# Slipping Through the Cracks: Detecting Manipulation in Regional Commodity Markets

Reid B. Stevens* and Jeffery Y. Zhang†

December 21, 2016

**Abstract**

Between 2010 and 2014, the regional price of aluminum in the United States (Midwest premium) increased threefold. We argue that the Midwest premium was likely manipulated during this period through the exercise of market power in the aluminum storage market. We first use a difference-in-differences model to show that there was a statistically significant increase of $0.07 per pound in the regional price of aluminum relative to the regional price of a production complement, copper. We then use several instrumental variables to show that this increase was driven by a single financial company's accumulation of an unprecedented level of aluminum inventories in Detroit. Since this scheme targeted the regional price of aluminum, regulators who monitored only spot and futures prices would not have noticed anything peculiar. We therefore present an algorithm for real-time detection of similar manipulation schemes in regional commodity markets. The algorithm confirms the existence of a structural break in the U.S. aluminum market in late 2011. Using the algorithm, regulators could have detected the scheme as early as December 2012, more than six months before it was publicized by an article in *The New York Times*. We also apply the algorithm to another suspected case of regional price manipulation in the European aluminum market and find a similar break in 2011, suggesting the scheme may have been implemented beyond the United States.

**Keywords:** Commodity Markets, Government Policy and Regulation

**JEL:** G1, G18

*Department of Agriculture Economics, Texas A&M University.
E-mail: stevens@tamu.edu

†Department of Economics, Yale University and Harvard Law School.
E-mail: jeffery.zhang@yale.edu

# 1 Introduction

When commodity prices rise, policy makers are quick to blame speculators, despite the lack of academic research supporting this viewpoint (U.S. Senate, 2014). Indeed, most academic investigations into energy (Knittel and Pindyck, 2013; Kilian and Murphy, 2014) and agricultural (Irwin et al., 2009) markets have been unable to find a measurable effect of speculation on commodity prices. Researchers consistently find that speculative inventories are not large enough to explain the commodity price booms between 2000 and 2015.

There are very few cases in which speculators have accumulated sufficiently large inventory positions to *manipulate* commodity prices, most recently, silver in 1980 (Williams, 1995) and soybeans in 1989 (Pirrong, 2004). However, these manipulative schemes were relatively short-lived. Once the manipulation was publicized, commodity exchanges changed rules regarding leverage and hedging to prevent ongoing manipulation, and the groups responsible for the manipulation suffered significant losses. This paper provides the first empirical examination of manipulation in the U.S. aluminum market that took place from 2010 through 2014. During this period, a single bank holding company restricted access to a massive aluminum stockpile, which they controlled, and caused the regional price of aluminum to rise threefold. This episode is notable not just for its duration and profitability, but also because it highlights the risks posed by bank holding company involvement in physical commodity markets.

An obstacle to analyzing the effects of manipulation on physical commodities is defining the term manipulation. It is more than simply speculation, which entails purchasing a commodity with the expectation that the price of the commodity will rise in the future. Speculation is perfectly legal. All investors speculate when deciding whether to purchase or sell an asset. In our case, we focus more narrowly on the extraordinary accumulation of physical inventories, which interferes with the normal operations of commodity markets through an exercise of market power and/or fraud (Fischel and Ross, 1991; Pirrong, 2010). The actors are no longer speculating when they purchase a commodity with the aim of gaining influential market power and subsequently leveraging that monopoly power to inflate prices.

In particular, our analysis begins by examining the causal impact of inventory accumulation in the U.S. aluminum market on the U.S. regional aluminum price, known as the *Midwest premium*. The Midwest premium measures the difference between the transaction prices paid by aluminum market participants and the aluminum cash settlement price on the London Metal Exchange (LME). This premium exists because commodity markets, unlike equity markets, involve physical goods, and physical goods in one regional market cannot be immediately and costlessly dispatched to other regional markets. When purchasers take delivery of a commodity, they must pay for transportation, and possibly storage, as the commodity is moved from an LME warehouse to the purchaser's storage facility. The Midwest premium reflects the cost of transporting aluminum out of LME warehouses as well as the variation in regional supply and demand for aluminum.

The Midwest premium, along with the other regional premiums, provide information about regional markets that is not available in the spot price. This regional aluminum price is constructed using a survey by S&P Global Platts of aluminum bids, offers, and transaction prices for delivery within the month. Similar regional premiums exist for all metals traded on the LME. By construction, these metal premiums only reflect regional determinants of price, since global factors that determine price are differenced out of the premium. Therefore, the causal models we present will only control for determinants of the regional metal price.

Between 2010 and 2014, financial institutions amassed substantial inventories and the regional aluminum price rose (Figure 1). At the peak, warehouses owned by Goldman Sachs in Detroit held over half of the total U.S. aluminum stock. We investigate whether the aluminum inventories stored in Goldman's warehouses caused, or were simply coincident with, the regional price increase.

For most of the twentieth century, financial market regulations prohibited banks from trading physical commodities. Beginning in the 1980s and culminating with the Gramm-Leach-Bliley Act of 1999, these restrictions were gradually repealed, and banks were allowed full access to physical commodities markets. In 2008, large investment banks—including Goldman Sachs, JPMorgan, and Morgan Stanley—formed financial holding companies and drastically increased their operations in physical commodity markets (U.S. Senate, 2014).

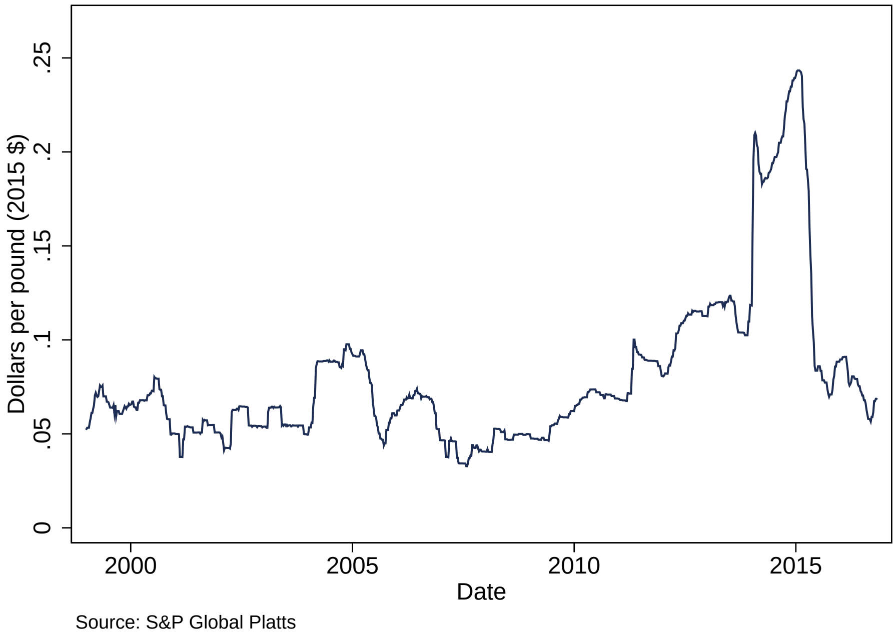



Figure 1: U.S. Aluminum Price (Midwest Premium)

The entry of these financial institutions changed metal markets. Prior to 1999, only minor metal trading took place on financial markets, including the Commodity Exchange (COMEX) and the LME. Less than 1 percent of total aluminum inventories were held through financial markets over this period; almost all were held by producers (USGS, 2014). Over the 15 years following Gramm-Leach-Bliley, total aluminum storage increased (up 80 percent) and the share held through LME and COMEX increased as well (nearly 70 percent was held in LME warehouses alone) (see Figure 2). As the amount of aluminum traded on financial markets grew, these markets began to have a large effect on physical aluminum markets.

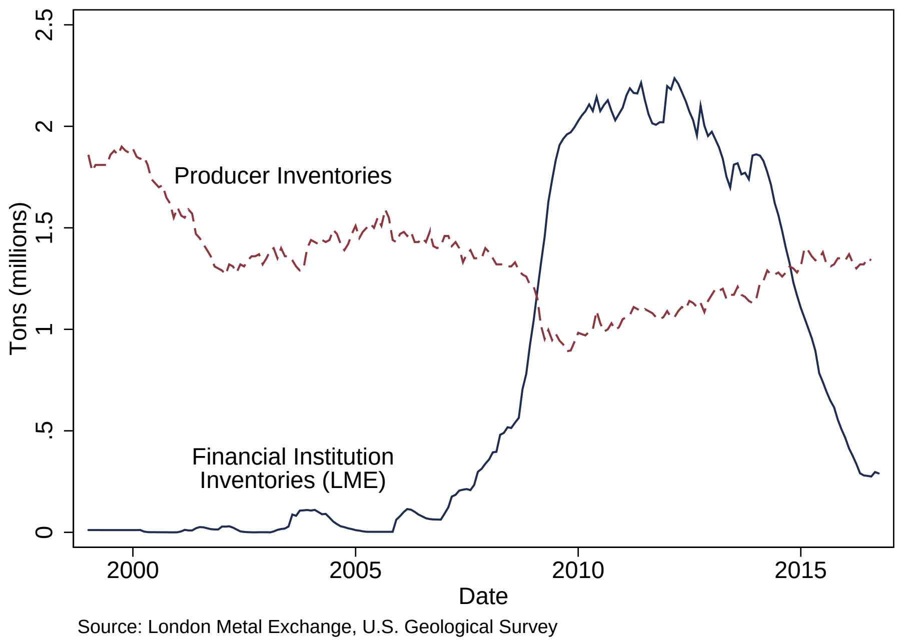



Figure 2: U.S. Aluminum Inventories (Producer and LME Levels)

Though less than 2 percent of cash trades are settled by physical delivery, the LME plays a large part in the physical aluminum market by providing a backstop option for aluminum users. If aluminum users are not satisfied by the price or quantity offered by aluminum producers, they can always purchase aluminum on the LME and take delivery of that aluminum at any time. For this reason, aluminum users view the LME as a supplier of last resort (LME, 2013). Aluminum producers sell at a discount to the LME price, and the LME price—along with the Midwest premium—is typically used as a reference price in aluminum contracts (U.S. Senate, 2014). An increase in the LME price will therefore increases the price for purchases taking place off of the LME.

In February 2010, Goldman Sachs purchased a network of warehouses in Detroit from Metro International Trade Services that was approved to hold LME inventories. Metro International is a warehouse operator that specializes in storing metals for LME in Europe and North America. The investment bank also increased its physical aluminum investments from under \$100 million in 2009 to over \$3 billion in 2012 (U.S. Senate, 2014). After Goldman purchased the warehouses, it aggressively solicited metal for its warehouses by offering steep discounts to metal owners. Goldman paid hundreds of millions of dollars in "freight incentives" (rebates) to attract aluminum to their warehouses. The incentives were so large, and attracted so much aluminum, that the Detroit warehouses quickly held more than twice the amount of aluminum held by aluminum producers in the United States. Though there were other warehouses in Detroit, Goldman's were the largest and they held almost all the aluminum in Detroit. By 2014, over 80 percent of U.S. inventories in LME were held in Detroit (Figure 3).

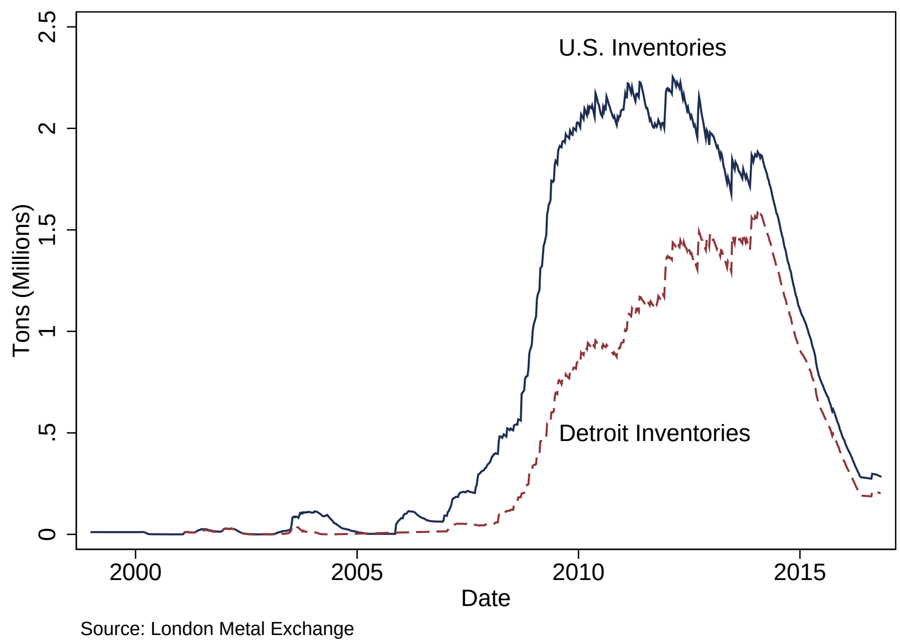



Figure 3: U.S. Aluminum Inventories (LME)

As the aluminum inventory increased, Goldman Sachs, along with several warehouse customers with large holdings, began transferring their metal between various Goldman Detroit warehouses. Goldman paid these metal owners to cancel the warrants associated with their inventory, which means they filed paperwork to remove their metal from the LME system. This action also required the warehouse owner to remove the cancelled-warrant inventory from the warehouse. As soon as the metal was removed from the warehouse, the aluminum owner would move the inventory to another warehouse and repeat the process by cancelling the warrants again. Since Goldman strictly limited the amount of metal they could remove from their warehouses each day, these "merry-go-round" transactions caused the queue length—the amount of time it takes an aluminum owner to remove their metal from a warehouse—to spike. At its peak, the queue length at the Detroit warehouses was nearly two years. This meant that aluminum users who purchased metal on the LME would have to wait two years before they could take physical possession of their metal. This extraordinary queue effectively removed 80 percent of the total U.S. aluminum stock on the LME from the spot market. Unsurprisingly, as the backstop provided by the LME became unavailable, the price of aluminum in the physical market spiked, as measured by the Midwest premium.

The aluminum merry-go-round operated by Goldman Sachs was widely publicized by *The New York Times* in 2013 (Kocieniewski, 2013). The article, and the public outcry it spawned, led to a year-long Senate investigation. Only after the Senate released a scathing report on the manipulation scheme did Goldman finally sell its warehouse business. The aluminum premium crashed immediately after the Goldman exited the aluminum market (Figure 4).

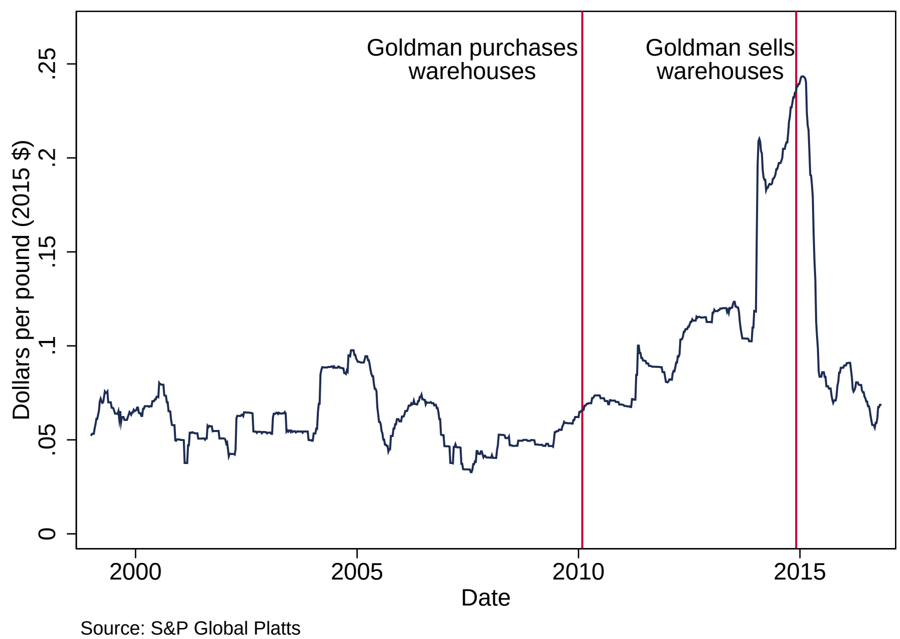



Figure 4: U.S. Aluminum Premium and Detroit Warehouse Purchase and Sale Dates

This paper examines the impact of Goldman Sachs' involvement in the aluminum market on the U.S. aluminum premium. We ask two questions regarding Goldman Sachs' involvement in U.S. aluminum markets. First, did the U.S. aluminum premium behave like similar U.S. metal premiums after Goldman Sachs entered the market? Second, did Goldman Sachs cause the U.S. aluminum premium spike by manipulating the aluminum storage market?

To answer our first question, we exploit the aluminum production process. Aluminum is rarely consumed in its pure form; it is typically combined with other metals through the alloying process. The metals that are combined with aluminum to form alloys are production complements to aluminum and are therefore subject to the same demand shocks.

Several metals traded on the LME are commonly used in aluminum alloys, including copper, nickel, and zinc (USGS, 2014) These metals were not stored in Goldman's Detroit warehouses between 2010 and 2014, and were not subject to the same long queues as aluminum. We use a difference-in-differences model to compare the premiums of aluminum and its production complements before and after Goldman enters the market. The identifying assumption of this approach is that production complements experience the same demand shocks as aluminum, and can therefore be used to estimate a counterfactual premium path.[1] Our estimates show the regional price of aluminum diverged from the regional price of the production complements after Goldman Sachs purchased the Detroit warehouses and began stockpiling aluminum. In other words, the aluminum premium increase was not a response to increased demand for aluminum.

To answer our second question, we use an instrumental variables model. We argue that Goldman increased the aluminum premium by increasing the queue length at its Detroit warehouses and trapping most of the U.S. aluminum inventory in them. Because of the extraordinarily long queue length, metal owners were unable to respond to the rising regional price of aluminum and remove their metal from Goldman's warehouses.

We use three instruments for Goldman's manipulation of warehouse queues. Our first instrument is an indicator for the Lehman Brothers bankruptcy. Following the September 2008 collapse of Lehman Brothers, Goldman Sachs became a bank holding company and aggressively pursued profits in commodity markets. Our second instrument is the number of Goldman Sachs employees on the warehouse company's (Metro's) board of directors. The cancelled-warrant scheme was complex and required significant operational oversight by Goldman Sachs. We assume the composition of the board of directors is correlated with the operational intensity of the market manipulation scheme, but not influenced by other determinants of the aluminum premium. Our third instrument is a lagged Detroit real estate price index. Like the financial crisis, we assume the (lagged) negative shock to real estate prices in Detroit was exogenous. Low real estate prices in Detroit allowed Goldman Sachs to profitably acquire and build warehouses.

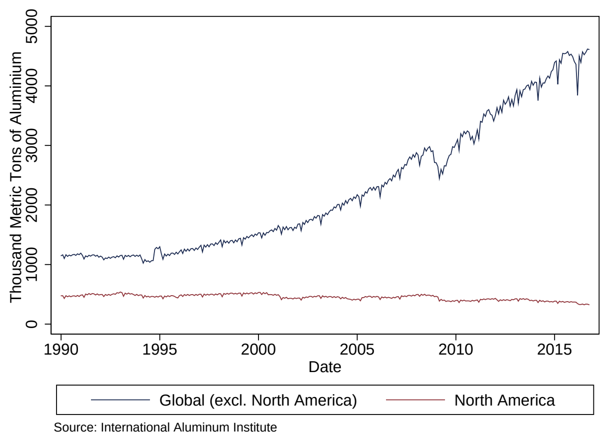
Using each instrument individually, and in combinations, we find a statistically significant, positive effect of the manipulation scheme on the aluminum premium. The extraordinary queue lengths at Goldman's Detroit warehouses, driven by an unprecedented surge in cancelled warrants, caused the Midwest premium to rise at least \$0.20 per pound. In other words, Goldman Sachs used their warehouses to create a supply disruption that increased the U.S. aluminum premium. We thus identify a situation where a substantial volume of inventories was accumulated and caused the regional price to rise significantly.

Importantly, we also find evidence that the spike in the aluminum premium increased costs for industrial aluminum users and consumers. When industrial users purchase and

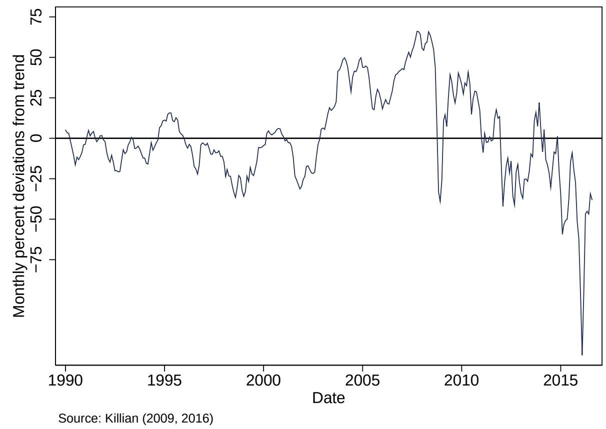
[1] We also investigate whether the premium spike was driven by an aluminum supply shock. Using a vector autoregression model of the U.S. aluminum industry, we find no evidence of a supply shock. The model includes monthly measures of U.S. aluminum supply, demand, and price.

sell aluminum, the contract price is typically based on the sum of the spot price and the regional premium. The substantial increase in the aluminum premium raised the costs of the upstream aluminum processors, who appear to have passed on the costs to the downstream aluminum manufacturers. A case study of the consumer carbonated beverage market provides evidence that these increased costs were eventually passed to consumers. Using a difference-in-differences model, with beverages in plastic bottles as a control, we find that the aluminum market manipulation caused a statistically significant increase of 1-2 percent in the the price paid by consumers for beverages in aluminum cans.

Though we focus in this paper on the regional price of aluminum and its determinants, we also note a few facts about the global aluminum market for completeness. The spot and futures prices trended downward between 2010 and 2014. The downward trend appears to have been caused by growing global supply and weakening global demand (see Figure 5). On the supply side, global aluminum production did not diverge from its long-term trend during this period. Indeed, global production increased at an increasing rate, while production in North America was essentially flat (USGS, 2011, 2012, 2014). An indicator that the U.S. aluminum market was not particularly tight when the manipulation scheme began is evinced by the aluminum anti-dumping case that the United States brought against China in 2010 (Bown, 2015). Evidently, the U.S. aluminum market had enough supply to warrant an anti-dumping complaint. On the demand side, global real economic activity in industrial commodity markets—a measure of global commodity demand—trended downward from 2010 to 2014 (Kilian, 2009). This decline in commodity demand was likely driven by weakening Chinese consumption (IMF, 2015). Therefore, regulators would not have detected any abnormalities in the aluminum market by focusing solely on global supply and demand shocks (or the lack thereof).



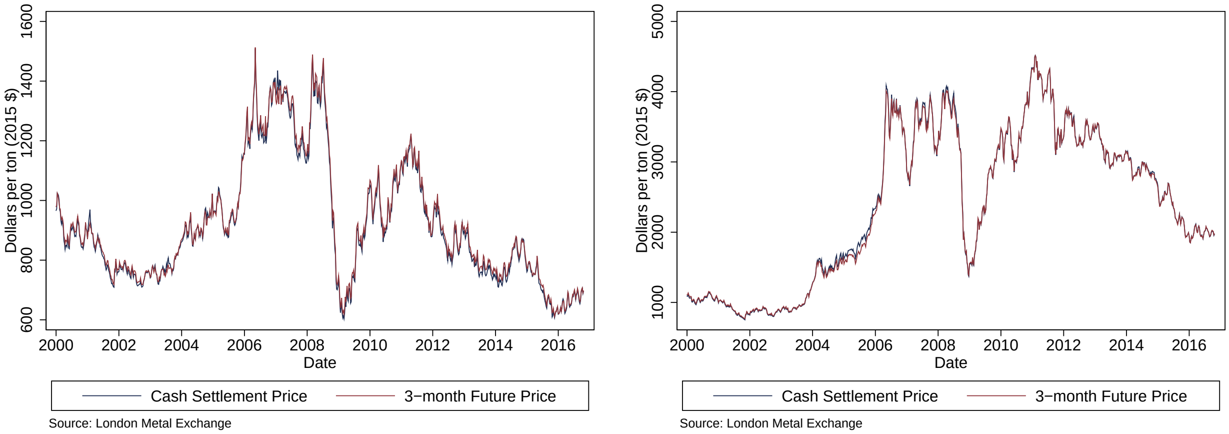

(a) Primary Aluminum Production




(b) Economic Activity in Commodity Markets

Figure 5: Global Aluminum Market

Moreover, regulators would not have detected this manipulation scheme by analyzing only the spot and futures markets for aluminum, regardless of whether they compared the prices across commodities or across time. From 2011 to 2014, the aluminum spot price and the aluminum futures price fell along with the spot and futures price of a production complement, copper (see Figure 6). Thus, regulators would not have noticed anything odd by comparing spot and futures prices across related commodities. Additionally, aluminum and copper prices rose in 2010, partially recovering from the trough reached during the Great Recession. This post-recession increase occurred in the early 2000s as well, following the 2001 recession. It is not surprising at all that commodity prices revert back to their pre-recession levels. Thus, regulators similarly would not have observed anything strange by comparing spot and futures prices across time.




(a) Aluminum

(b) Copper

Figure 6: LME Spot and Futures Prices, 2005-2016

The literature on the detection of financial market manipulation has focused primarily on spot and futures prices (Pirrong, 2004; Abrantes-Metz and Addanki, 2008; Öğüt et al., 2009; Pirrong, 2010), but does not address manipulation schemes that raise regional premiums without increasing prices in spot and futures markets. In the short run, spot and futures prices are not necessarily tied to regional premiums. A regional premium for a commodity could rise while spot and futures prices fall if, for example, high transportation costs prevent the commodity from flowing into the region. In the long run, we would expect that high regional premiums would attract investment that would facilitate the flow of commodities. Short-term manipulation schemes could raise the regional premium, the global spot and futures prices, or both.

In past cases of market manipulation—for example, silver in 1980 and soybeans in 1989—regional premiums were unaffected, so it was sufficient to only monitor spot and futures prices. The 2010-2014 aluminum manipulation scheme was successful because regulators did not monitor regional commodity markets as closely as they monitored the LME, and regional markets require much smaller inventories to manipulate. The spot and futures prices for aluminum traded on LME reflect the supply and demand on the world market, implying that manipulating LME contracts would require much larger aluminum inventories than those accumulated in the Detroit warehouses. Although the inventories accumulated in Detroit were likely insufficient for the purposes of moving the global market, they were sufficiently large to increase load-out queues in the region, causing prices paid by industrial aluminum users to significantly diverge from spot prices. Again, regulators who only monitored spot and futures prices would not have noticed anything peculiar occurring in the aluminum market. The trends in the spot and futures market for aluminum matched those of a production complement, copper. See Figure 6. Nothing would have appeared out of the ordinary to regulators looking at spot and futures prices.

This episode illustrates the challenge facing financial regulators. Financial market manipulation harms current market participants by increasing costs, and future market participants by decreasing liquidity (Cumming et al., 2011). Though the Dodd-Frank Wall Street Reform and Consumer Protection Act of 2010 increased the authority of financial regulators to prosecute manipulative practices in financial markets, regulators struggle to prevent manipulation (Abrantes-Metz et al., 2013). These regulators cannot simply ban all manipulative practices ex-ante, because financial institutions will always find new schemes that are not outlawed. Though the Senate investigation concluded that Goldman's activities in the physical commodities market "increased financial, operational, and catastrophic event" risk, the investigation did not find evidence of illegal activity and Goldman faced no financial penalties from regulators (U.S. Senate, 2014).

Since there were no fines associated with this aluminum market manipulation, Goldman and other financial institutions are more likely to view similar commodity market manipulation schemes as risk-adjusted profitable. This successful case of market manipulation will likely spawn imitators in other physical commodity markets. Real-time detection and investigation, not ex-post penalties, are the only reliable deterrents available to investigators. Regulators cannot rely on *The New York Times* to unearth these schemes every time, and need to improve in-house market monitoring.

We conclude by presenting an algorithm to detect similar behavior in the future. Our early detection algorithm uses complements for commodities and looks for breaks in the regional price and inventory levels relative to the complement. The model employs multivariate break tests, as univariate tests are not sufficient because their estimates are not as accurate and produce too many false positives. Regulators could have detected this break using our algorithm as early as December 2012, more than six months before the manipulation scheme was publicized by *The New York Times*. Note that this algorithm is designed to be a fire alarm. It is not a causal model. That is, it does not tell regulators that a certain institution or a set of institutions has manipulated the market. Rather, it alerts regulators to possible instances of manipulation. Regulators would still have to perform careful investigative work to determine the existence of causality.

We then apply our detection algorithm to a suspected case of manipulation in the European aluminum market over the same time period. In mid-2011, Glencore, a commodity trading firm that was paid by Goldman Sachs to participate in the aluminum merry-go-round in Detroit, purchased LME warehouses in Vlissingen, a port city in Netherlands. A few months after the purchase, these warehouses experienced massive aluminum warrant cancellations. Since Glencore, like Goldman Sachs, only loaded out the minimum metal tonnage required by the LME, these warrant cancellations caused the warehouse queue to spike. It eventually peaked at over 774 days in June 2014, nearly 3 months longer than the Detroit queue at the time. The European aluminum premium rose as the queues increased. Our detection algorithm estimates a break in the European aluminum market in late 2011, which could have been detected by regulators using only data available through late 2012. We emphasize again that this algorithm does not prove causation. These results do suggest that there is a high probability that the European aluminum market was manipulated in a similar fashion to the U.S. aluminum market. However, it is still the regulators' job to investigate the details of the case.

This paper makes three contributions to the literature. First, we use novel instruments to identify the cause of the 2010-2014 aluminum premium spike. Investigations of the Goldman Sachs' metal warehouses by journalists (Kocieniewski, 2013) and the U.S. Senate (U.S. Senate, 2014) found evidence of manipulation, but did not present a rigorous argument for causality. Second, we develop a new technique to identify manipulation in commodity markets using production complements. This approach can be used to identify manipulation in other physical commodity markets. Third, we present an algorithm to detect manipulation in commodity markets that incorporates inventories, regional premiums, and warehouse load-out wait times. Unlike other manipulation detection methods which rely only spot and futures market prices (Abrantes-Metz and Addanki, 2008; Öğüt et al., 2009), our algorithm accounts for inventory delivery backlogs and regional supply and demand shocks which raise transaction prices for industrial commodity users but do not effect commodity market prices. Given the success of the aluminum manipulation scheme, it is likely to be imitated. Regulators can use the algorithm developed in this paper to assist in their identification and enforcement of manipulation in physical commodity markets.

# 2 Analysis of the Aluminum Midwest Premium

We use two empirical models to investigate whether manipulative behavior by Goldman Sachs caused the Midwest premium spike. First, we use a difference-in-differences model to determine whether the spike in the U.S. aluminum premium was abnormal, or simply the result of a demand increase. This model exploits the aluminum alloy production process to identify a control group of metals that is subject to the same demand shocks as aluminum, but was not stored in Goldman's warehouses. After showing that the aluminum premium spike was abnormal, we use several instruments to show that long queues at Goldman's Detroit warehouses drove the premium increase.

## 2.1 Difference-in-Differences Model

In this section, we investigate whether the regional price of aluminum increased relative to the regional prices of other metals subject to similar demand shocks. To facilitate the examination, note the following facts about aluminum. Aluminum is the second most commonly consumed metal on earth, second only to iron. Aluminum is rarely consumed in its pure form; it is almost always alloyed, or combined with other metals, to achieve the desired conductivity, corrosion resistance, density, and strength. The properties of aluminum alloys depend on the metals used, and aluminum industry guidelines mandate that each alloy include specific proportions of the component metals (Aluminum Association, 2015). The metals cannot be substituted without changing the properties of the alloy. This strict industry regulation of alloys allows aluminum users to purchase alloys from any producer, knowing the alloy composition and properties are consistent.

The metals that are combined with aluminum to produce alloys form an ideal control group for a difference-in-differences regression. In particular, copper, nickel, and zinc are all used in aluminum alloys, traded on the LME, and were not stored in Goldman Sachs' warehouses. Copper is the most common aluminum alloying element, and is our primary metal of interest (Mondolfo, 2013).[2] Aluminum-copper alloys contain between 3 and 14 percent copper, and copper is also added in smaller amounts to other common alloys, including aluminum-silicon alloys (up to 5 percent copper) and aluminum-zinc alloys (up to 2.4 percent copper). Zinc and nickel are also regularly added to aluminum alloys (Aluminum Association, 2015). The U.S. premiums of those metals are used in robustness checks.

Figure 7 plots the Midwest premiums for aluminum and copper between November 1999 and December 2015, which are both based on actual transaction prices in the physical metal spot markets. Prior to Goldman Sachs' purchase of the aluminum warehouses in Detroit, the two premiums followed a parallel trend, and the average spread was \$0.02 per pound. After the warehouse purchase, the regional price of aluminum spiked, and the spread grew to \$0.18 per pound.

[2] Though industry standards specify the share of each metal in every aluminum alloy, production statistics by aluminum alloy are not available. While data on the total amount of aluminum and copper combined in alloy production each year are not published, industry sources consistently report that the most popular alloys are aluminum-copper (Aalco, 2005; Mondolfo, 2013).

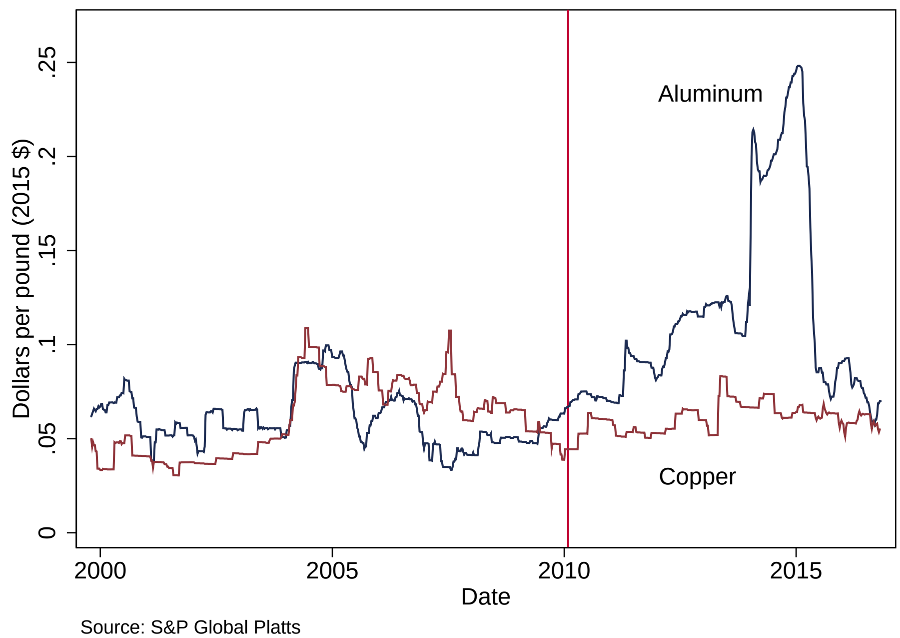



Figure 7: U.S. Metal Premiums

Our empirical approach compares the Midwest premium of aluminum to the Midwest premiums of its production complements:

$$P_{i,t} = \alpha + \alpha_t + \beta_1 Alum_i + \beta_2 Post_t + \beta_3 Alum_i \times Post_t + \varepsilon_{i,t} \tag{1}$$

where $i$ indexes metals and $t$ indexes weeks. The $Alum$ variable is an indicator for aluminum; $Alum = 1$ for aluminum and $Alum = 0$ for other metals. The $Post$ variable is an indicator for the dates after Goldman Sachs purchased Metro International; $Post = 0$ prior to February 2010 and $Post = 1$ after.

The explanatory variable of interest is the interaction between the aluminum indicator ($Alum$) and the warehouse purchase indicator ($Post$). The coefficient of this variable ($\beta_3$) represents the effect of Goldman's entry into the aluminum market on the aluminum premium.

The key identifying assumption of this difference-in-differences model is that the regional price of aluminum would have followed a similar trend to the complement metals absent Goldman's entry into the aluminum market. As shown in Figure 7, the pre-2010 premiums of aluminum and copper are very similar and do not deviate by more than a few cents. These data support our identifying assumption that the U.S. premiums for two metals commonly consumed together will not significantly diverge in a normally functioning market.

The results of this difference-in-differences model are presented in Table 1. The first column presents least squares estimates of equation 1. Using copper as a control, we find that the U.S. aluminum premium increased about \$0.057 per pound post-2010, essentially doubling the average premium from 1999 through 2010. Adding metal-specific time trends and month-of-sample controls increase the estimated effect to \$0.068 per pound post-2010 (Column 2). To put this premium increase in perspective, the Midwest premium accounted for less than 10 percent of the total cost of aluminum for U.S. consumers prior to 2010, but after 2010, the premium accounted for as much as 30 percent of the total aluminum price.

These results, established in the first two columns, are robust to including other, somewhat less common alloying metals—nickel and zinc—in addition to copper in the control group. With these three metals forming the control group, the estimated treatment effect is similar, about \$0.052 per pound (Column 3). Adding time trends for each metal and month-of-sample controls yields a slightly larger estimate, \$0.06 per pound. In all cases, the estimated effect of Goldman Sachs' warehouse purchase on the aluminum premium is statistically significant.

| Dependent Variable: U.S. Regional Metal Premiums ($/pound, Real) | | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Aluminum × Post | 0.0571*** | 0.0682*** | 0.0518*** | 0.0601*** |
| (“Goldman Effect”) | (0.00329) | (0.00251) | (0.00333) | (0.00357) |
| Aluminum | 0.000691 | 0.00834*** | 0.00706*** | 0.0127*** |
| (Treatment) | (0.00107) | (0.00156) | (0.000940) | (0.00181) |
| Post | 0.000599 | -0.0324*** | 0.00582*** | -0.0129 |
| (Post-Feb 2010) | (0.00101) | (0.0123) | (0.00112) | (0.0175) |
| Observations | 1,572 | 1,572 | 3,144 | 3,144 |
| R-squared | 0.430 | 0.774 | 0.282 | 0.562 |
| Controls Included: | | | | |
| Copper | YES | YES | YES | YES |
| Nickel & Zinc | NO | NO | YES | YES |
| Metal-Specific Time Trends | NO | YES | NO | YES |
| Month-of-Sample | NO | YES | NO | YES |

*Notes:* Each column contains the results for a separate regression. The unit of observation is week. *** denotes significance at the 1 percent level, ** denotes significance at the 5 percent level, * denotes significance at the 10 percent level.

Table 1: Difference-in-Differences Model Estimates

As a robustness check, we examine whether the estimated increase of the aluminum premium, relative to the copper premium, was unusually large for the U.S. metals market. Though aluminum and copper are regularly combined in aluminum alloys, the two metals are not *perfect* complements. Some aluminum alloys contain no copper, and some copper alloys contain no aluminum. We would therefore expect idiosyncratic supply and demand shocks to affect the aluminum-copper premium spread. To determine whether the “Goldman Effect” presented in Table 1 was significantly larger than the typical deviation between the aluminum and copper premiums, we estimate a distribution of the idiosyncratic shocks to the aluminum-copper premium spread between January 2000 and January 2010. To estimate this distribution, we replicate the difference-in-differences specification reported in Column (2) of Table 1 using each of the 417 weeks from January 2001 through December 2008 as placebo treatment dates. These placebo regressions use the same sample period as the regressions in

Table 1. The distribution of these placebo estimates is plotted in Figure 8. The estimated placebo treatment effects fall between -\$0.04 and \$0.04, with a median placebo treatment effect of zero. This analysis confirms that the estimated \$0.05-\$0.07 "Goldman Effect" is unusually large for the U.S. metals market.

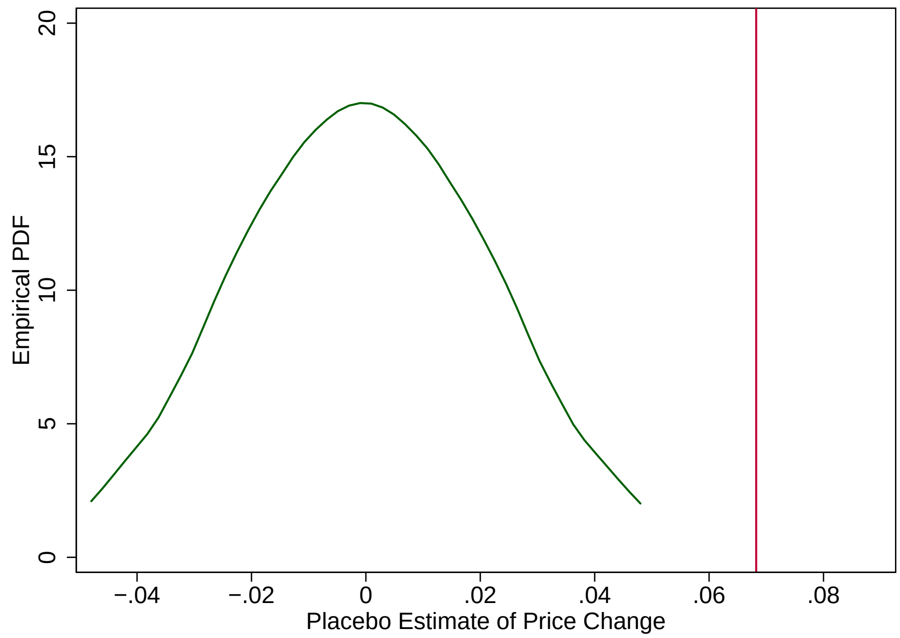



Figure 8: Distribution of Placebo Estimates

## 2.2 Instrumental Variables Model

Having shown the aluminum premium increase was abnormal—as it did not occur in metals subject to similar demand shocks—we now investigate the cause of the premium increase. The aluminum premium boom and bust that coincided with Goldman Sachs' entry and exit in the aluminum warehouse business (recall Figure 4) is compelling circumstantial evidence of manipulation, but Goldman could have anticipated an increase in aluminum demand—and hence, a premium increase—and invested accordingly.

When Goldman Sachs purchased the Metro International metal storage warehouses in February 2010, about 40 percent of the LME aluminum inventory was held in Detroit. When Goldman sold the warehouses in December 2014, over 80 percent of the inventory was held in Detroit (Figure 9). Over those four years, the queue length—that is, the time it takes to remove metal from the warehouse—in Goldman's Detroit warehouses increased from a few days to nearly two years. With such a long queue length, the aluminum in the Detroit warehouses was effectively removed from the market. Our identification strategy uses instruments for the queue length to demonstrate that Goldman used its Detroit warehouses to manipulate the U.S. aluminum premium.

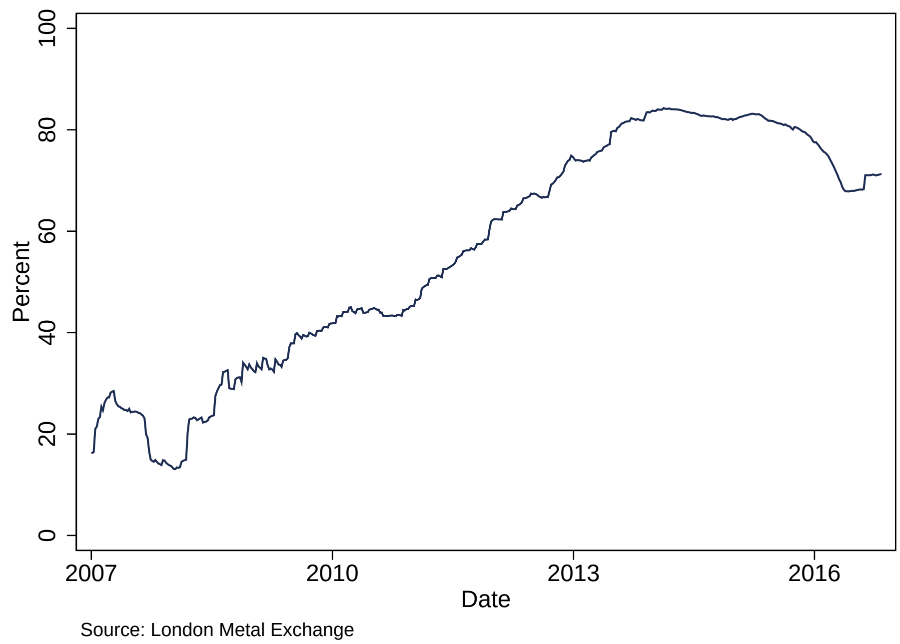



Figure 9: Share of Total U.S. LME Aluminum Inventory in Detroit

The scheme that Goldman Sachs used to manipulate the aluminum premium follows the Accumulation-Lift-Distribution (ALD) model of asset price manipulation (Lang, 2004; Klein et al., 2012). In an ALD scheme, the manipulator first accumulates the asset of interest through either long positions or physical inventory, then lifts the price through its newly acquired market power, and finally sells at the inflated price. The ALD model describes popular forms of securities fraud schemes like pump-and-dump and buy-tip-sell (Dalko, 2016). One can distinguish the ALD schemes from traditional investing strategies by the lift phase. In this phase, asset owners attempt to increase the price of an asset using an illegal method, like spreading misleading information about the asset to other investors, or accumulating a large position and exercising market power. In commodity markets, manipulation via market power is the most common mechanism used to lift prices (Coffee, 2009).

Goldman Sachs began the ALD scheme by accumulating inventory in their Detroit warehouses. After Goldman purchased the warehouses, they paid hundreds of millions of dollars on “freight incentives” (rebates) to attract aluminum to their Detroit warehouses (U.S. Senate, 2014). These rebates were so large and brought in so much aluminum that the LME actually investigated the warehouses for disrupting markets by “giving exceptional inducements" (LME, 2013). Goldman Sachs also increased the warehouse inventory by purchasing over $3 billion of aluminum and storing it in their warehouses. Thus, Goldman accumulated a vast inventory in Detroit by inducing holders of the aluminum to store their inventories in Detroit and also by directly purchasing aluminum on the market.

In the next step of the manipulation scheme, Lift, Goldman exploited the load-out requirements for LME warehouses. The LME required warehouse owners to load out a minimum of 1,500 tons of aluminum per day (LME, 2013).[3] Importantly, the minimum load-out requirement applied at the city-level for warehouse owners, not the warehouse level. This means that Goldman Sachs only needed to load out a total of 1,500 tons each day across all of its Detroit warehouses to meet the requirement. Information uncovered during the Senate investigation of Goldman Sachs' involvement in physical commodity markets suggests that their Detroit warehouses did not exceed the minimum required load-out rate. In other words, Goldman set the maximum load-out rate at the minimum required level.

As the warehouse inventory grew, Goldman paid a few large clients (the bank holding companies, Deutsche Bank and JPMorgan, and the commodity trading firms, Glencore and Red Kite) to transfer their aluminum between Goldman's Detroit warehouses. The transfer process had three steps. First, the client would cancel the warrants on their aluminum, which notified the LME that their metal was no longer available for trading. (Note that metal available for trading is referred to as "on-warrant.") Second, the cancelled-warrant aluminum would join the queue, thereby awaiting load out from the warehouse. The enormous amount of cancelled-warrant orders far exceeded the daily load-outs, so the queue grew in length. Third, after taking delivery of their metal, the clients would complete the transfer process by placing their aluminum on-warrant in another one of Goldman's Detroit warehouses and restarting the process, that is, canceling that warrant again and reentering the queue. Rinse and repeat. These large clients benefited from this scheme by receiving compensation from Goldman, and Goldman benefited from this scheme because the long queues gave Goldman control over an enormous aluminum inventory.

Goldman's approach to the Accumulation and Lift stages was novel because they controlled the aluminum inventory without owning all of it. By restricting the flow of aluminum out of their warehouses, Goldman prevented the LME participants from immediately selling their metal on the physical market or consuming it themselves. In essence, Goldman artificially created contractionary supply shocks in the aluminum market.

The Distribution step of Goldman's scheme was also innovative. Once prices rose, Goldman Sachs had at least two sources of profit: derivative contracts based on the aluminum

[3] In April 2012, LME increased the minimum load out rate to as much as 3,000 tons per day in response to complaints about queues at Goldman's warehouses.

premium and aluminum inventories. Prior to the increase in the Midwest aluminum premium, Goldman Sachs increased their exposure in the aluminum market by entering into contracts with the owners of the aluminum in the Detroit warehouses that required the aluminum owners to pay Goldman when the Midwest premium rose (U.S. Senate, 2014). This meant that Goldman Sachs directly profited from derivative contracts tied to the Midwest premium as queues at Goldman's warehouses caused the premium to rise. In addition, Goldman Sachs also profited from ownership of an enormous aluminum stockpile in Detroit, valued at $3.2 billion in 2012. As the Midwest premium rose, Goldman, "engaged in extensive aluminum trading" with their physical aluminum assets as the Midwest premium rose in 2013 and 2014 (U.S. Senate, 2014). Since the 2013-2014 Senate investigation of Goldman Sachs likely hastened the sale of the Detroit aluminum warehouses, it is possible that the Lift and Distribution steps were cut short.

This "merry-go-round of metal" (Kocieniewski, 2013) caused the queue length to peak at nearly two years in 2014 (Figure 10). Meaning, if an aluminum user purchased aluminum in the LME spot market in April 2014 and immediately filed the paperwork to remove the aluminum, they would not take physical possession of their metal until about March 2016. Recall that the growing queue made the LME inventories inaccessible to aluminum users, which allowed aluminum producers to raise prices knowing that their customers no longer had a nearby supplier of last resort. Without a supply backstop, the Midwest aluminum premium quadrupled between 2010 and 2014.

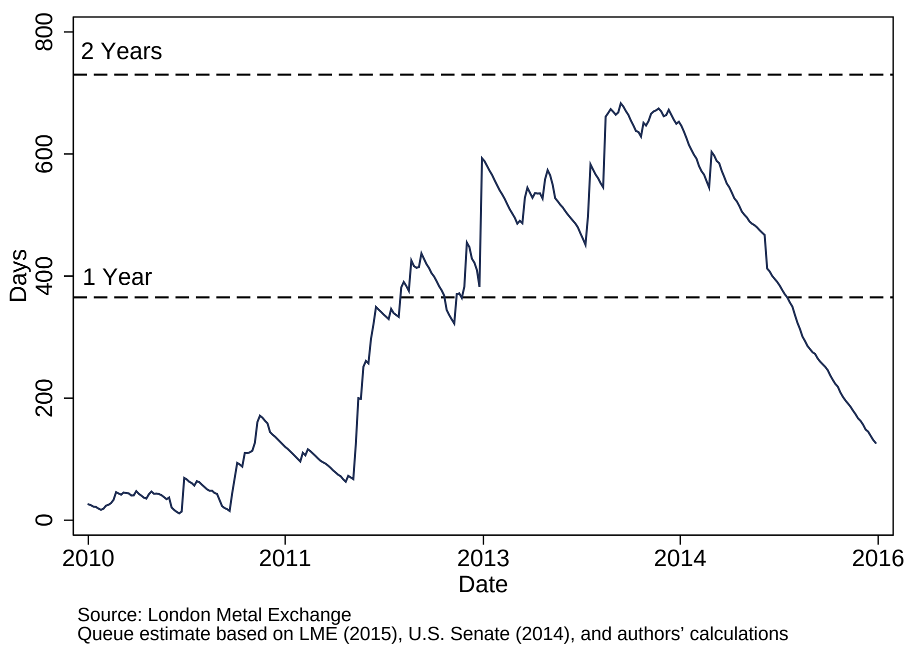



Figure 10: Queue Length at Goldman Sachs' Detroit Warehouses

In an Ordinary Least Squares (OLS) regression of the aluminum premium on cancelled warrants, which is our proxy variable for queue length,

$$P_t = \alpha + \alpha_t + \beta CW_t + \varepsilon_t \tag{2}$$

the coefficient on cancelled warrants, $\beta$, is positive and statistically significant (Table 2, column 5). Using this estimate, a 100,000 ton increase in cancelled aluminum warrants increases the Midwest premium by \$0.01 per pound. Thus, under a linearity assumption, the 1.2 million tons of cancelled warrants would have raised the aluminum premium by \$0.12 per pound, nearly double the 1999 through 2009 average premium of \$0.07. There are a couple of reasons to doubt this estimate. First, the causal direction is unclear. The rising aluminum premium could have caused the increase in cancelled warrants, not the other way around, as holders of aluminum may have wished to take their inventory off the market in order to sell for a higher price at a later date. Second, this simple regression omits many variables that are relevant in determining the aluminum premium. Therefore, we use instrumental variables model with three exogenous instruments to deal with these potential shortcomings.

Our first instrument is an indicator variable for the Lehman Brothers bankruptcy in September 2008 (Figure 11). Goldman Sachs formed a bank holding company and pursued profits in commodity markets during the financial crisis that followed the collapse of Lehman Brothers. Recall that under the Gramm-Leach-Bliley Act of 1999, sufficiently well-capitalized bank holding companies were no longer restricted from engaging in physical commodity transactions. We assume the financial crisis was an exogenous shock, an assumption also used by Bloom (2009) and Chodorow-Reich (2014). Cancelled warrants prior to the Lehman bankruptcy were near zero. The $R^2$ of a linear regression of cancelled warrants on the bankruptcy indicator is 0.42.

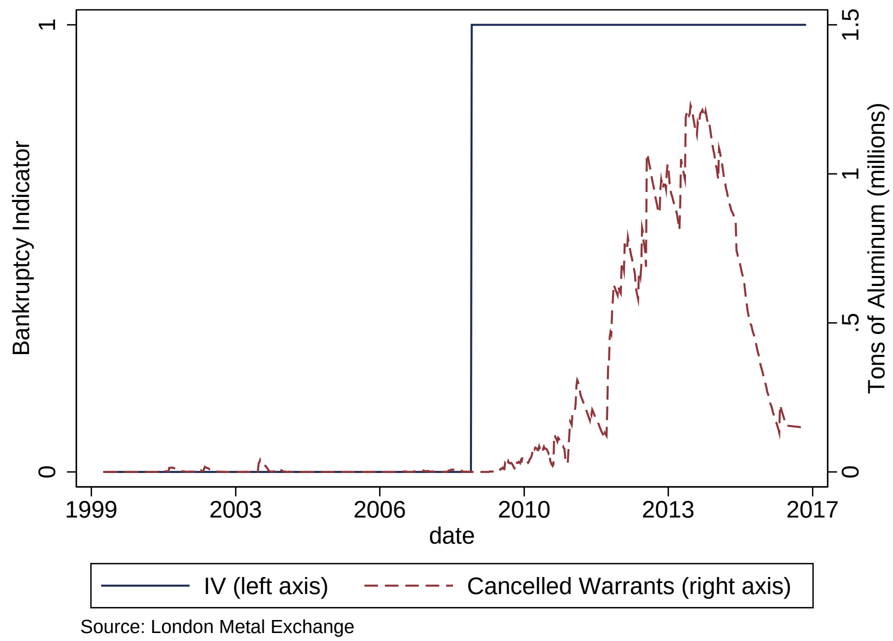



Figure 11: Lehman Brothers Bankruptcy Indicator

Our second instrument is the number of Goldman Sachs employees on the warehouse company's board of directors. The cancelled-warrant scheme was complex and required significant operational oversight by Goldman Sachs. Soon after purchasing Metro, Goldman used their own employees to staff the warehouse board of directors, which provided the needed operational control and oversight. We assume the number of Goldman employees on the warehouse board varied with the operational intensity of the market manipulation scheme. A linear regression of the cancelled-warrants level on the number of employees yields an $R^2$ of 0.39, which confirms our intuition that the two series would moves jointly (Figure 12).

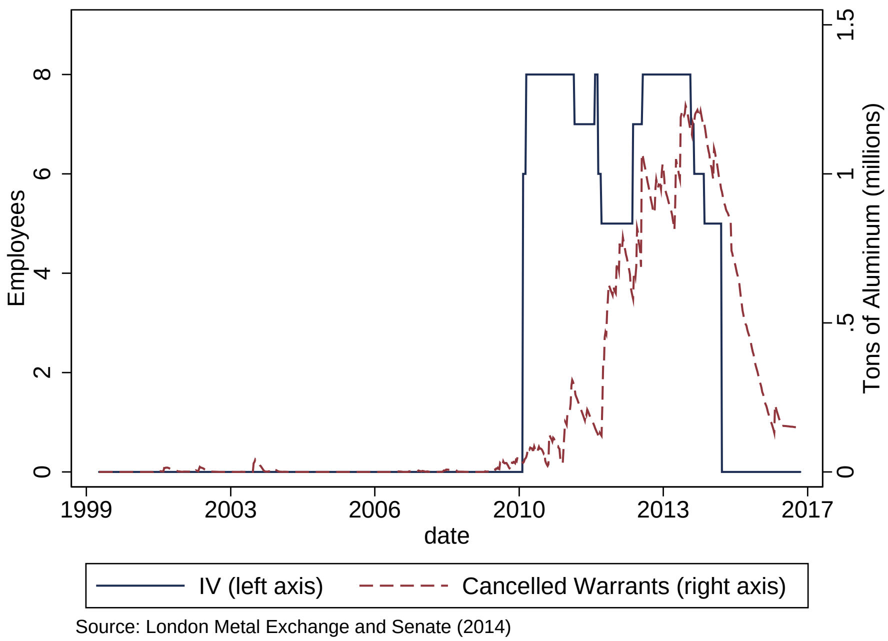



Figure 12: Goldman Sachs' Employees on Metro's Board

Our third instrument is the one-year lagged value of the Detroit real estate price index. This manipulation scheme required Goldman to cheaply store a substantial percentage of the LME aluminum stocks. Low real estate prices in Detroit allowed Metro, and hence Goldman, to inexpensively acquire and build warehouses to store this aluminum. The lagged value of the price index reflects the delays associated with purchasing land and constructing warehouses. Like the financial crisis, we assume the lagged negative shock to real estate prices in Detroit was exogenous. As shown in Figure 13, the real estate index is correlated with the cancelled-warrant level.

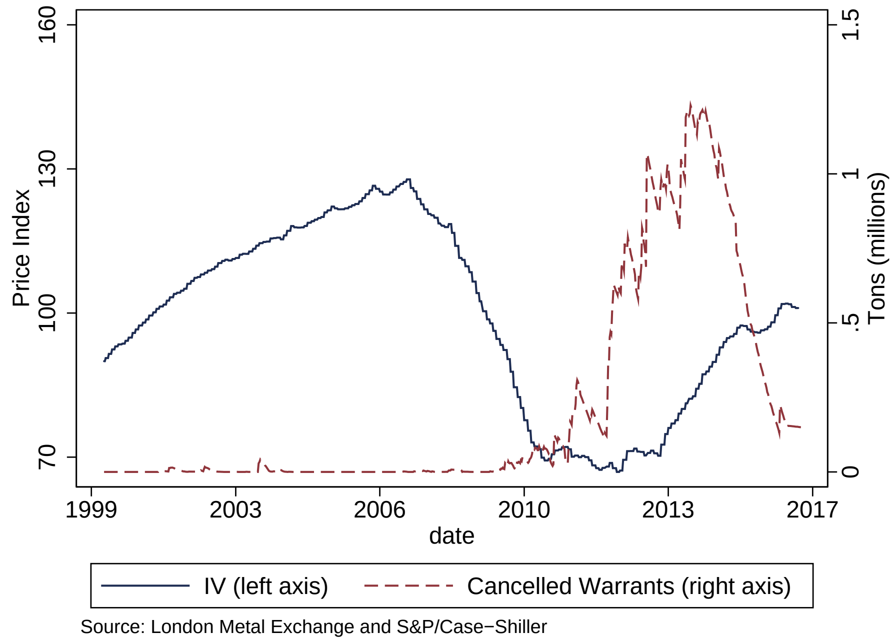



Figure 13: Detroit Real Estate Price Index

With these instruments in hand, we estimate a standard two-stage least squares model. In the first stage, we regress the cancelled-warrant level ($CW_t$) on the instrument ($IV_t$) and time fixed effects ($\alpha_t$),

$$CW_t = \alpha + \gamma IV_t + \eta_t + \varepsilon_t \tag{3}$$

In the second stage, we regress the aluminum premium on the predicted cancelled-warrant level ($\widehat{CW}_t$) and time fixed effects, $\eta_t$,

$$P_t = \theta + \beta \widehat{CW}_t + \eta_t + \epsilon_t \tag{4}$$

The results from the second-stage regression show that the surge in cancelled warrants, and the resulting queues, in Goldman's Detroit warehouses caused a statistically significant increase in the aluminum premium. Using the instruments individually in the first stage regression, the estimated effect of a 100,000 ton increase in cancelled warrants is about one cent (Table 3, columns 1-3), slightly smaller than the result from the OLS regression in column 4. The estimates from the instrumental variables model consistently demonstrate that the Midwest premium spike was caused by abnormally long queues at Goldman Sachs' Detroit warehouses. Goldman instigated this queue by paying clients to cancel warrants, and maintained it by restricting the daily load-out rate.

Dependent Variable: Aluminum Premium (Dollars per pound, Real)

| | Two-Stage Least Squares | | | OLS |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Cancelled Warrants | 0.00969*** | 0.00842*** | 0.00848*** | 0.0101*** |
| (100,000 tons) | (0.000364) | (0.000352) | (0.000311) | (0.000379) |
| $R^2$ | 0.705 | 0.686 | 0.689 | 0.707 |
| Observations | 882 | 882 | 882 | 882 |
| First Stage Instrument: | | | | |
| Lehman Bankruptcy | YES | NO | NO | NO |
| Employees on Board | NO | YES | NO | NO |
| Detroit Real Estate Index | NO | NO | YES | NO |
| Controls Included: | | | | |
| Month-of-Sample | YES | YES | YES | YES |

*Notes:* Each column contains the results for a separate regression. The unit of observation is week. *** denotes significance at the 1 percent level, ** denotes significance at the 5 percent level, * denotes significance at the 10 percent level.

Table 2: Instrumental Variable Model Estimates

While this particular set of instruments is useful in identifying the cause of the price increase in the U.S. aluminum market, these instruments would not necessarily be useful in identifying the cause of price increases in other commodity markets. Identification of commodity market manipulation using instrumental variables is context specific, and relies on details unique to particular commodity markets and knowledge of the mechanics of the manipulative scheme. Though we identified a set of instruments that are plausibly exogenous to other determinants of aluminum price, regulators might not always be able to find a set of such instruments. Fortunately, cleanly identifying the cause of manipulation is not necessary to detect and prevent manipulation. In Section 3, we discuss a statistical algorithm that regulators could use to detect manipulation that does not rely on instruments.

## 2.3 Commodity Price Manipulation and Industrial Users

The Midwest premium spike had a significant impact on the U.S. aluminum industry, which is composed of producers, processors, and manufacturers. Industrial aluminum processors stand between aluminum producers—who mine and refine raw material to produce primary aluminum[4]—and aluminum manufacturing firms, which use processed aluminum in products sold to consumers. As middlemen, industrial processors convert pure aluminum ingots into alloyed aluminum, extruded aluminum, or flat-rolled aluminum that is used in consumer goods and industrial applications.

Industrial processors in the United States typically purchase aluminum from aluminum producers using contracts that tie the purchase price to the "all-in" aluminum price. The all-in price is the sum of the spot price and the Midwest premium at the time of purchase. Aluminum processing takes place over the course of several weeks, after which the processors sell the processed aluminum to aluminum manufacturers at a markup to the all-in price at the time of sale. Since the industrial aluminum processors' purchase and sale contracts are based on the all-in aluminum price at different dates, these contracts leave firms vulnerable to changes in either the spot price or the Midwest premium that occur between the purchase of aluminum and the sale of processed aluminum. While aluminum processors can hedge against changes in the aluminum spot price with aluminum futures contracts, changes in the Midwest premium are not usually hedged with financial contracts.[5]

[4]Primary aluminum is produced by a refining process that converts bauxite ore into alumina, which is smelted into pure aluminum. Secondary aluminum is produced by recycling existing aluminum scrap into pure aluminum. The LME spot and futures prices, as well as the Midwest premium, are based on the price of primary aluminum.

[5]Midwest premium futures contracts were not available until August 2013, when the Commodities Mercantile Exchange (CME) began offering futures contracts based on the aluminum premium. The contract was illiquid and not regularly used by producers to hedge prior to 2013-2014.

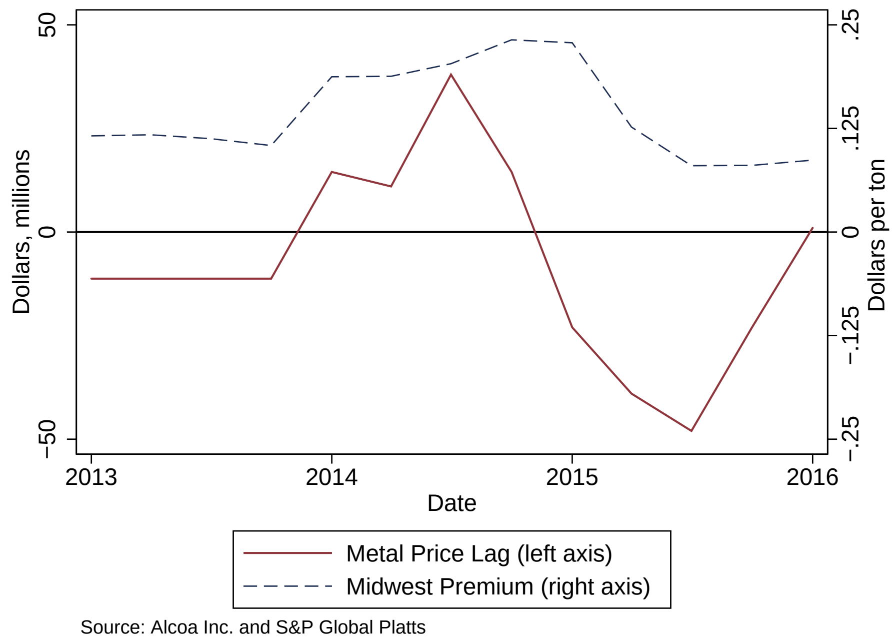



Figure 14: Alcoa Inc. Metal Price Lag Income

As the Midwest premium spiked, industrial aluminum processors began reporting losses in their SEC filings due to un-hedged exposure to the premium . The net income attributed to the difference between the price of metal at the time of purchase and sale is labeled the "metal price lag" in those filings.[6] As an example, the metal price lag for Alcoa Inc.—one of the largest firms in the aluminum industry—is plotted in Figure 14, along with the Midwest premium (Alcoa, 2015).[7] In 2013 and 2015, when the Midwest aluminum premium fell over the course of the year, Alcoa lost $45 and $155 million, respectively, due to the metal price lag. In 2014, when the Midwest premium was rising, Alcoa gained $78 million due to the metal price lag. The net income attributable to the metal price lag was relatively small, but not trivial, representing about 1 to 2 percent of total revenue from Alcoa's processed (flat-rolled) aluminum sales (Alcoa, 2015).

On the whole, aluminum processors appear to have passed on the increased aluminum cost to manufacturers, as is evident in the Producer Price Index (PPI) for aluminum sheet, plate, and foil manufacturing (BLS, 2016). This price index, which reflects the input costs of aluminum manufacturers, typically tracks the LME aluminum spot price closely. Between

[6]The metal price lag captures the effect of un-hedged exposure to all metals, not just aluminum. Since the Midwest premiums for other metals were relatively flat from 2010-2014 (see Figure 7), the metal price lag provides a reasonable measure of the effect of the aluminum premium on net income.

[7]The metal price lag was not regularly listed as a line item in SEC filings prior to 2013 because there was relatively little net income attributable to changes in regional metal premiums prior to the 2010-2014 aluminum premium spike.

2010 and 2014, however, the two series diverged significantly as the aluminum spot price fell and the Midwest premium rose (see Figure 15).

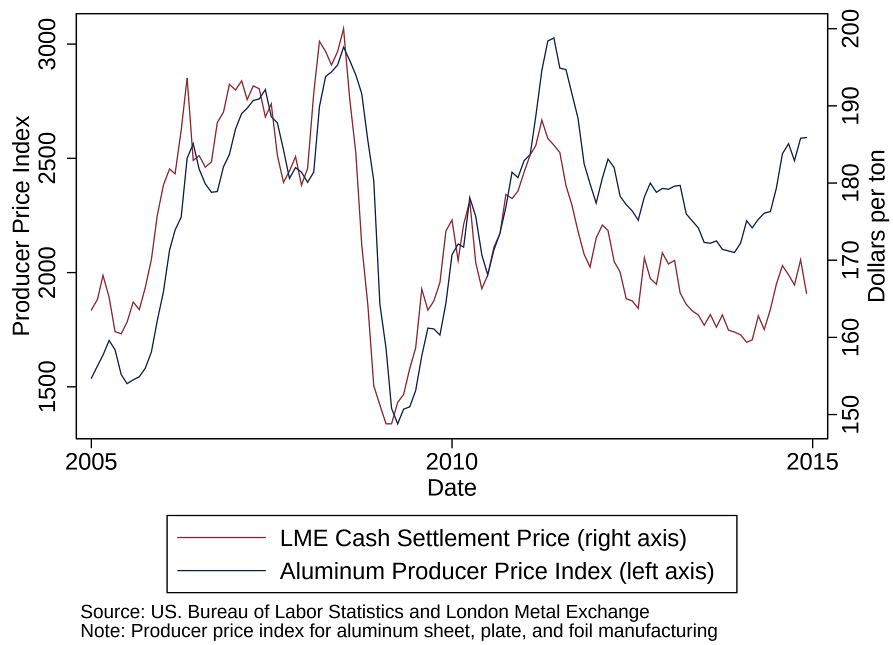



Figure 15: Aluminum Manufacturing Price Index and Spot Price

## 2.4 Commodity Price Manipulation and Consumers

Given that research has consistently shown that increases in PPI cause increases in CPI (Guglielmo Maria Caporale, 2002; Tiwari et al., 2014), we would expect the prices paid by consumers for goods that contain aluminum to reflect the increased aluminum costs paid by manufacturers. While an estimate of the total effect of aluminum price manipulation on consumers is beyond the scope of this paper, we provide a case study in the carbonated beverage market.

Consumers typically purchase carbonated beverages at retail stores in either aluminum or plastic containers.[8] For a given beverage, like Coca-Cola, the contents of the aluminum and plastic containers are identical. The only difference is the container size and number of containers in a package. Two-liter plastic bottles (67.6 ounces) are almost always sold in single units while aluminum cans (12 ounces) are most commonly sold in packages of 12, 20, or 24.

[8]Consumers can also purchase carbonated beverages in glass bottles, though glass bottles are significantly more expensive than either plastic of aluminum and represent a small fraction of the market. Using glass bottles instead of plastic bottles as the control group in the difference-in-difference regression does not have a qualitative effect on the results presented in Table 3.

Carbonated beverages provide an ideal setting to estimate the effect of aluminum price manipulation on consumer goods because we can compare prices of goods that are nearly identical except the packaging: one set of goods uses aluminum packaging and another set of goods uses plastic packaging. If the prices of bottled and canned carbonated beverages move in parallel and the differences in price are time invariant, a difference-in-differences model will allow us to estimate the effect of manipulation on the price of carbonated beverages sold in aluminum cans.

We use carbonated beverage price data from the Nielson Retail Scanner database. This database consists of weekly consumer goods prices from point-of sale systems at retail stores across the United States. In this analysis, we use prices scanned at the register for about 56 million Coca-Cola purchases. We narrow the focus of this case study to a single brand for computational ease, since including other brands, like Pepsi, yields similar results.

From 2006 to 2010, prices of Coca-Cola in cans and Coca-Cola in bottles increased in parallel (Figure 16), though there is more volatility in the can price (the dashed lines plot the 30-day moving averages in Figure 16). The longer-term trends are similar prior to Goldman Sachs' entry into the aluminum market in February 2010. After February 2010, the price of aluminum cans appears to increase somewhat more than plastic bottles, but the difference between the trends after February 2010 is subtle. Given that the container cost represents only a fraction of the total beverage cost, which includes ingredients, marketing, distribution, etc., we would not expect a large price response to an aluminum price increase.

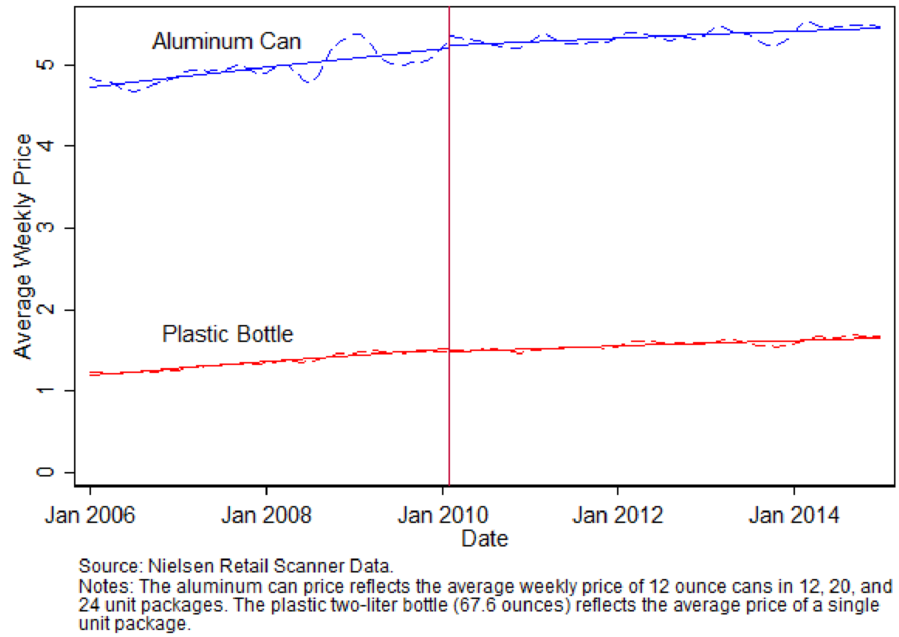



Figure 16: Coca-Cola Prices: Plastic Bottle and Aluminum Cans

We use a difference-in-difference model, similar to the model used in Section 2.1, to estimate whether there was a price increase in Coca-Cola in aluminum cans relative to Coca-Cola in plastic bottles following February 2010:

$$P_{i,t} = \alpha + \beta_1 Can_i + \beta_2 Post_t + \beta_3 Can_i \cdot Post_t + \beta_4 X_{i,t} + \varepsilon_{i,t} \tag{5}$$

where $i$ indexes beverage container and $t$ indexes the date. The $Can$ variable is an indicator for aluminum cans; $Can = 1$ for aluminum cans and $Can = 0$ for plastic bottles. The $Post$ variable is an indicator for the dates after Goldman Sachs purchased Metro International; $Post = 0$ prior to February 2010 and $Post = 1$ after. Additional controls are contained in $X_{i,t}$, including week by year fixed effects, month fixed effects, year fixed effects, and state fixed effects.

The results of this difference-in-differences model are presented in Table 3. The first column presents least squares estimates of equation 5 with no time or location fixed effects. Using Coca-Cola in plastic bottles as a control, we find that the average price of Coca-Cola in aluminum cans increased \$0.09 per multi-can package. Adding week of sample, month, year, and state fixed effects increases the estimated effect slightly to \$0.11 per multi-can package, which translates into about a half cent increase per can (column 2).

| Dependent Variable: Coca-Cola Price | | |
|---|---|---|
| | (1) | (2) |
| Aluminum Can x Post | 0.0933*** | 0.107*** |
| | (0.00683) | (0.00691) |
| $R^2$ | 0.979 | 0.984 |
| Controls Included: | | |
| State FE | NO | YES |
| Month and Year FE | NO | YES |
| Week-of-Sample | NO | YES |
| Sample Period: | | |
| Jan 2006 - Dec 2014 | X | X |

*Notes:* Each column contains the results for a separate regression. The data are reported weekly by county. The parentheses contain standard errors that are clustered by county. *** denotes significance at the 1 percent level.

Table 3: Coca-Cola Difference-in-Differences Model Results

Manipulation of the U.S. aluminum market increased the price of a can of Coca-Cola by 1 to 2 percent, and it is reasonable to assume that other consumer goods with aluminum packaging had similar, or even greater, increases. In 2015, aluminum packaging makes up only 20 percent of domestic aluminum consumption (USGS, 2016). Other categories, including automotive and consumer durables, account for a much larger share of domestic aluminum consumption, and could have had bigger increases. This points to a significant welfare loss caused by the price manipulation.

# 3 A Detection Algorithm

## 3.1 Manipulation in the U.S. Aluminum Market

The merry-go-round transactions in Goldman's Detroit warehouses were widely publicized by an article in *The New York Times* on July 20, 2013 titled, "A Shuffle of Aluminum, but to Banks, Pure Gold" (Kocieniewski, 2013). Though Goldman's activities in the aluminum market had been reported previously (Shumsky and Hotter, 2011) and market participants were aware of the merry-go-round (Wachtel, 2011), *The New York Times* article brought unprecedented attention to the issue. Three days later—on July 23, 2013—Goldman's aluminum market activities became the focus of the Senate Banking Subcommittee on Financial Institutions and Consumer Protection. During the following month, large aluminum consumers, including Eastman Kodak and Mag Instrument, filed more than a dozen lawsuits (U.S. Senate, 2014). The Senate's investigation continued through November 2014, when the committee released a detailed report on Goldman's aluminum market manipulation. This Senate report increased public scrutiny of the warehouse scheme, which only let up when Goldman sold the Metro International warehouses at the end of 2014. Aluminum premiums began falling within 4 weeks of the sale, but the aluminum market is still recovering (Figure 18). As of November 2015, the only LME warehouses with queues over 30 days were the warehouses in Detroit formerly owned by Goldman Sachs, where the queue stood at 206 days. Recall that, at its peak in late 2013, the queue length in Detroit reached nearly two years.

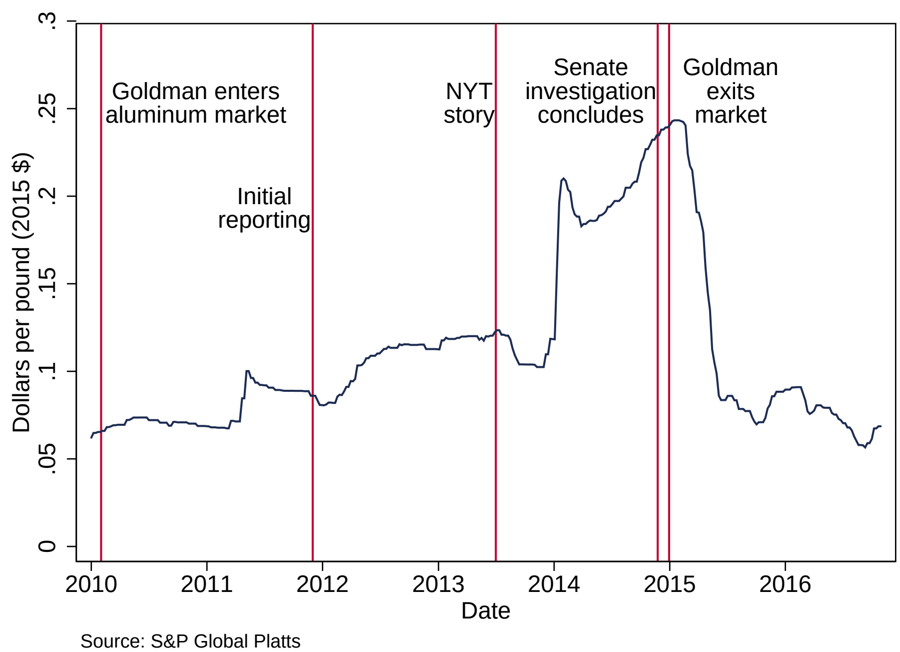



Figure 17: U.S. Aluminum Price (Midwest Premium)

Since regulators cannot foresee all manipulative practices, commodity markets remain susceptible. When facing these types of manipulative schemes, the best regulators can hope for is early detection. Though *The New York Times* publicized Goldman Sachs' aluminum scheme, regulators cannot expect the media to catch every case of manipulation. In this section, we present an econometric algorithm to aid regulators in detecting physical commodity market manipulation in real-time.

In particular, this algorithm is designed to identify Accumulation-Lift-Distribution (ALD) schemes—discussed in Section 2.2—that characterize manipulation. Ideally, regulators could use this algorithm to identify market manipulation in the Accumulation or Lift phases. This early warning signal would allow the regulator to thoroughly investigate the identified aberration and intervene, if necessary, to limit the damage done to markets.

The key to detecting Accumulation and Lift in a commodity market is identifying structural breaks in commodity inventory, queue, and premiums. A successful commodity manipulation scheme requires a trend break in each of these variables. In the Accumulation step, the market manipulator must acquire an unusually large inventory of the commodity, which can be detected as a break in the inventory trend. Likewise, the Lift step requires an abnormal spike in queue length and a subsequent increase in the regional price, which are captured by breaks in the cancelled-warrant and premium trends, respectively.

Following the intuition that underlies the difference-in-differences model in Section 2.1, the relevant inventory, queue, and premium series are the differences between those of the commodity of interest (aluminum in our case) and those of its production complement (copper). These series are plotted in (Figure 18). In other words, we are searching for a structural break in the differences between the inventory, queue, and premiums of aluminum and those of copper. Note that the existence of a statistically significant break across these three series does not prove causality, but rather indicates the possible existence of manipulation. Think of this detection algorithm as a fire alarm. If a fire alarm goes off in an office building, it does not necessarily mean that an office is on fire, but it does mean that people in the area should be on alert and call the firefighters.

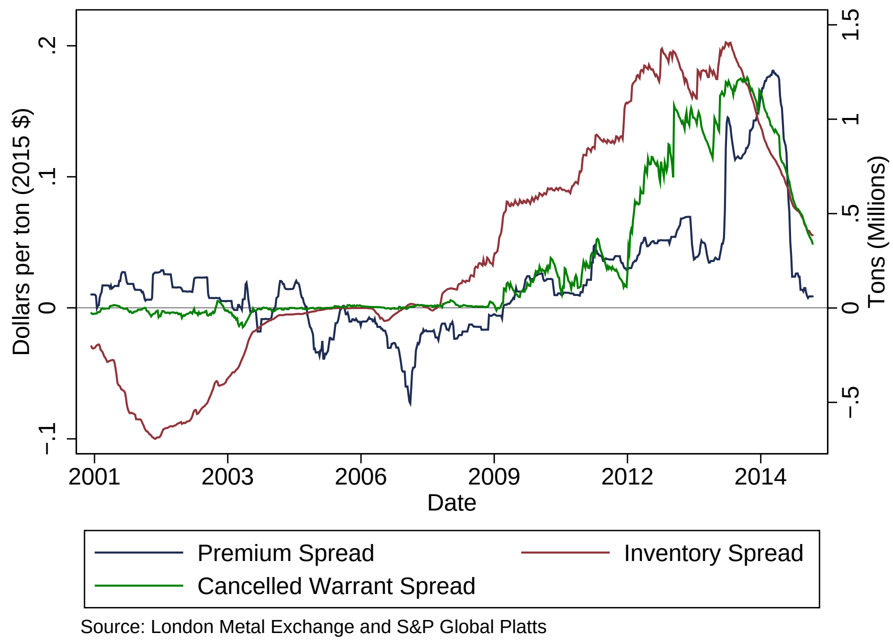



Figure 18: U.S. Aluminum-Copper Inventory, Queue, and Premium Spreads

This approach has two primary advantages. First, testing for a simultaneous break across multiple series improves the estimate by giving a tighter confidence interval around the estimated break date, relative to testing for a break with a single series. As shown in Bai et al. (1998), the confidence interval of a break estimate only decreases with the number of variables, not the sample size. The confidence interval is helpful for regulators who are interested in both the date of the break and the uncertainty. Second, using the difference between a commodity and its complement eliminates the effect of demand shocks. A detection algorithm is only useful if it generates a relatively small number of false positives. Since the differenced variables will not vary with demand shocks, the model should detect fewer spurious breaks.

We use the model developed by Bai et al. (1998)—and employed by Hansen (2001) and Bekaert et al. (2002)—to test for and date a structural break across multiple time series. Specifically, we estimate a vector autoregression (VAR) of the form

$$y_t = \alpha + \sum_{i=1}^{4} A_i y_{t-i} + \varepsilon_t \tag{6}$$

where $y_t$ is a $3 \times t$ vector containing the premium, inventory, and queue length variables. We estimate the model using weekly data on inventories, cancelled warrants, and premiums from November 1999 through December 2015.[9] The model tests whether there exists a date, $\gamma$, such that

$$\alpha + A_j = \begin{cases} \alpha_1 + A_2 \\ \alpha_2 + A_2 \end{cases} \tag{7}$$

In other words, for every week in the data set, we split the data into two sample periods: the sample period before the selected week and the sample period after the selected week. We then estimate the coefficients in the VAR model in equation 6 using each sample period and test whether there is a statistically significant difference between the coefficients estimated using the two different samples. The week for which the difference in model parameters is most statistically significant is the structural break date.

Importantly, break dates too close to the beginning or end of the selected sample cannot be identified, because there are too few observations at the end points to identify the model parameters. Thus, we use a trimming value of 5 percent—meaning that if there are 100 days in the sample, the model only tests for possible structural breaks dates between day 5 and day 95—to get around the problem.

Over the full sample period, November 1999 through December 2015, the model estimates a break date on of January 8, 2012, with the 90 percent confidence interval beginning on January 1 and ending on January 15, 2012 (Figure 19). In reality, regulators do not have the luxury of looking for structural breaks using the full sample period because they do not know when a manipulative scheme is occurring. To better simulate a real scenario, we run our algorithm only using the data available to regulators while the manipulation scheme was running. For instance, if we only use data available up until December 2012, we estimate the same break date, January 8, 2012, and confidence interval. This means that a regulator using our algorithm in late 2012 or early 2013 would have seen a statistically significant break in the physical aluminum market in late 2012, more than six months before the scheme was publicized by *The New York Times*.

[9]The model has four weekly lags. The lag length was determined by the Akaike Information Criterion, which is a lag selection procedure that tends to produce the most accurate models using small time series data sets (Ivanov and Kilian, 2005).

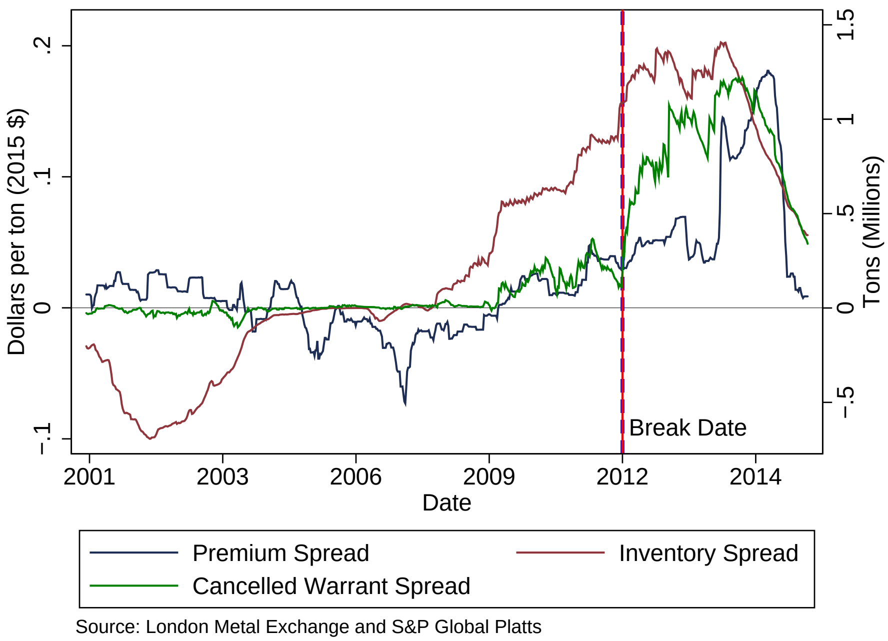



Figure 19: Test for Common Break in the U.S. Aluminum Market, Estimate and 90 Percent Confidence Interval

In sum, the model presented in this section provides a tool to assist regulators with monitoring physical commodity markets for manipulation. The model will identify breaks across premium, queue, and inventory levels associated with physical market manipulation, controlling for the effects of demand shocks. Though the algorithm does not prove market manipulation occurred, it is a useful first step in distinguishing between normal and abnormal commodity market fluctuations in inventories and premiums. Moreover, the algorithm is particularly useful because it is robust to changes in the sample window.

## 3.2 Manipulation in the European Aluminum Market

In late 2011, as the Midwest aluminum premium rose in response to record levels of cancelled-warrant inventories in the Detroit LME warehouses, a similar pattern emerged in the European aluminum market. Prior to December 19, 2011, the cancelled-warrant inventory levels for aluminum were extremely low at the LME warehouses in the port city of Vlissingen, Netherlands. Over the previous twelve months, only 0.008 percent of the total LME aluminum inventory in Vlissingen were cancelled-warrant inventories. In fact, during the first week of December 2011, the cancelled-warrant inventory level was zero for aluminum in the Vlissigen LME warehouses. That changed between December 19 and December 31, 2011, when the cancelled-warrant level for aluminum exploded from five tons to five hundred thousand tons (Figure 20). Aluminum warrant cancellations continued to grow throughout 2012, and the cancelled-warrant inventory represented an average of 49 percent of the total aluminum stock in the Vlissingen LME warehouses during that year.

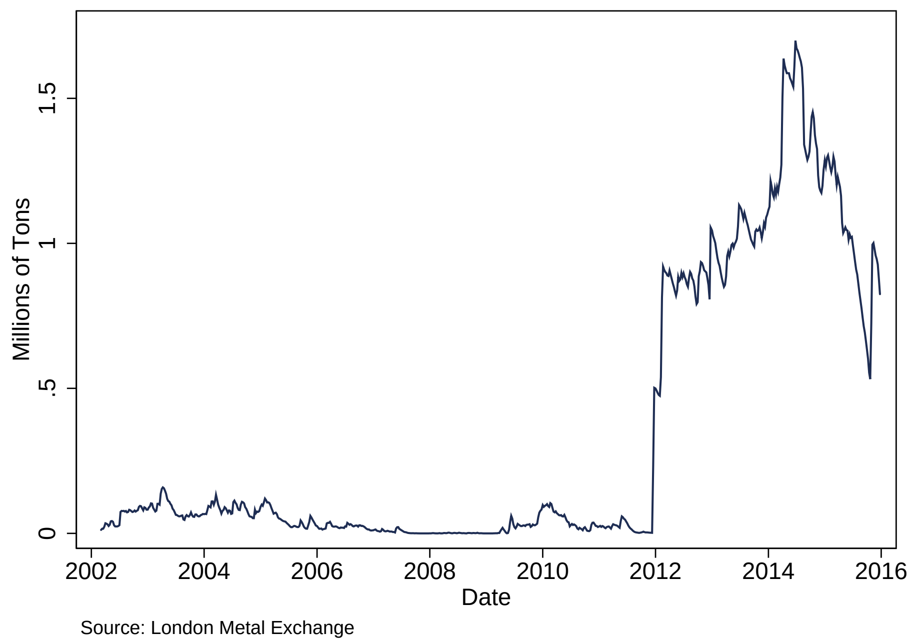



Figure 20: Cancelled-Warrant Inventories in Vlissingen LME Warehouses

In August 2011, six months before the enormous warrant cancellations, 27 of the 29 LME warehouses in Vlissingen were purchased by by Glencore, a commodity trading firm that was involved in large cancelled-warrant transactions at Goldman Sachs' Detroit warehouses. In Vlissingen, Glencore appeared to follow the same Accumulation-Lift-Distribution scheme employed by Goldman Sachs in Detroit. First, Glencore attracted record levels of aluminum to the Vlissingen warehouses by paying incentives to aluminum stockholders. These rebates more than doubled the aluminum in Glencore's warehouses, which eventually held as much as 93 percent of the total European LME aluminum inventory (Figure 21).

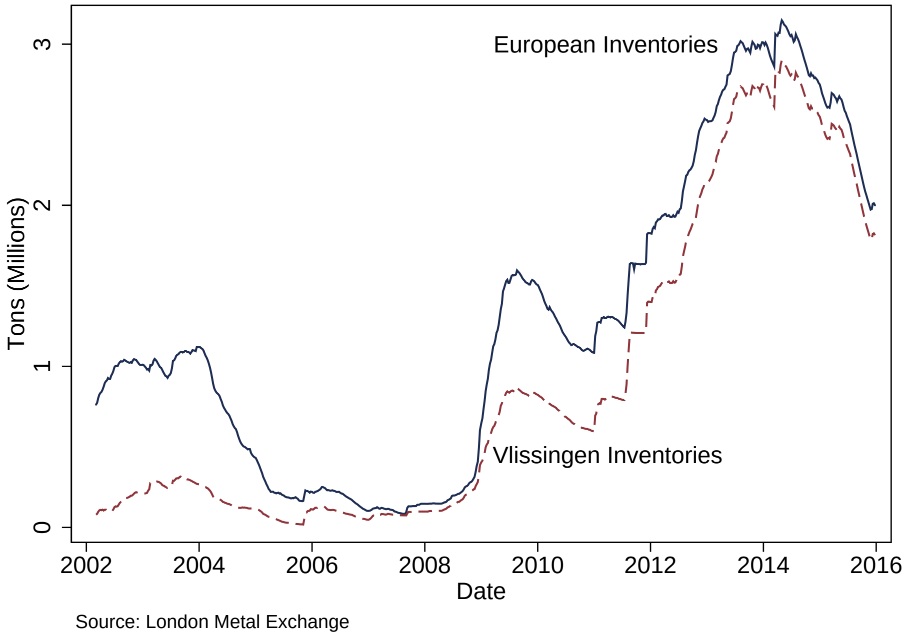



Figure 21: European Aluminum Inventories (LME)

Following a substantial buildup of aluminum inventory in the Vlissingen warehouses, aluminum warrants were cancelled at a record level. Unlike the U.S. Senate's investigation of Goldman Sachs, there has been no public investigation of Glencore's purchase and management of the Vlissingen warehouses. Without the data revealed by a public investigation, we do not know which firm(s) cancelled the aluminum warrants. Therefore, we do not address whether the practices at Glencore's Vlissingen warehouses caused the European aluminum premium to rise. Rather, we investigate whether our detection algorithm would have indicated the possible existence of manipulation. Our answer is yes: the fire alarm detects smoke in the area. Notably, this is supported by reporting at the time by the Financial Times (Farchy, 2012) and Reuters (Angel and Burton, 2012), which raised concerns about Glencore's practices.

Glencore responded to the cancelled warrants by only loading out the daily minimum tonnage required required by LME. This caused the load-out queue at the Vlissingen warehouses to rise with the level of cancelled warrants. The queue length at Vlissingen peaked in June 2014 at 774 days, over three months longer than the queue at Goldman's Detroit LME warehouses at the time. During this disruption of the local aluminum market, the European aluminum premium rose as the queue length restricted access to LME aluminum inventories which act as a backstop option for industrial aluminum consumers (Figure 22).

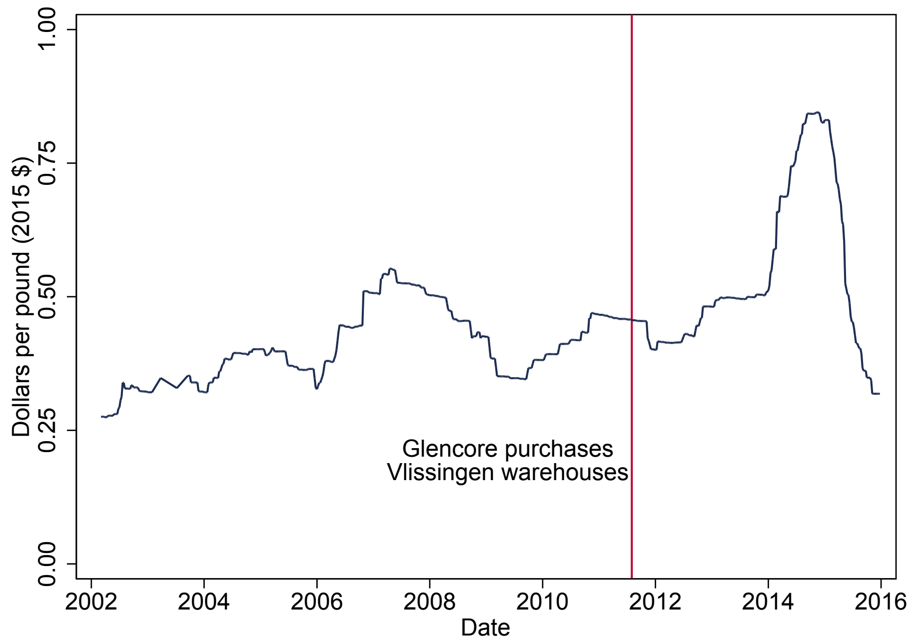



Figure 22: European Aluminum Premium

As in Section 3.1, we use copper as a complement for aluminum, and estimate whether there was a statistically significant structural break in the aluminum-copper European premium spread, the aluminum-copper European LME inventories, and the aluminum-copper European cancelled-warrant levels (a proxy for warehouse queues). Over the full sample period, March 2002 through December 2015, our model estimates a break occurred on December 15, 2011, with the 90 percent confidence interval beginning two weeks prior to the break and ending two weeks after the break (Figure 23). Not surprisingly, this break coincides with the jump in cancelled warrants that occurred six months after the Glencore purchased the warehouses. To better approximate the problem facing regulators, we again run our algorithm using only data available to regulators in real time to determine whether the scheme could have been detected earlier. If we use only data through December 2012, we estimate the same break date and an almost identical confidence interval. Once again, a regulator using our algorithm in late 2012 would have found that a statistically significant break occurred in the European aluminum market in late 2012.

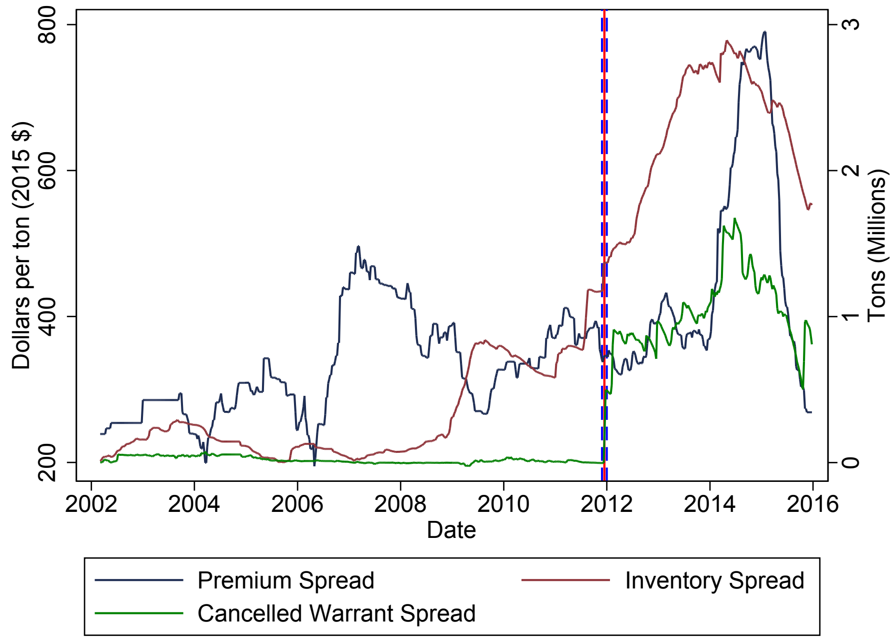



Figure 23: Test for Common Break in the European Aluminum Market, Estimate and 90 Percent Confidence Interval

Given that our detection algorithm estimates a statistically significant break in both the U.S. and European aluminum markets in late 2011, an objection could be that our algorithm is detecting shifts in a global aluminum market and not market manipulation. However, this phenomenon of skyrocketing cancelled warrants and extraordinary queue lengths did not occur throughout the global aluminum market. Of the 139 LME warehouses operating in June 2014, only the warehouses owned by Goldman Sachs in Detroit (which had a 681 day queue) and Glencore in Vlissingen (which had a 774 day queue) had non-zero load-out queues. At every other warehouse in the world, there was a zero-day wait time for aluminum; and this was true for every other metal traded on the LME. Given that about half of global LME inventories were not stored at warehouses owned by Goldman Sachs and Glencore, we would expect to see cancelled warrants and queues rising at other locations if our algorithm were simply detecting a global shift in the aluminum market.

# 4 Conclusion

Using data on the U.S. metal markets, we have examined the impact of Goldman Sachs' aluminum storage warehouses in Detroit on the regional price of aluminum. We show that following Goldman's entry into the aluminum market in 2010, the Midwest aluminum premium diverged sharply from the regional premium of its production complement, copper. In a difference-in-differences model, we estimate the regional price of aluminum rose \$0.07 per pound between 2010 and 2014 relative to copper, which shows that the aluminum premium increase was not the result of increased aluminum demand. Several instrumental variables for Goldman Sachs' purchase and operation of the Detroit warehouses demonstrate that Goldman likely caused the aluminum premium increase by restricting access to aluminum inventories in their Detroit warehouses. The spike in the regional price was followed by increased costs reported by industrial aluminum users, which appear to have been passed on to consumers.

Regional markets are an attractive target for manipulation because they require relatively small inventories to manipulate and they are not monitored as closely by regulators as the commodity exchanges, like the LME. Although the inventories accumulated in Detroit were likely insufficient for the purposes of moving the global market, they were sufficiently large to increase load-out queues in the region, causing prices paid by industrial aluminum users to significantly diverge from spot prices. This allowed the manipulation to continue without a perceptible impact on the global market, so regulators who only followed the LME spot and futures prices would not have noticed anything peculiar occurring in the aluminum market.

Though the particulars of Goldman's manipulation scheme were innovative, their scheme generally followed the Accumulation-Lift-Distribution model of market manipulation (Lang, 2004; Klein et al., 2012). We use this model of manipulation to develop a detection algorithm to help regulators identify future cases of regional physical commodity market manipulation. The algorithm uses time series methods to identify trend breaks in variables associated with Accumulation (physical inventories) and Lift (load-out queue length and regional price) of a metal relative to its production complement. Using only data available to regulators in 2012, our algorithm detects manipulative behavior in the aluminum market over six months before the scheme was broadly publicized by *The New York Times*. We also apply our detection algorithm to a suspected case of aluminum market manipulation in Europe. The results suggest that a similar manipulation scheme may have artificially raised the European aluminum premium between 2011 and 2014. Again, we emphasize that this algorithm does not prove causality, but it does strongly indicate that manipulation may have occurred.

These results have important policy implications. We show that despite the regulatory framework introduced by the Dodd-Frank Wall Street Reform and Consumer Protection Act of 2010, regional physical commodity markets remain vulnerable to manipulation. Financial institutions like bank holding companies and hedge funds can manipulate regional markets and cause a large, sustained impact on regional prices. Manipulation of regional commodity markets can leave spot and futures prices, which are closely monitored by regulators, untouched. Since regulators cannot foresee all manipulation schemes, early detection and investigation of these schemes is the best way to limit harm to commodity markets. The detection algorithm outlined in Section 3 can be an effective tool in detecting and deterring similar schemes in the future.

In addition, our work supports recent efforts by the Federal Reserve to limit the activities of financial holding companies with respect to physical commodities (Federal Reserve Board, 2016). In particular, the Federal Reserve has proposed a rule that would prohibit a financial holding company from owning, operating, or investing in facilities for the storage or distribution of commodities. Notably, the scheme studied in this paper would not have occurred if Goldman Sachs were not allowed to purchase and control the warehouses of Metro International. The proposed rule would also require increased reporting by financial holding companies that participate in physical commodity markets. This would likely improve the detection of manipulation in these markets.

Finally, some point out that this manipulation detection algorithm is susceptible to being gamed by the regulated parties, just like all other methods used by regulators to monitor financial markets. That is, if a firm wishes to manipulate the price of a commodity, and if it knows the complements used by regulators in the structural break test, then that firm could avoid detection by manipulating the inventory and price of both commodities. While manipulating the price of multiple commodities simultaneously is possible, it would be much more expensive and much more operationally difficult than manipulating the price of a single commodity. Moreover, regulators could make this prohibitively expensive by expanding the range of complements used in the detection algorithm. Therefore, at the very least, this algorithm can help regulators combat market manipulation by significantly increasing the cost of manipulation.

# References


Aalco (2005), *Aluminium - Specifications, Properties, Classifications and Classes, Supplier Data*. Aalco Metals Limited.

Abrantes-Metz, Rosa, Gabriel Rauterberg, and Andrew Verstein (2013), "Revolution in Manipulation Law: The New CFTC Rules and the Urgent Need for Economic and Empirical Analyses." *University of Pennsylvania Journal of Business Law*, 15, 357–418.

Abrantes-Metz, Rosa M and Sumanth Addanki (2008), "Is the Market Being Fooled? An Error-Based Screen for Manipulation." *SSRN Working Paper Series*.

Alcoa (2015), "Annual report, december 31, 2015." *, 2015*.

Aluminum Association (2015), *International Alloy Designations and Chemical Composition Limits for Wrought Aluminum and Wrought Aluminum Alloys*. The Aluminum Association.

Angel, Maytaal and Melanie Burton (2012), "Glencore Profits from Metals Backlog in Dutch Port." *Reuters*, April 27. Electronic.

Bai, Jushan, Robin Lumsdaine, and James Stock (1998), "Testing For and Dating Common Breaks in Multivariate Time Series." *The Review of Economic Studies*, 65, 395–432.

Bekaert, Geert, Campbell R Harvey, and Robin L Lumsdaine (2002), "Dating the Integration of World Equity Markets." *Journal of Financial Economics*, 65, 203–247.

Bloom, Nicholas (2009), "The Impact of Uncertainty Shocks." *Econometrica*, 77, 623–685.

BLS (2016), *Producer Price Index*. U.S. Bureau of Labor Statistics.

Bown, Chad (2015), "Global Antidumping Database." *The World Bank*, available at http://econ.worldbank.org/ttbd/gad/.

Chodorow-Reich, Gabriel (2014), "The Employment Effects of Credit Market Disruptions: Firm-level Evidence from the 2008-9 Financial Crisis." *The Quarterly Journal of Economics*, 129, 1–59.

Coffee, John, Jr. (2009), *Testimony of Professor John C. Coffee, Jr., Adolf A. Berle Professor of Law at Columbia University Law School*. Joint Meeting of the Securities and Exchange Commission and the Commodity Futures Trading Commission on "Harmonization of Enforcement".

Cumming, Douglas, Sofia Johan, and Dan Li (2011), "Exchange trading rules and stock market liquidity." *Journal of Financial Economics*, 99, 651–671.

Dalko, Viktoria (2016), "Perception alignment hypothesis: causality of herding?" *Qualitative Research in Financial Markets*, 8, 194–204.

Farchy, Jack (2012), “Warehousing Problem Hangs Over LME.” *Financial Times*, January 19. Print.

Federal Reserve Board (2016), “Proposed Rule Implementing Strengthened Prudential Requirements, including Risk-based Capital Requirements, for Physical Commodity Activities and Investments of Financial Holding Companies.” Proposed Rule by Board of Governors of the Federal Reserve System.

Fischel, Daniel R and David J Ross (1991), “Should the Law Prohibit “Manipulation” in Financial Markets?” *Harvard Law Review*, 105, 503–553.

Guglielmo Maria Caporale, Nikitas Pittis, Margarita Katsimi (2002), “Causality links between consumer and producer prices: Some empirical evidence.” *Southern Economic Journal*, 68, 703–711.

Hansen, Bruce E (2001), “The New Econometrics of Structural Change: Dating Breaks in US Labor Productivity.” *The Journal of Economic Perspectives*, 15, 117–128.

IMF (2015), “Commodity special feature.” Technical report, International Monetary Fund.

Irwin, Scott, Dwight Sanders, and Robert Merrin (2009), “Devil or Angel? The Role of Speculation in the Recent Commodity Price Boom (and Bust).” *Journal of Agricultural and Applied Economics*, 41, 377–391.

Ivanov, Ventzislav and Lutz Kilian (2005), “A Practitioner’s Guide to Lag Order Selection for VAR Impulse Response Analysis.” *Studies in Nonlinear Dynamics & Econometrics*, 9.

Kilian, Lutz (2009), “Not All Oil Price Shocks Are Alike: Disentangling Demand and Supply Shocks in the Crude Oil Market.” *American Economic Review*, 99, 1053–69.

Kilian, Lutz and Daniel Murphy (2014), “The Role of Inventories and Speculative Trading in the Global Market for Crude Oil.” *Journal of Applied Econometrics*, 29, 454–478.

Klein, Lawrence R, Viktoria Dalko, and Michael H Wang (2012), *Regulating Competition in Stock Markets: Antitrust Measures to Promote Fairness and Transparency through Investor Protection and Crisis Prevention*. John Wiley & Sons.

Knittel, Christopher and Robert Pindyck (2013), “The simple economics of commodity price speculation.” Technical report, National Bureau of Economic Research.

Kocieniewski, David (2013), “A Shuffle of Aluminum, but to Banks, Pure Gold.” *The New York Times*, July 20. Print.

Lang, Xianping (2004), *Manipulation*. Oriental Press.

LME (2013), “London Metal Exchange, Summary Public Report of the LME Warehousing Consultation.”

Mondolfo, Lucio F (2013), *Aluminum Alloys: Structure and Properties*. Elsevier.

Öğüt, Hulisi, M Mete Doğanay, and Ramazan Aktaş (2009), "Detecting Stock-Price Manipulation in an Emerging Market: The Case of Turkey." *Expert Systems with Applications*, 36, 11944–11949.

Pirrong, Craig (2004), "Detecting Manipulation in Futures Markets: The Ferruzzi Soybean Episode." *American Law and Economics Review*, 6, 28–71.

Pirrong, Craig (2010), "Energy Market Manipulation: Definition, Diagnosis, and Deterrence." *Energy Law Journal*, 31, 1–20.

Shumsky, Tatyana and Andrea Hotter (2011), "Wall Street Gets Eyed in Metal Squeeze." *The Wall Street Journal*, June 17. Print.

Tiwari, Aviral Kumar, KG Suresh, Mohamed Arouri, and Frédéric Teulon (2014), "Causality between consumer price and producer price: Evidence from mexico." *Economic Modelling*, 36, 432–440.

U.S. Senate (2014), "Wall Street Bank Involvement with Physical Commodities." Text from: *Committee Reports*. U.S. Government Printing Office.

USGS (2011), *Minerals Yearbook: Aluminum 2011*. U.S. Geological Survey for Mineral Industry Surveys Annual Reviews. U.S. Government Printing Office for Minerals Yearbook.

USGS (2012), *Minerals Yearbook: Aluminum 2013*. U.S. Geological Survey for Mineral Industry Surveys Annual Reviews. U.S. Government Printing Office for Minerals Yearbook.

USGS (2014), *Minerals Yearbook: Aluminum 2014*. U.S. Geological Survey for Mineral Industry Surveys Annual Reviews. U.S. Government Printing Office for Minerals Yearbook.

USGS (2016), *MineralCommodity Summary 2016*. U.S. Geological Survey for Mineral Industry Surveys Annual Reviews. U.S. Government Printing Office for Minerals Commodity Summaries.

Wachtel, Katya (2011), "Coca-Cola: Goldman Sachs Is Manipulating Metals Prices." *Business Insider*, June 20.

Williams, Jeffrey (1995), *Manipulation on Trial: Economic Analysis and the Hunt Silver Case*. Cambridge University Press.