Report of the Market Simulation Group on Competitive Supply/Demand Balance in the California Allowance Market and the Potential for Market Manipulation

By Severin Borenstein, James Bushnell,
Frank A. Wolak, and Matthew Zaragoza-Watkins*

June 2014

* This research was performed under a contract with the California Air Resources Board. Borenstein, Bushnell, and Wolak are members of the Market Simulation Group that advised ARB. Zaragoza-Watkins worked with the MSG as a researcher. We thank Elizabeth Bailey, David Kennedy, Ray Olsson, Billy Pizer, Rajinder Sahota, and Emily Wimberger for their input. We also thank participants in the 18th Annual POWER research conference in March 2013 and the NBER Energy & Environmental Economics program conference in January 2014 for valuable comments. The opinions in this paper do not represent those of the California Air Resources Board or any of its employees. Email addresses: Borenstein: severinborenstein@berkeley.edu; Bushnell: jbbushnell@ucdavis.edu; Wolak: wolak@zia.stanford.edu; Zaragoza-Watkins: mdzwatkins@berkeley.edu.

EXECUTIVE SUMMARY

California's Cap and Trade market in greenhouse gases (GHG) is now in its third calendar year, with the first allowance auction taking place on November 14, 2012 and compliance obligations commencing on January 1, 2013. A key design element of the system is its limited price-collar mechanisms that place soft lower and upper bounds of allowance prices. To date the market prices have held at or near the lower bound "floor" prices established by the allowance auction reserve price. However, the market will be entering important new phases over the next 18 months. The first firm information on covered emissions during the first compliance phase (2013-2014) will emerge in November of 2014. Coming near the end of this compliance phase, the release of this information is the first opportunity for the market to confirm expectations of the supply and demand balance of allowances during the first phase. Starting in 2015, the market will expand to include several new sectors, most significantly transportation fuels and the bulk of natural gas consumption in the state. It is therefore important to anticipate any possible shocks to the market that can arise as it matures and expands over the course of the next two years.

One central issue is the status of the price-collar mechanisms. While the details of California's price-collars are described in regulations developed by the California Air Resources Board (ARB), recently approved regulatory changes would alter the exact manner in which the price ceiling – known as the *allowance price containment reserve* (APCR) mechanism – would be applied and the degree to which it could mitigate uncertainty over prices.[1] A key question relating to this issue is the extent to which either the auction reserve price or APCR price is likely to be relevant, that is, the probabilities that market prices may be near the price

[1] The regulations are available at: http://www.arb.ca.gov/cc/capandtrade/september_2012_regulation.pdf. See also the ARB Board resolution dated October 18, 2012 at http://www.arb.ca.gov/cc/capandtrade/final-resolution-october-2012.pdf and an issue analysis from the Emissions Market Assessment Committee dated September 20, 2012 at http://www.arb.ca.gov/cc/capandtrade/emissionsmarketassessment/pricecontainment.pdf.

floor or the APCR soft price ceiling.[2] A second key question is whether some market participants may be able to strategically change the allowance price, in particular by buying more allowances than they need and withholding them from the market in order to sell a portion later at a higher price.

In this report, we simulate distributions of possible market outcomes in order to address these questions. We first develop estimates of the distribution of competitive allowance prices and the probabilities that one of the price containment mechanisms may be binding. A key factor driving these probabilities is the amount by which GHG-producing entities will reduce their emissions. This reduction is likely to be a highly non-linear function of allowance price. Specifically, we find that a large quantity of emissions reductions are mandated by programs auxiliary to the cap and trade mechanism, and will therefore be available at or below the auction reserve price. Other businesses can reduce their need to purchase allowances at a cost that is below or only slightly above the auction reserve price. Relatively little additional emissions abatement is likely to be available as the price climbs, at least before the price rises high enough to trigger additional supply of allowances from the price containment reserve.

Our key simulation findings are

1) The steeply rising cost of emissions abatement between the auction reserve price floor and the price containment reserve ceiling, along with relatively inflexible supply of abatement below the price containment reserve, implies a bi-modal distribution of prices with most of the probability mass at either low or high price outcomes.

2) Under most scenarios, the most likely 2020 market price will be very close to the auction reserve price floor.

3) However, under all scenarios, there is a smaller but significant risk that the allowance price containment reserve will be exhausted at or before 2020.

[2] As described below, the APCR makes a limited number of extra allowances available if the price hits certain levels.

For scenarios in which there is low or medium availability of carbon offsets and relatively little reshuffling of electricity imports, this probability ranges from as 4%-25%.

4) The probability of reaching, but not exhausting, the APCR by 2020 falls between 8% and 31% under low and medium abatement scenarios. We find there is low risk of exhausting the APCR *before* the third compliance phase, which begins in 2018.

5) There are small but significant probabilities that the market could reach the APCR during one of the first two compliance phases. Under our low and medium abatement scenarios, there is a 2%-4% chance of reaching the APCR (assuming no strategic withholding behavior by market participants, which we investigate later) during the first compliance phase. Under our low and medium abatement scenarios, the probability of reaching the APCR (assuming no withholding) during the second compliance phase ranges from 4%-17%.

6) There is a straightforward mechanism in which firms can withhold allowances from one phase of the market by banking them into compliance accounts for future compliance phases. We study the risks that such strategies could inflate prices during the first and second compliance phases. The largest risk is that one (or more) of a small number of large firms acquires significantly more allowances than it requires for the first compliance phase, and deposits these extra allowances into compliance accounts for use in later periods. This could result in 10% of available allowances or more being removed from an approximately 330 million metric tons (MMT) market in the first compliance period. Such a strategy, if attempted, would increase the probability of reaching the APCR from 2% (absent withholding) up to about 7% or higher with medium abatement and from 4% to about 13% or

higher under a low abatement scenario.[3]

This strategy would be most likely to affect the allowance price for the first compliance period during 2015, after the first compliance period ends but before the final surrender of allowances for this period.

7) During the second compliance phase, a similar strategy could increase the risk of needing to access the APCR from around 15% to as much as 30% or higher.

We provide several recommendations to reduce the risk of very high allowance prices due to either the competitive supply/demand balance or a withholding strategy. It is important to emphasize that the higher prices are allowed to rise, the more potentially profitable a withholding strategy becomes. Therefore an unambiguous policy that credibly limits the maximum allowance price is important to market stability *and* a strong deterrent to attempts at market manipulation. Our major recommendations are:

1) *Establish policies that reinforce the viability of the allowance price containment reserve.* The recently adopted rule changes that make adjustments to the APCR only address transient shortfalls and therefore do not address the threat that there could be a supply/demand mismatch for the entire 8-year program.[4] If there were not enough allowances over the 8-year period to cover the cumulative emissions under the cap, then there would be no policy in place to further restrain prices. It is likely that an ad-hoc government intervention into the market might occur under such a circumstance. This would prove to be extremely disruptive to both the market and to the broader policy goals of AB 32.

[3]The impact of withholding depends very much on the exact strategy a firm would pursue, which is difficult to predict. It is also difficult to know how large a postion a firm could acquire without raising the acquistion price it would face. As a result, we report ranges of the possible impact of withholding, but do specify precise estimates.

[4]See http://www.arb.ca.gov/regact/2013/capandtrade13/capandtrade15dayattach1.pdf.

We therefore recommend that a policy be established to ensure that the APCR could not be exhausted. The Air Resources Board should stand ready to expand the pool of allowances in order to maintain the market price at or below the highest price step of the APCR. Two alternatives that could achieve this goal are allowing sales of post-2020 compliance period allowances or allowing direct or indirect use of compliance instruments from other GHG markets such as the European Union Emissions Trading System (EU-ETS) or the Regional Greenhouse Gas Initiative (RGGI) under such circumstances.

2) *Allow Conversion of Allowance Vintages.*

Currently, market participants are not allowed to use allowances from later vintages for compliance in earlier phases. For example a vintage 2015 allowance cannot be used for compliance obligations in phase I, which concludes on December 31, 2014, but for which final surrender doesn't occur until late in 2015. This boundary between phases creates the prospect of transitional shortages in which allowance prices in the expiring phase rise to the APCR while current vintage allowance remain near the price floor. As we demonstrate, the potential for withholding increases the probability of such an outcome.

A second concern with the current design is the potential that allowances could end up inefficiently owned *ex post*. As firms acquire allowances according to their expectations of needs, shocks to individual firms or even sectors could result in too few allowances from a current phase being available to some sectors while others hold a surplus they are unable to sell, if their surplus is held in compliance accounts rather than holding accounts.

Conversion – for a fee – of the vintages of allowances held by market participants would greatly reduce the risk and consequence of both problems. Under this proposal firms would be allowed, for instance, to purchase 2015 or

later vintage allowances during 2015 and convert them to meet their phase I obligations, for which final allowance surrender would occur in November 2015. To prevent stakeholders from casually undertaking such conversions, a cost in terms of either a conversion fee or the number of allowances surrendered could be imposed. For example, ARB could require 1.25 vintage 2015 allowances be converted to yield 1.0 2014 vintage allowance, thereby imposing an implicit 25% cost on the conversion. Alternatively, a conversion fee, for example $2.50 or $5 per allowance, could be applied to each converted allowance. Firms would only avail themselves of this option if the allowance price in the expiring phase rises above the price of the later vintage allowance by an amount greater than the conversion fee. At the same time, this option would bound the extent to which prices in the expiring phase could rise above later vintage allowance prices. This would greatly reduce the incentive to attempt to raise prices in the expiring phase by withholding allowances from that phase.

The proposal would also address accidental over-compliance by some participants, as well as strategic withholding, either of which could create an artificial shortage at the end of a compliance period. Firms would be able to purchase future vintages (at a premium) to meet their needs.

3) *Maximize the Timeliness and Quality of Information Available to the Market.* Currently the market suffers from opacity in several important areas. First, there is almost no way to observe, even indirectly, the emissions associated with electricity imports and the only source of official information will arrive with up to a nearly two-year lag on the market. Second, current proposals would limit the public availability of information on the allowance holdings of individual firms. We recommend steps be taken to increase the frequency with which key emissions figures, particularly from electricity imports, be provided to the market. We also recommend that if individual allowance holdings must be held confidential, statistics on the overall con-

centration of allowance holdings be made available. In this way, market participants would be able to detect attempts by one firm to acquire a substantial long position and take measures to defend themselves against any attempts at withholding allowances from the market.

I. INTRODUCTION

California's Cap and Trade market in greenhouse gasses (GHG) is now in its third calendar year, with the first allowance auction taking place on November 14, 2012 and compliance obligations commencing on January 1, 2013. The quantity of available allowances has been set for the first eight years, through 2020, after which the future of the program is uncertain. This market is a modified cap and trade system with a limited price-collar mechanism. There is an auction reserve price, managed through adjustments to the supply of allowances to the periodic auctions that sets a soft floor price for the market. This price floor rises each year. There is also an allowance price containment reserve (APCR) designed to have a restraining effect on prices on the high end by adding a pre-specified number of allowances to the pool when prices exceed pre-specified levels.

While the details of California's price-collars are described in regulations developed by the California Air Resources Board (ARB), recently approved regulatory changes would alter the exact manner in which the price ceiling – known as the *allowance price containment reserve* (APCR) mechanism – would be applied and the degree to which it could mitigate uncertainty over prices.[5] A key question relating to this issue is the extent to which either the auction reserve price or APCR price are likely to be relevant, that is, the probabilities that market prices may be near the soft price floor or the APCR soft price ceiling.[6] A second key question is whether some market participants may be able to strategically change the allowance price, in particular by buying more allowances than they need and withholding them from the market.

In this report we first develop estimates of the distribution of competitive allowance prices and the probabilities that one of the price containment mechanisms

[5] The regulations are available at: http://www.arb.ca.gov/cc/capandtrade/september_2012_regulation.pdf. See also the ARB Board resolution dated October 18, 2012 at http://www.arb.ca.gov/cc/capandtrade/final-resolution-october-2012.pdf and an issue analysis from the Emissions Market Assessment Committee dated September 20, 2012 at http://www.arb.ca.gov/cc/capandtrade/emissionsmarketassessment/pricecontainment.pdf.

[6] As described below, the APCR makes a certain number of extra allowances available if the price hits certain price levels.

may be binding. A key factor driving these probabilities is the amount by which GHG-producing entities will reduce their emissions. The amount of these actions that are ultimately undertaken is a highly non-linear function of allowance prices. We find that a large quantity of emissions reductions are mandated by programs auxiliary to the cap and trade mechanism, and will therefore be available even at or below the auction reserve price. Other business activities can reduce their need to purchase allowances at a cost that is below or only slightly above the auction reserve price. While these programs and business activities will substantially lower the demand for allowances even at very low allowance prices, we show that, in part because of the design of the program, relatively little additional emissions abatement is likely to be available as the price climbs, at least before the price rises high enough to trigger an additional supply of allowances from the price containment reserve.

This steep supply of emissions abatement between the effective price floor and the price containment reserve, along with a relatively inflexible supply below the price containment reserve, implies a bi-modal distribution of prices with most of the probability mass at either low or high price outcomes. A primary factor determining where in that distribution the market will equilibrate is the "business as usual" (BAU) emissions level that would result if there were no GHG reduction activities. BAU emissions are substantially the result of economic activity driving electricity consumption and vehicle travel, as well as the emissions intensities of those activities, and emissions from natural gas combustion in the residential and commercial sectors and industrial processes. In this paper we develop estimates of these drivers of emissions utilizing forecasting techniques adapted from time-series econometrics, which we apply to emissions and economic data from 1990-2011, in order to forecast future emissions and the uncertainty of emissions.

Our empirical assessment of the potential demand for, and supply of, emissions allowances, as well as the offsets that augment this supply, suggests that the most

likely 2020 market price will be very close to the auction reserve price floor.[7] In all of the scenarios we examine, we also find a low probability that the price will be in the intermediate range above the auction reserve price floor and below the containment reserve price. Thus, most of the remaining probability weight is on outcomes in which some or all of the allowances in the price containment reserve are needed. Moreover, for all scenarios that we consider likely, there is a non-trivial probability that allowance prices will be above the highest price in the price containment reserve.

We also analyze the supply/demand balance in the allowance market during the first compliance period in isolation and during the first and second compliance periods combined but in isolation from the third compliance period. This analysis addresses the probability that the market could experience limited supply relative to demand in earlier compliance periods even if the supply is less tight over the full 8-year life of the program.

We then turn to the possibility that a market participant might be able to exert market power or manipulate the market for emission allowances. For the ends of the first and second compliance periods, we examine whether any market participant might have an incentive to buy more allowances than it needs and withhold supply from the market. By doing so, the entity may be able to profitably sell some of these allowances at a much higher price than the one it bought them at. Although we find this is not the most likely outcome, there is significant probability that market participants that are allowed to hold large long positions in the allowance market could be in a position to influence the price in this way.

The remainder of this analysis proceeds as follows. Section II gives an overview of the possible outcomes in the market for California emissions allowances given the characteristics of the supply and demand for GHG emissions abatement. Sec-

[7] Throughout this paper we will refer to an "allowance market." The trading of allowances and their derivatives will be arranged through several competing and coexisting platforms including quarterly auction of allowances by the State. We assume that prices between these markets will be arbitraged so that all trading platforms will reflect prices based upon the overall aggregate supply and demand of allowances and abatement.

tion III describes how we model the Business As Usual (BAU) drivers of GHG emissions over the 2013-2020 life of the program using a Vector Autoregression (VAR) model that imposes the restrictions implied by the existence of cointegrating relationships among the elements of the VAR. In Section IV we explain how we incorporate into the simulations the major additional California GHG reduction programs, known in California as "complementary policies," though they may not be complements to the cap-and-trade program in the economic sense. These include a renewables portfolio standard (RPS) that will increase electricity generation from renewable sources, a fuel economy standard that will reduce fuel use per vehicle mile traveled, a low-carbon fuel standard (LCFS) that will reduce the measured emissions intensity of the transport fuel used, and additional programs to improve non-transport and transport energy efficiency. Even though the impacts of these programs should be largely independent of allowance prices, the effects of these programs, as with the allowance market, will be highly dependent on the economic and emissions variables that we model in the VAR.

Section V analyzes the reduction in reported emissions related to other programs and activities in California, including both consumer response to higher prices for electricity, transport fuels, and natural gas, and two other important activities, reshuffling and offsets. Reshuffling, also known as "contract shuffling" or "resource shuffling", occurs when output of an energy product is reallocated among buyers in different regions so that the entities covered by the cap and trade program are buying the lower-carbon version and uncovered entities are buying the higher-carbon version, but no reduction in total emissions results.[8] Due to the California cap and trade market, there is likely to be significant "reshuffling" of electricity purchases among buyers and sellers across state lines. Offsets are emission reductions from sources not covered by the cap and trade program. Production of such offsets can then be credited to offset buyers against their allowance

[8]We distinguish between reshuffing and classical leakage, because reshuffling typically involves no change in the emissions producing activities in and outside of the region or industry covered by the cap-and-trade program.

obligation. As explained below, offsets are envisioned to augment significantly the supply of allowances in the California market, but there is a great deal of uncertainty as to how much offset supply will ultimately occur. In section VI, we bring together the analysis of abatement pathways with the previous estimates of emissions to forecast the possible supply/demand balance in the market and the probabilities of different price outcomes over the entire 8 years of the program as currently established, and during the earlier compliance periods. In section VII, we turn to the issue of market manipulation. We analyze the ability and incentive of a market participant to establish a large long position in the market and then withhold supply in order to drive up the price. We also consider the impact such a strategy would have on other market participants and end-use consumers. In section VIII, we move from analysis to policy recommendations. We put forth a number of possible policy changes that help to avert price volatility, high allowance prices, and the potential for market manipulation. We conclude in section IX with a broader discussion of our findings for the use of cap and trade programs to address climate change.

II. THE CALIFORNIA CAP AND TRADE MARKET

We focus on estimating the potential range and uncertainty in allowance prices over the entire 8-year span of the market.[9] The underlying source of demand for allowances will be emissions of GHGs from the covered entities, which will be a function of the levels and intensities of their emissions-producing activities. Banking and limited borrowing of allowances is permitted between the years of each compliance period and banking is permitted between compliance periods. Because of the relatively generous allowance budgets in the earlier years and a

[9] In late 2013, the ARB finalized plans to link California's cap and trade market with the market in Quebec, Canada as of January 1, 2014. Our analysis does not include Quebec, though it could easily be extended to do so if comparable data were available for Quebec. Quebec's total emissions were roughly 1/7 that of California. The supply-demand balance of allowance in this province could alter the probabilities presented in this paper.

policy change adopted in 2014,[10] under nearly all scenarios, emissions during the first two compliance periods (ending 12/31/14 and 12/31/17) will not exceed the caps, so the eight years of the market are likely to be economically integrated. As a result, we examine the total supply and demand balance over the entire eight years of the program (2013-2020). Because there is a large degree of uncertainty around the level of BAU emissions, we pay particular attention to establishing confidence intervals for the time path of annual emissions from 2013 to 2020.

The number of allowances available in the California GHG cap and trade program derives from the allowance cap, a portion of which is allocated to the APCR.[11] Of the 2,508.6 million metric tonnes (MMT) of allowances in the program over the 8-year period, 121.8 MMT of allowances are assigned to the price containment reserve to be made available in equal proportions at allowance prices of $40, $45, and $50 in 2012 and 2013. In later years, these price levels increase by 5% plus the rate of inflation in the prior year.

The supply of abatement is multi-faceted and features several elements that are either unique, or present in a more extreme form, in California. These elements combine to create an extremely steep abatement supply curve, which we will demonstrate implies the potential for a very wide distribution of price outcomes. Abatement of capped emissions will flow through two mechanisms: a direct effect in which firms or consumers reduce emissions in response to a level of allowance prices, and an independent effect in which emissions are reduced due to additional "complementary policies" outside the cap and trade program.

The supply of relatively price-independent abatement comes from (a) complementary policies that abate GHGs independent of the price in the market, (b) activities that reduce measured GHGs due to the process of accounting for

[10] See the ARB Board resolution dated October 18, 2012 at http://www.arb.ca.gov/cc/capandtrade/final-resolution-october-2012.pdf and an issue analysis from the Emissions Market Assessment Committee dated September 20, 2012 at http://www.arb.ca.gov/cc/capandtrade/emissionsmarketassessment/pricecontainment.pdf. For the recently adopted changes, see http://www.arb.ca.gov/regact/2013/capandtrade13/capandtrade15dayattach1.pdf.

[11] A recently policy change that was recently approved by the ARB Board will allow reallocation of a large number of allowances from later compliance periods to earlier periods if the allowance price reaches the highest step of the price containment reserve.

electricity imports ("reshuffling" and "relabeling"[12]), and (c) offsets, which we discuss later (and which might be considered a form of lessening demand rather than increasing supply, but the analysis would be unchanged). While incentives for reshuffling and offsets are affected by the price of allowances, previous analyses suggest that the bulk of this activity would be realized at prices below or just slightly above the auction reserve price.[13]

In its revised scoping plan of 2010, ARB's preferred model projects that 63% of emissions abatement would arise from complementary policies rather than from responses to the cap (four additional sensitivity models project between 30% and 63% of emissions abatement would arise from complementary policies).[14] It is important to recognize that these reductions are not costless, indeed many may impose costs above the allowance price. Rather, these reductions, and the accompanying costs, will occur *approximately independently* of the level of the allowance price. Therefore, while these policies provide reductions, and contribute to the goal of keeping emissions under the cap, they do not provide the price-responsive abatement that can help mitigate volatility in allowance prices.

In this paper, we treat the impact of these complementary policies as influencing the distribution of the supply of abatement. For example, aggressive vehicle fuel-efficiency standards should lead to slower growth in the emissions from the transportation sector, which we represent as a change in the rate at which the emissions intensity of vehicles declines over time. Similarly mandates for renewable energy production decrease the amount of electricity demand that needs to be served by more carbon intensive sources, thereby reducing emissions.

As described below, the supply of price-responsive mitigation is limited by some

[12]Relabeling describes the practice of reselling out-of-state power that comes from a high-emissions source such that the buyer can then import the power into California at the administratively determined default emissions rate. Relabeling might be considered a type of reshuffling. We consider them in combination.

[13]The potential levels of reshuffling and relabeling are examined in Bushnell, Chen, and Zaragoza-Watkins (2014). The offset market is discussed below. Some offset supply may be available at prices somewhat above the auction reserve price.

[14]See http://www.arb.ca.gov/cc/scopingplan/economics-sp/updated-analysis/updated_sp_analysis.pdf at
page 38 (Table 10).

of the allocation policies that have been implemented under AB 32. The large amount of allowances allocated through mechanisms that are likely to reduce the price impact of allowance prices to consumers – output-based updating for many industrial emitters and allocations to utilities that will use them to limit the impact of allowance prices on consumer prices – will limit the amount of price-responsive emissions mitigation.[15] Most of the remaining emissions reductions in response to allowance prices would therefore come from consumer responses to changes in energy prices, namely transportation fuels (gasoline and diesel), natural gas, and, possibly, electricity consumption. Compared to the aggregate level of reductions needed and expected under AB 32, we show that the reductions from these energy price effects are relatively small.[16] This is due in part to a feature of the program that will use the revenues from the sales of allowances to fossil fuel electricity suppliers to limit the magnitude of potential retail electricity price increases. A similar policy is under consideration at ARB for the retail natural gas sector. If implemented they would further increase the slope of abatement supply curve.

The combination of large amounts of "zero-price" abatement, and relatively modest price-responsive abatement creates a hockey stick shaped abatement-supply curve (See Figure 1). Analysis undertaken by ARB indicates that the marginal abatement cost curve rises sharply after the relatively low-cost abatement options are exhausted. ARB states in its updated Scoping Plan dated March 2010 that "...GHG emissions in the model show limited responsiveness to

[15] Output-based updating describes allocation of allowances to a company based on the quantity of output (not emissions) that the firm produces. Output-based updating reduces the firm's effective marginal cost of production and, thus, reduces the incidence of the allowance price on firms and consumers, while retaining the full allowance price incentive for the firm to adopt GHG-reducing methods for producing the same level of output (see Meredith Fowlie, "Updating the Allocation of Greenhouse Gas Emissions Permits in a Federal Cap-and-Trade Program," in Don Fullerton and Catherine Wolfram, ed. *The Design and Implementation of U.S. Climate Policy*, University of Chicago Press. 2012). If applied to a large enough set of industries or fraction of the allowances, the effect can be to inflate allowance prices as higher prices are necessary to offset the diluted incentive to pass the carbon price through to consumers. See Bushnell, James and Yihsu Chen. "Regulation, Allocation, and Leakage in Cap and Trade Markets for CO2." Resources and Energy Economics. 34(4), 2012.

[16] Offsets and reshuffling/relabeling may also be sensitive to allowance prices, but are considered separately.

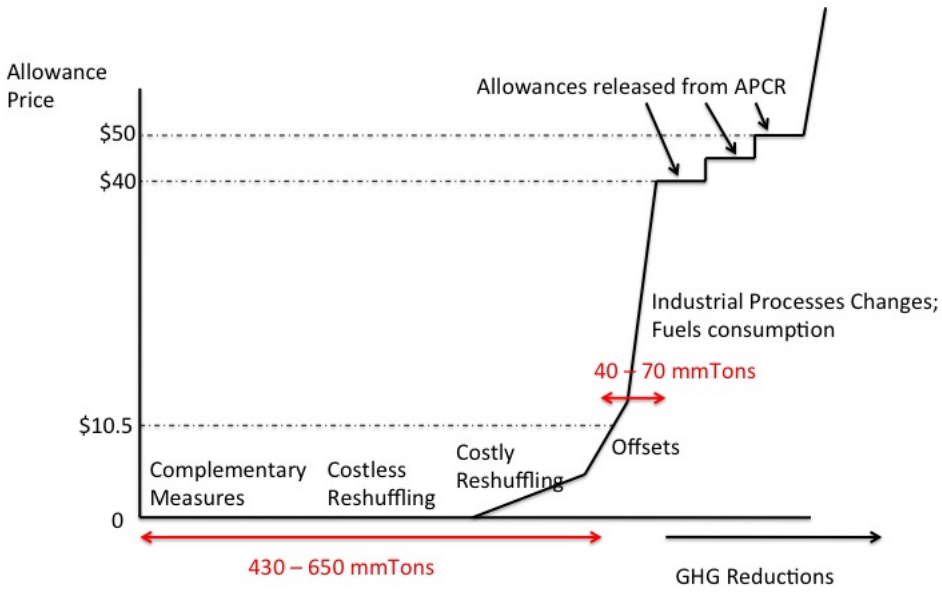

Figure 1. Supply of Abatement

allowances prices...This lack of responsiveness results from the limited reduction opportunities that have been assumed to be available in the model."[17]

One implication of this is that allowance prices are more likely to be either at or near the level of the auction reserve price or at levels set by the APCR policy than they are to be at some intermediate level. When one considers an uncertain range of BAU emissions, even if strongly centered on the expected level, the probabilities of prices falling at either the APCR ceiling or auction reserve price floor constitutes a large fraction of the overall distribution of potential emissions outcomes.

This intuition is illustrated in Figure 2, which superimposes a hypothetical

[17]Available at: http://www.arb.ca.gov/cc/scopingplan/economics-sp/updated-analysis/updated_sp_analysis.pdf. See also, the ARB analysis contained in Appendix F: Compliance Pathways Analysis available at: http://www.arb.ca.gov/regact/2010/capandtrade10/capv3appf.pdf.

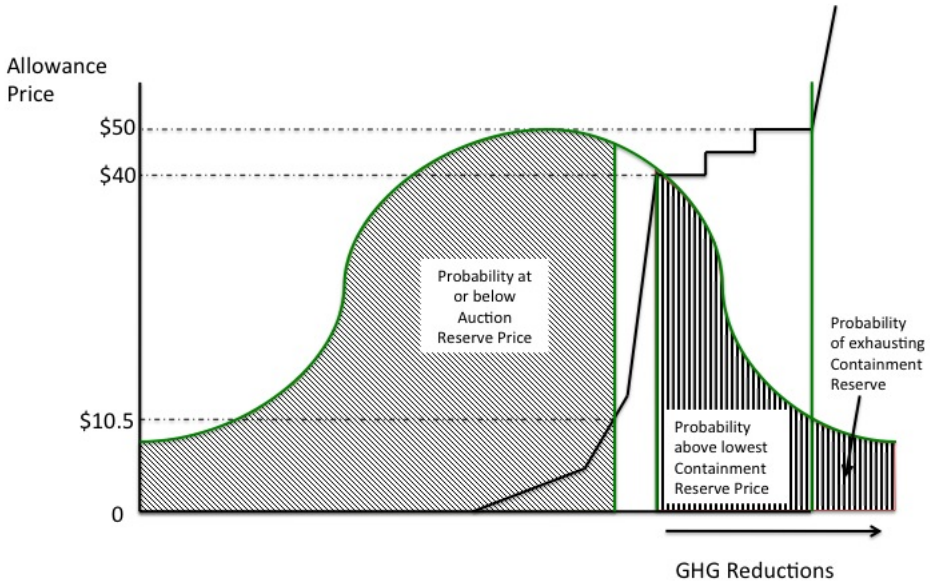

Figure 2. Hypothetical Distribution of Abatement Demand (BAU minus allowances outside price containment reserve) versus Abatement Supply

symmetric distribution of the amount of abatement needed (BAU emissions less the total amount of available allowances) onto the same horizontal axis as our supply curve. Note from Figure 2 that the range of abatement quantity that falls between the auction reserve price ($10.50/tonne in this illustration) and the first-step of the price containment "ceiling" ($40/tonne in this illustration), which is the area with no pattern, is relatively small.

The implications of California's abatement supply curve is that the vast majority of probability for a given price outcome falls either at the auction reserve price or in the range in which the price containment policy is likely to be triggered. Rather than the intuitive bell-shaped distribution of possible prices, it is more appropriate to think of the probabilities as distributed according to the dashed line of Figure 3, which has the same mean as the solid line, but this mean is

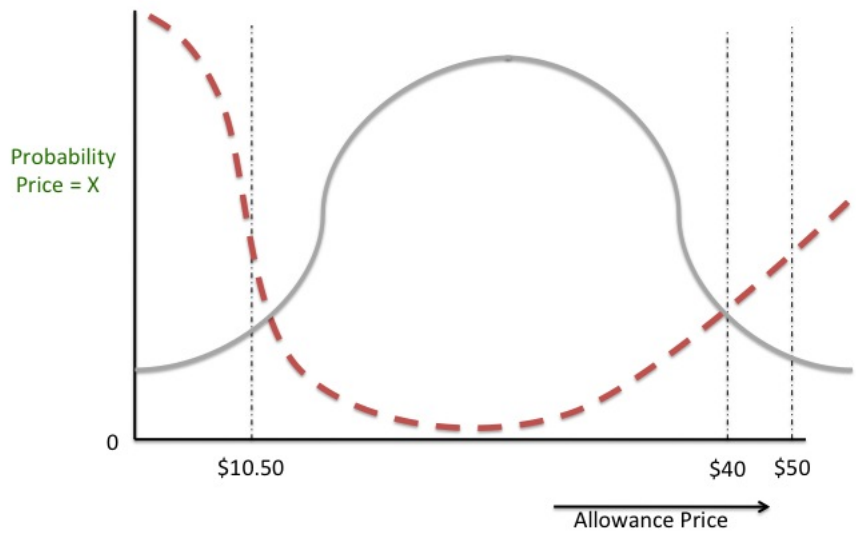

Figure 3. Possible Density Functions of Allowance Price

generated by a high probability of a "low" (auction reserve) price balanced by a somewhat lower probability of a "high" (price containment reserve) price.

A. Price Evolution and Estimated Equilibrium Price in the Market

The analysis we present here models abatement supply and demand that evolves over time and is then aggregated over the 8-year span of the market. We calculate the equilibrium as the price at which the aggregate demand over the 8 years is equal to the aggregate supply. We analyze this program alone, assuming that the market is not continued after the 8 years or integrated into some other program. At this point there is not clarity about how the program will evolve after 2020.

At any point in time, two conditions will drive the market price, an intertemporal arbitrage condition and a market equilibrium condition. If the markets for

emissions at different points in time are competitive and well integrated, then intertemporal arbitrage enabled by banking and borrowing will cause the *expected* price change over time to be equal to the nominal interest rate (or cost of capital).[18] At the same time, the price *level* will be determined by the condition that the resulting expected price path – rising at the nominal interest rate until the end of 2020 – would in expectation equilibrate the total supply and demand for allowances.[19]

Throughout the market's operation, new information will arrive about the demand for allowances (*e.g.,* weather, economic activity, energy prices and the energy intensity of Gross State Product (GSP) in California and the supply of abatement (*e.g.,* supply of offsets, response of consumers to higher fuel prices, and the cost of new technologies for electricity generation). These types of information will change expectations about the supply/demand balance in the market over the length of the program and thus change the current equilibrium market price. The price at any point in time reflects a weighted average of all the possible future prices that may occur in order to equilibrate supply and demand.

For instance, while high allowance prices are a possibility if the economy grows rapidly and abatement efforts are less effective than anticipated, early in the market operation that would be only one of many possible future outcomes that the market price would reflect. Over time, however, if economic growth were stronger and abatement weaker than expected, this would become an increasingly likely scenario and price would rise faster than had been anticipated. Thus, if lower-

[18]This is the outcome envisioned when banking was first developed (Kling and Rubin, 1997). See also Holland and Moore (2013), for a detailed discussion of this issue.

[19]Because of lags in information and in adjustment of emissions-producing activities, supply and demand will not be exactly equal at the end of the compliance obligation period (December 31, 2020). At that point, the allowance obligation of each entity would be set and there would be no ability to take abatement actions to change that obligation. The supply of allowances would have elasticity only at the prices of the APCR where additional supply is released and the level at which a hard price cap is set, if one is enacted. Thus, the price would either be approximately zero (if there is excess supply) or at one of the steps of the APCR or a hard price cap (if there is excess demand). Anticipating this post-compliance inelasticity, optimizing market participants would adjust their positions if they believed the weighted average post-compliance price outcomes were not equal to the price that is expected to equilibrate supply and demand. Such arbitrage activity would drive the probability distribution of post-compliance prices to have a (discounted) mean equal to the equilibrium market price in earlier periods.

probability outcomes were to occur over time, their impact would become evident gradually in the adjustment of the market price. In that case, an extremely high market price would probably not occur until the later years of the program.

Market participants are likely to employ an analysis similar to ours to decide the allowance price that they should use when choosing how much GHG to emit and whether an investment to abate emissions is likely to be cost effective. Analyses like this will also determine the price at which participants' are willing to buy and sell in the allowance market.

Table 1—Aggregate Emissions from Key California Sectors in 2010 (MMT)

Source	1990 Emissions	2011 Emissions
Electricity (domestic)	44.76	38.25
Electricity (imports)	29.65	46.13
Transportation (on road)	134.70	147.10
Industrial	79.77	75.40
Nat. Gas and Other	69.94	67.90

III. ESTIMATING THE BUSINESS AS USUAL EMISSIONS

Perhaps the largest factor driving the supply/demand balance in the GHG market will be the level of emissions that would take place under business as usual (BAU). There is, however, considerable uncertainty about BAU emissions over the period 2013 to 2020. The scope of the cap-and-trade program is very broad, and will be implemented in two phases. The first phase, which began January 1, 2013 covers large stationary sources, which are dominated by power plants, oil refineries, and other large industrial facilities. The second phase, to begin January 1, 2015, will expand the cap to include emissions associated with the combustion of transportation fuels and natural gas at non-industrial facilities. Table 1 summarizes the aggregate emissions from the key sectors from 1990 through 2011.

Historically, there has been considerable variability in the level of economic activity in each of these sectors, which in turn implies considerable uncertainty

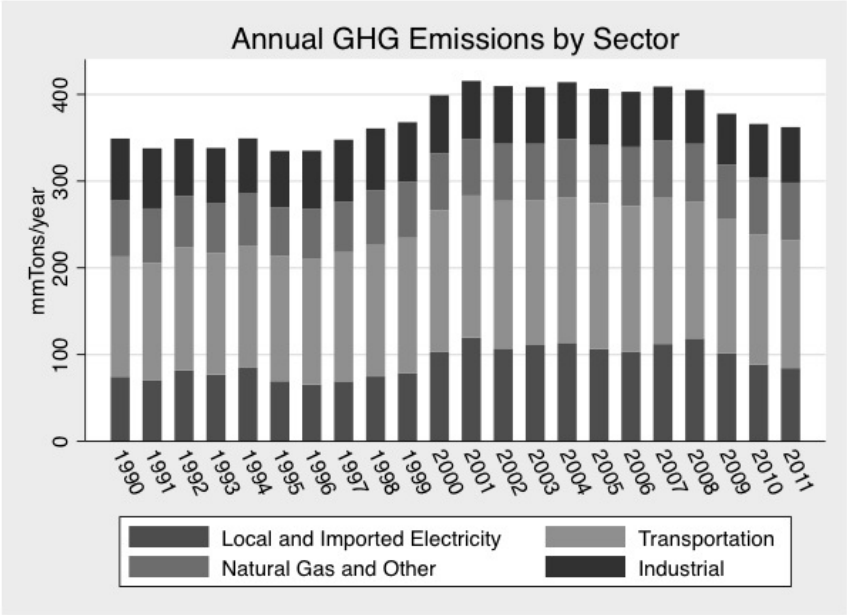

Figure 4. California Emissions Data 1990-2011

in the production of GHG emissions from these activities. Figure 4 presents the annual emissions from each sector over a 22-year period beginning in 1990. Predicting the level of economic activity from each of these sectors only one year in advance has the potential for significant forecast errors. Forecasting the level of economic activity and GHG emissions nine years into the future involves even greater forecast errors, which implies a greater potential for very low or high allowance price realizations.

An important category of emissions to highlight is those associated with imported electricity. Although these emissions are substantial, because they are from sources located outside of California, their measurement is uncertain and subject to potential avoidance through reshuffling or relabeling of sources. As described below, we apply ARB-derived emissions levels from imports as BAU and

consider scenarios of reshuffling in determining the net value of GHG emissions from electricity imports.

To derive estimates of the expected future time path of GHG emissions and the uncertainty associated with this forecast, we estimate a seven-dimensional Vector Autoregression (VAR) model with determinants of the three major components of state-level GHG emissions that are covered under the program and the key statewide economic factors that impact the level and growth of GHG emissions.[20] Due to the short time period for which the necessary disaggregated GHG emissions data have been collected, the model estimation is based on annual data from 1990 to 2011. Because data are available for 2012 on real Gross State Product (GSP), in-state electricity production by source, and the real price of gasoline in California, we condition on these values in forecasting the expected future time path of GHG emissions and the computing the uncertainty in the future time path of GHG emissions.

The short time series puts a premium on parsimony in the model. As a result, we use a 7-variable model that includes the three drivers of GHG emissions–in-state fossil-fuel electricity production, vehicle-miles traveled (VMT), and non-electricity natural gas combustion and industrial process GHG emissions–and the two economic factors that influence those drivers–real GSP and the real price of gasoline in California. To facilitate forecasting the future time path of GHG emissions in the transportation and electricity sectors under different sets of complementary policies for reducing GHG emissions in these sectors, we also model the behavior of the emissions intensity of the transportation and electricity sectors in California. Our approach is to estimate a VAR for these seven variables, simulate them through 2020 and apply a range of emissions intensities to the economic drivers of transportation and electricity emissions in order to simulate future GHG emissions under different complementary policies in these two sectors.

[20] Vector Autoregressions are the econometric methodology of choice among analysts to construct short to medium-term (from 1 to 10 time periods into the future) forecasts of macroeconomic variables and for this reason are ideally suited to our present task. Stock and Watson (2001) discuss the successful use of VARs for this task in a number of empirical contexts.

Several features of our VAR model are chosen to match the time series relationships between the seven variables implied by economic theory and existing state policies to limit GHG emissions. We allow for the fact that all seven variables exhibit net positive or negative growth over our sample period and model them as stochastic processes that are second-order stationary in growth rates rather than second-order stationary in levels. The results of unit root tests reported in the Appendix for each of individual time series are consistent with this modeling assumption. We also impose restrictions on the parameters of the VAR model implied by the cointegrating relationships between these seven variables that are supported by the results of these hypothesis tests. Engle and Yoo (1987) show that imposing the parameter restrictions implied by cointegrating relationships between variables in a VAR improves the forecasting accuracy of the estimated model.

A. Model

Let $X_t = (X_{1t}, X_{2t}, ..., X_{7t})'$ denote the vector composed of the seven annual magnitudes included in the VAR for year t, $t = 1990, 1991, ..., 2011$. The elements of X_t are:

$X_{1t} =$ CA electricity production net of hydroelectric generation (TWh)
$X_{2t} =$ Total vehicle-miles traveled (thousands of miles)
$X_{3t} =$ Industrial GHG & other nat. gas emissions (MMT)
$X_{4t} =$ Real Retail Gasoline price ($2011/gallon)
$X_{5t} =$ Real Gross State Product ($2011)
$X_{6t} =$ Emissions Intensity of In-State Thermal Gen. (metric tonnes/MWh)
$X_{7t} =$ Emissions Intensity of VMT (metric tonnes/thousand miles)

All real dollar magnitudes are expressed in 2011 dollars. All GHG emissions are in metric tonnes of CO2-equivalents. As noted above, we include real GSP in the model to capture the empirical regularity observed both over time and

across jurisdictions that a higher level of economic activity leads to greater energy consumption and GHG emissions. The price of gasoline reflects the fact that movements in transport fuel prices change the energy intensity of economic activity and the value of VMT.

Estimating this VAR produces parameters that allow us to construct simulated realizations of the elements of $X_t = (X_{1t}, X_{2t}, ..., X_{7t})$ from 2013 to 2020. Note X_{3t} is already in terms of metric tonnes of GHG. However, in order to get the total GHG emissions covered under the program, we do two further calculations. First, from X_{1t}, the simulation of the production of electricity in California net of hydroelectric generation, we subtract the anticipated amount of renewable and nuclear energy, described in more detail below. The remaining residual production is assumed to be provided by thermal generation and it is this residual amount that is multiplied by the thermal intensity, X_{6t}. Emissions from in-state electricity generation are included in the cap and trade program in all years from 2013 to 2020. Second, we parse X_{3t} – industrial GHG and other natural gas emissions – for 2013 and 2014 into the portion of these emissions that are and are not covered by the program during those years. Essentially, industrial processes and natural gas combustion by large industrial sources are covered in the first two years of the program, while off-road diesel consumption, and residential and small business emissions from natural gas consumption are not covered until 2015.

We do not include the GHG emissions from electricity imports in the VAR because this is largely an administratively determined number. All that can actually be measured is the aggregate GHG emissions outside of California and total electricity produced outside of California. The specific energy deemed to be "delivered" to California is largely the choice of the importing firm. Because incentives for this choice will change dramatically with the start of the cap and trade program, historical data on imports are not predictive of future values. We instead take the ARB's forecast for emissions from electricity imports and then adjust total electricity emissions for reshuffling, as described later.

Define $Y_{it} = ln(X_{it})$ for $i = 1, 2, ..., 7$ and $Y_t = (Y_{1t}, Y_{2t}, ..., Y_{7t})'$. In terms of this notation a first-order autoregression or VAR that is stationary in first-differences can be written as

$$\Theta(L) \cdot Y_t = \mu + \epsilon_t \tag{3.1}$$

where L is the lag operator which implies, $L^k Y_t = Y_{t-k}$, I is a (7x7) identity matrix, $\Theta(L)$ is (7x7) matrix function in the lag operator equal to $(I - \Theta_1 L)$ where Θ_1 is a (7x7) matrix of constants, μ is a (7x1) vector of constants, and ϵ_t is a (7x1) white noise sequence with (7x1) zero mean vector and (7x7) covariance matrix Ω. Recall that white noise series are uncorrelated over time. In terms of the lag operator notation $(1 - L) = \Delta$, so that $\Delta Y_t = Y_t - Y_{t-1}$.

Although model (3.1) allows each element of Y_t to be non-stationary, reflecting the fact that each element exhibits net positive or negative growth over the sample period. A linear time series process that is stationary in first-differences is also called an integrated process with the order of integration equation equal to 1. For each of the elements of Y_t we performed a Dickey-Fuller (1979) test of the null hypothesis that the time series contained a unit root and was unable to reject that null hypothesis at $\alpha = 0.05$ level of significance for each series.[21] These hypothesis testing results are consistent with our decision to model the vector ΔY_t as 2nd-order stationary process.

It is often the case that stationary linear combinations of non-stationary economic time series exist because of long-run economic relationships between these variables. This logic suggests that linear combinations of the elements of Y_t are likely to be 2nd-order stationary in levels. Times series processes that are 2nd-order stationary in first-differences (*i.e.,* ΔY_t is 2nd-order stationary) and have stationary linear combinations of their elements are said to be cointegrated.[22] For a k-dimensional VAR in first-differences of Y_t, the number of stationary linear combinations of the elements of Y_t is called the cointegrating rank of the VAR.

[21] Dickey and Fuller, 1979. Results of the Dickey-Fuller tests are shown in the Appendix.
[22] See Engle and Granger, 1987, for a complete discussion of this concept and its implications.

The cointegrating rank is also equal to the rank of the matrix $(I - \Theta_1)$. The existence of cointegrating relationships among elements of Y_t imposes restrictions on the elements of Θ_1. Suppose that the rank of the matrix $(I - \Theta_1)$ is equal to r $(0 < r < 7)$. This implies that the following error correction representation exists for Y_t:

$$\Delta Y_t = \mu - \gamma Z_{t-1} + \epsilon_t \tag{3.2}$$

where $Z_t = \alpha' Y_t$ is a (r x 1) vector of 2nd-order stationary random variables (these are the stationary linear combinations of Y_t) and γ is a (7 x r) rank r matrix of parameters and α is a (7 x r) rank r matrix of co-integrating vectors, and $(I - \Theta_1) = -\gamma\alpha'$.

Johansen (1988) devised a test of the cointegrating rank of a VAR that is 2nd-order stationary in first-differences. Following the multi-step procedure recommended by Johansen (1995) for determining the rank of a VAR, we find that the null hypothesis that the rank of $(I - \Theta_1)$ is equal to 1 can be rejected against the alternative that the rank is greater than 1 at 0.05 level.[23] However, the null hypothesis that the rank of $(I - \Theta_1)$ is 2 against the alternative that it is greater than 2 cannot be rejected at a 0.05 level. According to Johansen's procedure, this sequence of hypothesis testing results is consistent with the existence of 2 stationary linear combinations of the elements Y_t. We impose these co-integrating restrictions on the parameters of VAR model (3.2) that we estimate to forecast future GHG emissions. Imposing the restrictions implied by the two cointegrating relationships between the elements of Y_t reduces the number of free parameters in the (7x7) matrix $(I - \Theta_1)$ from 49 to 28 = (7x2) x 2, the total number of elements in γ and α.

We utilize Johansen's (1988) maximum likelihood estimation procedure to recover consistent, asymptotically normal estimates of μ, Ω, and Θ_1 with these co-integrating restrictions imposed. The coefficient estimates from this model written in the notation of equation (3.2) are given in the Appendix.

[23] Results of these tests are shown in the Appendix.

Using these parameter estimates we can then compute an estimate of the joint distribution of $(X'_{2013}, X'_{2014}, ..., X'_{2020})'$ conditional on the value of X_{2011} that takes into account both our uncertainty in the values of μ, Ω, γ, and α because of estimation error and uncertainty due to the fact that $(X'_{2013}, X'_{2014}, ..., X'_{2020})'$ depends on future realizations of ϵ_t for $t = 2012, ..., 2020$. Because we have 2012 data for instate electricity production net of hydroelectric generation (X_1), the real price of gasoline in California (X_4), and real State GSP (X_5), we compute our estimate of the distribution of $(X'_{2013}, X'_{2014}, ..., X'_{2020})'$ conditional on the values of these three elements of X_t for $t = 2012$ as well as the observed value of X_{2011}.

We employ a two-stage smoothed bootstrap approach to compute an estimate of this distribution.[24] The first step computes an estimate of the joint distribution of the elements of μ, Ω, γ and α by resampling from the smoothed empirical distribution of the (7x1) vector of residuals from the estimated Vector Autoregression (VAR) and re-estimating μ, Ω, γ, and α using Johansen's (1988) maximum likelihood procedure. We use the following algorithm. Let $\hat{\mu}$, $\hat{\Omega}$, and $\hat{\Theta}_1$ equal the estimates of the elements of the VAR imposing the cointegration rank restriction that $(1 - \Theta_1) = -\gamma\alpha'$. Compute

$$\hat{\epsilon}_t = Y_t - \hat{\mu} - \hat{\Theta}_1 Y_{t-1} \tag{3.3}$$

for t =1991 to 2011. Note that we can only compute values of $\hat{\epsilon}_t$ for t =1991 to 2011, because our sample begins in 1990 and the $(t-1)$th observation is required to compute the value of $\hat{\epsilon}_t$ for period $t = 1991$. Construct the kernel density estimate of the $\hat{\epsilon}_t$ as

$$\hat{f}(t) = \frac{1}{Th^7} \sum_{t=1}^{T} K\{\frac{1}{h}(t - \hat{\epsilon}_t)\} \tag{3.4}$$

where T is the number of observations, h is a user-selected smoothing parameter, and $K(t)$ is a multivariate kernel function that is everywhere positive and

[24]For a discussion of the smoothed bootstrap, see Efron and Tibshirani, 1993.

integrates to one. We use the multivariate normal kernel

$$K(x) = \frac{1}{(2\pi)^{7/2}} exp(-\frac{1}{2}x'x) \quad \text{where} \quad x \in \Re^7$$

and $h = 0.5$. We found that our results were insensitive to the value chosen for h, as long as it was less than 1.

We then draw $T = 21$ values from (3.4) and use the parameter estimates and these draws to compute re-sampled values of Y_t for $t = 1, 2, ..., T = 21$. Let $(\hat{\epsilon}_1^m, \hat{\epsilon}_2^m, ..., \hat{\epsilon}_{21}^m)'$ denote the mth draw of the 21 values of $\hat{\epsilon}_t$ from $\hat{f}(t)$. We compute the Y_t^m, the 21 resampled values of Y_t for t =1991 to 2011, by applying the following equation starting with the value of Y_t in 1990 ($Y_{1990}^m = Y_{1900}$ for all m)

$$Y_t^m = \hat{\mu} + \hat{\Theta}_1 Y_{t-1}^m + \hat{\epsilon}_t^m. \tag{3.5}$$

We then estimate the values of μ, Ω, and Θ_1 by applying Johansen's (1988) ML procedure using the Y_t^m and imposing the cointegration rank restriction that $(1 - \Theta_1) = -\gamma\alpha'$. Call the resulting estimates $\hat{\mu}^m$, $\hat{\Omega}^m$, and $\hat{\Theta}_1^m$. Repeating this process $M = 1000$ times yields the bootstrap distribution of $\hat{\mu}$, $\hat{\Omega}$, and $\hat{\Theta}_1$. This step accounts for the uncertainty in future values of Y_t due to the fact that true values of the of μ, Ω, and Θ_1 are unknown and must be estimated.

To account for the uncertainty in Y_{T+k} due to future realizations of ϵ_t, for each m and set of values of $\hat{\mu}^m$, $\hat{\Omega}^m$, and $\hat{\Theta}_1^m$, we draw nine values from $\hat{f}(t)$ in equation (3.4). Call these values $(\hat{\epsilon}_{T+1}^m, \hat{\epsilon}_{T+2}^m, ...\hat{\epsilon}_{T+9}^m)'$. Using these draws and $\hat{\mu}^m$, $\hat{\Omega}^m$, and $\hat{\Theta}_1^m$ compute future values Y_{T+k} for $k = 1, 2, ..., 9$ given Y_T using the following equation:

$$Y_{T+k|T}^m = \hat{\mu}^m + \hat{\Theta}_1^m Y_{T+k-1|T,T-1}^m + \hat{\epsilon}_{T+k}^m \quad \text{for} \quad k = 1, 2, ..., 9 \tag{3.6}$$

This yields one realization of the future sample path of Y_t for t =2012, 2013,...,

2020. The elements of Y_t are then transformed to X_t by applying the transformation $X_{it} = exp(Y_{it})$ to each element of Y_t to yield a realization of the future time path of X_t. The elements of X_t are then transformed to produce a realization of the future time path of GHG emissions by each covered sector. This two-step process of computing $\hat{\mu}^m$, $\hat{\Omega}^m$, and $\hat{\Theta}_1^m$ and then simulating $Y_{T+k|T}^m$ for $k = 1, 2, ..., 9$ replicated $m = 1$ to $M = 1000$ times produces 1,000 realizations from the simulated distribution of $(X'_{2012}, X'_{2013}, ..., X'_{2020})'$.

The procedure for simulating the value X_{2012} is slightly different from the procedure for simulating values for 2013 to 2020 described above because we know the values of X_1, X_4, and X_5 for 2012. Simulating the value of $(X'_{2013}, X'_{2014}, ..., X'_{2020})'$ conditional on the values of instate electricity production net of hydroelectric generation (X_1), the real gasoline price in California (X_4), and real State GSP (X_5) in 2012, requires constructing the smoothed conditional density of $(\hat{\epsilon}_{2t}, \hat{\epsilon}_{3t}, \hat{\epsilon}_{6t}, \hat{\epsilon}_{7t})'$ conditional on $(\hat{\epsilon}_{1t}, \hat{\epsilon}_{4t}, \hat{\epsilon}_{5t})' = (\hat{\epsilon}_{1,2012}, \hat{\epsilon}_{4,2012}, \hat{\epsilon}_{5,2012})'$, the elements of $\hat{\epsilon}_t$ corresponding to instate electricity production net of hydroelectric generation (X_1), the real price of gasoline in California (X_4), and real State GSP (X_5) in 2012 that reproduce the observed values of these variables in 2012 given the values of all of the elements Y_t in 2011. We draw $(\hat{\epsilon}_{2t}, \hat{\epsilon}_{3t}, \hat{\epsilon}_{6t}, \hat{\epsilon}_{7t})'$, the remaining elements of $\hat{\epsilon}_t$ from this conditional density for 2012 in computing the simulated value of Y_t for 2012. This re-sampling process ensures that the simulated value of instate electricity production net of hydroelectric generation, the real price of gasoline, and real GSP in California in 2012 are always equal to the observed value for each of these variables. It also ensures that the simulated value of $\hat{\epsilon}_t$ for 2012 is consistent with the smoothed joint distribution of $\hat{\epsilon}_t$ in (3.4) when drawing the remaining elements of this vector.

Although California's cap and trade program phases in the entities under the cap over time, our approach forecasts emissions from Phase I entities (narrow scope) and Phase II entities (broad scope) over the entire post-sample period. Phase I, in effect during the first compliance period of 2013 and 2014, covers

electricity generation and emissions from large industrial operations. Phase II, in effect for the second and third compliance periods, 2015-2017 and 2018-2020, expands the program to include combustion emissions from transportation fuels and emissions from natural gas and other fuels combusted at residences and small commercial establishments.

B. Data

To compute the GHG emissions intensities of the instate electricity sector and transportation sector from 1990 to 2011 that enter the VAR model, we require data on the annual emissions from instate electricity production and annual emissions from the transportation sector to enter the numerator of each of these intensities. Annual emissions from the large industrial processes and the residential and commercial natural gas sector from 1990 to 2011 is the final GHG emissions-related time series required to estimate the VAR.[25] To construct these data, we start with data on annual emissions for each covered sector in California for 1990 to 2011. The remaining data that enter the VAR come from a variety of California state and federal sources, discussed below.

Annual emissions levels for each covered sector are taken from the 1990-2004 Greenhouse Gas Emissions Inventory and the 2000-2011 Greenhouse Gas Emissions Inventory (hereafter, Inventory).[26] This is the longest series of consistently measured emissions data and the basis for developing the 1990 statewide emissions level and 2020 emissions limit required by AB 32. The annual Inventory data was prepared by ARB staff and relies primarily on state, regional or national data sources, rather than individual facility-specific emissions. The Inventory's top-down approach to quantifying emissions differs importantly from the bottom-up method of accounting for facility-specific emissions under the cap and trade program. In particular, the Inventory likely overstates emissions from industrial

[25] Emissions from the off-road consumption of diesel also comprises a small component of the "other" category.

[26] California's GHG emissions inventory is available at: http://www.arb.ca.gov/cc/inventory/inventory.htm.

activity relative to those covered in the first compliance period of the cap and trade program. That is, the Inventory methodology may attribute some emissions to the industrial sector, such as natural gas combustion from small industrial or commercial sources that are not covered until the second compliance period. We investigate the impact of this difference by comparing the Inventory data to annual data collected under the Mandatory Reporting Regulation (MRR), which is the methodology used to calculate an entity's compliance obligation under the cap-and-trade program.[27]

Table 2—Summary Statistics of Data for Vector Autoregression

	mean	S.D.	min	max	year min.	year max.
California Elec. Generation (TWh)	191.20	15.80	158.90	216.80	1991	2006
California Hydro. Gen (TWh)	34.60	9.30	20.20	49.50	1992	1998
Vehicle Miles Traveled (Billions)	300.60	26.84	257.98	329.27	1991	2005
Industry, Natural Gas & Other Emissions (MMT CO2e)	141.90	4.83	131.98	145.60	1995	1998
Gross State Prod.t (Nominal $Trillion))	1.36	0.43	0.77	2.00	1990	2012
Gasoline Price (Nominal $/gallon)	2.20	0.96	1.09	4.03	1990	2012
In-state Elec. Thermal Intensity (tons/MWh)	0.483	0.045	0.402	0.529	2011	1990
Vehicle Emissions. Intensity (tons/1000 VMT)	0.507	0.02	0.459	0.534	2011	1990

Note: Data are for 1990-2011

Comparing the 2008-2011 MRR and Inventory industrial emissions data series shows annual differences of 8.98 to 13.24 MMT, with Inventory industrial emissions fifteen percent higher than MRR industrial emissions, on average. We address this difference by forecasting industrial capped source emissions in the first compliance period using the Inventory industrial emissions data series adjusted downward by fifteen percent. We use the unadjusted Inventory data as our measure of industrial capped source emissions covered in the second and third compliance periods. This approach does not appear to impact either our expected time path or the degree uncertainty in the future time path. Because

[27]Information on the ARB mandatory reporting regulation is available at: http://www.arb.ca.gov/cc/reporting/ghg-rep/ghg-rep.htm.

our maintained assumption is that the first compliance period difference is due to differences in accounting, as opposed to classical measurement error, using the Inventory emissions estimates for the second and third compliance periods should not bias our emissions estimates upward.

California GSP is collected from the Bureau of Economic Analysis (BEA).[28] Gasoline prices are collected from the Energy Information Administration (EIA).[29] In-state electric generation is also collected from the EIA.[30]

Our primary measure of Vehicles Miles Traveled (VMT) is compiled from a series of state-level transportation surveys administered by the National Highway Transportation Safety Administration's (NHTSA) Office of Highway Information (OHI). These data capture on-road VMT and were independently constructed and reported by the states, rather than centrally calculated by OHI.

While these data measure on-road VMT, the cap and trade program caps emissions from all diesel and gasoline combusted as transportation fuel in California, regardless of whether the fuel is combusted on-road or off-road. To address this potential source of bias we deviate from ARB's emissions categorization of "transportation" by excluding GHG emissions from off-road vehicle activities, in favor of categorizing them into "Natural Gas and Other." Therefore, beginning with total transportation sector combustion emissions, we partition emissions into on-road and off-road activities using the more granular activity-based emissions values reported in the combined 1990-2004 and 2000-2011 Emissions Inventories. The emissions levels reported in Table 1 reflect this partition of on-road and off-road emissions.

Finally, to adjust the emissions from natural gas, off-road diesel, and industrial processes for partial coverage under the cap of these emissions in 2013-14, we multiply the value of $X^m_{3,T+k}$ for each simulation by $0.53 \cdot 0.85 (= 0.4675)$ for the values in 2013 and 2014. This adjustment reflects that over the last 20 years,

[28] Gross Domestic Product by State is available at: http://www.bea.gov/regional/index.htm#data.
[29] Retail fuel price by State is available at: http://www.eia.gov/dnav/pet/pet_pri_gnd_dcus_sca_w.htm.
[30] In-state California electric generation and consumption are available from the CEC at http://energyalmanac.ca.gov/electricity/index.html.

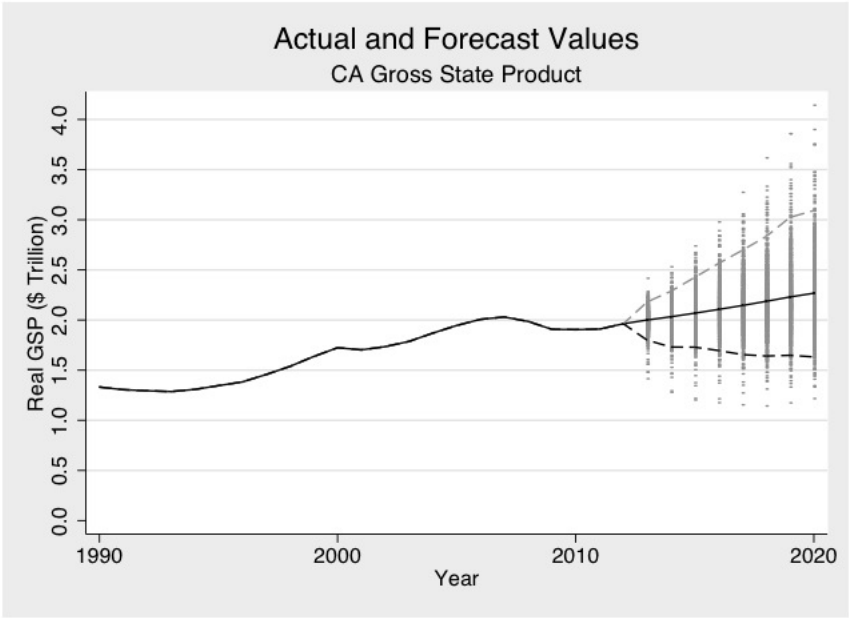

Figure 5. Forecast Results – Gross State Product

the industrial sector has consistently accounted for approximately 53% of emissions from non-electricity-generation natural gas combustion and other industrial processes (X_3) (min: 51.5% and max: 56.5%), and the Inventory accounting difference (discussed above), which leads us to attribute 85% of industrial emissions to sources covered under the first compliance period.

Summary statistics for all data of the VAR are in table 2.

C. Results

The parameter estimates from estimating the 7-variable VAR are shown in the Appendix. The parameter estimates are reported in the error-correction model notation of the VAR as:

$$\Delta Y_t = \mu + \Lambda Y_{t-1} + \epsilon_t \tag{3.7}$$

where Λ is (7x7) matrix that satisfies the restriction $\Lambda = -\gamma\alpha'$. Repeating the two-step procedure described above, yields 1000 simulations of the elements of X_t. Table 3 lists the means and standard deviations of simulated value of each element of X_t for each year from 2013 to 2020, as well as the annual and cumulative emissions resulting from those values. Figure 5 shows actual data (up to 2012) and forecast from VAR for GSP, with 95% confidence intervals for the forecast. The vertical dots show the distribution of simulation outcomes. The next section describes the details of our procedure for simulating future values of annual emissions covered by the program for each year from 2013 to 2020.

Table 3—Summary Statistics of Simulated VAR Variables and Emission

Year	California Electricity net of Hydro Twh	Vehicle Miles Traveled Million Miles	Nat. Gas, Ind. & Other MMT	Gasoline Price $2011	Gross St. Product $2011 Trillion	Therm. Intensity tons/ MWh	Trans. Intensity tons/1000 Miles	Broad Scope Emis. MMT	Cum. Emis. MMT
2013	170.06	321.41	153.90	4.36	2.00	0.37	0.47	402.73	168.17
	(27.00)	(10.66)	(12.23)	(0.95)	(0.10)	(0.05)	(0.03)	(18.82)	(15.91)
2014	175.49	323.81	154.35	4.37	2.03	0.37	0.46	402.04	334.76
	(26.96)	(12.57)	(14.74)	(0.95)	(0.14)	(0.05)	(0.03)	(20.21)	(29.69)
2015	175.68	326.95	154.49	4.55	2.07	0.37	0.46	399.82	734.55
	(29.74)	(14.11)	(16.91)	(1.21)	(0.18)	(0.05)	(0.03)	(23.02)	(47.88)
2016	178.90	330.05	154.59	4.71	2.11	0.36	0.46	400.58	1135.10
	(30.01)	(15.71)	(18.77)	(1.41)	(0.22)	(0.06)	(0.04)	(24.39)	(68.68)
2017	180.27	333.42	154.75	4.86	2.15	0.36	0.46	400.14	1535.21
	(32.49)	(17.44)	(20.68)	(1.58)	(0.26)	(0.06)	(0.04)	(27.55)	(93.25)
2018	182.67	336.82	154.72	5.07	2.19	0.35	0.45	400.80	1935.98
	(33.88)	(19.23)	(22.23)	(1.88)	(0.30)	(0.06)	(0.04)	(29.25)	(119.61)
2019	185.95	340.37	154.58	5.27	2.23	0.35	0.45	402.51	2338.46
	(36.96)	(21.10)	(23.75)	(2.08)	(0.34)	(0.06)	(0.04)	(31.54)	(148.30)
2020	187.51	343.46	154.51	5.42	2.27	0.35	0.45	403.19	2741.62
	(37.61)	(22.80)	(25.44)	(2.33)	(0.38)	(0.07)	(0.05)	(33.61)	(178.53)

Note: Estimates are mean values of 1000 draws, values in parenthesis are Std.Dev. of 1000 draws.

IV. ACCOUNTING FOR COMPLEMENTARY POLICIES IN FORECASTS

While the Air Resources Board (ARB) has identified many categories of complementary policies and stated the reductions in GHG emissions that are expected to result from each policy, it is unclear how the baseline from which such estimates are claimed relates to the simulations we obtain from the VAR. Thus, rather than incorporating potential reductions from an uncertain baseline, we proceed by applying emissions intensities of electricity generation and VMT that reflect the likely outcomes of the complementary policies. That is, the effects of complementary policies are incorporated into our simulations of GHG emissions from 2013 to 2020 through changes in the ratios we use to translate forecasts of X_{1t} and X_{2t}, instate electricity production minus hydroelectric energy production and vehicle miles traveled respectively, into GHG emissions.

In the case of electricity, the main complementary policies are energy efficiency (EE) investments and the Renewables Portfolio Standard (RPS). We treat both of these measures as impacting the *quantity of non-zero* carbon-emissions-producing power generation, rather than the intensity of overall generation.

In the case of the RPS, two important recent changes imply that historical trends of zero-carbon-emissions generation are not satisfactorily predictive of future supply. These two changes are the imposition of the 33% RPS and the recent unexpected retirement of the San Onofre Nuclear Generation Station (SONGS) in Southern California. To get from a simulation of X_{1t} for 2013-2020 to a simulation of GHG emissions from in-state thermal electricity generation, we first subtract off estimates of *future* renewable and nuclear power generation from each simulation of X_{1t}. These values are taken from external data sources rather than generated within the VAR. What remains is a simulation of instate fossil fuel electricity generation. We then multiply this number by the simulated value of the emissions intensity of in-state fossil-fuel generation from our two-step procedure.

For the RPS, we apply a California Public Utilities Commission (CPUC) fore-

cast of new renewable generation (MWh) taken from the 2012 Long-term Procurement Planning process.[31] These estimates of renewable power generation incorporate the impact of the 33% target for the RPS by 2020. We then add this annual quantity of new renewable energy to the average level of renewable generation (taken from EIA) over the last 20 years of about 24 TWh.[32]

Table 4—Assumed Zero-Carbon Electricity Output and Vehicle Emissions Intensities

Year	Zero-Carbon Power RPS TWh	Zero-Carbon Power Nuclear TWh	Low VMT Intensity tons/1000 miles	Medium VMT Intensity tons/1000 miles	BAU Forecast VMT Intensity tons/1000 miles
2013	30520	17530	0.482	0.492	0.467
2014	41369	17530	0.471	0.484	0.465
2015	48217	17530	0.457	0.472	0.462
2016	50586	17530	0.438	0.456	0.460
2017	54268	17530	0.419	0.440	0.457
2018	56054	17530	0.400	0.423	0.455
2019	56054	17530	0.382	0.407	0.453
2020	56151	17530	0.364	0.391	0.450

For in-state generation of nuclear power, we assume that the Diablo Canyon Nuclear Power Plant will continue to operate during 2013-2020 and that it will produce an average of 17.53 TWh per year, which is its average production for the 10-year period 2003-2012. These values are summarized in the second and third columns of Table 4. The remaining in-state generation is assumed to be from fossil fuel generation sources.

We then multiply this simulated value of instate fossil-fuel electricity production by X_{6t}, the emissions intensity factor produced by the simulation of future values from the VAR, to translate the simulation of instate electricity production into

[31] Specifically, we utilize the annual forecast of additional renewable energy from the RPS Calculator developed by E3 for the LTPP process found at http://www.cpuc.ca.gov/PUC/energy/Procurement/LTPP/-2012+LTPP+Tools+and+Spreadsheets.htm. This forecast shows increased renewable energy to provide an additional 32 TWh of renewable energy per year by 2020.

[32] Note that the EIA value of 24 TWh of renewable energy is lower than the official current level of RPS compliant energy. The difference is due to certain existing hydro resources that qualify under current rules. The EIA lists this energy as "hydroelectric" rather than renewable.

GHG emissions. More formally, we calculate electricity emissions from instate electricity production to be

$$ElecGHG_{m,T+k} = (TWH_N hydro_{m,T+k} - RPS_TWH_{T+k} - Nuke_TWH_{T+k}) \cdot EI_{m,T+k}$$

where $TWH_N hydro$ is the realization of $X_{1,T+k}$ for simulation draw m of the in-state production of electricity net of hydro production. The variables RPS_TWH and $Nuke_TWH$ are the values of renewable and nuclear annual TWH described in Table 4 and $EI_{m,T+k}$ is $X_{6,T+k}$, the realization of emissions intensity for thermal generation in California for simulation draw m.

Reflecting California's longstanding commitment to energy efficiency, there is a strong pre-existing trend of efficiency improvements already present in the time-series data we used to forecast the BAU emissions. Total emissions per unit of GSP declined at an average rate of about 1.83% per year from 1990 to 2011. We are therefore concerned that further reductions from our forecast to account for energy efficiency improvements would double count the reductions that are already part of the forecast. Indeed, as table 3 indicates, emissions per unit of GSP decline under our BAU forecast by about 1.74% per year from 2013 to 2020. We therefore make no further adjustments in addition to energy efficiency effects already integrated into our forecasts.

To incorporate the impact of complimentary policies targeting the transportation sector, we interact the forecast of VMT from the VAR with three possible values of emissions intensity per mile. The first value, essentially a business-as-usual intensity, takes $X_{7,T+k}$, the VMT intensity forecast by the VAR without any further adjustment. The second and third emissions intensities we use are based upon expectations of the impacts of AB 32 transportation policies derived from EMFAC 2011, the ARB tool for forecasting fleet composition and economic activity in the transportation sector. Our derivations are summarized here but described in more detail in the Appendix.

Using EMFAC, we derive anticipated emissions intensities (essentially fleet average miles per gallon) under two assumptions about transport policy. The first scenario assumes that all LCFS and miles per gallon (MPG) standards are met. This reduces emissions-per-mile both through improved MPG and through a higher percentage of biofuels, which are treated as having zero GHG emissions for the purchases of program, in the transportation fuel mix. The second scenario assumes that the mileage standards for new vehicles are met, but that the penetration of biofuels remains at 10%.[33] Thus, under this scenario the emissions per mile are reduced solely due to the increased fuel-efficiency of vehicles.

The EMFAC 2011 model provides, for each of our transportation policy scenarios, a point estimate of fleet average emissions intensity. Columns 4-6 of table 4 summarize these two values, along with the mean transport intensity value forecast by the VAR, for each year. However, even though the standards may be fully complied with, considerable uncertainty remains as to the emissions intensity of the full transportation emissions. Among other factors, a substantial minority of transport emissions come from commercial trucking and other heavy-duty vehicles that will not be subject to the same kind of binding fuel economy standards as the passenger vehicle fleet.

In order to reflect the underlying random aspects of vehicle emissions, even with successfully implemented complementary policies, we model the effect of these policies as a shift in the distribution of emissions intensity from a BAU level to a level achieved, on average, by the policies. This is accomplished by shifting each VMT intensity realization, $X_{7,T+k}$, by an amount equal to the difference between the BAU mean intensity level and the EMFAC forecast of the policy-induced point estimate. This adjusted emissions intensity is then multiplied by the coinciding VMT realization for the same VAR simulation draw to calculate total transport sector emissions for year t. More formally, transport emissions

[33]The carbon content of that 10% of biofuels may in fact be lower due to the LCFS, but from an emissions cap perspective that does not matter, because all biofuels are treated equally as zero emissions under the cap, and the current level of biofuels is already around 10%.

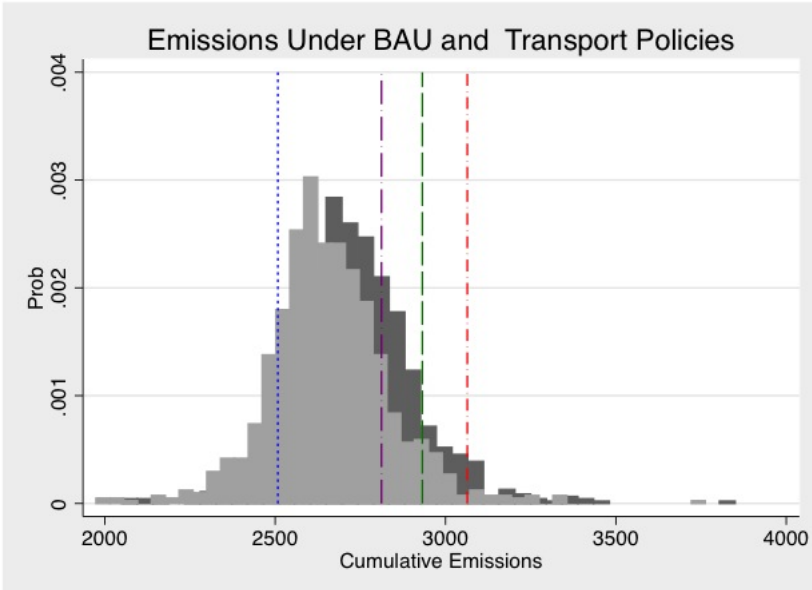

Figure 6. Targeted Transportation Policies Shift Emissions Distribution

can be expressed as

$$TransportCO2_{m,T+k} = VMT_{m,T+k} \cdot (TI_{m,T+k} - (E_j(TI) - TI_{policy}))$$

where $VMT_{m,T+k}$ and $TI_{m,T+k}$ are the vehicle miles traveled and transport emissions intensity from simulation draw m of the VAR during year t, respectively, and TI_{policy} is the transport emissions intensity derived by EMFAC 2011 for the given policy assumption. This effect is illustrated in Figure 6, which shows the distribution of transportation sector emissions for 2020 under the BAU intensity forecast (dark), as well as the shifted distribution (light) that incorporates the "low" vehicle intensity values from table 4. The four vertical lines are, from left to right, the total allowance budget, followed by the allowance budget plus the total abatement available at a price at the top of the APCR under low, medium,

and high abatement scenarios, which we discuss in the next section.[34]

Both of these adjustments–shifting MWh of in-state electricity generation and adjusting the intensity of VMT emissions–yield estimates of the emissions that will result from the three sectors covered in the California economy. These reductions will be independent of the price of allowances. Three other adjustments are necessary, however, before comparing this demand for allowances with the supply that is available under the cap and trade program: the impact of imported electricity, emissions offsets, and changes in the price of allowances. We incorporate these effects in the next section.

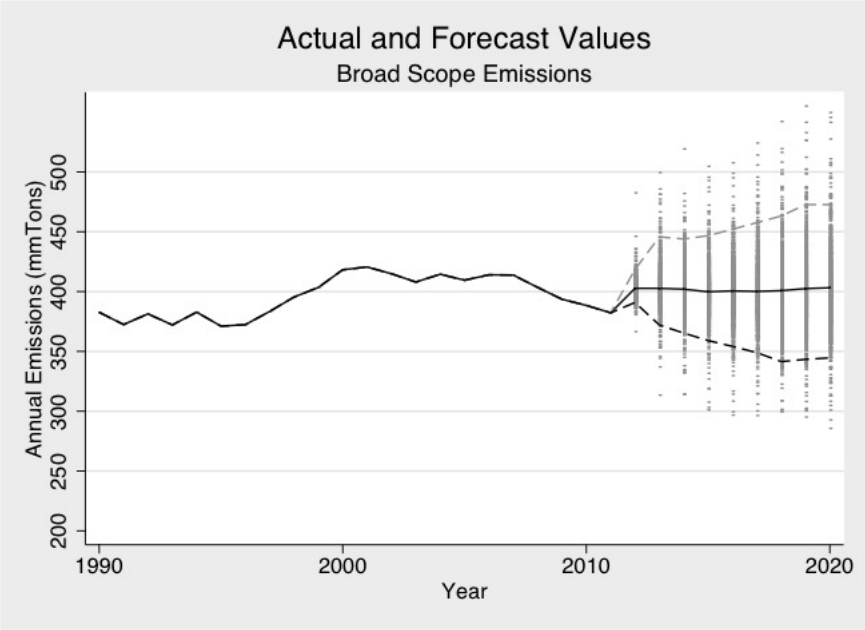

Figure 7. Forecast Results – Broad Scope Emissions

Figure 7 shows actual data (up to 2011) and forecast from VAR for Broad Scope Emissions, with 95% confidence intervals for the forecast. The vertical dots show

[34]The lines are all for cases with more stringent fuel economy standards.

the distribution of simulation outcomes. Figure 8 shows the forecast cumulative covered emissions – narrow scope for 2013-2014, broad scope for later years – along with pointwise 95% confidence intervals for the value for each year from 2013 to 2020.

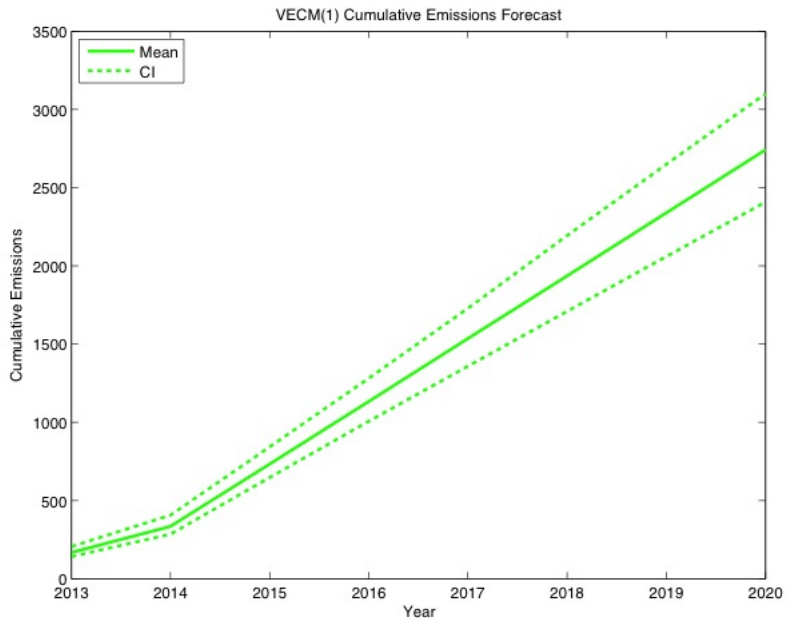

Figure 8. Forecast Results: Cumulative Covered Emissions

V. ADDITIONAL SOURCES OF EMISSIONS ABATEMENT

While the VAR estimation and simulations described in the previous section account for the growth in emissions levels and changes in transport emissions intensities, the price of allowances and other government policies will also affect total emissions. In this section we analyze these other sources of emissions abatement and compliance opportunities.

A cap and trade system is based on the presumption that as the allowance price

rises, the implied increased production costs will change consumer and producer behavior. In order to assess the impact of the change in the emissions price on quantity demanded in the allowance market, we first analyze such price-elastic demand for allowances in four areas on the consumer side: demand for gasoline, diesel, electricity, and natural gas. For each of these areas, we calculate the emissions reduction that would occur with the price at the auction reserve price floor, at the price to access the first (lowest) tier of the allowance price containment reserve (APCR), and at the price to access the third (highest) tier of the APCR.[35] We also consider responses of industrial emissions to allowance prices.

It is important to recognize that the actual allowance price path will evolve over time as more information arrives about whether the market is likely to have insufficient or excess allowances over the life of the eight-year program, as discussed in section II. Prices at these very low or high levels may not be observed until much later in the program, when participants are fairly certain of whether the market will be short or long allowances. Furthermore, there may be considerable uncertainty about future prices throughout the program. Thus, to the extent that response to high allowance prices involves irreversible investments, there may be significant option value in waiting to make those investments until more of the uncertainty is resolved. For these reasons, while we use the APCR price levels to calculate potential responses to high prices in every year, we consider low to medium elasticities in recognition that APCR-level prices are very unlikely until later years and delayed responses of market participants – due to uncertainty and option value – may reduce responses to those prices.

A. Demand for Fuels

The potential impact of the allowance price on consumption of transportation fuels – gasoline and diesel – is a function of short-run effects, such as driving less

[35] Each of these price levels escalates over time in real terms, so we calculate the price-sensitive abatement for each year separately.

and switching among vehicles a family or company owns, and longer-run effects, such as buying more fuel-efficient vehicles and living in areas that require less use of vehicles. If, however, fuel-economy standards have pushed up the average fuel-economy of vehicles above the level consumers would otherwise voluntarily choose (given fuel prices), then raising fuel prices will have a smaller effect, because the fuel-economy regulation has already moved some customers into the vehicle fuel economy they would have chosen in response to higher gas prices. For this reason, in jurisdictions with effective fuel-economy standards, such as California, the price-elasticity of demand for transportation fuels is likely to be lower. Short-run price elasticity estimates are generally -0.1 or smaller.[36] Long-run elasticities are generally between -0.3 and -0.5.[37] Furthermore, the fuel-economy standards would reduce the absolute magnitude of emissions reductions in another way: by lowering the base level of emissions per mile even before the price of allowances has an effect. Recall that we incorporate the direct impact of fuel-economy standards on emissions holding constant vehicle miles traveled when we account for transport emissions intensities in the VAR simulation.[38]

We recognize that improved fuel-economy standards will phase in gradually during the cap and trade compliance periods. To balance these factors, we assume that the base level of vehicle emissions is unchanged from 2012 levels in calculating the price response, and we assume that the price elasticity of demand will range from -0.1 to -0.2.[39] Our fuel price elasticity value is linked to our assumption about the effectiveness of the fuel-economy regulations. If these regulations move consumers into the higher-MPG vehicles they would have bought in response to higher fuel prices, then that emissions savings occurs regardless of the price of allowances. If fuel prices then rise, we wouldn't expect as great a quantity response, as consumers have already purchased cars that are optimized for higher

[36]See Hughes, Knittel and Sperling, 2008.
[37]See Dahl, 2012
[38]The VAR also accounts for estimates of uncertainty in the change in gasoline prices absent GHG costs.
[39]We also assume that the cost of tailpipe CO2 emissions is passed through 100% to the retail price.

fuel prices.

At the highest price in the price containment reserve in each year (which, in 2012 dollars, is $50 in 2013 going up to $70.36 in 2020),[40] the result using a -0.1 elasticity is a reduction of **10.6 MMT** over the life of the program from reduced use of gasoline and diesel. Assuming an elasticity of -0.2 about doubles the reduction to **21.1 MMT**.[41] We also consider the potentially more-elastic response if vehicle fuel economy standards are not separately increased; assuming an elasticity of -0.4 yields a reduction of **44.1 MMT**.[42] (Note the fuels will be under the cap only in 2015-2020, so we calculate reductions for only these six years.) We combine this last case with the business-as-usual transport emissions intensity described in the previous section, essentially assuming this higher price elasticity if higher fuel-economy standards have not been effectively implemented.

If policy is changed to give free allowances to refiners with output-based updating, to incent them not to pass along allowance prices in the price of gasoline, then this source of abatement elasticity will be reduced or eliminated as we discuss in section VII.

B. Demand for Electricity

The impact of a rising allowance price on emissions from electricity consumption depends primarily on the pass-through of allowance costs to retail prices of electricity. As noted earlier, three large regulated investor-owned utilities (IOUs) that serve the vast majority of load in California receive free allocations of allowances that they must then sell in the allowance auctions, resulting in revenues to the utilities. Those revenues must then be distributed to customers. They can be used to reduce the retail rate increases that would otherwise occur due to higher wholesale electricity purchase prices caused by generators' allowance

[40] These allowance prices translate to an increase of about $0.45 to $0.63 per gallon at the pump in 2012 dollars.

[41] Each of these estimates assumes that the LCFS has already raised the biofuel share of retail gasoline to 15%.

[42] This calculation also assumes that biofuels remain at 10% of retail gasoline.

obligations for their GHG emissions. Publicly-owned utilities are not obligated to sell their allowances, but are effectively in the same position of deciding how much of the value of the free allowances will be used to offset rate increases that would result when wholesale prices rise.

Based on a resolution from the CPUC in December 2012,[43] a best guess seems to be that the revenues from utility sales of allowances will be used first to assure that cap and trade causes no price increase to residential consumers. In addition, the revenues will be allocated to dampen price increases for small commercial customers and likely greatly reduce them for energy-intensive trade exposed large industrial and commercial customers. Remaining revenues will be distributed to residential customers through a semi-annual lump-sum per-customer credit. It appears that most electricity sold to commercial and industrial customers will see the full pass-through of energy price increases due to allowance costs.[44]

The CPUC estimates that 85% of revenues will go to residential customers, who make up about 34% of demand.[45] Conversely, 15% of revenues will go to non-residential customers, that is, customers who comprise 66% of demand. If the total allocation of allowances is about equal to 100% of a utility's associated indirect (*i.e.,* through power providers) obligation, and the utility is allowed to cover its cost of compliance, this means that the 66% of demand that is not residential will bear associated costs equal to 85% of the total cost of allowances that cover the utility's obligation.

With a statewide *average* GHG intensity of 0.350 metric tonnes per MWh (based on the 2011, most recent, GHG inventory), this means that the price of electricity per MWh would increase for non-residential customers by an aver-

[43]http://docs.cpuc.ca.gov/PublishedDocs/Published/G000/M040/K841/40841421.PDF. The full decision is at http://docs.cpuc.ca.gov/PublishedDocs/Published/G000/M039/K594/39594673.PDF.

[44]It is worth noting that it is far from straightforward once the program begins for a regulator to know what the counterfactual price of electricity would have been if allowances had sold for a different price or for a price of zero. The price of allowances has a complex impact of wholesale electricity expenditures depending on the emissions intensity of the marginal supplier versus the average supplier and the competitiveness of the wholesale electricity market. Thus, it is not clear how the CPUC would make good on a promise not to pass through the cost of allowances without a detailed study of the impact that cost on equilibrium wholesale electricity prices.

[45]The 34% figure is based on 2012 EIA data for all of California.

age of $(0.85/0.66) \cdot 0.350 \cdot$ allowance price. At an allowance price of \$50/tonne, this raises average non-residential rates by \$22.54/MWh and at \$70.36/tonne by \$31.55/MWh.[46] We apply these increases to the state average retail rates for commercial and industrial customers, based on EIA data, to get a percentage price response. Commercial and industrial electricity demand elasticity estimates are few and not at all consistent. The only study we found in the last 20 years is Kamerschen and Porter (2004), which estimates a long-run industrial price elasticity of demand of -0.35 when controlling for heating and cooling degree-days. We use this figure, though we recognize that it could be too large because the long-run assumption imparts an upward bias to the impact if price is actually increasing over time and we calculate the elasticity based on same-year average price.[47] On the other hand, some earlier studies–reviewed in Taylor (1975) –find much larger long-run elasticities, in some cases above 1 in absolute value.

The -0.35 elasticity is then applied to the share of IOU-served demand subject to this price change, which we take to be 66%, to calculate the resulting reduction in demand. Because the resulting impact on electricity consumption would be a reduction at the margin, we multiply the demand reduction by an assumed *marginal* GHG intensity–which we take to be 0.428 tonne/MWh–to calculate the reduction in emissions at different prices. The result is a reduction of **7.7 MMT** when the price is at the auction reserve throughout the program, **27.3 MMT** when price is at the lowest step of the containment reserve, and **33.4 MMT** when price is at the highest step of the containment reserve.[48]

[46]The 0.350 MT/MWh figure is arrived at by taking total 2011 GHG electricity emissions measured for in-state (38.2 MMT) and assumed for imports (53.5 MMT) and dividing by total consumption (261.9 MMWh). Two assumptions are implicit in this calculation. First, we calculate the impact by spreading the cost of the allowances over all non-residential customers, rather than calculating a slightly higher increase for a slightly smaller set of customers by excluding trade exposed large customers and reducing the obligation of small customers. This is unlikely to make a noticeable difference. Second, we assume that the wholesale price obligation is increased by the cost of the allowances, when it could be more or less depending on the GHG intensity of the marginal versus the average producer and the share of long-term supply contracts with prices set prior to or independent of the impact of GHG costs on market price.

[47]In particular, because the price at any time should reflect all expectations of future changes, the increase in price over time, if it were to occur, would be due to a series of unpredicted upward shocks. Thus, one would not expect market participants to behave as if they had foreseen these shocks.

[48]For an elasticity of -0.2, the reductions are, respectively, 4.6, 15.8, and 19.3 MMT, while for an

Electricity prices, however, are likely to rise for all customers over the years of the program for reasons independent of the price of allowances–increased renewables generation, rising capital costs, and replacement of aging infrastructure, among others–and these increases will reduce consumption.

Taking an average statewide retail electricity price of $149/MWh in 2012,[49] we assume that this price will increase by 2.15% (real) per year due to exogenous (to cap and trade) factors.[50] Again assuming a long-run demand elasticity of -0.35 and a marginal CO2e intensity of 0.428 tonne/MWh, yields a reduction of **24.1 MMT** (if the allowance price is at the highest price in the price containment reserve) over the life of the program.[51]

Thus, at the highest level of the price containment reserve we estimate total abatement from electricity demand reduction of **57.5 MMT** over the life of the program. Both the price elasticity we assume and the marginal CO2e intensity figures may be on the high side. Using an elasticity of -0.2 reduces the impact of electricity demand reduction to **33.2 MMT** at the highest price of the containment reserve. The marginal GHG intensity of 0.428 is based on a combined-cycle gas turbine generator. If some of the reduction comes out of renewable, hydro or nuclear generation the marginal intensity will be lower. The impact scales linearly with the assumed marginal GHG intensity.

C. Demand for Natural Gas

It appears very likely that the ARB will vote in 2014[52] to give natural gas suppliers (who are virtually all investor-owned regulated utilities in California)

elasticity of -0.5 the reductions are, respectively, 10.9, 38.6, and 47.2 MMT. We use these elasticities as a high and low case. The baseline price on which all price increases are calculated is the average price over the life of the program assuming a 2.15% annual real increase in electricity prices during this period, as discussed next.

[49] http://www.eia.gov/electricity/monthly/epm_table_grapher.cfm?t=epmt_5_6_a

[50] This increase is based on a projected real increase from $144/MWh in 2012 to $211/MWh in 2030, an average increase of 2.15% per year. See Energy & Environmental Economics (2014).

[51] Ito (forthcoming) estimates a medium-long run price elasticity for residential electricity demand of -0.2. The reduction from the exogenous price increase drops to 13.9 MMT at an elasticity of -0.2.

[52] See http://www.arb.ca.gov/regact/2013/capandtrade13/capandtrade13isor.pdf. At the October 2013
ARB Board meeting, a decision on these proposals was postponed.

free allowances equal to the obligation associated with their 2011 supply, but then declining at the cap decline factor. If this were done, then nearly all of the suppliers' obligations could be covered with the free allowances (or the revenue from selling them in the allowance auction). From discussions with industry participants and CPUC staff, it appears the most likely outcome is there would be almost no impact of emissions pricing on retail natural gas price, and therefore almost no price-responsive emissions reduction by consumers in this sector. That outcome is not certain, however, so we also explore the impact of emissions prices being passed through to consumers. "Consumers" in this case include all emissions sources not covered in the industrial categories. (Large industrial customers, which are in the program beginning with the first compliance period, are discussed in subsection e.)

If the cost of natural gas emissions were fully passed through to these consumers, then an allowance price at the auction reserve would raise natural gas prices by an average of $0.71/MMBTU (in 2012 dollars) over the 2015-2020 period. At the lowest price in of the APCR, the allowance cost would raise the price of natural gas by an average of $2.71/MMBTU and at the highest price of the APCR, the effect would be to raise the natural gas price by an average of $3.40/MMBTU. We assume an average retail price of $8.49/MMBTU across all nonindustrial types of natural gas customers[53] before allowance costs, and 100% pass-through of the allowance cost to retail. It's difficult to know the elasticity of retail demand for natural gas. We take a low-end estimate of -0.2 and a high-end estimate of -0.4 over the 6-year time frame of natural gas in the program.[54] We assume a baseline emissions rate of 49.7 MMT/year for each of the six years that non-industrial customers are in the program. Based on these assumptions, at the highest price

[53] According to the EIA (http://www.eia.gov/dnav/ng/ng_pri_sum_dcu_SCA_a.htm) in 2012 residential averaged $9.22/MMBTU, commercial about $7.13/MMBTU for the about half of commercial customers in their data. These are likely the smaller customers because larger customers probably have proprietary contracts, which the price data don't cover. The $8.49/MMBTU price is the quantity-weighted average based on EIA estimated quantities.

[54] Though some estimates of the price elasticity of gas and electricity demand are higher than those we use here, such estimates generally include substitution from gas to electricity and vice versa, which would have a much smaller net impact on emissions.

in the price containment reserve, the low-elasticity estimated abatement is **19.4 MMT** and the high-elasticity scenario is **37.5 MMT**. If policy is indeed changed to give free allowances to utilities with the effect of reducing or eliminating the associated retail price increase, then this source of abatement elasticity will be approximately zero.

D. Abatement from Out-of-State Electricity Dispatch Changes

To the extent that some high-emitting out-of-state coal plants are not reshuffled or declared at the default rate, there is possible elasticity from higher allowance prices incenting reduced generation from such plants. We considered this, but the most recent ARB policy suggests that short-term energy trades would fall under a safe harbor and would not be considered reshuffling. If that is the case, then an operator would be better off carrying out such trades than actually reducing output from the plant. This suggests that allowance price increases might incent some changes in reported emissions. In any case, we consider that as part of the reshuffling and relabeling analysis.

E. Industrial Emissions

For the industries covered under output-based updating, there may still be some emissions reductions as the allowance price rises. This could happen in two ways. First, once a baseline ratio of allowances to output is established, these firms have an incentive to make process improvements that reduce GHG emissions for a given quantity of output. It is unclear how much of such improvement is likely to occur. At this point we have no information on this. Our current estimates assume this is zero. ARB's analysis of compliance pathways suggests that at a price of up to $18/tonne (25% of the highest price of the APCR in 2020), the opportunity for industrial process reduction is at most 1-2 MMT per year.[55]

[55]See figures F-3 through F-9 of Appendix F, "Compliance Pathways Analysis," available at http://www.arb.ca.gov/regact/2010/capandtrade10/capv3appf.pdf.

Second, because the output-based updating is not 100%, additional emissions that result from marginal output increases do impose some marginal cost on the firms. That impact is likely to be small, however, because the effective updating factors average between 75% and 90% over the program, which implies that the firm faces an effective allowance price of 10% to 25% of the market price for emissions that are associated with changes in output. At this point, we have not incorporated estimates of this impact, but it seems likely to be quite small.

F. Imported Electricity, Reshuffling, and Relabeling

The ARB has attempted to include all emissions from out-of-state generation of electricity delivered to and consumed in California under the cap and trade program's GHG accounting framework. ARB projects annual BAU emissions from imported electricity of 53.53 MMT, during the period 2013-2020.[56] However, due to the nature of the Western electricity market, it is generally impossible to identify the spescific generation resource supplying imported electricity. Electricity importers therefore have an incentive to engage in a variety of practices that lower the reported GHG content of their imports, a class of behaviors broadly labeled reshuffling. While reshuffling would not yield aggregate emissions reductions in the Western Interconnection, it could be a major source of measured emissions reductions under the California cap and trade program.

Under one extreme, importers could reshuffle all imports to GHG free resources, resulting in no demand for allowances to cover imported electricity. ARB has tried to limit reshuffling by focusing on imports from coal plants partially owned by California utilities. Given the current information, we project emissions associated with imports from these plants to account for 109 MMT during the eight-year period. We treat this as a lower bound on emissions from imports, assuming that all other imported energy is sourced from zero carbon generation.

[56]This comes from the ARB's 2012-2020 California GHG Emissions Forecast. http://www.arb.ca.gov/cc/inventory/data/tables/2020_ghg_emissions_forecast_2010-10-28.pdf

In 2010 there were about 85 net TWh of electricity imported into California. If we assume imported electricity remains at this level during the 8 years, this implies 680 TWh over the 8 years of the cap.[57] Taking the 109 MMT, associated with roughly 109 TWh of electricity imports as a baseline, we consider two other possibilities for the remaining 571 TWh. The first is that all the remaining energy is imported at an emissions rate of 0.428 tons/MWh. This is the "default" emissions rate applied to any imports that do not claim a specific source for the power. Another scenario assumes roughly half the remaining energy is imported at zero emissions, while the other half is imported at 0.428 tonnes/MWh. The result is an average emissions rate of 0.214 tonnes/MWH for this remaining 571 TWh of energy.

Under the three scenarios for the residual (non-utility-owned coal) energy, we have cumulative emissions of either 109.5, 232, or 354 MMT of GHG associated with power imports over the 8 years of the cap. Given that the 2013 cap was based upon emissions of 53.53 MMT from imports, we treat $53.53 \cdot 8 = 428.24$ as the BAU level from imports. The low, medium, and high "reductions" in carbon from power imports would therefore be **74**, **197**, or **319 MMT**.

G. Offsets

The cap and trade program permits a covered entity to meet its compliance obligation with offset credits for up to eight percent of its annual and triennial compliance obligations. This means that over the 8-year program up to 218 MMT of allowance obligations could be met with offsets.

Thus far, ARB has approved four categories of compliance offset projects that can be used to generate offsets: U.S. Forest and Urban Forest Project Resources Projects; Livestock Projects; Ozone Depleting Substances Projects; and Urban Forest Projects. Each individual offset program is subject to a rigorous verifica-

[57] California Energy Commission. http://energyalmanac.ca.gov/electricity/electricity_generation.html. The net total includes roughly 90 TWh of imports and 5 TWh of exports.

tion, approval, and monitoring process. The ARB has approved two offset project registries – American Carbon Registry[58] and the Climate Action Reserve[59] – to facilitate the listing, reporting, and verification of specific offset projects. The ARB reports there are approximately 5.3 million offsets have been listed with ARB under a voluntary early action offset program that are eligible for conversion to cap and trade program compliance offsets.

Offsets are expected to be a relatively low-cost (though not free) means for a covered entity to meet a portion of its compliance obligation.[60] The number of offsets expected to be available in the cap and trade program is subject to a high degree of uncertainty and best guesses put the estimate substantially below the potential number of offsets that could be used (*i.e.,* 8% of compliance obligations). One third-party study from September 2012 estimates the number of offsets available under all four protocols between 2013 and 2020 at 66 MMT, only 30% of the 218 MMT of offsets that theoretically could be used to satisfy compliance obligations.[61] ARB, however, is considering adding at least two additional offset protocols – Rice Cultivation and Mine Methane Capture and Destruction. The addition of these two protocols is estimated to make an additional 100 MMT of offsets available (for an estimated total of 130 MMT) between 2013 and 2020.[62]

For the purposes of our analysis, we consider three scenarios for offsets, one based on the existing protocols (**66 MMT**), one that adds in estimates for rice cultivation and coal mine methane (**130 MMT**), and one that assumes the full allowed **218 MMT** of offsets are approved and utilized for compliance.[63] These offsets enhance the effective supply of allowances. Most estimates of the price

[58]See http://americancarbonregistry.org/carbon-accounting/california-compliance-offsets.
[59]See http://www.climateactionreserve.org/.
[60]http://www.arb.ca.gov/regact/2010/capandtrade10/capv3appf.pdf.
[61]http://americancarbonregistry.org/acr-compliance-offset-supply-forecast-for-the-ca-cap-and-trade-program.
[62]http://americancarbonregistry.org/acr-compliance-offset-supply-forecast-for-the-ca-cap-and-trade-program.
[63]The analysis described in this document assumes a single eight-year compliance time horizon. As a result, the analysis does not address the fact that current rules do not allow a shortfall of offsets in an earlier compliance periods to be recaptured in later time periods, and thus results in a permanent shortfall in offsets from the theoretical potential.

at which offsets would be available put their cost at below or just above the auction reserve price. For all three scenarios we assume that the offsets utilized are available below the auction reserve price. In reality, studies suggest that some may require a price slightly above the auction reserve price, but still likely below $20/tonne. We group these with the abatement available at or slightly above the auction reserve price.

H. Aggregating Scenarios for Emissions Abatement

Table 5 summarizes the analyses of emissions abatement. For each abatement source and scenario, the number shown represents the total abatement that would occur over the life of the program at an allowance price equal to the highest price of the APCR for each year.[64] For each source, we also highlight what seems to be the most likely abatement scenario.

From Table 5 we then aggregate the scenarios for emissions. By summing the minimum, medium, and maximum abatement figures by for each source, we create the "Very Low", "Medium," and "Very High" estimated abatement shown in Table 6. The Very Low and Very High aggregates, however, would require extreme outcomes for each of these sources, which is extremely unlikely. So, we create Low and High scenarios as the average between the medium and the extreme outcomes. This is obviously somewhat arbitrary, but it allows us to show the sensitivity of allowance prices to the abatement level that is attained. These scenarios are shown in Table 6.

[64]Table 6 shows figures at an allowance price equal to the auction reserve price, the lowest price of the APCR, and the highest price of the APCR.

Table 5—Summary of Emissions Abatement Assumptions

Price-responsive Allowance Demand Reduction	Elasticities		Range of Energy Price Changes At Different Levels of Allowance Price Over 8 years ($2012):			Abatement over 8 years at highest APCR step each year (MM tons)	
Sector	Low	High	Auction Reserve	Lowest step of APCR	Highest step of APCR	Low	High
Electricity most C&I ($/MWh)	-0.20	-0.50	$4.74/$6.66	$18.04/$25.38	$22.55/$31.73	19.3	47.2
Transportation ($/Gallon)	-0.10	-0.40	$0.10/$0.12	$0.37/$0.47	$0.46/$0.59	10.6	44.1
Natural Gas ($/MMBTU)	-0.20	-0.40	$0.60/$0.80	$2.39/$3.06	$2.99/$3.82	19.4	37.5

Notes: All energy price changes assume 100% passthrough.
Range of price changes shown are for first and last year covered by cap and trade program
Range of price changes for Transportation and Natural Gas are for 2015-2020 only, electricity for 2013-2020
Range of Transportation price changes based on weighted average of gasoline and diesel
Transportation abatement impact is for tailpipe emissions only, does not include associated upstream emissions
Low case considered for natural gas is zero abatement, like result if utilities awarded allowances to cover this liability
GHG intensities assumed are explained in the text

Other Abatement Not Responsive to Allowance Price	Low	High
Electricity - Exogenous Price Increases for all customers	13.9	34.1
Transport price increase due to Low Carbon Fuel Standard	0	10
Offsets	66	218
Reshuffling	74.6	318.7

All "other abatement" assumed available at auction reserve price except offsets
Two-thirds of Offsets assumed available at auction reserve price, remainder as just above auction reserve price

Table 6—Summary of Abatement Supply Scenarios

More Stringent Fuel Economy Standards

	Very Low			Low			Medium			High			Very High		
	ARP ARP	Low APCR	High APCR	ARP ARP	Low APCR	High APCR	ARP ARP	Low APCR	High APCR	ARP ARP	Low APCR	High APCR	ARP ARP	Low APCR	High APCR
Electricity Elasticity	4.6	15.8	19.3	6.1	21.6	26.4	7.7	27.3	33.4	9.3	33.0	40.3	10.9	38.6	47.2
Transport Elasticity	2.2	8.5	10.6	2.8	10.6	13.2	3.3	12.7	15.8	3.9	14.8	18.4	4.4	16.8	21.1
Natural Gas Elasticity	0.0	0.0	0.0	2.0	7.8	9.7	4.1	15.5	19.4	6.0	22.8	28.5	7.9	30.0	37.5
Offsets	66.0	66.0	66.0	98.0	98.0	98.0	130.0	130.0	130.0	174.0	174.0	174.0	218.0	218.0	218.0
Exogenous Elec. rate effects	13.9	13.9	13.9	19.0	19.0	19.0	24.1	24.1	24.1	29.1	29.1	29.1	34.1	34.1	34.1
Transport LCFS	0.0	0.0	0.0	2.5	2.5	2.5	5.0	5.0	5.0	7.5	7.5	7.5	10.0	10.0	10.0
Resource Shuffling	74.6	74.6	74.6	135.6	135.6	135.6	196.6	196.6	196.6	257.7	257.7	257.7	318.7	318.7	318.7
Total Abatement	161.3	178.7	184.4	266.0	294.9	304.3	370.7	411.2	424.3	487.4	538.7	555.4	604.0	666.3	686.5

:

Current Trend Fuel Economy Standards

Transport Elasticity	9.3	35.3	44.1	9.3	35.3	44.1	9.3	35.3	44.1	9.3	35.3	44.1	9.3	35.3	44.1
Total Abatement	168.3	205.5	217.9	272.5	319.7	335.2	376.7	433.8	452.6	492.8	559.2	581.1	608.8	684.7	709.6

It is immediately clear from Table 6 that the greatest uncertainty in abatement supply to the market is in the use of offsets and the amount of reshuffling that will occur. Unfortunately, we currently have no way to narrow this uncertainty, which will be driven by both unknown factors – such as the willingness of utilities out of California to sell cleaner power and buy coal-generated power – and by endogenous policy decisions – such as the speed of approval and stringency of new offset protocols and the degree of oversight and intervention to minimize reshuffling. Instead, we present results for a range of aggregate abatement figures and discuss scenarios that might result in those levels.

VI. SUPPLY/DEMAND BALANCE UNDER ALTERNATIVE SCENARIOS

In order to compute the probabilities of different price outcomes in California's GHG market, we combine the emissions simulations generated from the VAR models we estimated in Sections II and III with scenarios for abatement supply, offsets and reshuffling. We consider four mutually exclusive and exhaustive potential market clearing price ranges: (1) at or near the auction reserve price, with all abatement supply coming from low-cost abatement and offset supply, (2) noticeably above the auction reserve price, though without accessing any of the allowances in the allowance price containment reserve (APCR), with marginal supply coming from price-elastic sources, (3) above the lowest price at which allowances would be available from the APCR, but at or below the highest price of the APCR, and (4) above the highest price of the APCR.

We characterize price range (1) as "at or near" the auction reserve price for two reasons. First, the mechanism of the auction reserve price implies an uncertain economic price floor. The auction reserve price was set at $10 per tonne for 2012 and then rising at 5% per year plus inflation. Setting aside the uncertainty of inflation, if investors' real cost of capital differs from 5%, then the effective economic price floor will not be the auction reserve price. If, for instance, in-

vestors' real cost of capital were 3% per year for an investment such as this, then the effective price floor today would be the present discounted value of the price floor in the last auction in which allowances are sold.[65] Thus, in any one year the effective economic price floor may differ somewhat from the auction reserve price. Second, we recognize that some offsets may require a price slightly above the auction reserve price.

As of this writing, the ARB is expected to implement new policies to address the possibility of the price containment reserve being exhausted. We do not address how high the price might go in case (4). This would be difficult to do even in the absence of this policy uncertainty, because it will be greatly influenced by the ARB's policy decisions scheduled to occur this year. We simply report the estimated probability of reaching this case and note that prices could go extremely high.

Our analysis is in terms of real 2012 dollars, so there is no need to adjust for inflation, but the price trigger levels for the price containment reserve will, under current policy, increase at 5% in real terms every year. Thus, while the containment reserve is made available at prices from \$40-\$50 in 2013, the range escalates to \$56.28-\$70.35 in 2020 (in 2013 dollars). As we show below, the containment reserve prices are only likely to occur if BAU GHGs grow faster than anticipated over many years, so the most relevant containment reserve prices are those that will occur in the later years of market operations. Nonetheless, for the price-responsive abatement, we calculate response (for a given elasticity) as if the price is at the relevant step of the APCR in each year of the program. This will tend to overstate price-responsive abatement and understate the probability of exhausting the reserve.

We consider emissions forecasts from the VAR under the three different estimation approaches described in Section III: first with the VAR-forecasted trans-

[65]For example, if inflation were anticipated to be 2% per year, the nominal auction reserve price in 2020 would be \$17.18. If investors anticipated some new sales of allowances in 2020 and their cost of capital was 3% per year, then the effective economic price floor in 2012 would be \$17.18 discounted back to 2012 at 5% per year, or \$11.63, rather than \$10.

portation emissions intensity and then with two different adjustments that lower the assumed transport emissions intensity to reflect the impact of stricter fuel economy standards and a greater biofuels share of retail fuel. We combine each scenario with the low, medium and high abatement scenarios that were described at the end of the last section.

We consider the medium availability scenario a good center of the possible outcomes. It is unlikely that all the low or all the high cases for abatement and offset factors would occur, so we consider less extreme low and high cases as described above.

We put these together with the predetermined allowance supply available (not counting allowances in the price containment reserve) to determine the supply through 2020 at prices below the lower trigger price for the containment reserve. At prices between the lower and upper trigger price for the containment reserve, we also added in the available supply from the containment reserve. We then combine the supply scenarios with the distribution of demand for greenhouse gas allowances under the three VAR estimation approaches discussed in Section III to determine the probabilities that the market outcome will fall in each of the four price ranges discussed above. Figure 9 shows these probabilities using each of the three demand response estimation methods and each of the three supply scenarios.

Focusing on the middle bar of the graph – using the VAR with adjustment to the higher transport intensity from the EMFAC model and with medium abatement – the bar suggests that by 2020 there is a 76% probability that the allowance price will be at or near the auction reserve price, a 7% probability that it will be substantially above the auction reserve price, but still below the lowest price at which the containment reserve allowances can be sold, a 12% probability that the price will be within the range of the containment reserve, and an 6% probability that the containment reserve will be exhausted.

In contrast, if low, but plausible, abatement outcomes occur, then even with

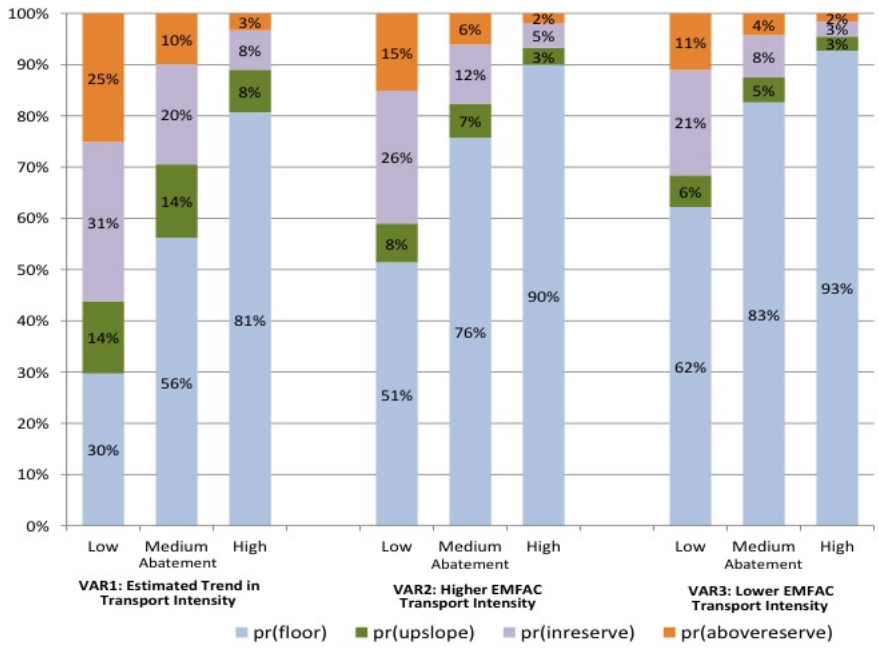

Figure 9. Allowance Price Probabilities By Scenario

the assumed moderate improvements in transport emissions intensity of the VAR2 case, we estimate a 15% probability that the APCR will be exhausted and, absent other government intervention, the price would climb to above the levels of the APCR. The probability of triggering the APCR is 41% in that case. If the state is very successful in reducing transport emissions intensity, the VAR3 case, then the low abatement scenario still leaves an 11% probability of exhausting the APCR and a 32% probability of triggering the APCR.

The results make clear the importance of accomplishing high levels of what we have termed abatement, but the previous section and Table 6 make clear that the greatest variation in that category will come from offsets and reshuffling. Both of these reduce the need for abatement by covered entities. Over the range of prices from the auction reserve to the top of the APCR, price-responsive abatement, while not inconsequential, is likely to play a smaller role.

The three different VAR scenarios with different transport emission intensity paths also demonstrate that the effectiveness of the state in lowering transport emissions intensities will play a major role in determining the ultimate supply/demand balance in the market. If the state achieves the full range of planned policies in improving transport emissions intensities (the VAR3 cases in figure 9), then the probability of exhausting the APCR is below 10% under nearly all scenarios of other abatement methods. But if the state were to just maintain the existing trend in transport intensities, as estimated in the VAR1 case, then other abatement will need to be successful in order to keep the probability of exhausting the APCR in a low range.

Finally, the results demonstrate that the relationship between these scenarios of transport emissions intensities and abatement on the one hand and the allowance market outcome on the other hand is not at all deterministic. There is quite a bit of variation in the business as usual emissions, as shown in figure 9, resulting from uncertainty in GSP, fuel prices, and related factors. Without accounting for this BAU uncertainty, it is not possible to recognize the range of possible outcomes and how other policies change the probabilities of those outcomes.

A. Analysis of Supply/Demand Balance in First Two Compliance Periods

We next estimate the probability distribution of emissions and construct possible abatement supply curves for the first two compliance periods. Our approach for the probability distribution of emissions is the same as in the estimates for 2020. Separate analysis for the early compliance periods is important, because although allowance banking is permitted without restriction, allowance borrowing from later compliance periods is not permitted. Thus, a tight market and high price could potentially occur in an earlier compliance period even if the 8-year market equilibrium would occur at a low price. These results should be thought of as estimates of what could happen in earlier compliance periods if they could not be arbitraged with later periods.

The 1000 simulations used to construct the forecasts for 2013-2020 also yield 1000 simulated forecasts for 2013-2014 and for 2013-2017, as shown in table 3. For each year, we use the same assumption or range of assumptions for RPS, nuclear power, emission intensity of thermal generation, and emissions intensity of transportation as were used for the same year in the previous section, shown in table 4.

For the most part, construction of abatement supply follows the same assumptions as were used in construction of the data in tables 5 and 6. The analysis is simplified for the first compliance period, because transportation fuels and non-industrial natural gas are not included in the program, and therefore in the abatement. As in the prior analysis, reshuffling and offsets are the largest drivers of abatement supply. While reshuffling is calculated with reference to specific out-of-state plants, as described in section V.F, we have less detailed data on yearly offset supply. Offset supply for the first compliance period is set at 26 MMT, based on discussions with individuals at ARB and Climate Action Reserve. However, we were unable to find further detail for the years 2015-2020. So, offsets for the second compliance period are a linear interpolation between the assumed 26 MMT for the first compliance period and the end-of-program aggregate assumptions shown in table 6.

For the other abatement supply sources that are present in the first compliance period, we assume that the abatement occurs uniformly over the 8-year program. For the sources not under the cap in the first compliance period, we assume that abatement takes place uniformly over the 6 years from 2015 to 2020. These assumptions likely overstate somewhat the supply of abatement in the first two compliance periods, but these sources contribute relatively less to total abatement than reshuffling and offsets, so we proceed with these simplifications for now. The summary of assumed abatement in the three compliance periods is shown in table 7.[66] We present only the low and medium abatement cases. The high abatement

[66] The phase I cumulative allowances includes 10 MMT of so-called True-Up Allowances. These are

Table 7—Allowance Availability and Abatement by Compliance Period

Compliance Phase	Low			Medium		
	Phase I (2013-14)	Phase II (2015-17)	Full Program	Phase I (2013-14)	Phase II (2015-17)	Full Program
Average Holding Limit	6.2	13.0	10.9	6.2	13.0	10.9
Cumulative allowances	329.3	1420.7	2386.8	329.3	1420.7	2386.8
Cumulative Offsets	26.0	62.0	98.0	26.0	78.0	130.0
Cumulative Reshuffling	29.7	81.1	135.6	44.3	118.7	196.6
Floor Price						
Responsive Abatement	6.3	19.3	32.4	7.9	26.0	44.1
Low APCR Price						
Responsive Abatement	10.1	35.7	61.3	12.8	48.7	84.6
Total Floor Price	62.0	162.4	266.0	78.2	222.7	370.8
Total APCR Low Price	65.8	178.8	294.9	83.1	245.4	411.2
Changes with status quo fuel economy standards						
Floor Price Transport	0.0	4.6	9.3	0.0	4.6	9.3
Ceiling Price Transport	0.0	17.6	35.3	0.0	17.6	35.3
Total Floor Price	62.0	165.7	272.5	78.2	225.7	376.7
Total APCR Low Price	65.8	191.2	319.7	83.1	256.7	433.8

Note: Phase I Cumulative Allowances Include 10 mmTon of True-up Allowances (see footnote 66).

cases in the previous section that covered eight years seemed plausible only as a result of greatly expanded abatement towards the end of the program – long-run adjustments to higher energy prices, expanded offsets, and more aggressive reshuffling. For these shorter time horizons, the abatement levels implied by the high scenario seem very unlikely.

Based on these assumptions, we combine abatement during the first and second compliance periods with our estimates of BAU emissions over those time periods to calculate whether the market during each of these periods could see a short-run supply constraint at the end of the compliance periods. These constraints would not be reflected in the 8-year probability distribution shown above, but instead would reflect a tight market within the compliance period that cannot be relieved through arbitrage with later periods due to restrictions on borrowing.

allowances given to entities as part of adjustments to output-based updated allocations if the entity had greater production output than was anticipated. Such entities will be given 2015 vintage allowances for this difference, but will be allowed to use those particular allowances for the surrender that covers the first compliance period, 2013-14. The 10 MMT is based on discussions with ARB staff. We do not make a similar adjustment for phase II, instead assuming that the reduction in phase II allowances from the true-up for phase I approximately balances the increase in allowances eligible for phase II surrender when the equivalent true-up adjustment is made at the end of phase II.

For these earlier compliance periods, we also assume that reallocation of allowances from later compliance periods is feasible at the highest price of the APCR and that these are sufficient to ensure that price never exceeds the highest APCR level in either of the first two periods. Thus, our analysis of the end of each of the first two compliance periods focuses on the probability of reaching the APCR low price, but not on the probability of exceeding the APCR high price.

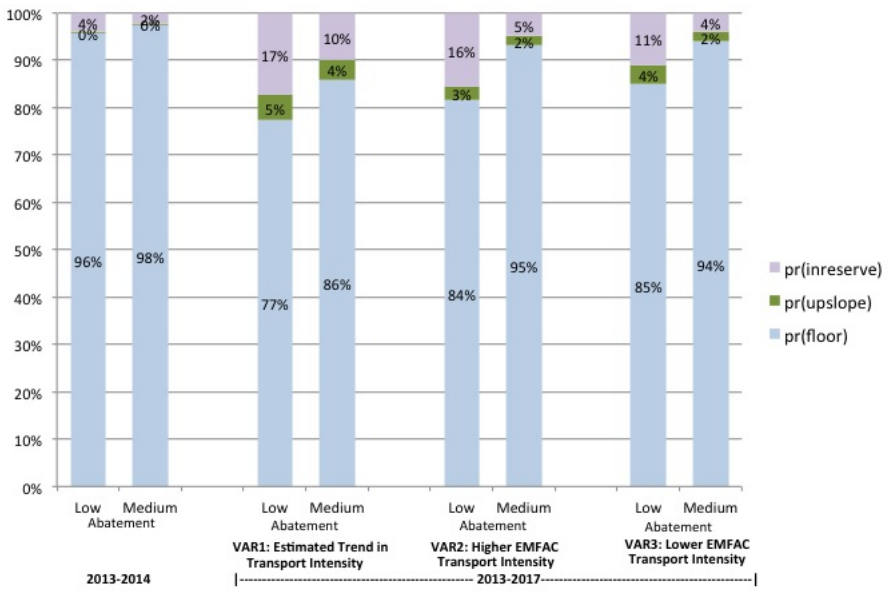

Figure 10. Allowance Price Probabilities By Scenario

The results for the early compliance phases are summarized in Figure 10. It is important to note that these distributions should not be interpreted as predicted allowance prices. Due to the option to bank allowances, prices could be higher than the shorter term supply/demand balance would suggest if longer term projections suggest a tighter market and encourage banking. These are the prices one would expect assuming that no allowances are banked for later use. Implic-

itly, this analysis assumes that all available allowances would be used for the compliance period if the price were above the floor.

As before, the degree of reshuffling plays a significant role in driving allowance prices, making up most of the difference between the low and medium abatement scenarios in table 7. In the first compliance period, our analysis suggests a probability of 2%-4% that the allowance demand would exhaust available supply before the APCR. In the first and second compliance period combined, Figure 10 suggests a 4%-17% probability of price reaching the APCR low price and another 2%-5% of price being significantly above the floor, but not up to the APCR.

These results are based on the assumption that allowance prices are competitively determined, *i.e.,* that no firm attempts to change the price of allowances by withholding supply from the market. We next turn to examine whether any market participants might have an incentive to act non-competitively.

VII. ANALYSIS OF UNILATERAL WITHHOLDING OF ALLOWANCES

In this section we analyze the potential for a firm or firms to manipulate the allowance market in order to create artificial scarcity and raise allowance prices. We combine the analysis of abatement supply and demand from the previous sections with the market rules on allowance holdings to study whether firms acting within the current restrictions could still profitably withhold allowances in order to drive up allowance prices.

To simplify the analysis and apply it to our study of potential market manipulation below, we calculate the potential price-elastic response at the auction reserve price for the first compliance period (for just electricity) and for the second compliance period (for electricity, fuels and natural gas), and then we add this quantity to the horizontal abatement supply available at the price floor. We do this because we are investigating the situation in which the price remains at or near the price floor under competitive conditions while a firm is able to purchase

up secretly large quantities of allowances. Such market manipulation would be a surprise to the market so allowance prices would not anticipate an artificial shortage of allowances caused by such withholding of supply.

In order for the manipulation strategy to succeed, the entity would have to buy allowances at a relatively low price and then withhold some of the allowances from the market while selling the remaining allowances at a higher price. One aspect of the California allowance market – the fact that the allowances are bankable – makes this strategy more attractive. As a result of bankability, a firm could, for instance, buy 2013-2014 vintage allowances and withhold many or most of them from use for 2013-14 compliance, but still get value from the withheld allowances in later years by using them in later compliance periods. In fact, because the price floor rises at 5% above the inflation rate, the withheld allowances themselves, if purchased at or near the price floor, could potentially be attractive, low-risk, long-term investments even absent a manipulation motive. This lowers the cost of attempting to manipulate the market.[67]

The attractiveness and success of a manipulation strategy depends on being able to withhold a sufficient quantity of allowances to significantly raise price. This depends on how many allowances an entity can own, as we will discuss below, but it is also a function of the elasticity of demand and the elasticity of alternate supply in the market. In our analysis, we have considered all price responsiveness as part of "abatement supply" including changes that aren't technically emissions abatement within the program, such as reshuffling, offsets and allowances from the APCR. Figure 1 illustrated that over a large range of abatement, the supply is extremely elastic due to complimentary policies, but where the effects of those policies are exhausted, the supply is likely to be quite steep up to the APCR. This steep jump from near the price floor to the lowest tier of the APCR increases the potential for a profitable manipulation strategy, because withholding a relatively

[67]Of course, allowance banking also has enormous positive impacts on the market, allowing intertemporal arbitrage that encourages cost-efficient abatement. Our discussion of the impact of banking on market manipulation strategies in no way suggests that banking should not be permitted.

small number of allowances may cause a large price increase. In analyzing the earlier compliance periods, as we do below, the upward sloping part of the supply curve is likely to be especially steep because there is less opportunity for price-responsive adjustment over a shorter period of time.

The shape of the abatement supply curve, however, also suggests that this manipulation strategy would require that the supply-demand balance fall into a fairly narrow intermediate region. If the emissions trajectory were very low, then the likelihood of prices being at the floor would be very high and even a shock such as the removal of tens of millions of allowances would not be enough to trigger an increase in prices. If the emissions trajectory (and therefore allowance demand) is quite strong and price rises into the APCR even without any withholding, then this strategy would be unnecessary to reach the APCR price levels. Withholding allowances in order to move price from near the floor to the APCR would be effective only if the market were on a trajectory that were low enough that it would cause prices to be at or near the floor, but high enough that prices could plausibly rise rapidly if a fairly small share of all allowances were removed from the market.

Once demand has pushed the market into the APCR, there is much less to gain from withholding in order to make the smaller jumps between tiers of the APCR, though in certain circumstances that could also be a profitable strategy. In those cases, however, the opportunity cost of withholding allowances may be very high if the tight allowance market that led to the high price were not anticipated to continue in the next compliance period. If that were the case, then the potential loss from a price drop on the withheld allowances would more likely be greater than the potential gain from the small price increase on those the firm could sell at a higher tier within the APCR.

We do not consider here the potential for withholding that would drive price above the highest price of the APCR. The measures recently approved by the Board to insulate the market from exhaustion of the APCR during the first two

compliance periods make such an outcome less likely in the first two compliance periods of the program. By removing allowances from the later years of the program and making them available at the highest price of the APCR in the auction just prior to allowance surrender,[68] the containment policy provides the ability for a sizable amount of borrowing, but only at the highest price of the APCR. This "front-loading" of allowances in the reserve does not increase the number of allowances available at lower prices. Therefore the possibility exists that earlier phases of the program may reach into the containment reserve, even if front-loading policy substantially reduces the probability that they will exceed the highest price in the APCR.

We also do not consider withholding strategies for the third compliance period, which ends in 2020. While the supply/demand analysis in the previous sections suggests that there could be a tight market for allowances at the end of 2020 that could create possibilities for market manipulation, there are likely to be significant changes before then – including adoption of a post-2020 plan – that make analysis of withholding strategies for 2020 too speculative to be useful at this point.

These facts lead us to focus our manipulation analysis on the ends of the first two compliance periods. The first period ends December 31, 2014 with final allowance surrender for the period taking place in November of 2015 and the second period ends December 31, 2017 with final surrender taking place in November 2018. We evaluate the extent to which a firm, operating within the current holding limits, could profitably remove eligible allowances from a compliance period and thereby impact prices.

Our concern here is market manipulation or withholding, not speculation. A speculator buys allowances as a bet that the price will rise, but has no direct ability or intent to influence that price. We do not consider that to be troublesome or undesirable behavior. In fact, such speculative behavior can play a valuable role in price discovery and liquidity of commodity markets, including markets for

[68] See http://www.arb.ca.gov/regact/2013/capandtrade13/capandtrade13isorappe.pdf at page 180.

tradable emissions allowances. The behavior that we are concerned with would be a firm that buys up allowances at low prices with the *intent* of withholding some of them from the market in order to drive up price and then profitably sell only some of those allowances at higher prices. That would be manipulative behavior and measures should be taken to both discourage and mitigate the risks of such behavior.

It is also important to recognize that market manipulation is fundamentally dependent on asymmetric information and deception. If other market participants recognize that a firm is attempting to acquire a large long position they will at the least attempt to cover any short position of their own and may possibly also attempt to acquire a long position in order to benefit from another firm's withholding strategy. Such widespread knowledge would drive up demand immediately and drive up price, which would raise the cost to the potential market manipulator and reduce the incentive to carry out such manipulation. Thus, our analysis examines primarily cases in which a market participant with intent to manipulate the market is able to acquire up to their holding limit without tipping off other firms to their intent.

Assumptions of the Withholding Analysis

As we described above, our focus is on the possible emissions trajectories that would imply prices at or near the floor, but demand somewhat close to the steeply rising segments of the abatement supply curve. Note that if the market is at or near a trajectory to remain along the floor price, this would have several implications for a withholding strategy: First, allowances would be available to purchase at or very near the floor price. Second, the "removal" of allowances (purchased at or near the price floor) from one phase of the program (say 2013-2014) by banking of those allowances would be very low cost, because they could be held until a later period while appreciating in price by 5% above inflation. Third, a strategy implemented towards the end of a compliance period, or after the period ends, but before final allowance surrender, would engender little or

no price-responsive abatement of emissions. If other firms believed that there would be sufficient allowances for the compliance period and one or more entities implemented a purchase and withholding strategy late in the period, then even as the allowance price responded by rising there would be little or no time in which abatement could increase.

We therefore focus on the following scenario.

1) A firm is able to purchase allowances at the floor price up to its holding limit.

2) That firm removes some current vintage allowances from the current compliance period by depositing them into their compliance account for a future compliance period.

3) The firm then sells the remaining allowances, which remain in the firm's holding account, at the price that results after the market becomes aware that these allowances have been removed.

4) These actions happen within a short-enough time frame, or late enough in the compliance period, that there is little or no ability for abatement in response to the higher allowance prices.

We recognize that attempts to purchase large quantities of allowances would eventually drive up the market price. It's very difficult, however, to know how large a long position that a firm (or set of firms, as discussed below) would be able to secretly build up before its acquisition price would rise. As a result, we show result for a range of possible withholding quantities, reporting the associate price increase that would be likely to result if a firm could buy the quantity without raising allowance prices.

Under these assumptions, there is almost no cost to purchasing a given quantity of allowances from the current compliance period and depositing them into a compliance account for the next compliance period, as explained above. The

potential revenue from such a strategy would depend upon how many allowances a firm has left in its holding account and the change in price that such an action would stimulate. On any small scale this strategy would yield no revenue gains. Withdrawing 1000 tons from a first phase market of 329 million tons would almost certainly have no effect. On the other hand, withdrawing 30 million tons, almost 10 percent of all first phase allowances has a much higher probability of raising price.

The likelihood of a price change will depend upon the cumulative BAU emissions relative to the amount of allowances released. For example, if BAU emissions on their own took the market into the APCR, then withholding could have little additional effect, as discussed above, assuming that the APCR is not exhausted. Conversely, if phase I emissions totaled only 270 million tons, well short of the allowance plus abatement supply available at the price floor, then withdrawing even 30 million more would still be insufficient to impact the price. This discussion illustrates that a withholding strategy will only impact scenarios in which BAU emissions are close to, but not into the APCR. We therefore focus our analysis on the probabilities that BAU emissions would fall in such a range, and examine how widening or narrowing this range would affect these probabilities.

In order to assess these probabilities, it is therefore necessary to consider what the relevant range of holding limits might be. This is the subject of the following subsection.

A. Holding limits in the California market

The specific limits on the number of allowances a firm can own vary over time and by firm. Allowance ownership limits vary by firm only in the sense that firms can pre-comply with future emissions obligations, and can therefore deposit allowances into an irrevocable compliance account for that purpose.

Holding Limits and Pre-compliance: The degree to which firms can pre-comply in this way is pegged to their historic and anticipated emissions. As described

below, only some level of pre-compliance is exempt from the holding limits. This exemption is set roughly at the currently accrued aggregate allowance obligation. In other words, in addition to the basic holding limits, firms can own allowances roughly equal to their total emissions up to and including the current year. This amount is known as the "limited exemption" to the holding limits. This extra amount can only be held in an irrevocable compliance account however, and cannot therefore be resold to other parties.

General Holding Limits: Absent a specific compliance obligation (*i.e.,* emissions), the absolute amount of allowances that can be held by a single firm are differentiated into "current" allowances and "future" allowances. There is a separate holding limit for allowances in each category. Limits change each year, but the average holding limits for each compliance phase are summarized in Table 7 above.

Current allowances include any allowances with vintages up to and including the current year. For example, during the calendar year 2015, the allowance vintages of 2013-2015 would be considered current. Allowances with post-2015 vintages, some of which will have been auctioned in prior years, would be subject to a separate holding limit. There is a separate limit for each future vintage year. The total allowed would therefore be the sum of the limit on "current" allowances (*i.e.,* current and all previous year vintages) and the individual limits on each future year allowance.

Compliance Accounts and Limited Exemptions to Holding Limits: Firms with compliance obligations can also deposit allowances into a compliance account. A certain quantity of allowances held in compliance accounts do not count toward the general holding limit. This quantity is known as the limited exemption. The limited exemption has several nuances but in general the quantity firms can hold in excess of their holding limit is pegged to their expected annual emissions. In October of each year, the limited exemption increases by a quantity equal to the previous year's emissions. This amount serves as a proxy for the expectation of the emissions for the future year. Thus the amount of allowances a firm may own

in a compliance account grows each year of a compliance period. At the end of a compliance period, after all obligations have been settled through the allowance surrender (e.g. November 2015 or November 2018), the limited exemption declines by an amount equal to the outstanding obligation that had just been surrendered. In November of all other years, a firm must surrender allowances equal to 30% of its emissions in the prior year, but its limited exemption is not reduced by this amount. This is one way in which firms can own (in compliance accounts) more allowances then they are expected to need for compliance.

This policy means that firms with large obligations can hold large amounts of allowances, but only in their compliance accounts. Two elements of the policy reduce, but do not eliminate, manipulation concerns. First, firms are unable to resell allowances held in compliance accounts. This reduces manipulation concerns somewhat as firms can still sell their possibly substantial allowance balance that is in their holding accounts. As we discuss below, there are also possibly harmful implications to restricting the resale of allowances held in compliance accounts.

The second factor limiting the withholding concerns with such large holdings is that frequently these holdings are owned by firms with even larger obligations. It is the net *long* (holdings less obligations) position of firms that would be central in determining their incentive to attempt to raise allowance prices. Absent this incentive the ability to raise prices is not a concern.

We therefore focus our withholding analysis on the hypothetically largest net long position a firm could possess. For non-compliance entities, this would be the holding limit. For compliance entities, the largest net long position is possible during the transition from one compliance period to the next. During this period, firms are able to purchase a quantity of allowances equal to their previous year emissions that is in addition to their current obligation that is coming due in November of that year.

Following the logic of this discussion, we utilize estimates of a given firm's annual obligations as a reasonable proxy for the maximum long position that

firm could have under the limited exemption. During 2014, a firm can hold up to 6.4475 million allowances above its limited exemption. That number increases to 13.37 million allowances for 2015 and declines by around .35 million allowances for each year after that (Table 7).

Critically, a firm with a compliance obligation has the opportunity to acquire a substantial long position in allowances of the vintage necessary to meet the obligation in the previous compliance period when the limited exemption is increased at the end of a compliance period. For instance, from January 2015 until November 2015, a firm could hold – under the limited exemption – allowances equal to the sum of their obligation for 2013-14 *and* their estimated obligation for 2015. Table 8 summarizes ARB approximations of the top broad scope annual emissions obligations based upon 2012 data. In addition, each entity could hold about 6.4 million current vintage allowances in their holding account prior to January 1, 2015 and about 13.4 million after that date. This means that the six largest compliance firms could each own up to about 37 to 59 million 2013-2014 allowances in excess of their 2013-2014 obligations. A similar opportunity arises at transition from the second to third compliance periods in late 2017 through late 2018.

B. Analysis of Withholding in First Two Compliance Periods

In order to assess the potential for profitable withholding at the end of either of the first two compliance periods, we utilized the probability distribution of emissions and the possible abatement supply curves for the periods 2013-2014 and for 2013-2017 that were described in section VI.A. Our approach for utilizing the probability distribution of emissions is the same as in the estimates for 2020, 2017 and 2014 described above. As before, construction of abatement supply follows the same assumptions as were used in construction of the data in tables 5 and 6 and summarized in table 7. For the first compliance period, we drop the vehicle intensity variations, because fuels are not under the cap in 2013-14. We

Table 8—Largest Compliance Obligations

Firm	Approximate 2012 (MMT)		
	Narrow Scope Emissions	Broad Scope Emissions	Total
Chevron U.S.A. Inc.	9.74	32.16	41.90
Tesoro Refining & Marketing Co.	8.10	26.71	34.81
BP West Coast Products		23.88	23.88
Phillips 66 Company	4.52	18.91	23.42
Southern California Gas Co.	0.17	22.55	22.71
Pacific Gas and Electric Co.	3.36	18.90	22.26
Valero Marketing and Supply Co.	3.73	14.00	17.73
Shell Energy North America	4.24	10.71	14.95
LADWP	12.91	0.00	12.91
Exxon Mobil Co.	3.39	8.61	12.00
Southern California Edison Co.	9.96	0.00	9.96
Calpine Energy Services	9.41	0.00	9.41

also fix the offset quantity at 26 MMT. We consider two reshuffling scenarios, one using the "medium" amount of reshuffling described in previous sections, and one with relatively low levels. The low reshuffling scenario assumes that all non-coal imports are imported at 0.428 tons/MWh (the current default emissions rate). The medium reshuffling scenario assumes that coal commitments as of January 2014 remain through their current contractual lifetimes, and that the remaining energy is imported at an average intensity of 0.214 tons/MWh, which is assumed to be one-half the default emissions intensity.

To understand the incentive to withhold allowances, we start from a simple model in which

1) A firm can purchase allowances at the floor price without driving up the price.

2) The abatement supply curve is flat at the price floor up to a pre-determined quantity (set by the complementary policies, low-cost reshuffling, and offsets), then vertical up to the lowest price of the APCR. There are then three steps of the APCR with no abatement elasticity between the steps.

3) The highest price of the APCR has unlimited supply. We are not focusing on exceeding the APCR supply during the first two compliance periods for the reasons discussed earlier.

4) A firm considering a withholding strategy knows with certainty the total emissions of the market, availability of offsets, and degree of reshuffling.

We recognize that this is a simplified model. In reality, a firm engaging in such a strategy would face some uncertainty about all aspects of supply and demand. Modeling the firm's uncertainty and how that changes incentives would be a very complex task that we leave for future research.

To examine the impact of withholding we look at a range of potential allowances that could be withdrawn from the market and calculate the probability that such a removal would elevate the price into the APCR. This analysis is summarized in figure 11 for 2014 and in figures 12 and 13 for 2017. The 2017 results are presented for each of the three vehicle intensity scenarios that were utilized in previous results and summarized in Table 4.[69]

Since transportation is not under the cap during phase I, assumptions about transportation intensities are not required for the 2014 results. As can be seen from figure 11, assumptions about abatement, which are dominated by assumptions about reshuffling, are critical to the analysis. The largest firms could significantly increase the probability of pushing the market into the APCR by withholding extra quantities that they would be eligible to purchase after January 1, 2015 under current rules. Low abatement or reshuffling would greatly increase this opportunity compared to the medium scenario.

For 2017, at low levels of abatement, there is about a 19% chance of reaching the APCR with no withholding from the market, as shown in figure 12. As the

[69]The three possible vehicle emissions intensities include a business as usual trajectory resulting from the VAR estimates, a value based upon the EMFAC model assuming compliance with auto standards but fuels with 10% alternative fuels ("High EMFAC") and a case where auto standards are met as are EMFAC's projections of expanded alternative transportation fuel fuels due to the LCFS ("Low EMFAC").

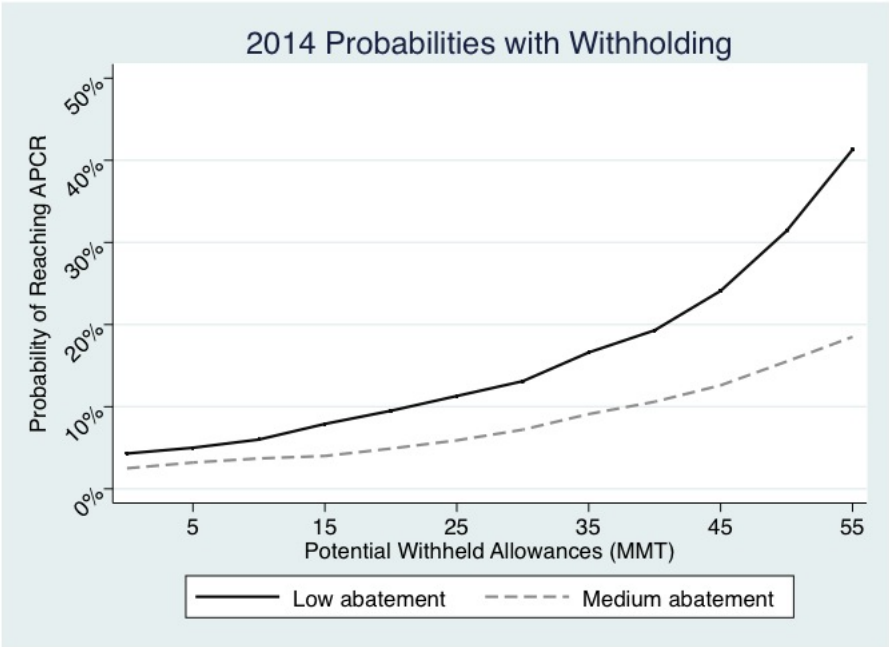

Figure 11. Allowance Price Probabilities By Scenario

level of withholding rises, the probability of prices reaching the APCR rise as well. If 30 MMT of allowances are banked past the phase II period, the probability of the phase II price reaching the APCR under low abatement rises to as high as 35%. If we assume medium amounts of reshuffling the probabilities are quite a bit lower but non-trivial, and the impact of withholding is also less dramatic, as shown in figure 13. If 30 MMT of 2013-13 allowances are banked under a medium withholding scenario, probabilities of reaching the APCR range from about 9% to 16%, depending on the degree to which emissions intensity of transportation decline.

Withholding has a large incremental effect in 2014, because the possible single-firm long positions (up to 58 MMT) constitute a larger share of the total cap for this period (330 MMT). Thus a firm that withholds about 30 MMT from the market would remove nearly 10% of all available allowances during the first

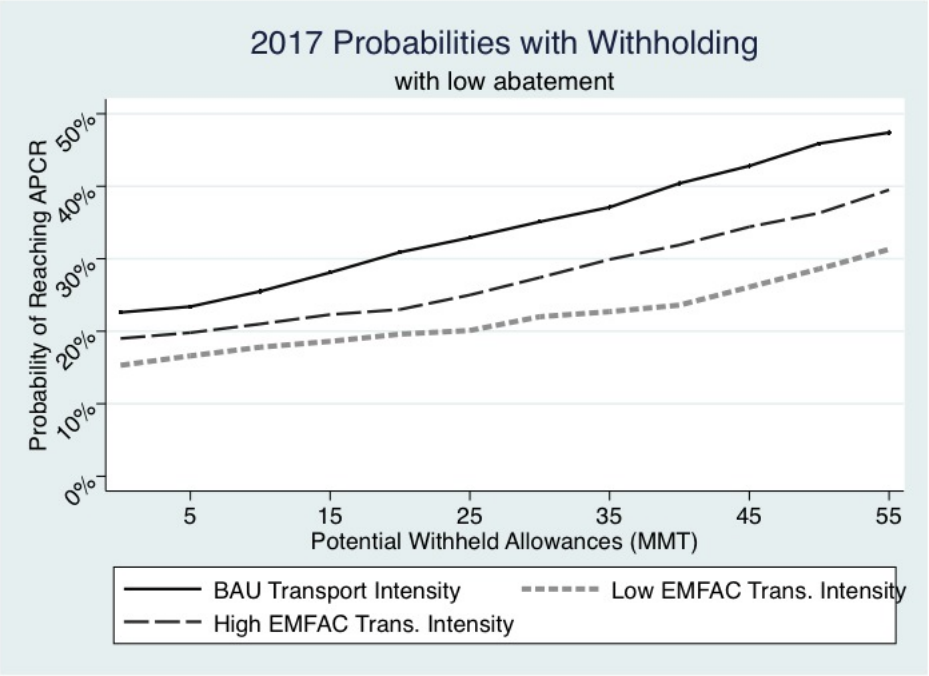

Figure 12. Allowance Price Probabilities By Scenario

compliance period. In a low abatement scenario, 30 MMT of withholding can nearly triple the probability of reaching the APCR up to around 13% from slightly more than 4% with no withholding.

The probabilities for discrete levels of withholding are summarized in Table 9. The upper panel lists the probabilities assuming low levels of abatement and reshuffling, while the lower panel gives probabilities assuming medium levels of abatement and reshuffling. In 2017, overall probabilities of reaching the APCR are higher for both reshuffling cases, but the impact of different levels of withholding are less dramatic. This table highlights the fact that even without any withholding of allowances, there is a non-trivial chance of reaching into the APCR by 2017, although smaller than those we forecast for 2020.

In closing this section, we note that preventing and detecting withholding is extremely challenging because many actions that look like prudent allowance

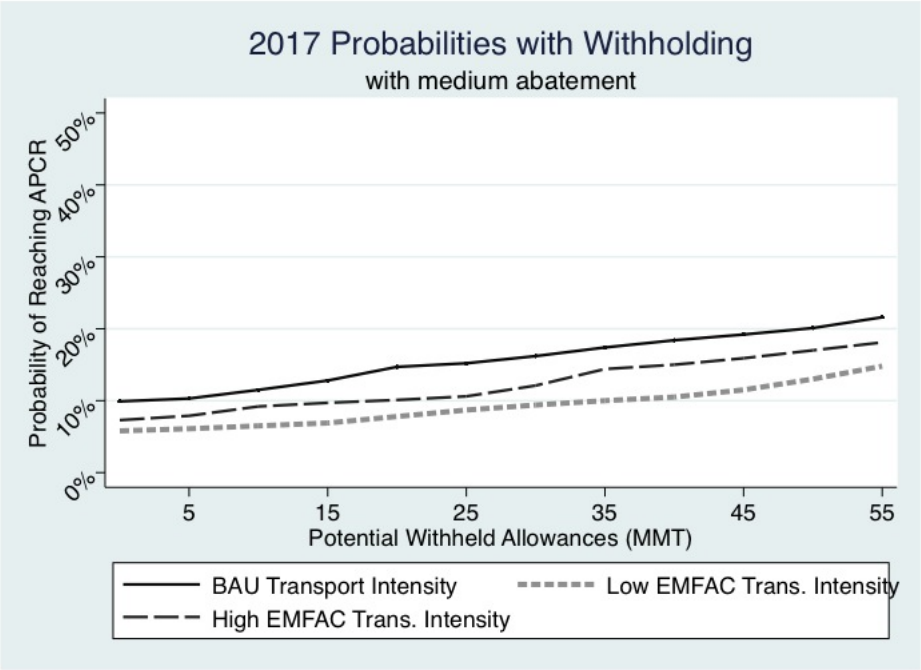

Figure 13. Allowance Price Probabilities By Scenario

purchases strategies can also be employed as part of a profitable withholding strategy.

C. *Implications of a withholding strategy*

While the risk of a profitable withholding opportunity is concerning, the impact of such a strategy would be narrower than one might at first think. To the extent that a withholding strategy would be a surprise to the market that occurs near or after the end of a compliance period, the impact of the strategy on allowance price would have little or no impact on retail prices. This is because those incurring the compliance obligation would not recognize the higher allowance price as the marginal cost of emissions until near the end of the compliance period – in which case the higher cost would impact retail prices for only a short time – or

Table 9—Probabilities of Reaching the APCR

With Low Abatement

Year	2014	2017		
Amount Withheld (MMT)	Trend Intensity (probability)	Trend Intensity (probability)	Emfac low Intensity (probability)	Emfac high Intensity (probability)
5	5%	23%	17%	20%
10	6%	26%	18%	21%
15	8%	28%	19%	22%
20	10%	31%	20%	23%
25	11%	33%	20%	25%
30	13%	35%	22%	27%
35	17%	37%	23%	30%
40	19%	40%	24%	32%
45	24%	43%	26%	34%
50	31%	46%	29%	36%
55	41%	47%	31%	40%

With Medium Abatement

Year	2014	2017		
Amount Withheld (MMT)	Trend Intensity (probability)	Trend Intensity (probability)	Emfac low Intensity (probability)	Emfac high Intensity (probability)
5	3%	10%	6%	8%
10	4%	12%	7%	9%
15	4%	13%	7%	10%
20	5%	15%	8%	10%
25	6%	15%	9%	11%
30	7%	16%	9%	12%
35	9%	17%	10%	14%
40	11%	18%	11%	15%
45	13%	19%	12%	16%
50	16%	20%	13%	17%
55	19%	22%	15%	18%

after the compliance period has ended – in which case the higher allowance cost for the then-*previous* compliance period would not be relevant for going-forward production costs.

For example, consider the possibility that a firm with a large limited exemption were to quietly buy up 2013 and 2014 vintage allowances at the end of 2014 and, especially, after January 1, 2015, when its limited exemption grows significantly due to the expansion of the program to fuels and end-use natural gas. This could result in a price spike for 2013 and 2014 vintages, but after 1/1/15 the price of those earlier vintage allowances would be irrelevant to production costs of gasoline or any other emissions source if they were above the price of 2015 allowances. Emissions after 1/1/15 – whether from tailpipes, electricity generation, or any other source – would incur a compliance cost equal to the least expensive vintage allowance that could fulfill the obligation. Thus, a jump in the price of 2013 and 2014 vintage allowances after 1/1/15 that is caused by withholding those vintages from firms that need them for the first compliance period would not impact the cost of emissions that occur after 1/1/15. If the withholding became evident and caused an allowance price spike before 12/31/14, then the cost would be reflected in retail prices until the end of that year (and would potentially cause some abatement). But the largest risk of profitable withholding is from the expanded limited exemptions and increase in the holding limit that occurs on 1/1/15.

This is not to say that a withholding strategy couldn't still cause significant wealth transfers, but that the transfers are likely to be among those with compliance obligation and others that trade allowances, not from retail customers. Even these transfers, however, would be only for the allowances transacted after the price jumps, which would likely be after the compliance period ends. A firm with a compliance obligation could avoid the risk of being a victim of such manipulation by making sure that they cover all or most of their allowance needs as their obligation grows during the compliance period rather than waiting until near (or after) the end of the compliance period to make significant purchases. Still, such a

sudden price spike – even if it were at the end of a compliance period, affected few allowances, and did not change retail prices for gasoline or other goods – would likely still create political concerns about the fairness of the wealth transfers and the reliability of the market.

It is worth pointing out that this limited impact of a price spike due to manipulation is in contrast to the potential impact we estimate from a real supply/demand imbalance that – even without manipulation – leads to high allowance prices. Such a real scarcity price (in contrast to the artificial scarcity from manipulation) could occur much earlier in the compliance period and would thus impact retail prices for a longer period. For the same reason, the impact of real scarcity on price is likely to affect more allowance transactions, lead to larger wealth transfers, and have a greater impact on retail prices.

D. Multi-Firm Withholding

The analysis presented here considers only the potential impacts of unilateral withholding by a single firm. This analysis is directly analogous to the *single pivotal supplier* test that is applied in the context of electricity market monitoring. In essence we calculate the probabilities of the largest firm being pivotal in one of the multi-year compliance phases. Note that the single pivotal supplier standard is one of the the most extreme forms of market power, where a single firm is able to unilaterally raise prices. In many cases a broader standard is applied as justification for market power mitigation in wholesale electricity markets. For example, the California ISO uses a three-pivotal-supplier standard for its application of local market power mitigation.

Consider a simple example with no uncertainty. Assume that all firms in the market knew that BAU emissions would be 60 MMT below the point where the supply jumps vertically from the price floor to the APCR. Additionally assume that there are two firms that each has the ability to withhold 40 MMT, partially in their compliance accounts and partially in their holding accounts. If one firm

allowed information to become known that it had banked 31 MMT of 2013-14 allowances in its compliance account, beyond its compliance obligation, then the second firm's best response would likely be to withhold an additional 29 MMT in order to push the price from the floor to the APCR. The feasibility of such multi-firm strategies depends on the ability of the actors to acquire excess allowances without alerting other market players and driving up the price prematurely. If firm 1 was considering the potential impacts of withholding allowances, their strategy would also depend upon their expectations of other firms increasing allowance sales if firm 1 withheld them. Firms can credibly signal to others that they will not increase sales of allowances if their allowances are held in compliance accounts which, under current policy, do not allow resale.

Economic models that represent a multi-firm pivotal supplier situation usually result in multiple equilibria, particularly when the demand is perfectly inelastic as would be the case for allowances after the compliance phase ends. While economic simulations do not offer a clean single prediction to these cases, the risks of multi-firm market power are recognized in wholesale electricity markets through mechanisms such as 2-firm and 3-firm pivotal supplier tests. A multi-firm pivotal supplier test would combine the potential size of the 2 or 3 largest firms and test for the probabilities that they would jointly be pivotal. For large compliance entities, two or three firms could combine to remove nearly 100 mmTon from the market position. Even if long positions were determined only by the holding account balance, during later phases a 3-firm pivotal supplier test would consider the possibility of over 39 MMT of allowances being removed from the market. In this sense our estimates here can be considered conservative estimates of the potential for withholding to represent a profitable strategy.

E. Other factors that affect the probabilities

The analysis we have presented omits two factors that could affect the probability of high prices at the end of the first or second compliance periods.

As mentioned earlier, our analysis ignores the role of the recently-announced linkage of California's market with that of Quebec. Some press and industry analysis suggests that Quebec will be a net buyer of allowances from California.[70] If that is the case, then our analysis understates the probabilities that the competitive supply and demand could result in the allowance prices that trigger access to the APCR. For small net purchases from Quebec, our analysis also understates the probability that strategic withholding could lead to prices in the APCR when the competitive price would be much lower. If Quebec is a large net buyer, however, then the increased probability of competitive supply/demand prices in the APCR reduces the additional impact withholding could have. That is, if under most BAU emissions outcomes the competitive price would be in the APCR, then the incremental addition to that probability from withholding is smaller. Quantifying these effects would require an analysis of Quebec's supply and demand similar to the analysis we have done here. We have not been able to extend the analysis to include Quebec, because we do not have access to comparable data to what we have for the California market to incorporate Quebec into our model of the distribution of future GHG emissions.

In addition, a perception of excess 2013 and 2014 allowances could lead a firm to procure extra allowances with the intent of banking some of them in its compliance account (once its holding account is full) with no intention of market manipulation. Nonetheless, because those allowances would then no longer be tradable, such prudent-appearing banking could create a situation in which a pivotal strategic firm would have to withhold far fewer allowances than the overall supply/demand balance would suggest in order to drive price to the APCR.[71]

[70] To date, allowance auctions held by Quebec have not sold all of the allowances on offer. Given the very early stages of the market, it is difficult to interpret the implications of this.

[71] In the most extreme case, if many firms banked allowances in their compliance account, an accidental price spike to the APCR could occur because the supply of believed-to-be-excess allowances has been deposited with no ability to withdraw and sell to other market participants who need them for 2013-14 compliance.

VIII. RECOMMENDATIONS FOR POLICY PROPOSALS

In this paper, we have attempted to analyze the impact on the cap and trade market of California's greenhouse gas policies as they are currently written. A number of proposals, however, are under consideration to modify the policies. We discuss some of these and also recommend some other options for reducing the risk of a transitory price-spike during a given phase, due to either withholding, physical short-term market shocks, or lags in information about key factors such as emissions associated with imported electricity. Based upon our analysis we make the following recommendations.

1) *Reinforce the Price Containment Reserve*

   One issue that has a strong influence on the overall performance of the market is the policy regarding the allowance price containment reserve (APCR).

   The APCR was established to help to mitigate undue volatility in allowance prices. It is accompanied by an associated soft floor price that will be enforced in the allowance auctions. We strongly support the role of the price containment reserve, but we believe that it should be strengthened further to reduce uncertainty about the likely policy response if the price rises to the highest tier of the APCR and all available allowances are exhausted. The ARB should stand ready to expand the pool of allowances in order to maintain the price of allowances at or below the highest price of the APCR. This can accomplished by either borrowing allowances from a post-2020 compliance period or purchasing equivalent compliance instruments from other GHG emissions allowance markets such as the RGGI or EU-ETS.

   The changes that the Board recently approved permit allowances from later years (of the 2013 to 2020 program) to be shifted to earlier years if the price rises to a sufficiently high level.[72] This is a useful response to the concern

[72] See http://www.arb.ca.gov/regact/2013/capandtrade13/capandtrade15dayattach1.pdf

that the first compliance period (2013-14) could have a shortage of supply at the highest price of the APCR. The Board's action, however, does not address the threat that there could be a supply/demand mismatch for the entire 8-year program. If market participants thought that there were not enough allowances over the 8-year period to cover the cumulative emissions under the cap, then the price of allowances for all remaining periods would continue to rise. If that happened, moving allowances from one year to another is akin to trying to raise a bathtub level by taking water from one end of the tub and pouring it into the other.

Our results in section VI directly address this question by examining the possibility that an 8-year trajectory of emissions might exceed the 8-year budget of allowances. Under the low abatement scenarios, these probabilities reach 10-15%. Recall that the largest difference between our low and high abatement scenarios is the degree of reshuffling, over which very little is known at this point. If a reinforced reserve is established along the lines described above, this probability drops to zero.

Left unchanged, current regulations suggest that if the demand for allowances exceeds supply at the highest price of the APCR, the allowance price would be allowed to rise to any level that is necessary to ratchet down allowance demand to meet the capped supply. However, we believe that it is highly unlikely that the political and regulatory process would allow the market to continue to operate freely at very high allowance prices, such as above the highest tier of the APCR. Such an intervention in the California GHG market is currently not well defined in the regulation and would almost certainly be more disruptive when taken under duress. A far superior solution is to have a transparent and credible process for limiting allowance prices established in advance rather than rely on ad hoc emergency measures during a period of crisis.

A strong defense of the price containment reserve is crucial to enhancing the

integrity of the market for several reasons. First, some negative shock to the market, such as severe drought or long-lasting power plant outages, could cause a short-run disruption to the market. If this shock happens late in the process, say 2019, there may not be time for the market to recover without a sharp increase in allowance prices that otherwise could have been borrowed from a post-2020 reserve. The result would be allowance prices at levels that could negatively impact both the California economy and the integrity of the cap-and-trade program. Second, if allowance prices truly had no ceiling, speculative trading would likely account for these low-probability, high-price events, which would raise average prices overall. Third, the possibility of extremely high allowance prices raises the potential rewards to a strategy to manipulate the market, and therefore raises the risk of manipulation. If allowance prices had a credible and transparent (to market participants) ceiling, some costly actions necessary to attempt to manipulate the market would likely not be worth trying in the first place. The price ceiling therefore plays an important role as a deterrent to detrimental trading behavior. Such deterrence is only effective if it is credibly and transparently established in advance.

While there may be some concern that the expansion of the allowance price containment reserve would harm the environmental integrity of the program, we believe that, by strengthening confidence in the allowance market, it will ultimately enhance the integrity of California's cap-and-trade system.

2) *Allow Conversion of Allowance Vintages*

Currently, market participants are not allowed to apply allowances from future vintages to compliance in previous phases. For example a vintage 2015 allowance cannot be used for compliance with phase I obligations. This boundary between phases creates the prospect of transitional shortages in which allowance prices in the expiring phase rise to the APCR while

current vintage allowance remain inexpensive. As the analysis in section VII.B demonstrates, the potential for withholding substantially increases the probabilities of such an outcome, up to as high as 30%-40%.

A second concern with the current design is the potential that allowances could end up inefficiently owned, ex-post. As firms acquire allowances according to their expectations of needs, shocks to individual firms or even sectors could result in too few allowances from a current phase being available to some sectors while others hold a surplus they are unable to sell. While we do not specifically model these probabilities, we do observe in our results that many draws of our simulation that yield high total emissions also feature low emissions from some sectors coinciding with high emissions from other sectors. In our simulation results there is little contemporaneous correlation between the year-to-year growth rates in emissions in the transportation sector and those of the natural gas, industrial, and other sectors. We therefore believe that it is plausible that firms, if they purchase allowances to cover their expected emissions, could find themselves with too many or too few allowances at the end of a compliance period.

Permitting conversion of allowance vintages would greatly reduce the risk and consequence of both problems. Under this proposal, for example, firms would be allowed to purchase 2015 allowances during 2015 and apply those allowances to their phase I obligations, which must be settled by November 2015. This option would bound the extent to which prices in the expiring phase could rise above current vintage allowance prices. This would greatly reduce the incentive to attempt to raise prices in the expiring phase by withholding allowances from that phase. The proposal would also help to address accidental over-compliance by some participants creating a shortage at the end of a compliance period. While firms would still be unable to buy allowances held in other firm's compliance accounts, they would at least be able to purchase future vintages (possibly at some premium) to meet their

needs. Firms with excess allowances would simply be able to defer further purchases.

In order to ensure that the decision to convert allowance vintages is not taken lightly, we recommend some cost be associated with this action. The conversion cost could take one of two forms, either a surrender ratio for out-of-phase allowances that is greater than one, or a conversion "fee." With the surrender ratio, firms would have to surrender a larger number of later vintage allowances in order to apply them to an expiring phase. For example a firm that needs 100 MMT of phase I (2013-14) allowances to meet its obligations, could apply 125 MMT of 2015 allowances as an alternative. This essentially imposes 25% charge on the shifting of vintages.[73] Another alternative would be a conversion fee, for example $2.50 or $5.00 per allowance, that would be applied to each converted allowance.[74] Firms would only avail themselves of this option if allowance prices in the expiring phase rises above the price in the current vintage by an amount greater than the conversion fee.

The Board has recently approved a related change that allows for limited shifting of future vintages to current periods, but only at the highest price of the APCR. While that proposal greatly reduces the probability of a transient price spike above the APCR level for earlier vintage allowances at the time of surrender, it would not affect price outcomes up to that highest APCR level. For example, Phase I prices (2013 and 2014 vintages) could reach $50/ton, while Phase II allowance prices remained around $12/ton. Such an outcome could arise through either strategic withholding or accidental over-allocation of extra Phase I allowances into the compliance accounts of some entities. Conversion at, for example a fee of $5.00/ton would keep

[73]There may be concern that this option might reduce the overall number of allowances. If this is a concern, the extra allowances associated with the conversion charge could be returned to circulation either through a subsequent auction or as an addition to the APCR.

[74]Setting a dollar-denominated fee, rather than requiring a higher number of allowances per ton, would mean that the cost of conversion would not change as the price of allowances changes.

Phase I prices from rising above \$12.00 + \$5.00 = \$17.00 in this example, and thereby also greatly reduce the potential payoff to withholding, assuming that there are plenty of Phase II allowances available relative to the expected Phase II compliance obligations of all market participants.

We see vintage conversion as a relatively straightforward policy shift that could address the problem of transient price spikes without impacting the market mechanism at any other times. If the ARB cannot adopt this change, however, we believe there is a suite of other policy changes that, while not as straightforward to implement, could accomplish many of the same goals.

3) *Modify the limited exemption.*

There are two changes to the limited exemption that would greatly reduce the ability of certain firms to obtain large long positions. One concerns the vintages covered by the limited exemption to holding limits and the other change concerns the treatment of surrendered allowances. Specifically, we recommend that firms be allowed to purchase only allowances from 2015 and later years when the limited exemption is expanded at the end of the first compliance period in January 2015, and only be allowed to purchase allowances from 2018 and beyond when the limited exemption is expanded at the end of the second compliance period in October 2017. For example, starting January 1, 2015 a firm would only be allowed to count 2015 or later allowances against its new compliance account headroom. This restriction can be removed after the allowance surrender date for the previous compliance phase. Once the market for Phase I valid allowances has settled, there is no further concern that a firm that is long in Phase I allowances could raise prices in the Phase I market. For example, after November 2015, when Phase I obligations have been settled, firms could be allowed to own vintage 2013 and 2014 allowances and have them treated exactly as Phase II allowances in the limited exemption calculation.

This adjustment helps to reduce the largest risk for a price-shock by preventing large amounts of expiring-phase allowances from being sequestered in future-year compliance accounts. At the same time, this rule would allow firms the ability to purchase allowances equivalent to their ongoing obligations as long as the vintage of the allowances coincides with that of the obligation.

Another issue with the limited exemption is that it currently does not adjust for the mandatory surrender of allowances to match at least 30% of annual emissions. As a result, firms can acquire positions in excess of their emissions obligations that would reach 60% of annual emissions during the phase II period. We recommend that the limited exemption decline by an amount equal to the amount surrendered by a compliance entity during a given compliance phase.

One might be concerned that large compliance entities can still acquire large transitory "long" positions, but with the exception of the compliance phase transitions, these long positions are temporary and offset by pending future obligations. Thus a long position in 2015 vintage allowances acquired in January of 2015 would be offset by expectations of the need for those allowances during the next three years. However a purchase of 2014 vintage allowances in January of 2015 could constitute a pure long position in the phase I compliance period.

As with recommendation #2, this proposal would address the risk of transitory spikes in allowance prices at the end of compliance periods. This is the problem simulated in section VII.B. This proposal would reduce the quantity that any given firm can withhold from the market during these transition periods. Instead of the risk of 30 MMT up to 50 MMT of expiring vintage allowances being shifted into a future compliance period, the amount that could be shifted would be reduced to around 10 MMT at the end of phase I.

4) *Modify the Compliance and Holding Accounts.*

While recommendation #3 helps address one of the ways in which firms can acquire very long positions. It does not eliminate the risk of transitory price spikes nor does it address other potential problems such as the ex-post misallocation of allowance holdings. We believe recommendation #2 would address both those risks effectively. As a potential substitute for recommendation #2, the risk of transitory price spikes – whether strategic or not – could be mitigated through a collection of other changes to compliance and holding accounts. If recommendation #2 were adopted, we do not believe the following changes would be necessary.

4a. Allow resale of allowances from compliance accounts. Current policy has divided allowance holdings into two categories, those held in holding accounts that are freely exchangeable and those held in compliance accounts, that are not tradable. The intention of this policy was to allow firms with large obligations to own large amounts of allowances, but to discourage their attempts at withholding allowances by limiting how many those firms can sell. However, current holding limit levels, soon to exceed 10 million allowances, provides a large scope for allowance sales despite this policy.

At the same time, the compliance account design could inadvertently facilitate withholding by providing a means for firms to credibly withdraw allowances permanently from the market. If several firms were holding extra allowances and prices begin to rise, each would be tempted to take advantage of that price increase by selling into the market, driving down prices. However, if these firms can credibly demonstrate to others that they will *not* sell all of their allowance holdings, this can increase price expectations and make the withholding strategy more effective. Putting allowances into a compliance account does just that.

A second concern with the current design is the potential that allowances

could end up inefficiently owned, ex-post. As firms acquire allowances according to their expectations of needs, shocks to individual firms or even sectors could result in too few allowances being available to some sectors while others hold a surplus they are unable to sell. For example, if one industry (e.g. refining) experienced a surge in demand, while another (e.g. electricity) had unexpectedly low compliance needs, then the fact that many allowances may be "stuck" in the compliance accounts of electricity firms can be inefficient and raise compliance costs. Similar problems could emerge if one firm had a negative shock (say a refinery outage or nuclear plant retirement) that raised or lowered their compliance obligation relative to other firms in the industry. Again, this is only an issue if firms have pre-funded their expected obligations in compliance accounts. Then they would have less ability to adjust to these new conditions.

The manipulative scheme we describe above could be prevented by relaxing the compliance account "one way" restriction. In our view, the key issue is not whether a firm owns a large quantity of allowances but whether it has a large long position in excess of its compliance needs. We therefore suggest modifying the focus of the current framework to allow for additional flexibility for compliance account transactions while *at the same time* further limiting the ability of firms to accumulate substantial long positions. The problems this proposal addresses are not explicitly simulated in this report. Our analysis is at the sector level so we do not predict the emissions of specific firms. Further very strong assumptions about the purchasing strategies of firms would be necessary to predict whether one firm over or under procures as a result of surprisingly strong or weak emissions.

4b. Reduce the holding account holding limits. The limits for holding accounts define the lower bound on how long firms can be in the allowance market. Current restrictions on compliance accounts contributed to the belief that firms needed relatively large holding limits to be able to respond

to fluctuations in their own emissions. However, additional flexibility with compliance accounts could be balanced with a reduction in the holding account limit. While we see a need for such flexibility, we note that 13 MMT exceeds the annual obligation of all but 5 firms in the market. We recommend that a smaller holding limit be adopted as an accompaniment to a relaxation in compliance account restrictions. We propose that the current limit of 6.4 MMT be maintained into later phases rather than expanded as would happen under current policy.

The combination of the recommendations #3 and #4 – modifying the limited exemption, relaxing restrictions on trading from compliance accounts, and tightening holding limits on holding accounts – would greatly reduce the risk of transitory price spikes due to misallocation of allowances across firms, whether strategic or not. An equivalent outcome could be achieved by eliminating the distinction between the two accounts and simply applying the current limited exemption rule to the overall holding account totals. In this way, larger compliance entities would continue to be able to purchase allowances on a scale commensurate with but not more than 6.4 MMT in excess of their obligations, while smaller and non-compliance entities would similarly be limited to no more than a 6.4 MMT long position.

Unlike recommendation #2, however, the collection of changes in this recommendation would not avoid transitory price spikes in the case of actual high emissions during an earlier compliance period. If, for instance, emissions were much higher than expected during 2013-14, but were expected to drop significantly in later years, the price of allowances for 2013-14 would still rise very significantly for 2013-14 compliance while the price of later-vintage allowances would remain low.

We prefer recommendation #2 in part because it is difficult to see how a transitory price spikes – known by the market participants to be transitory

– would trigger efficient investments in abatement, as the cap and trade mechanism was intended to do.

5) *Auction Frequency and Emissions Information Disclosure*

Our conclusion that the abatement supply curve is likely to be flat over a wide range and then fairly steeply upward sloping has potentially important implications for the impact that new information might have on the market. If at some point in the program the supply and demand balance is thought to be close to exhausting the supply of cheap (or required by complementary policies) abatement and offsets, then small changes in beliefs about abatement demand or supply could cause very large price movements. Some parties have suggested that more frequent auctions be held and/or that market or sector level emissions information be released more often. Both of these changes could potentially reduce large information shocks to the market and thus help to mitigate the price volatility that could result from a steep abatement supply curve.

The modeling approach taken in this report assumes all market participants are well informed about the supply-demand balance of the market, which in reality might not be the case. In our withholding analysis we do assume that one firm could acquire substantial allowances without significantly increasing the price. The ability to do this would be reduced by greater transparency to the market. Our methodology cannot quantitively estimate the impacts of transparency, but the long history of regulatory experience with commodity markets confirms that transparency is an effective tool in combating potential manipulation.

Frequent public auctions would result in timely and transparent allowance price information available to all market participants at no cost. Current policy is for quarterly auctions. Whether more frequent auctions would reduce volatility is in part a function of how quickly new information about

the supply/demand balance becomes available. One potential concern about frequent auctions is liquidity. To increase participation, as well as reduce transaction costs for market participants, it could be helpful to conduct two-sided auctions in which participants are permitted to submit sell offers as well as purchase bids for allowances.

Even with frequent auctions that result in transparent prices, there is concern that some confidential information may become available only intermittently and could cause volatility. Measures of industry activity may allow analysts to predict emissions fairly well from many sectors. One area, however, where this is less certainty is electricity imports. Emissions from electricity imports will depend on the declared source of the power, and could vary from zero for renewable sources, to the default emissions rate for unspecified sources, to more than twice the default rate for power from a coal-fired plant. There seems to be a real potential for the annual ARB GHG emissions inventory reports to have substantial impact on prices.

This problem is exacerbated by the timing of such reports, generally planned to cover a calendar year and be released in November of the following year. This means, for instance, that the first ARB report on emissions during the program period will be in November 2014, covering 2013. No more information will be released before December 31, 2014 when the first compliance period ends. At that point, all demand elasticity and virtually all supply elasticity for the first compliance period disappears with the exception of elasticity from allowances that were to be banked for future compliance periods. More frequent information releases would seem to have substantial value, particularly if they were done with a shorter lag from the time period covered. We believe that timeliness is very important, and that even preliminary data can be useful, even if it is later revised.

There are, of course, very real administrative costs of more frequent auctions and more frequent and timely information releases. These must be weighed

against the potential benefits. The benefits, however, may be substantial, particularly in light of the steep abatement supply curve once low-cost offsets, complimentary policies, and other exogenous emissions reductions are exhausted.

The ARB originally proposed restricting holding account information but providing full firm-level detail on the allowance balances held in compliance accounts. The proposal to make full firm-level detail on compliance account balances is being reconsidered. The ARB is considering sector-level aggregation with quarterly updates. We have been concerned that too high a level of aggregation would harm transparency and make the acquisition of a dominant position less costly, while providing little public benefit.

We also see a real need for timely disclosure to all market participants if one (or more) entities builds up a large long position in the market. Establishing such a long position potentially gives the entity an incentive to withhold allowances from the market and drive up the price while selling a subset of their holdings at the artificially high price, as discussed in section VII. Note that firms with large gross holdings of allowances will not have an incentive to raise prices if the net position of that entity is still "short." In other words, entities with large holdings that are nevertheless smaller than their expected obligation are much less likely to attempt to withhold allowances from the market.

Confidential monitoring by the ARB is not sufficient to curb potential for abuses of a long position. If the ARB found that an entity had acquired a large long position, but not in violation of holding limits, and ARB could not disclose that information, then there is little else that ARB could do. The natural market forces that would take place in response to the information – an upward price adjustment and greater interest on the part of entities with short positions to cover their shortfall – would not occur.

While we see benefits to disclosing the net positions of all parties, stakeholders are concerned that such disclosure could put an individual firm at a disadvantage when attempting to buy (or sell) allowances. We therefore propose a measure that balances the desire to preserve a degree of anonymity of balances but also conveys information about the aggregate net positions of firms. The general idea would be to provide an index of the concentration of net positions in the market. This proposal has been described in detail in the January 2014 memo on the subject by the Emissions Market Assessment Committee.

IX. CONCLUSION

California has now embarked on a plan to reduce greenhouse gas emissions through a cap and trade program. For the program to succeed in California, and as a model for the rest of the world, it is important that the outcomes of the market are reasonable and understandable. In this report, we have modeled supply and demand in the market – including the incentives of the market participants – in order to forecast the range of possible outcomes and the factors that could drive those outcomes.

We have shown that there is significant uncertainty in both the demand and supply in this market. Furthermore, it seems likely that the great majority of available abatement supply will occur independently of the allowance price or at prices near the price floor. As a proportion of the market, our analysis indicates that fairly little additional supply will be forthcoming at prices substantially above the floor, but still below the price that will trigger releases of additional allowances from the allowance price containment reserve (APCR). Combined with the uncertainty in the demand for allowances, this suggests that the market price is unlikely to fall in an intermediate range substantially above the auction reserve price, but still below the level at which allowances from the APCR would be made available. A significant driver of this outcome is the fact that several program

design features that enhance the political viability of the program also steepen the supply curve of abatement at prices between the auction floor and first step of the APCR.

Our analysis also suggests that there is a small, but not insignificant, chance that the demand for emissions allowances over the 8-year program could exceed the available supply after accounting for abatement activity and the supply of emissions offsets. This possibility supports the view expressed by ARB in October 2013 that it is prudent to pursue further policies that would prevent the price from rising to unacceptable levels if demand for emissions allowances turned out to be much stronger than expected.

It is important to note that the scenarios under which the price for emissions could climb very high by 2020 may not produce high prices in 2013. High prices towards the end of the program would result from unexpectedly strong demand and/or low abatement/offset supply over the years 2013-2020. Our analysis suggests that such outcomes are plausible, but are not the most likely outcome. The price of allowances in 2013 reflects the full distribution of potential supply/demand outcomes that could occur over the life of the program. If demand for allowances turned out to be higher than expected over the subsequent years (owing most likely to stronger than expected economic growth in the state) or the supply of abatement/offsets were lower than expected (owing to smaller effects of complementary policies than anticipated, smaller offset supply than anticipated, or other factors) then we would expect that the market price would gradually increase over these years to reflect the increased probability that a shortage of allowances could occur by the end of the program.

We have also examined the probabilities of price spikes at the ends of the first and second compliance periods. Due to recent regulatory changes, it is very unlikely that the APCR could be exhausted in these periods, but we still find significant probabilities that the supply/demand balance could drive price into the range of the APCR, particularly at the end of the second compliance period.

We then examine the incentive of larger market participants to manipulate the market by withholding allowances that are needed by others to meet surrender obligations at the end of these compliance periods. Due to the structure of the compliance process, we find that there are significant probabilities such market manipulation could occur if other participants allow themselves to get behind in purchasing allowances relative to their compliance obligation.

Finally, we propose a collection of possible rule changes that we believe could mitigate the chances of both price spikes due to real scarcity of allowances and price spikes that could result from market manipulation. With some relatively straightforward rule changes, we believe that the risk of very high prices could be greatly reduced, bolstering the reliability of California's market and assuring that it will be seen as a successful model of market-based greenhouse gas reduction.

References

Aldy, Joseph E., Alan J. Krupnick, Richard G. Newell, Ian W.H. Parry, and William A. Pizer. "Designing climate mitigation policy," National Bureau of Economic Research Working Paper #15022, June 2009.

Burtraw, Dallas, Karen Palmer, and Danny Kahn. "A Symmetric Safety Valve." Discussion Paper 09-06. Washington D.C.: Resources for the Future.

Bushnell, James, and Yihsu Chen. "Allocation and leakage in regional cap-and-trade markets for CO2." *Resource and Energy Economics* **34**(4), (2012): 647-668.

Bushnell, James, Yihsu Chen, and Matthew Zaragoza-Watkins. "Downstream regulation of CO2 emissions in California's electricity sector." *Energy Policy* **64**, (2014): 313-323.

California Air Resources Board, "Appendix F to the California Cap on Greenhouse Gas Emissions and Market Based Compliance Mechanisms: Compliance Pathways Analysis." (2010).

California Air Resources Board, "Board Resolution 12-51." (2012).

California Air Resources Board, "California Cap on Greenhouse Gas Emissions and Market Based Compliance Mechanisms." (2013).

California Air Resources Board, "Greenhouse Gas Emissions Inventory Technical Support Document." (2010).

California Air Resources Board, "Initial Statement of Reasons: Amendments to the California Cap on Greenhouse Gas Emissions and Market Based Compliance Mechanisms." (2013).

California Air Resources Board, "Regulation for the Mandatory Reporting of Greenhouse Gas Emissions." (2011).

California Air Resources Board, "Updated Economic Analysis of California's Climate Change Scoping Plan." (2010).

California Energy Commission, "California Installed Capacity and Generation." (2012).

California Energy Commission, "Energy Consumption Database." (2013).

California Public Utilities Commission, "Renewable Portfolio Standard Calculator for the Transmission Planning Process." (2013).

California Public Utilities Commission, "Decision Adopting a Cap and Trade Greenhouse Gas Allowance Revenue Allocation Methodology for the Investor-Owned Electric Utilities." (2012).

Dahl, Carol A., " Measuring global gasoline and diesel price and income elasticities." *Energy Policy* **41**, (2012): 2-13.

Dickey, David A., and Wayne A. Fuller. "Distribution of the estimators for autoregressive time series with a unit root." *Journal of the American Statistical Association* **74**(366a), (1979): 427-431.

Efron, Bradley, and Robert Tibshirani. *An introduction to the bootstrap*. Vol. 57. CRC press, 1993.

Ellerman, A. Denny, and Barbara K. Buchner. "Over-allocation or abatement? A preliminary analysis of the EU ETS based on the 2005-06 emissions data." *Environmental and Resource Economics* **41**(2), (2008): 267-287.

Energy & Environmental Economics. *Investigating a Higher Renewables Portfolio Standard in California*, January 2014.

Engle, Robert F., and Granger, Clive W., "Co-integration error correction: Representation, estimation and testing." *Econometrica* **55**(2), (1987): 251-276.

Engle, Robert F., and Byung Sam Yoo. "Forecasting and testing in co-integrated systems." *Journal of Econometrics* **35**(1), (1987): 143-159.

Fowlie, Meredith. "Updating the allocation of greenhouse gas emissions permits in a federal cap-and-trade program." in *The design and implementation of US climate policy*, Don Fullerton and Catherine Wolfram, editors (2011).

Holland, Stephen and Moore, Michael R. "Market design in cap and trade programs: Permit validity and compliance timing," *Journal of Environmental Economics and Management*, **66**(3), (2013): 671-687.

Hughes, Jonathan E., Christopher R. Knittel and Daniel Sperling, "Evidence of a shift in the short-run price elasticity of gasoline." *The Energy Journal*, **29**(1), (2008): 113-34.

Ito, Koichiro. "Do consumers respond to marginal or average price? Evidence from nonlinear electricity pricing." *American Economic Review*, (Forthcoming).

Johansen, Søren. "Statistical analysis of cointegration vectors." *Journal of Economic Dynamics and Control*, **12**, (1988): 231-254

Johansen, Søren. *Likelihood-based inference in cointegrated vector autoregressive models*. New York: Oxford University Press, (1995).

Kamerschen, David R., and David V. Porter. "The demand for residential, industrial and total electricity, 1973-1998." *Energy Economics*, **26**(1), (2004): 87-100.

Kling, Catherine L., and Jonathan Rubin. "Bankable Permits for the Control of Environmental Pollution." *Journal of Public Economics* 64(1), (1997): 101115.

Pizer, William A. "Combining price and quantity controls to mitigate global climate change." *Journal of Public Economics*, **85**(3), (2002): 409-434.

Stock, James H., and Mark W. Watson. "Vector autoregressions." *The Journal of Economic Perspectives*, **15**(4), (2001): 101-115.

Taylor, Lester D. "The Demand for Electricity: A Survey." *The Bell Journal of Economics*, **6**(1), (1975): 74-110.

U.S. Bureau of Economic Analysis, "Real Gross Domestic Product by State." (2012).

U.S. Energy Information Administration, "California Average Retail Price of Electricity to Ultimate Customers by End-Use Sector." (2013).

U.S. Energy Information Administration, "California Natural Gas Prices." (2013).

U.S. Energy Information Administration, "California Weekly Retail Gasoline and Diesel Prices." (2012).

APPENDIX

Parameter Estimates and Unit Root/Cointegration Tests for VAR

This appendix describes the results of the unit root tests for each of the individual elements of the vector Y_t, the results of the cointegrating rank tests for the vector autoregressive model for Y_t, and presents the parameter estimates of the error correction vector autogressive model that is used to perform our simulations.

The following variable definitions are used throughout this appendix.

ln_twh_p_hydro = Natural logarithm of instate electricity production net of instate hydroelectric generation (terawatt-hours (TWh))]

ln_vmt = Natural logarithm of total vehicle-miles traveled (thousands of miles)

ln_ngother_industrial = Natural logarithm of emissions from non-electricity-generation natural gas combustion and other industrial processes (millions of metric tons (MMT) of GHGs)]

ln_real_gas_price = Natural logarithm of Real Retail Gasoline price ($2011/gallon)

ln_real_gsp = Natural logarithm of Real Gross State Product ($2011)

ln_thermal_intensity = Natural logarithm of Emissions Intensity of In-State Thermal Generation (metric tons/MWh)

ln_transport_intensity = Natural logarithm of Emissions Intensity of Vehicle Miles Traveled (metric tons/thousand miles)

We perform three versions of the unit root test for each element of Y_t and report two test statistics for each hypothesis test. Let Y_{it} equal the ith element of Y_t. The first version of the unit root test, the zero mean version, assumes Y_{it} follows the model,

$$Y_{it} = \alpha Y_{it-1} + \eta_{it} \qquad (ZeroMean)$$

meaning that Y_{it} is assumed to have a zero mean under both the null and alternative hypothesis. The hypothesis test for this model is H: $\alpha = 1$ versus K: $\alpha < 1$. We report two test statistics for this null hypothesis

Variable	Type	$\hat{\rho}$	Pr < $\hat{\rho}$	$\hat{\tau}$	Pr < $\hat{\tau}$
ln_twh_p_hydro	Zero Mean	0.01	0.67	0.44	0.80
	Single Mean	-6.77	0.24	-1.86	0.34
	Trend	-15.04	0.08	-2.26	0.43
ln_vmt	Zero Mean	0.01	0.67	1.45	0.96
	Single Mean	-2.37	0.72	-2.36	0.16
	Trend	0.26	0.99	0.09	0.99
ln_ngother_industrial	Zero Mean	0.00	0.67	-0.07	0.65
	Single Mean	-19.04	0.00	-3.00	0.05
	Trend	-18.56	0.02	-2.87	0.19
ln_real_gas_price	Zero Mean	0.79	0.86	1.27	0.94
	Single Mean	0.01	0.95	0.00	0.95
	Trend	-10.60	0.29	-2.30	0.42
ln_real_gsp	Zero Mean	0.03	0.68	1.30	0.95
	Single Mean	-2.73	0.67	-1.72	0.41
	Trend	-18.46	0.02	-2.01	0.56
ln_thermal_intensity	Zero Mean	0.35	0.76	1.84	0.98
	Single Mean	0.44	0.97	0.27	0.97
	Trend	-5.15	0.77	-1.55	0.78
ln_transport_intensity	Zero Mean	0.01	0.67	1.02	0.91
	Single Mean	1.01	0.98	0.28	0.97
	Trend	-2.35	0.95	-0.62	0.97

Table A-1: Unit Root Test Statistics

$$\hat{\rho} = T(\hat{\alpha} - 1) \qquad \text{and} \qquad \hat{\tau} = \frac{\hat{\alpha} - 1}{SE(\hat{\alpha})}$$

where $\hat{\alpha}$ is the ordinary least squares (OLS) estimate of α and $SE(\hat{\alpha})$ is OLS standard error estimate for $\hat{\alpha}$ from a regression without a constant term and T is the number of observations in the regression. The column labeled "$Pr < \hat{\rho}$" is the probability that a random variable with the asymptotic distribution of the $\hat{\rho}$ under the null hypothesis is less than the value of the statistic in the column labeled "$\hat{\rho}$". The column labeled "$Pr < \hat{\tau}$" is the probability that a random variable with the asymptotic distribution of the $\hat{\tau}$ under the null hypothesis is less than the value of the statistic in the column labeled "$\hat{\tau}$".

The second version of the unit root test assumes a non-zero mean. In this case the assumed model is:

$$Y_{it} = \mu + \alpha Y_{it-1} + \eta_{it} \qquad (SingleMean)$$

where $\mu \neq 0$. The hypothesis test is still H: $\alpha = 1$ versus K: $\alpha < 1$. The two test statistics for this null hypothesis are

$$\hat{\rho} = T(\hat{\alpha} - 1) \qquad \text{and} \qquad \hat{\tau} = \frac{\hat{\alpha} - 1}{SE(\hat{\alpha})}$$

Table A-2: Cointegration Rank Test Using Trace				
H0:	H1:	Eigenvalue	LR(r)	5% Critical Value
Rank=r	Rank>r			
0	0	0.9095	155.59	123.04
1	1	0.8448	105.15	93.92
2	2	0.6599	66.02	68.68
3	3	0.6077	43.37	47.21
4	4	0.4727	23.72	29.38
5	5	0.2937	10.28	15.34
6	6	0.1323	2.98	3.84

Table A-2: Cointegration Rank Test Using Trace

where $\hat{\alpha}$ is the ordinary least squares (OLS) estimate of α and $SE(\hat{\alpha})$ is OLS standard error estimate for $\hat{\alpha}$ from a regression that includes a constant term and T is the number of observations in the regression. The test statistics and probability values are reported in the same manner as for the zero mean version of the test statistic.

The third version of the test assumes that the mean of Y_{it} contains a time trend so that the assumed model is:

$$Y_{it} = \mu + \nu t + \alpha Y_{it-1} + \eta_{it} \qquad (Trend)$$

where $\mu \neq 0$ and $\nu \neq 0$. The hypothesis test is still H: $\alpha = 1$ versus K: $\alpha < 1$. The two test statistics for this null hypothesis are again

$$\hat{\rho} = T(\hat{\alpha} - 1) \qquad \text{and} \qquad \hat{\tau} = \frac{\hat{\alpha} - 1}{SE(\hat{\alpha})}$$

where $\hat{\alpha}$ is the ordinary least squares (OLS) estimate of α and $SE(\hat{\alpha})$ is OLS standard error estimate for $\hat{\alpha}$ from a regression that includes a constant term and a time trend, and T is the number of observations in the regression. The test statistics and probability values are reported in the same manner as for the zero mean version of the test statistic.

For all three versions of the unit root test and two test statistics, there is little evidence against the null hypothesis for all seven elements of the Y_t. In all but a few cases, the probability value is greater than 0.05, which implies no evidence against the null hypothesis for a size 0.05 test of the null hypothesis. Although there are a few instances of probability values less than 0.05, this to be expected even if the null hypothesis is true for all of the series, because the probability of rejecting the null given it is true for a 0.05 size test is 0.05.

Table A-2 presents the results of our cointegrating matrix rank tests. In terms of the notation of our error correction model

$$\Delta Y_t = \mu + \Lambda Y_{t-1} + \epsilon_t \qquad (A-1)$$

where Λ is (7x7) matrix that satisfies the restriction $\Lambda = -\gamma\alpha'$ and γ and α are (7 x r) matrices of rank r. Hypothesis test is H: $Rank(\Lambda) = r$ versus K: $Rank(\Lambda) > r$, where r is less than or equal to 7, the dimension of Y_t. Each row of the table presents the results of Johansen's (1988) likelihood ratio test of the null hypothesis that $Rank(\Lambda) = r$ against the alternative that $Rank(\Lambda) > r$, for a given value of r. Johansen (1995) recommends a multi-step procedure starting from the null hypothesis that $Rank(\Lambda) = r = 0$ and then proceeding with increasing values of r until the null hypothesis is not rejected or all null hypotheses are rejected in order to determine the rank of Λ. Rejecting the null hypothesis for all values of r would imply that the elements of Y_t are not cointegrated.

The column labelled "LR(r) " is Johansen's (1988) likelihood ratio statistic for the cointegrating rank hypothesis test for the value of r on that row of the table. The column labelled "5% Critical Value" is the upper 5th percentile of the asymptotic distribution of the LR statistic under the null hypothesis. The column labelled "Eigenvalue" contains the second largest to smallest eigenvalue of the estimated value of Λ. Let $1 > \hat{\lambda}_1 > \hat{\lambda}_2, ... > \hat{\lambda}_K$ equal the eigenvalues of the maximum likelihood estimate of Λ ordered from largest to smallest. The LR(r) statistic for test H: $Rank(\Lambda) = r$ versus K: $Rank(\Lambda) > r$ is equal to

$$LR(r) = -T \sum_{j=r+1}^{K} ln(1 - \hat{\lambda}_j)$$

Following Johansen's procedure, we find that the null hypothesis is rejected for r = 0 and r = 1, but we do not reject the null hypothesis at a 0.05 level for r = 2 or for any value larger than 2. For this reason, we impose the restriction that rank of Λ is equal to 2 in estimating and simulating from our error correction vector autoregressive model.

Table A-3 presents the results of estimating our error correction vector autoregressive model in the notation in equation (A-1). The prefix "Δ" is equal to $(1 - L)$, which means that the dependent variable in each equation is the first difference of variable that follows. The variable Λ_i_j is the (i,j) element of Λ and μ_j is the jth element of μ.

Equation	Parameter	Estimate	Std Err	Variable
Δ_ln_twh_p_hydro	μ_1	-1.6518	2.0290	1
	Λ_1_1	-0.7333	0.2648	ln_twh_p_hydro(t-1)
	Λ_1_2	0.7430	0.2762	ln_vmt(t-1)
	Λ_1_3	-0.2489	0.0995	ln_ngother_industrial(t-1)
	Λ_1_4	0.4276	0.2695	ln_real_gas_price(t-1)
	Λ_1_5	-0.2619	0.0962	ln_real_gsp(t-1)
	Λ_1_6	0.6383	0.5114	ln_thermal_intensity(t-1)
	Λ_1_7	-0.0706	0.0405	ln_transport_intensity(t-1)
Δ_ln_vmt	μ_2	-0.3245	0.3854	1
	Λ_2_1	0.0481	0.0503	ln_twh_p_hydro(t-1)
	Λ_2_2	-0.0265	0.0525	ln_vmt(t-1)
	Λ_2_3	0.0239	0.0189	ln_ngother_industrial(t-1)
	Λ_2_4	-0.0733	0.0512	ln_real_gas_price(t-1)
	Λ_2_5	0.0194	0.0183	ln_real_gsp(t-1)
	Λ_2_6	-0.1375	0.0971	ln_thermal_intensity(t-1)
	Λ_2_7	0.0110	0.0077	ln_transport_intensity(t-1)
Δ_ln_ngother_industrial	μ_3	-1.3596	0.7855	1
	Λ_3_1	-0.1799	0.1025	ln_twh_p_hydro(t-1)
	Λ_3_2	0.2300	0.1069	ln_vmt(t-1)
	Λ_3_3	-0.0449	0.0385	ln_ngother_industrial(t-1)
	Λ_3_4	0.0078	0.1043	ln_real_gas_price(t-1)
	Λ_3_5	-0.0595	0.0372	ln_real_gsp(t-1)
	Λ_3_6	-0.0485	0.1980	ln_thermal_intensity(t-1)
	Λ_3_7	-0.0036	0.0157	ln_transport_intensity(t-1)
Δ_ln_real_gas_price	μ_4	-9.2245	2.5068	1
	Λ_4_1	0.3356	0.3272	ln_twh_p_hydro(t-1)
	Λ_4_2	0.1628	0.3412	ln_vmt(t-1)
	Λ_4_3	0.2840	0.1229	ln_ngother_industrial(t-1)
	Λ_4_4	-1.2189	0.3329	ln_real_gas_price(t-1)
	Λ_4_5	0.1695	0.1188	ln_real_gsp(t-1)
	Λ_4_6	-2.4524	0.6318	ln_thermal_intensity(t-1)
	Λ_4_7	0.1766	0.0501	ln_transport_intensity(t-1)
Δ_ln_real_gsp	μ_5	-2.9867	0.4974	1
	Λ_5_1	-0.1127	0.0649	ln_twh_p_hydro(t-1)
	Λ_5_2	0.2522	0.0677	ln_vmt(t-1)
	Λ_5_3	0.0084	0.0244	ln_ngother_industrial(t-1)
	Λ_5_4	-0.2151	0.0661	ln_real_gas_price(t-1)
	Λ_5_5	-0.0266	0.0236	ln_real_gsp(t-1)
	Λ_5_6	-0.4948	0.1254	ln_thermal_intensity(t-1)
	Λ_5_7	0.0288	0.0099	ln_transport_intensity(t-1)
Δ_ln_thermal_intensity	μ_6	0.9922	1.0852	1
	Λ_6_1	-0.0808	0.1416	ln_twh_p_hydro(t-1)
	Λ_6_2	0.0223	0.1477	ln_vmt(t-1)
	Λ_6_3	-0.0476	0.0532	ln_ngother_industrial(t-1)
	Λ_6_4	0.1683	0.1441	ln_real_gas_price(t-1)
	Λ_6_5	-0.0348	0.0514	ln_real_gsp(t-1)
	Λ_6_6	0.3263	0.2735	ln_thermal_intensity(t-1)
	Λ_6_7	-0.0249	0.0217	ln_transport_intensity(t-1)
Δ_ln_transport_intensity	μ_7	-0.5789	0.6728	1
	Λ_7_1	-0.0813	0.0878	ln_twh_p_hydro(t-1)
	Λ_7_2	0.1019	0.0916	ln_vmt(t-1)
	Λ_7_3	-0.0210	0.0330	ln_ngother_industrial(t-1)
	Λ_7_4	0.0077	0.0894	ln_real_gas_price(t-1)
	Λ_7_5	-0.0271	0.0319	ln_real_gsp(t-1)
	Λ_7_6	-0.0131	0.1696	ln_thermal_intensity(t-1)
	Λ_7_7	-0.0022	0.0134	ln_transport_intensity(t-1)

Table A-3: Error Correction Vector Autoregression Parameter Estimates

Transportation Emissions

The California data were reportedly constructed by the California Department of Transportation (CalTrans) from a mix of in-road traffic monitors (*e.g.*, from the California Performance Measurement System (PeMS)) and traffic counts conducted by CalTrans. Figure

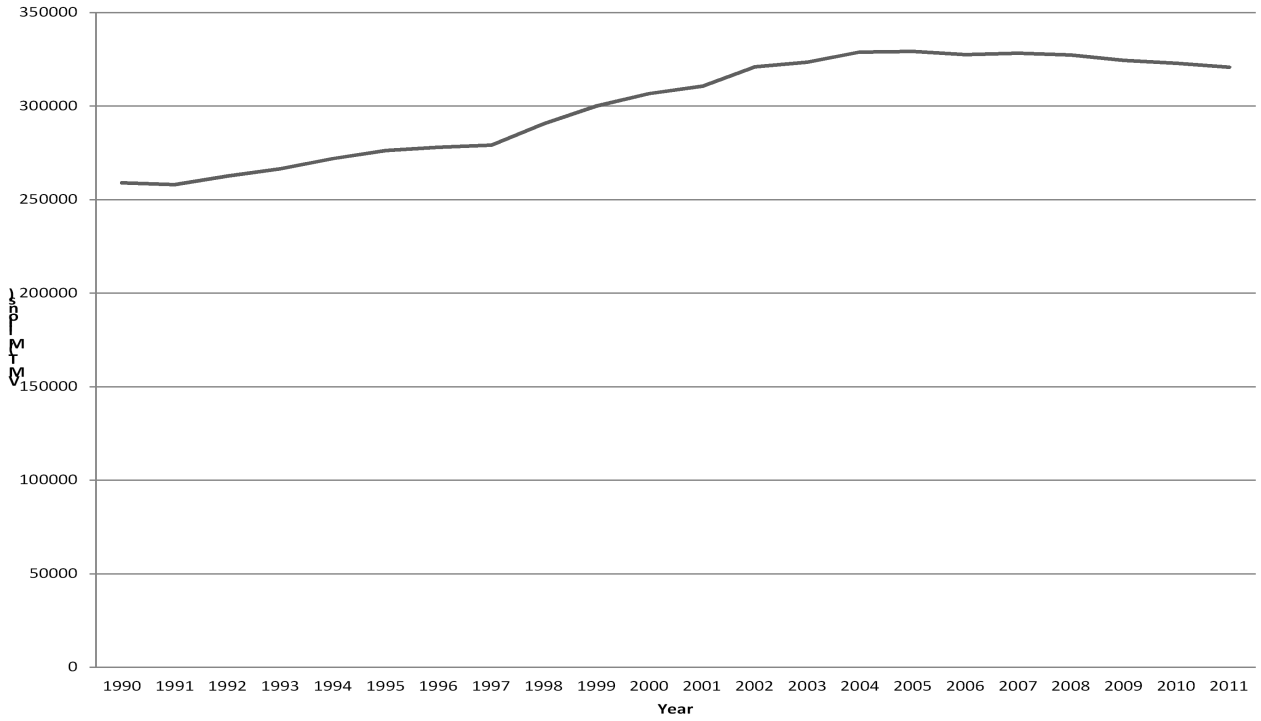

Figure A-1: Annual California On-road VMT 1990-2011

A-1 displays the series of annual California on-road VMT as reported in these surveys.

While these data measure on-road VMT, the cap and trade program caps emissions from all diesel and gasoline combusted as transportation fuel in California, regardless of whether the fuel is combusted on-road or off-road. Therefore, this measure of on-road VMT understates the total VMT covered under the cap and (when carried through our calculations) overstates average emissions factors for on-road VMT. Critically, because certain complementary policies target vehicle emissions factors, an overstated measure of "business-as-usual" emissions factors could lead us to conclude that complementary policies should be expected to achieve a larger impact than might realistically be feasible.

To address this potential source of bias we deviate from ARB's emissions categorization by excluding GHG emissions from off-road vehicle activities from the transportation sector, in favor of categorizing them into "Natural Gas and Other". Therefore, beginning with total transportation sector combustion emissions, we partition emissions into on-road and off-road activities using the more granular activity-based emissions values reported in the combined 1990-2004 and 2000-2011 Emissions Inventories. Table A-4 reports the results of this partition, revealing the contribution of off-road emissions to be small and somewhat

weakly correlated with total transportation sector emissions, ranging from a low of 2.57% in 1993 to a high of 4.52% in 2006, around a mean of 3.55%.

As described above, our approach to forecasting emissions from the transportation sector is to decompose GHG emissions into its VMT component and an average emissions factor per mile of travel. Separating emissions into VMT and an average emissions factor allows us to more accurately capture the underlying drivers of GHG emissions trends and to better model the effects of complementary policies that may cause these emissions drivers to deviate from their preexisting trends. Essentially, our data are derived from the basic identity relating annual GHG emissions to annual VMT and an annual average emissions factor per mile:

$$GHG_t = VMT_t \cdot \bar{E}I_t.$$

To decompose transportation sector GHG emissions into VMT (miles) and an average emissions factor per mile (grams/mile), we take our adapted series of transportation sector GHG emissions (described above) as given, and divide annual GHG emissions by our measure of VMT, the ratio of which is our implied average emissions factor per mile of travel. Table A-5 reports our adjusted transportation sector emissions, OHI VMT, and the calculated average annual emissions factors for on-road activity over the period 1990-2011.

Transportation Complimentary Policies

To incorporate the impact of complimentary policies targeting the transportation sector, we use EMFAC 2011, the ARB's tool for forecasting fleet composition and activity in the transportation sector. The advantage of explicitly modeling on-road vehicle fleet composition and activity is that we can more precisely simulate the impact of complimentary policies that are designed to directly target specific segments of the vehicle fleet. Moreover, because vehicles are long-lived durable goods, it is advantageous for a model to be capable of carrying forward the effects of earlier policies as the composition of the vehicle fleet evolves through time.

EMFAC 2011 is an engineering-based model that can be used to estimate emissions factors for on-road vehicles operating and projected to be operating in California for calendar years 1990-2035. EMFAC2011 uses historical data on fleet composition, emissions factors, VMT, and turnover to forecast future motor vehicle emissions inventories in tons/day for a specific year, month, or season, and as a function of ambient temperature, relative humidity, vehicle

Year	Off-road (MMT)	On-road (MMT)	Share On-road
1990	6.09	137.96	95.77%
1991	6.18	134.45	95.61%
1992	5.15	141.73	96.49%
1993	3.68	139.40	97.43%
1994	4.77	140.42	96.71%
1995	4.97	143.53	96.65%
1996	4.78	145.00	96.81%
1997	4.54	148.31	97.03%
1998	4.23	151.25	97.28%
1999	4.30	155.80	97.31%
2000	5.33	163.48	96.84%
2001	5.54	163.58	96.72%
2002	6.17	169.88	96.49%
2003	6.50	166.35	96.24%
2004	6.95	167.45	96.02%
2005	7.62	167.69	95.66%
2006	7.94	167.65	95.48%
2007	7.40	167.56	95.77%
2008	6.23	157.04	96.18%
2009	5.22	153.28	96.71%
2010	5.40	149.19	96.51%
2011	5.67	146.08	96.26%

Table A-4: On-road and Off-road Transportation Emissions 1990-2011

population, mileage accrual, miles of travel and speeds. Emissions are calculated for forty-two different vehicle classes composed of passenger cars, various types of trucks and buses, motorcycles, and motor homes. The model outputs pollutant emissions for hydrocarbons, carbon monoxide, nitrogen oxides, particulate matter, lead, sulfur oxides, and carbon dioxide. EMFAC 2011 is used to calculate current and future inventories of motor vehicle emissions at the state, air district, air basin, or county level. Accordingly, the model can be used to forecast the effects of air pollution policies and programs at the local or state level.

For our purposes, EMFAC 2011 generates adjusted estimates of average VMT and annual GHG emissions for each on-road vehicle-class by model-year. From the EMFAC2011 outputs, we calculate annual average emissions factors for on-road VMT by taking the ratio of the sum of GHG emissions over the sum of VMT across vehicle-classes and model-years within each calendar year. A known weakness of the EMFAC 2011 model is that it does not accurately reflect the effects of the Great Recession on new light-duty vehicle sales, emissions factors or fleet VMT for the years 2009-present. In terms of new vehicle sales, EMFAC 2011 figures there to have been approximately 30% more new vehicle sales in California in 2009 than were actually recorded by the California Board of Equalization. This difference has declined, approximately linearly, over time as sales of new vehicles have

Year	Emissions (MMT)	EF (kg/mi)	VMT (1,000)
1990	137.96	0.53	258,926,000
1991	134.45	0.52	257,976,000
1992	141.73	0.54	262,548,000
1993	139.40	0.52	266,408,000
1994	140.42	0.52	271,943,000
1995	143.53	0.52	276,371,000
1996	145.00	0.52	278,043,000
1997	148.31	0.53	279,096,000
1998	151.25	0.52	290,630,000
1999	155.80	0.52	300,066,000
2000	163.48	0.53	306,649,000
2001	163.58	0.53	310,575,000
2002	169.88	0.53	320,942,000
2003	166.35	0.51	323,592,000
2004	167.45	0.51	328,917,000
2005	167.69	0.51	329,267,000
2006	167.65	0.51	327,478,000
2007	167.56	0.51	328,312,000
2008	157.04	0.48	327,286,000
2009	153.28	0.47	324,486,000
2010	149.19	0.46	322,849,000
2011	146.08	0.46	320,784,000

Table A-5: On-road Emissions, Emissions Factors, and VMT 1990-2011

slowly rebounded, and are on track to return to pre-recession levels in 2015. Additionally, EMFAC 2011 has VMT growing steadily through the recession, while in reality VMT sharply declined in 2009 and has declined modestly ever since.

To account for these differences we adjust new vehicle sales and total (not per-capita) VMT for model-years 2009-2014. Beginning with a 30% reduction in sales and VMT for model-year 2009, we reduce the adjustments to sales and VMT in each subsequent model-year by five percentage points, so that 2014 is the last model-year impacted by our adjustment. Importantly, as the impact of the Great Recession on the size of each model-year fleet can reasonably be expected to persist over time, these adjustments are imposed across all calendar years 2009-2020. That is, because fewer model-year 2009 vehicles were sold in 2009, there will accordingly be fewer model-year 2009 vehicles in the fleet in future years. While the decline in VMT was almost certainly not purely driven by the decline in new vehicles sales, the reduction in VMT resulting from the sales adjustment causes EMFAC 2011's measure of VMT to closely mimic the actual path of VMT over the same time period. In the absence of better information about the distribution of changes to VMT across model-years, we make this simplifying assumption, noting the goodness of fit.

To account for the impact of complementary policies, we calibrate average emissions factors and emissions intensities of transportation fuel over the period 2012-2020 using our

adjusted EMFAC 2011 model.

To account for CAFE, a policy that proposes to drive the average emissions intensity of new light-duty cars and trucks from 26.5 in 2011 to 54.5 in 2020, we calculate average emissions factors by model-year and vehicle class from the adjusted EMFAC2011 forecasts and force new light-duty vehicles in model-years 2012-2020 to match the fuel-economy standards established by CAFE. We then calculate annual average emissions factors for calendar years 2012-2020, by taking the VMT weighted sum over the set of all model-year by vehicle-class emissions factors.

To account for the LCFS, a policy that proposes to reduce the average carbon content of all on-road vehicle transportation fuel sold in California by an additional 10% between now and 2020, we adjust the emissions intensity of gasoline and diesel according to the incremental share of zero-GHG fuel that must be sold in order to achieve the LCFS. Here it is worth noting an important difference between the cap and trade program and EMFAC 2011 methods of accounting for GHG emissions from biofuels. While the cap and trade program does not assign a compliance obligation to emissions from ethanol, EMFAC 2011 includes combustion emissions from fossil and bio-fuels in the measure of GHG emissions. Therefore, our adjustment of emissions intensity of gasoline and diesel must take into account not only the incremental contribution of the LCFS, but also the preexisting levels of biofuels in California transportation fuel.

We model the full implementation of the LCFS as a linear decline in GHG emissions intensity of on-road gasoline VMT as beginning at 89% in 2012 and falling to 81% in 2020. For diesel, the share of preexisting biofuels is quite small, so we model the decline in GHG emissions intensity of on-road diesel VMT as beginning at 98% in 2012 and falling to 90% in 2020. These declines are taken after the implementation of CAFE, so in practice they are implemented as reductions in the annual average emissions factors calculated above. In light of recent court challenges, we also consider an alternative implementation of LCFS where the regulation is not fully implemented. In this scenario GHG emissions intensity of on-road gasoline VMT is held steady at 89% through 2020 and no penetration of biodiesel is modeled. Table A-6 reports annual average emissions factors and implied average MPG under the combinations of full implementation of CAFE with full and partial implementations of the LCFS. The combined impact of the full implementation of these policies and the preexisting trend in VMT emissions intensity takes average emissions factors from 0.49kg/mi in 2012 down to 0.36kg/mi in 2020.

Unlike our VAR, EMFAC 2011 only provides point estimates for the emissions intensity

	CAFE & Partial LCFS		CAFE & Full LCFS	
Year	EF (kg/mi)	MPG (mi/gal)	EF (kg/mi)	MPG (mi/gal)
2012	0.48	18.36	0.48	18.60
2013	0.48	18.68	0.47	19.04
2014	0.47	19.02	0.46	19.52
2015	0.46	19.51	0.44	20.16
2016	0.44	20.24	0.42	21.07
2017	0.42	21.06	0.40	22.07
2018	0.41	21.91	0.38	23.13
2019	0.39	22.80	0.37	24.25
2020	0.37	23.80	0.35	25.50

Table A-6: Adjusted EMFAC 2011 Average Emissions Factors and MPG 2012-2020

of VMT. We believe that taking the point estimates of VMT intensity from EMFAC 2011 could eliminate an important source of variance in our VAR. To account for the uncertainty in VMT intensity we incorporate the EMFAC 2011 point estimates for each of the adjusted EMFAC 2011 cases into the VAR framework. We treat the impact of complimentary policies as varying with the realization of VMT coming from our VAR. Here, we calculate the annual emission reduction of the complimentary policies targeting the transportation sector as the product of the realized random draw of VMT from our VAR and the difference between mean VTM emission intensity from the VAR and the relevant EMFAC 2011 annual point estimate of VMT emission intensity.