profeta, T. 2010. Allowance Price Containment. Workshop to Discuss Cost Containment and Offsets. CARB. June 22. http://www.arb.ca.gov/cc/capandtrade/meetings/062210/allowance_price_containment_profeta.pdf

Putnins, Talis J. 2012. Market Manipulation: A Survey. Stockholm School of Economics in Riga. Discussion Paper.

Shobe, W., D. Burtraw, C. Holt, J. Goeree, K. Palmer, and E. Myers. 2010. An Experimental Analysis of Auctioning Emissions Allowances under a Loose Cap. *Agricultural and Resource Economics Review,* Vol. 39, No. 2: 162-175.

Stavins, Robert N. 1998. What Can We Learn from the Grand Policy Experiment? Lessons from SO_2 Allowance Trading. *Journal of Economic Perspectives*, Vol. 12, No. 3, 69-88.

University of California (UC) Market Simulation Group. 2012. Options for Emissions Market Simulation. June 7. http://www.arb.ca.gov/cc/capandtrade/simulationgroup/simulationgroup.htm

UC Market Simulation Group. 2012. Market Simulation Group. Potential Issues for Analysis. June 7. http://www.arb.ca.gov/cc/capandtrade/simulationgroup/simulationgroup.htm

Weitzman, M.L. 1974. Prices Vs. Quantities. *The Review of Economic Studies*, Vol. 41, No. 4, 477-491.

Western Climate Initiative (WCI). 2012. Discussion Draft Economic Analysis Supporting the Cap-and-Trade Program - California and Québec - Prepared by the WCI Economic Modeling Team. May 7. http://www.westernclimateinitiative.org/document-archives/Markets-Committee-Documents/

WCI. 2010. Report on Holdings Limits to the Western Climate Initiative Markets Committee. Prepared by Jeffrey H. Harris, Alfred Lerner College of Business and Economics. University of Delaware. May 6. http://www.westernclimateinitiative.org/document-archives/Markets-Committee-Documents/

Technical Appendix 1:
Complementary Analysis: Experimental Economics and Economic Modeling

Technical Appendix 1:
Complementary Analysis: Experimental Economics and Economic Modeling

Economists use many different analytical tools to investigate the performance of public policies. So the question naturally arises as to what are the appropriate tools for a given investigation and whether, on a given question, one tool is more or less appropriate than another. Generally speaking, since the different modes of analysis all arise from roots in economic theory, one should view different modes of analysis as complementary, although some will be easier to apply to some questions than to others. Here, we wish to compare the use of experimental economics with two other possible approaches to the questions addressed in this study: structural modeling of individual behavior and general equilibrium modeling.

Structural modeling of individual behavior uses mathematical representations of how individuals make choices when confronted with a set of choices and obligations to draw general conclusions about patterns of response one would expect from a given policy. The conclusions from the direct modeling of individual behavior can lead to powerful predictions, in the sense that these predictions can be tested against actual data to confirm or reject the hypothesis that people will respond in a certain way to a set of incentives. A hypothesis that receives sufficient confirmation can also be used to predict future responses in situations consistent with the underlying structural model. In order for a structural model to be analytically tractable, the mathematical representation of behavior must be of a carefully simplified (or abstract) representation of individual behavior; otherwise the model will not be solvable and cannot be used for testing hypotheses. The process of developing a structural model can be quite time-consuming, and the data needed for testing the resulting hypotheses may not be available. In some cases, the behavior of interest is sufficiently complicated that researchers have not yet been able to come up with ways to generate structural models of the given circumstances. It is our view that such is the case in this case, as we will argue shortly.

General equilibrium models focus less on testing elements of individual behavior but rather on examining the relationships between different sectors of the economy, taking into account the essential reciprocal nature of cause and effect. A policy change that affects one industry or set of individuals will ultimately feed forward into other sectors of the economy by changing supply and demand conditions in those other sectors, which, in turn, feedback to behaviors in the sectors where the change originated. Modeling this fully circular causal chain is essential for understanding the true economy-wide effects of any substantial policy change, the implementation of a cap on GHG gases being a prime example.

To accomplish this feat, general equilibrium models must abstract from individual behaviors and aggregate to sectors of the economy, using our best understanding of individual behavior to build equations that mimic the expected response of a group of individuals to the new incentives they face. It is sometimes possible to use these models to examine whether the outcomes observed are consistent with certain behavioral assumptions, but it is difficult to arrange for tests of this sort that have strong statistical power because of the many assumptions that are necessary to build the solvable equilibrium model.

This requires the specification of greatly simplified preferences, technology, and institutional frameworks. Simplifying markets means putting aside detailed features about how markets are structured because it is difficult to characterize such detail in a solvable mathematical representation.

Equilibrium models may fall short in cases where it is necessary to know in detail how firms and institutions react to complex policies.

Experimental economics takes a third approach, using an abstract representation of the policy regime of interest and individual laboratory subjects with carefully structured choices and information. By establishing two laboratory settings that are identical except for one key variable of interest and repeatedly testing subject responses to incentives under one or the other of these "treatments," experimenters can look for differences in subject behavior between treatments, and, with a sufficient number of repetitions can draw statistically valid conclusions about the effect of the treatment variable on subject behavior. While the internal statistical validity of the test may be uncontroversial, the applicability to the actual policy regime—the external validity—depends on how well the laboratory setting captures the essential elements of the economic environment in which actual individuals will make choices in response to the policy. There is no simple statistical test to assure high external validity.

Experimental methods do have a couple of particular advantages in investigating the questions addressed in this report. First, while experiments do test individual behavior, they do not depend on an explicit detailed structural model of behavior. As we describe in the methodology section, the set of decisions made by participants in the allowance market are quite complex, consisting of several separate choices, each of which may be difficult to capture with a structural model. In fact, for the auction alone, it is not possible at this time to build an explicit structural model of the bidding behavior in emission allowance auctions that can be used to predict participant behavior. The decisions are sufficiently complex that many, quite different outcomes are consistent with a given set of conditions. Second, experiments have been widely used for investigating questions of the sort addressed in this study. These experiments have been subjected to critical review and have been implemented under a fairly wide variety assumptions and laboratory environments. Substantial consistency across different experiments and confirming repetitions of some results offers some measure of confidence in their usefulness. Third, designing and performing informative experiments may take less time than the sequence of tasks required for structural modeling accompanied by the testing of the model against available data. In this case, the availability of relevant data seems an almost insurmountable task.

Technical Appendix 2:
Compliance Planning in the Professional Simulation –
Detail on Structure of ET-Sim

Technical Appendix 2:
Compliance Planning in the Professional Simulation –
Detail on Structure of ET-Sim

We adapted the ET-Sim[24] application to simulate the key features of the California cap and trade market. Our intent was to have professional participants begin the market simulation process by developing a compliance strategy, based on their current expectations of California carbon allowances/offsets prices and availability and their initial take on their position relative to other players in the AB 32 market. We also tried to heighten the correspondence of the market simulation to actual trading situations by enhancing the coverage of the various alternative markets, including offsets, wholesale power, and energy efficiency projects

Offsets Markets: In an efficient carbon market, carbon offsets with identical compliance value to carbon allowances will trade at or near the same market price as allowances. Discounts to offset pricing relative to allowance prices are often the result of increased risk, usage limitation, and increased transaction costs. Discounts are indeed apparent in the California market, where offset prices are dramatically lower than allowances prices. Contributing factors to a higher risk, higher cost offset market in California include usage restrictions in the form of quantitative usage limits (QUL) (8% of a participant's total compliance obligation), after-the-fact invalidation risk, lack of a secondary market for a commoditized form of California Carbon Offset (CCO), and high bi-lateral transaction costs. What has emerged is an uncertain, highly fragmented market, with little standardization. As a result, the market has responded with transactional structure and trading products which reflect the myriad risks and which are priced accordingly.

In order to represent the market for offsets in California, ET-Sim defined offset products by their risk profile rather than by project methodology. Offset procurement in ET-Sim is defined at the point of transaction settlement with pricing indexed to allowance pricing. When a participant considers investing in an offset project, ET-Sim presents a range of costs and a range of potential offsets the project will generate. Only after the participant chooses to invest will they learn the true cost and actual offsets created. Projects that are not approved still incur project design and development costs.

Our analysis of the market identified three general categories of offsets, with different types of market risk, resulting in different pricing levels for each change relative to allowance price.

- Guaranteed delivery of offsets, priced at 85% current CO_2 price
    - Lowest risk (primarily counterparty risk) and thus lowest discount to allowance price
- Unit contingent, priced at 70% CO_2 price
    - Medium risk category with volumetric and counterparty performance risk
- Pre-compliance spot, priced at 55% current CO_2 price
    - Highest risk (regulatory risk), highest cost (economic, contractual, administrative), and offered at the greatest relative discount to allowance pricing.

[24] ET-Sim is a publicly-available training application developed by the U.S. Environmental Protection Agency to teach emissions trading to interested parties; see http://www.epa.gov/captrade/etsim.html.

At settlement, volumes delivered, net pricing, and regulatory performance issues are factored in and deliveries and payments may differ from expectations to reflect real risks and expenses such as offset counterparty credit, capital costs, invalidation, regulatory approvals, transactional, and delivery risks. After each successive offset block is purchased, the chance of approval for the next project is reduced, the range of costs is increased, and the range of offsets generated declines. The "supply curve response" is designed to reflect the concept that the most cost-effective projects are often selected first.

Wholesale Power: In the professional simulations using ET-Sim for this report, participants were offered choices between power purchase options that could be selected as part of their hypothetical California cap and trade compliance portfolio. Since utilities in California must serve existing load requirements, these resources must be chosen as part of the participant's compliance strategy to replace power generated by marginal generation resources in their portfolio's initial baseline. For the professional simulation, we assigned the merchant generator the responsibility for the emissions associated with the plants generating the selected power purchases.

Thus, the portfolio manager's compliance strategy must weigh the 1) cost-effectiveness of reducing their emissions compliance obligation, and 2) benefit of locking in the current GHG/CO_2 allowance price against their expectations for GHG allowance prices. For Compliance Period 1 (CP-1) simulations two wholesale power purchase options were offered.

1. *An existing natural gas-fired steam generator* at an implied market heat rate of 10,250 Btu/kWh and an emissions rate of 0.541 per metric ton CO_2/MWh, and
2. *Unspecified gas-fired power purchases* at a nominal heat-rate of 8,260 Btu/kWh and an emissions rate of 0.436 per metric ton CO_2/MWh (an ARB-assigned default emission rate for power delivered from unspecified sources).

Although the two CP-1 power purchase options included in ET-Sim are more limited than the actual options available to the California power market, they were selected to be representative. For validation purposes, the estimated initial wholesale power purchase prices were benchmarked against recent Western electricity and natural gas prices at various market hubs. The implied market heat rate is consistent with data derived from trade publications for on-peak electric and natural gas price data. For these simulations, natural gas delivered prices to California generators were assumed to be $4.00/MMBtu in 2013–14, $4.75 in 2015–17 and $5.25 in 2018–2020 (in nominal dollars). These values are consistent with a review of recent economic projections and through 2018 with reported forward curves.

The price of wholesale power in ET-Sim is updated by the cost of associated GHG/CO_2 allowances. At the outset, portfolio managers input a starting allowance price that is generally consistent with the most recent closing price of the highest volume exchange-traded December 2013 California Carbon Allowance (CCA) futures contract. At the close of each compliance period in the PEAR auction simulation exercises, the CO_2 allowance price is updated to reflect the most recent auction clearing price, with wholesale power thereby adjusted accordingly for the next period's trading.

Energy Efficiency – Electric Sector: In order to more accurately reflect the options load-serving entities have for managing their carbon exposure, the professional simulation allows market participants simulating the portfolios of publicly-owned and investor-owned utilities to reduce the loads they are required to serve by investing in energy efficiency (EE) projects. In the professional simulation implemented in the ET-Sim application, information about the cost of the EE project and the percent reduction in demand is displayed as a range of values. A participant will learn the true cost and actual demand reduction resulting from the EE option only after they choose to invest in the EE option.

EE investments thus indirectly reduce greenhouse gasses. Correspondingly, while the energy efficiency option within ET-Sim does not directly reduce the carbon compliance obligation of the load-serving entity, it does reduce the amount of generation the entity must dispatch or purchase to meet its load obligation. By reducing the load obligation, these market participants are able to either decrease the dispatch of their generation fleet or purchase less replacement power, or both.

In order to realize corollary emission reductions, participants will need to either reduce their electricity purchases or their generation at one or more electric generating units. When viewing the screen detailing the efficiency project, auction participants will see a cost-per-ton reduced range that assumes an average cost per ton of GHG reduction appropriate for that participant's entire fleet of units. Thus, by lowering generation at a high-emitting unit (usually coal-fired units), a participant may be able to reduce its emissions at a cost lower than that displayed in the simulation. Conversely, lowering generation at a low-emitting unit may mean that emission reductions will cost more than the amount displayed in the efficiency project details.

Thus load-serving entities have the potential to benefit by either:

1. Replacing more expensive energy purchases with less costly EE projects—this benefit only occurs if energy/load reduction costs are less than replacement energy (RE) costs (EE $/MWh < RE $/MWh),
2. Reducing operating costs by lowering the amount of offsets or allowances needed—this benefit only occurs if it is less costly to install an EE project than it is to operate the marginal generation that the project would replace (EE $/MWh < [Production Costs + cost of allowances and/or offsets necessary to cover the emissions from the marginal generation])

In the first case, the decision is identical to historic procurement decisions load-serving entities have had to make about investment in EE (albeit replacement energy may be more costly on the wholesale market with the addition of the costs of carbon allowances and/or offsets). As indicated above, the merchant generator keeps its responsibility for the emissions associated with its plant that is being dispatched (i.e., this simulation does not contemplate other forms of power purchase agreements).

In the second case, there is an additional value to EE that has not previously been monetized—the costs of carbon allowances associated with the generation being dispatched. Of course, this benefit is proportional to the carbon intensity of the generating units the EE project is displacing. Therefore, with the addition of a price on carbon, utilities with higher-carbon intensity are likely to see the value of energy efficiency increase more rapidly than will lower-carbon intensive utilities.

In general, market participants have two separate types of projects they may purchase—residential or commercial EE. For the purpose of the simulation, both cases work in a similar fashion. Specifically, the cost to purchase incremental EE improvements rises as additional projects are acquired (i.e., the second residential project is more expensive than the first project). In ET-Sim the market participant does not know the exact outcome of the EE project until after it is purchased.

However, auction simulation participants are given a range for each project before they make their purchase decision.

Technical Appendix 3:
Applicable Regulations – California Cap and Trade Program Simulation Treatments

Technical Appendix 3: Applicable Regulations – California Cap and Trade Program Simulation Treatments

California Cap and Trade Holding Limits

1. The ARB Holding Limit Formula

Precedent: The regulation's holding limits formula primarily reflects concern about traders' holdings relative to market size and accordingly, whether there is sufficient liquidity. Two key concerns related to the California market were:

- although the market has more than 350 covered entities in 2013—rising to 600 in 2015—it is still small relative to the originally anticipated regional WCI, or national, markets; and
- the state lacks a financial regulator comparable to the Securities and Exchange (SEC) or Commodity Futures Trading (CFTC) commissions.

Alternatives to holding limits include having multiple auction cycles or applying holding limits to financial intermediaries and speculators (Linklaters 2011). But as ARB noted, having more allowances available to the market does not necessarily guarantee greater liquidity (WCI 2010, 10).

Accordingly, ARB followed the holding limits precedent from the CFTC rule designed to address systematic risk in commodity futures markets. Derivative markets are different from emission markets and as a result are likely to function differently. The CFTC applies a two-tier structure in setting federal limits on agricultural products: Ten percent of the lagged open interest in contracts up to 25,000 contracts and 2.5 percent of the open interest, with the limits reset each year by reviewing a market's size and liquidity.

The ARB Holding Limits formula is consistent with the CFTC framework, which allows for higher market share in less-liquid markets. In practice, limits on the amount of allowances that covered entities may hold also may reduce incentive for them to trade actively. In other words, taking allowances out of the market for holding accounts in theory also may reduce liquidity.

As discussed earlier, California's GHG allowances are given a vintage for a particular emission budget year. Current year allowances can be applied to satisfy an annual emission compliance obligation for the current or any future year, or the triennial compliance obligation for any year in a current or future compliance period. Advance allowances, however, cannot be applied to cover any emissions that occur prior to the vintage year specified for that particular allowance. In an allowance's designated vintage year an Advance/Future Allowance becomes a Current Allowance.

California's two-tier Holding Limits formula is applied to these two categories of allowances: Current year/prior year and Advance/future year allowances, where the latter have separate holding limits for each future vintage. The ARB formula corresponds to the CFTC definition with "lagged open interest" as the Base = 25,000,000 allowances and "open interest thereafter" as (Annual Allowance Budget – Base). In each year the formula is applied to all the combined allowances held in an entity's accounts that are of current and prior year vintages, along with allowances from any vintage purchased from the Allowance Price Containment Reserve (APCR) and Advance

Allowances that have become Current after purchase. The Annual Allowance Budget is the Budget for the current year shown in Table 1[25].

Table 1. California's Annual Allowance Budget

	Budget Year	Annual Allowance Budget (Million Tonnes of CA GHG Allowances)
First Compliance Period	2013	162.8
	2014	159.7
Second Compliance Period	2015	394.5
	2016	382.4
	2017	370.4
Third Compliance Period	2018	358.3
	2019	346.3
	2020	334.2

The holding limit declines each year except in the transition to Compliance Period (CP-) 2, as the Annual Allowance Budget, i.e., the "cap" declines.

For all Advance Allowances the same formula is applied, but the formula applies to each single future year's vintage of allowances using each year's Annual Allowance Budget from Table 1. Hence, an entity may not hold more than the formula's specified amount of Advance Allowances for each future year's vintage. When an Advance Allowance becomes a Current Allowance, it will be included in the Holding Limit calculation for Current and prior year vintages.

The ARB formula is:

Holding Limit = 0.1*Base + 0.025*(Annual Allowance Budget – Base)

In which:

"Base" equals 25 million metric tons of CO_2e.
"Annual Allowance Budget" is the number of allowances issued for the current budget year.[26]

Table 2. Cap and Trade Holding Limits Without the Limited Exemption

Year	Holding Limit (million tonnes)
2013	5.945
2014	5.868
2015	11.738
2016	11.435
2017	11.135
2018	10.833
2019	10.533
2020	10.230

[25] The budget levels are for a California-only market; linkage with Quebec would increase holding and purchase limits by between 10 and 15% depending on compliance period.
[26] Article 5, Subarticle 11: Trading and Banking, §95920 (d) (1) and (e).

The Holding Limits apply to all allowances combined in the General Holding Account, the Compliance Account and the Exchange Account.[27] Penalties may be assessed when the holding limit is exceeded.

2. The ARB Limited Exemption Adds to the Effective Holding Limit

Certain large emitters have annual compliance obligations that are significantly in excess of the specified Holding Limits shown in Table 2. For example, Chevron and Calpine have historic emissions on the order of ~12 and ~8 million tonnes/year, respectively, and in 2010, PG&E's total GHG emissions were about 15.6 million tonnes/year from self-generation and purchased power.[28]

Allowances that have been put into the Compliance Account will qualify for a Limited Exemption (LE) from the Holding Limit. However, allowances placed into the Compliance Account are not removable, except when they are surrendered to meet the triennial compliance deadlines by November 1, 2015, 2018, and 2021, or annual compliance deadlines in other years. Thus, while the LE is essential for larger emitters to acquire sufficient allowances over time in order to comply, each allowance that contributes to the LE must be sequestered and made unavailable for trading in the allowance market.

The LE is calculated based on Article 5 of the regulations:[29]

> "(A) The limited exemption is the number of allowances which are exempt from the holding limit calculation after they are transferred by a covered entity or an opt-in covered entity to its compliance account.
>
> (B) On June 1, 2012 the limited exemption will equal the annual emissions most recent emissions data report that has received a positive or qualified positive emissions data verification statement.
>
> (C) Beginning in 2013 on October 1 of each year the limited exemption will be increased by the amount of emissions contained in the most recent emissions data report that has received a positive or qualified positive emissions data verified statement during that year.
>
> (D) If for any year ARB has assigned emissions to an entity in the absence of a positive or qualified positive emissions data verification statement the calculation of the limited exemption will use the assigned emissions.

---

[27] Allowances in an Exchange Clearing Holding Account will be included in the calculation of the Holding Limit for the entity listed as the purchaser in the transfer request. Purchases or transfers that would cause Holding Limits to be exceeded will not be approved by the Executive Officer. Allowances held in Limited Use Holding Accounts will not be included in calculating each entity's holding limit.

[28] PG&E 2010 Power/Utility Reporting Protocol Report. Note that PG&E, SCE, and SDG&E's suppliers that deliver power to these EDUs must acquire allowances in order to satisfy their compliance obligations as first deliverers. Thus, supplier emissions make up a significant fraction of the above-reported emissions for PG&E. Updated reports of GHG emissions are expected to be available in early 2013. However, some contracts with merchant generators are being renegotiated to give the IOU the responsibility to procure some or all of the allowance needs associated with power deliveries under a particular contract. In addition, as the allowance cap tightens and as renewable power resources continue to grow, fossil-fired GHG emissions should be displaced, so that IOU allowance needs in a normal future year should be below GHG emission levels in the last decade.

[29] Article 5, Subarticle 11: Trading and Banking, §95920 (d) (2).

> (E) For the first compliance period all reported emissions or assigned emissions used to calculate the limited exemption will include only the emissions associated with the scope for the program during the first compliance period.
>
> (F) Beginning in 2015, all reported emissions or assigned emissions used to calculate the limited exemption will include the emissions associated with the change in scope taking place in 2015.
>
> (G) On January 1, 2015 the limited exemption will be increased by the amount of emissions included in the emissions data report received during 2014 but not yet included in the limited exemption pursuant to section 95920(d)(2)(E).
>
> (H) On December 31 of the calendar year following the end of a compliance period, the limited exemption will be reduced by the sum of the entity's compliance obligation over that compliance period."

As of June 30, 2012, one historical year's verified GHG emissions can contribute to the LE. After October 1, 2013 two years' verified emissions could contribute, for example, emissions from 2011 and 2012. Up to four years' emissions can, in principle, contribute to the LE by October 2015, but on December 31, 2015, GHG emissions from 2013 and 2014 will be removed from the LE calculation. In the year following the next two compliance periods, the LE will be reduced on December 31 by the sum of each entity's compliance obligation over the completed compliance period.

Notably, verified emissions from both 2011 and 2012 can continue to contribute to the LE throughout all three compliance periods. Nevertheless, sufficient allowances must remain in the Compliance Account to maintain the LE, which may fluctuate in size as allowances are transferred in to increase the LE and out for compliance. Important deadline dates in the last quarter of the year are likely to cause some additional variations in allowance market liquidity and prices.

Moreover, purchasing a quantity of allowances to establish an allowance bank/compliance cushion equal to two years' emissions may not be economically or administratively feasible, particularly given the large quantities needed by large emitters. Given that covered entities other than EDUs may only purchase 15% of the allowances in each auction, it could be difficult for industrial firms to acquire a sufficient bank through auction purchases. In addition, until the LE is created and grown by adding allowances to the Compliance Account up to the allowable LE, the Holding Limits shown in Table 2 will apply.[30]

Even though the investor-owned utilities may purchase up to 40% of the allowances in an auction, the CPUC has imposed annual purchase limits that might restrict the rate at which allowances and offsets are acquired. Moreover, it is likely that utilities will prefer to procure allowances in order to create and maintain a bank that will assure their compliance, rather than to increase market liquidity by keeping allowances in their general holding accounts to make more allowances available for trading in the market.

[30] During 2013 and 2014, EDUs can purchase up to 40% of the Current Allowances offered in an auction, while non-covered entities, such as financial and investment firms, may only buy up to 4% of the Current Allowances in an auction, relying on the secondary market for larger transactions. The purchase limit for the Advance auction is 25% of the allowances offered for auction for all entities.

In 2015, the AB 32 cap and trade market scope and allowance budget expands to cover suppliers and users of natural gas, RBOB,[31] distillate fuel oils, LPG, and transportation fuel distributors. Although Chevron's Richmond oil refinery is already covered in CP-1 for process emissions, emitting about 4.5 MT GHG annually, in 2015 the actual carbon content of Chevron and other refiners' transportation fuel products will become subject to the cap and their compliance obligations will increase substantially. Note that gas and electric utilities, large fuel distributors, and other CP-2 entities will not be able to take advantage of an LE associated with CP-2 obligations during CP-1, and therefore cannot "pre-load" by purchasing significant numbers of allowances prior to the jump in their compliance obligations in 2015.[32]

However, for all covered entities with compliance obligations in CP-2, the potential LE will be increased on January 1, 2015 to include 2013 emissions associated with the expanded scope of coverage in CP-2. On October 1, another prior year's emissions can be added to the LE. It will take some time for CP-2 covered entities to take advantage of their individual LEs in CP-2.

In order to systematize the foregoing analysis, we developed a spreadsheet tool[33] providing for user inputs of CO_2 emissions and expected allowance prices, which facilitates the iterative search for an approximation of the least-cost acquisition and distribution of allowances.

3. The CPUC Purchase Limit

The CPUC's Direct Compliance Obligation Formula calculates the maximum quantity of compliance instruments, including CA GHG allowances and offsets, which an IOU would be allowed to purchase in the current year for all vintages of allowances combined.[34] The purchase limit formula is estimated for each year assuming perfect forecasting of emissions out to 2020. It incorporates a forecasting approach that increases base case expected emissions forecasts by using a higher than expected market heat rate for production simulations. This approach is intended to allow some cushion for procuring allowances above the anticipated need. Although the higher emission forecasts under the CPUC's methodology are probably about 20% to 30% higher than base case forecasts, low hydro-year scenarios, continuing nuclear plant outage scenarios, and high electric demand growth scenarios could all project even higher future emissions. The forecast can be increased by providing updated analysis and submitting an advice letter to the CPUC.

Under the CPUC framework, the IOUs would not be allowed to purchase allowances or offsets with vintages more than 3 years from the current year.[35]

The formula is as follows:

---

31 Reformulated blendstock for oxygenate blending.

32 In an August 28, 2012 teleconference with ARB, ARB explained that "a CP-2 entity should not be registered as a covered entity in CP-1 unless they also have a CP-1 obligation. If a CP-2 entity is registered in CP-1, their CP-2 covered emissions are not considered part of the limited exemption in CP-1. For example, a company may own a refinery that has an obligation under CP-1. They may also be considered the fuel distributor in CP-2. The CP-2 fuel obligation is not allowed as part of the limited exemption until CP-2 unless this facility has chosen to opt-in, in which case they would have an obligation in CP-1."

33 a free download available from www.peartree.com.

34 The CPUC defines "purchase" as taking title of the instrument when it is delivered.

35 CPUC Decision D.12-04-046, Appendix 1. April 2012.

$L_{cy} = A + (100\% * FD_{cy}) + (60\% * FD_{cy+1}) + (40\% * FD_{cy+2}) + (20\% * FD_{cy+3})$, where:
"L" is the maximum number of GHG compliance instruments an IOU can purchase for purposes of meeting its direct compliance obligation.

"A" is the utility's "net remaining compliance obligation to date," calculated as the sum of the actual emissions for which the utility is responsible for retiring allowances (or purchasing on behalf of a third party) up to the Current Year, minus the total allowances or offsets the utility has purchased up to the Current Year that could be retired against those obligations. This term in the calculation ensures the IOUs are always able to buy sufficient allowances to cover any prior years' shortfalls, given that actual emissions may end up being less than forecast and/or prior decisions about how much procurement to do.

"FD" is the utility's "forecasted compliance obligation," the projected amount of emissions for which the utility is responsible for retiring allowances, or responsible for purchasing on behalf of a third party, calculated using an implied market heat rate (IMHR) that is two standard deviations above the expected IMHR consistent with an approach described by PG&E.

"cy" is the current year, i.e., the year in which the utility is transacting in the market.

Should the above equation result in a negative number in a given year, the utility's Direct Compliance Obligation Purchase Limit for that year should be set at zero.

The CPUC also promulgated a Financial Exposure Purchase Limit to limit the quantity of GHG compliance instruments that can be purchased to hedge financial exposure. This formula requires that an estimate of the utility's financial exposure to GHG costs be embedded in the price of energy over and above the costs of meeting the Direct Compliance Obligation. As such, the allowable purchases under this formula are expected to be less than under the Direct Compliance Obligation formula. The sum of both formulae constitutes the CPUC annual purchase limit. Since the A term incorporates prior purchases, the CPUC limits also act as a de facto holding limit.

I. California Cap and Trade Allowance Price Containment Reserve

In its regulatory impact analysis, the ARB recognized that conditions could develop that could lead to higher than anticipated program costs, and determined that it was appropriate to examine whether additional features could enable the program to adapt to changing conditions in a way that reduces the risk of extraordinary price excursions.

ARB found in review of predecessor initiatives a precedent in creating an allowance reserve that could be an "effective means to reduce the risk of higher than anticipated costs while maintaining the environmental integrity of the program."[36] To offer cost containment, reserve allowances could be made available for sale at quarterly intervals with participants bidding at pre-established prices across a three-tiered auction structure. The resulting system, known as the Allowance Price Containment Reserve (APCR) is structured to provide the flexibility necessary to meet surges in allowance demand without exceeding the cap.

While providing a measure of price certainty, the ARB noted that the size of the reserve was not unlimited, and that under extreme conditions it could be exhausted. As a result, there has been concern that in order to guarantee that the state will not be subject to extraordinary carbon price volatility, the APCR must be replaced or augmented by a hard price cap.

The APCR is created by taking a number of allowances from the program's allowance budget across all three compliance periods. The share of allowances reserved from auction for each period[37] to fill the APCR is:

- 1% for years 2013-2014
- 4% for years 2015-2017
- 7% for years 2018-2020

for a total of 121.8 million, or 4.9 percent of the total 2.5 billion allowances allocated from 2013–2020. Once committed to the reserve, allowances are no longer restricted to surrender in a particular vintage year; the entire stock is available at any reserve sale. In endeavoring to limit the role of the allowance reserve to support compliance obligations, ARB restricted participation in the reserve sales to covered sources and specified that purchases from the reserve be placed directly in the compliance account.[38]

Sales of allowances from the APCR are divided into three equal-sized tiers, each containing 40.6 million allowances. In 2013, one of these tiers will become available at each of the following prices: $40, $45, and $50. After 2013, the offering price of reserve allowances increases by 5 percent annually plus inflation. The first reserve sale will occur on March 8th, 2013, with each subsequent sale being conducted six weeks after every quarterly auction.[39]

---

[36] Appendix G, Allowance Price Containment Reserve Analysis, p. G-6; http://www.arb.ca.gov/regact/2010/capandtrade10/capv3appg.pdf

[37] Section 95870

[38] In the lab sessions, allowances purchased from the APCR actually went into the subject's holding account but, apart from one allowance out of the 64 purchased from the APCR, these APCR allowances were immediately transferred to the subject's compliance account to get out from under the holding limit, which has the same effect.

[39] Section 95913

One of the questions posed regarding the APCR is whether it puts potential bidders at perpetual risk of "buyer's remorse" since its structure requires covered sources to forecast whether auction prices will be above or below reserve sale prices. If below and sufficient allowances will be available to the entity from the auction or secondary market, then one would expect the entity not to bid for reserve allowances. If higher, then entities must choose between trying to compete in the auction or run the risk of buying either too many or too few allowances in the reserve sale trying to beat the market. Given the sequencing of these transactional events, bidders must decide which to do before bidding in the reserve sale, and thus the APCR structure "could introduce considerable uncertainty into the bidding process."[40]

If the total number of allowances requested by covered sources participating in the sale is less than the number of allowances in the tier, each entity receives what they requested. Unsold bids from one tier can be sold to bidders for the next higher tier using a "roll-down" process described below. However, if the total number of allowances requested from a tier exceeds the number of allowances in the tier, each entity receives the proportion of the tier that they contributed to the total demand.

Under this approach, suppose we assume for simplicity that five allowances were offered at each tier. If there are no bids at the third tier price, eight bids at the second tier price, and three bids at the first tier price, under the proportional plus roll-down system we would use the following steps[41] to determine the number of allowance sold from each tier.

1. The purchases are fulfilled from the lowest tier to the highest. Following the fulfillment order means that the three allowances at the first tier price would be filled first, which leaves two allowances available at the first tier price.
2. The two remaining first tier allowances are next sold (for the tier 1 price) based on random numbers assigned to the eight that were bid at the second tier price. This leaves six allowances bid for at the second tier price.
3. There are only five allowances available, so the bidders each receive a proportionate share, or 5/6 of a unit at the second tier price. None would be sold at the third tier price.

So, the end result is five allowances sold at the first tier price and five sold at the second tier price. The original bidders at the first tier price get the three they requested. The original bidders at the second tier price only get seven of the eight they wanted, but two of those are sold at the first tier price.

[40] "Experimental Evidence on the Properties of the California's Cap & Trade Price Containment Reserve" R. Bodsky, D. Donato, K. James, D. Porter, Chapman University, May 2012

[41] The example is re-stated from ARB email responses (dated September 6, 2012) follow-up to a teleconference discussion with the Project Team (held on August 28, 2012).