

Exogenous Shocks and Misattribution of Responsibility for Economic Performance: Results From Survey Experiments*

Daniela Campello
Assistant Professor
Getúlio Vargas Foundation
daniela.campello@fgv.br

Cesar Zucco Jr.
Assistant Professor
Getúlio Vargas Foundation
cesar.zucco@fgv.br

This version: September 17, 2015

Abstract

Recent models of economic voting assume that citizens can discount exogenous factors when assessing governments' economic performance. Yet there is evidence that Latin American voters do not behave in such way, and attribute to presidents outcomes that are beyond their control. This paper presents three survey experiments designed to explore mechanisms that could potentially correct such misattribution, and therefore contribute to debiasing individual behavior towards government evaluation. Our results provide individual-level evidence of the misattribution found in aggregate studies of electorate behavior, and reinforce psychologists' skepticism towards prospects of mental decontamination, as we found very scant evidence that providing information, raising awareness, or increasing motivation to correct biases influenced individuals' evaluation of presidents' performance.

There is reasonable scholarly consensus that the economic vote can work as a mechanism of democratic accountability, as long as voters are capable of correctly assigning responsibility for economic performance. With economies progressively more integrated, this presupposes recognizing and discounting factors that influence economic outcomes but are beyond governments' responsibility. According to most recent theories, one way that voters develop this capacity is by observing their country's economic performance over time, and comparing it to other countries that experience common cycles (Duch & Stevenson 2008). Relative performance works, in that sense, as a *proxy* for government's competence.

Numerous studies, most of them in OECD countries, suggest that voters actually behave in accordance with these expectations. Duch & Stevenson (2008) show that economic voting is less likely to occur in small open European economies, in which voters perceive governments as limited in their capacity to influence domestic economic outcomes. Along the same lines, Kayser & Peress (2012) show that voters punish and reward governments based on their country's relative economic

*Prepared for delivery at the Annual Meeting of the American Political Science Association, San Francisco, CA, September 3-6, 2015. This is a very early draft, and our (mostly null) results are tentative. We thank Eduardo Andrade for suggestions and insights into the first versions of our experimental designs, and Evelyn Medrano for help in designing the survey.

performance. Based on data from the U.K., they contend that the media may play an important role in this process. Hellwig & Samuels (2007) find that greater economic integration weakens economic voting in a large sample of countries, and Ebeid & Rodden (2005) demonstrated that the connection between macroeconomic performance and incumbent governor success is weak in states in the U.S. whose economy is based on natural resources and farming, but strong elsewhere in the country.

Yet recent research has challenged these findings. In Latin America, for example, Campello & Zucco Jr. (2015) demonstrate that voters reward presidents that rule under high commodity prices and low U.S. interest rates, and punish those who govern in an opposite scenario, even though these factors are unequivocally exogenous to policymaking. The authors conjecture that low exposure to information about global outcomes — associated with a tradition of inward-looking development strategies, relatively low economic and political integration and limited access to media — might explain Latin American’s tendency to misattribute economic outcomes caused by exogenous shocks, and therefore to overestimate the impact of government policies.¹ Still, the exact mechanisms that explain voters’ incapacity to discount “chance” — a crucial element to understand voting behavior and democratic accountability — have yet to be established.

In this paper, we build on the psychology literature on misattribution and mental contamination (Wilson & Brekke 1994)² to design three experimental debiasing studies, which are aimed at identifying whether voters’ behavior can be credited to lack of information or to mental contamination. The first and second studies focus on lack of information as an explanation for voters’ misattribution; we presume that, if voters attribute chance to competence simply because they do not have access to information that allows them to appropriately assign responsibility for economic performance, once this information is available they will be able to correct their misattribution.

In the first study, we prime respondents with (positive and negative) information about their country’s relative economic performance, and observe whether this affects their assessment of government competence relative to a control group. The second study follows the same rationale, but leverages on the fact that some countries in Latin America are largely dependent on exports of a single commodity. Rather than relative performance, thus, we prime respondents with information about commodity price fluctuations, and observe whether this influences their evaluation of

¹Alcañiz & Hellwig (2011) also show that, even though economic integration is associated with an increased propensity to blame non-elected agents for economic performance, government policies are still largely perceived as the most important determinant of economic outcomes in the region.

²Wilson & Brekke define mental contamination as the process whereby a person has an unwanted response because of mental processing that is unconscious or uncontrollable. Here, we propose that attributing chance to competence is an unconscious unwanted response, in the sense that it is inaccurate and distorts the incentives for good policymaking. For that reason, we suppose that voters would not chose to act in such way once they were made conscious of their bias. Our strategy is to verify whether debiasing happens once voters are provided with the relevant information, or made conscious of their mental contamination.

presidents' competence relative to a control group. Finally, the third study focuses on individuals' cognitive process; we want to know whether, rather than lack of information, misattribution occurs due to mental contamination. We rely on recent studies in psychology to examine whether making respondents aware of a potential misattribution bias, indicating the direction of the bias, and motivating them to fix this bias prompts individuals to correct their misattribution. In our analyses, this means to discount the impact of exogenous shocks when assessing government competence.

We conduct Study #1 in Brazil and Studies #2 and #3 in both Chile and Ecuador, economies heavily dependent on a single commodity — copper and oil, respectively. Our results at the individual level are compatible with Campello & Zucco Jr.'s (2015) findings at the macro-level, as there is very scant evidence that any of our manipulations induce discounting exogenous factors when assessing their government's performance. Neither information about the bias, nor awareness of it or motivation to correct seem to affect respondents' evaluations. This is, in principle, coherent with psychologists' recent findings that, once biases are formed, they are very unlikely to change. If so, a promising venue for future research lies in exploring whether respondents are able to discount “chance” in settings in which they do not have any prior about government competence.

This paper is organized as follows. The next section reviews political science research on attribution of responsibility for economic performance, and in the subsequent one we examine the problem from the standpoint of psychological theories of mental contamination. The following three sections report each of the experimental studies, and the final section discusses tentative conclusions and explores potential venues for advancing this research.

1 Misattribution of Responsibility for Economic Performance

Voters' capacity to evaluate and sanction elected politicians is a central topic in the study of democracy. This debate has been traditionally anchored on two conflicting perspectives. The Michigan school, on the one hand, asserts voters' lack of both knowledge about political issues and coherent ideological structure (Campbell, Converse, Miller & Stokes 1960, Converse 1969). In contrast, the retrospective voting literature (Ferejohn 1986, Fearon 1999, Torsten, Roland & Tabellini 1997, Canes-Wrone, Herron & Shotts 2001) posits that, by using information at hand and shortcuts from like-minded citizens and parties, voters are capable of sanctioning and selecting the most competent and honest leaders. As Healy & Malhotra (2013) point out, recent research suggests real voters occupy a “middle ground” where they sometimes, but not always, make mistakes when observing economic outcomes, assigning responsibility for them, and deciding to punish or reward incumbents accordingly. Arguably, these mistakes only matter as long as they distort the incentives for good policymaking.

In this context, scholarly work has progressively focused on establishing the conditions under which the economic vote can actually work as a mechanism for improving democratic accountability. Along these lines, the literature on clarity of responsibility proposes that economic voting is more likely to promote accountability the more citizens can identify the party responsible for economic conditions (Powell & Whitten 1993). In the particular case of presidential systems, scholars have explored how electoral laws (Benton 2005) and separation of powers (Samuels 2004, Johnson & Schwindt-Bayer 2009) affect voter's capacity to attribute responsibility and hold politicians accountable, and have concluded that the presidency is the arena where the economic vote most often manifests in the region (Samuels 2004).

Increasing levels of economic internationalization motivated another strand of the literature to investigate the conditions that allow voters to punish and reward governments exclusively for economic outcomes they can control or influence. For this to happen, voters should be able to identify (and discount) exogenous components of their country's economic performance. Alesina & Rosenthal (1995) offered a theoretical foundation for this process by modeling economic growth as a function of a natural rate plus unanticipated shocks that are caused either by incumbents' competence or by an exogenous element. In this model, voters cannot directly identify the components of economic shocks but, by discounting the variance of these shocks, they can correctly assess incumbent competence.

Duch & Stevenson (2008) elaborated on this model by stressing that both elected and non-elected decision makers influence the domestic economy. The voters' conundrum, thus, is to identify and separate competence shocks — those associated with elected officials — from everything else that can be considered exogenous shocks. In countries where non-elected decision makers have relatively large influence over economic outcomes, the observed variance of exogenous shocks is substantially larger than that of competency shocks. In these settings, voters should not evaluate their governments based on the state of the economy, a proposition that finds support in a sample of European countries. Yet whereas Alesina & Rosenthal assume that voters know the distribution of exogenous shocks, Duch & Stevenson stress that they learn it by “observing global economic outcomes” (p. 150). This means that, even when the competence shock is substantial, voters can only extract the information necessary to reward or punish leaders by making a relative assessment of their country's economic performance.

The logic that relatively large economic shocks weaken the economic vote finds echo in the results of several other recent studies. Hellwig & Samuels (2007), for instance, find that greater exposure to trade and capital flows decreases economic voting in a large sample of countries. Alcañiz & Hellwig (2011) show that economic integration is also associated with an increase propensity to

blame non-elected agents (such as the IMF) for economic performance in Latin America. Ebeid & Rodden (2005) demonstrated that the connection between macroeconomic performance and incumbent governor success is weak in states in the U.S. whose economy is based on natural resources and farming, but strong elsewhere in the country. Finally, Kayser & Peress (2012), show that voters punish and reward governments mostly based on their country's *relative* economic performance, which works as a proxy for government competence. Based on data from the U.K., Kayser & Peress further contend that the media may play an important role in enabling this process.

Nevertheless, the same rationale that explains Europeans' capacity to assess relative performance suggests that voters not exposed to the same levels of economic integration, and therefore to information about global economic outcomes, should be less capable of discounting exogenous shocks. In such a context, voters will overestimate the impact of policymaking on economic outcomes, punishing and rewarding incumbents for factors they can not control. In South America, for example, where economies are heavily reliant on commodity exports and on foreign savings, Campello & Zucco Jr. (2015) have shown that presidents that govern in periods of high commodity prices and low U.S. interest rates are far more popular and successful in electing their successors than those that preside under opposite conditions. Considering that neither factor falls under executives' control, this evidence reveals the occurrence of misattribution, and that voters' assessment of government competence are biased. These results are similar of those reported by Leigh (2009) in a sample of 268 countries, in which citizens are shown to vote according to *absolute*, not *relative* economic performance.

Campello & Zucco Jr. conjectured that voters' misattribution might result from insufficient information. Due to the prevalence of inward-looking developing models during much of the 20th century, very limited levels of economic or political integration, and citizens' low access to media,³ voters might not have enough information to benchmark their country's economic performance and, for that reason, they overestimate the impact of policymaking. Yet this is one out of many possibilities.

Scholarly work has documented numerous instances in which irrelevant events seemed to have a substantial impact on voters' support for the government. In an influential article Achen & Bartels (2006) show that Americans are less likely to vote for incumbent governors in the aftermath of

³Average circulation of daily newspapers in the region is only about 54 per 1,000 people, compared to 289 in the United Kingdom, 267 in Germany, and 313 in the Netherlands (*Total average circulation per 1,000 inhabitants*, UIS Data Centre and UNESCO Institute for Statistics, available at <http://data.un.org/>). Internet usage in Latin American countries stands, on average, at 48% of the population whereas in the European Union this figure is at 75%, and reaches 90% in some Western European countries (*Internet users per 100 inhabitants*, World Development Indicators, available at <http://databank.worldbank.org>).

events like shark attacks and droughts, completely unrelated to policymaking. Gasper & Reeves (2011) and Cole, Healy & Werker (2012) also document voters’ punishment of incumbents after severe weather damage in the U.S. and in India, respectively. Similar results were also obtained in experimental settings Huber, Hill & Lenz (2012).

Healy, Malhotra & Mo (2010) also report the effect of irrelevant events on voting behaviour; they show that a win in local college football games before election day increases the vote for the incumbent in Senate, gubernatorial, and presidential elections, with the effect being larger for teams with stronger fan support. Their design is particularly compelling as, contrary to natural disasters, no government action would be expected in response to game outcomes and the public would almost certainly not relate them to incumbent performance. In a second experiment, Healy, Malhotra & Mo (2010) go a step further to examine mechanisms that could contribute to voters’ debiasing. Based on Schwarz & Clore’s (1983) findings that the effect of externally-induced mood on political judgments can be eliminated when subjects are explicitly exposed to the irrelevant information, they prime respondents into thinking about the results of the games before asking them about support for the government, and find that the treatment reduces the impact of wins on support for the government.

Our case is in many ways comparable to those mentioned above. Even though the effects of exogenous shocks on the economy are harder to disentangle, or discount, than the well-being effects of college games or shark attacks, they are arguably similarly unrelated to incumbent competence. Conceiving voters’ attitude of not discounting exogenous factors as a specific case of misattribution allows us to tap into a vast body of scholarly work in psychology that has examined sources and possible solutions to cognitive biases, which we discuss in the next section.

2 Correcting Misattribution

Wilson & Brekke (1994), in a comprehensive categorization of the literature on cognitive biases, distinguish those that arise from not knowing or not applying some established inferential rule from those that arise from “mental contamination,” a situation in which a person ends up with an “unwanted judgement, emotion, or behavior because of mental processing that is unconscious and uncontrollable” (p.119).

The first type of bias can be dealt with by providing information about the rule that should apply to the situation at hand. Granted, simply knowing the rule does not guarantee that it will be used correctly, but knowledge of the rule alone should reduce the incidence of misattribution. The second type of error, however, is harder to deal with because avoiding contamination implies clearing successive hurdles that are hardwired into the basic nature of our cognitive system,

namely one needs to be aware of the “unwanted mental processing,” motivated to correct the bias, aware of the direction and magnitude of the bias, and able to control the process that produces the response (i.e. it cannot be an unconscious process). Because all hurdles need to be cleared in order to avoid misattribution, studies that attempted to debias participants’ judgements have arrived at contradictory results. In the words of the authors,

Some studies have shown that an increase of people’s awareness eliminates mental contamination; some have found that awareness causes people to adjust insufficiently, leading to under correction; some have indicated that awareness causes people to adjust their responses too much, leading to overcorrection; and some have shown that awareness does not cause people to adjust their responses (Wilson & Brekke 1994, p. 130).

Overall, the authors express pessimism about people’s ability to avoid mental contamination: “due to lack of awareness of mental processes, the limitations of mental control, and the difficulty of detecting bias, it is often very difficult to avoid or undo mental contamination” (p.117).

In order to study misattribution of responsibility for economic performance, we interpret a lack of information about the actual connections between exogenous shocks and observed economic outcomes as a bias originating from not knowing or not applying some established inferential rule that could, in principle, be learnt and applied. In contrast, the urge to automatically attribute the current economic performance to the president even if one knows, in principle, that such attribution should not happen, is treated as a case of mental contamination, the second type of bias. As discussed above, exogenous fluctuations in export prices are as unrelated to government competence as college athletic results or droughts. Yet whereas any voter would agree that she should not chose incumbents based on the weather or on game wins, the reasons not to support a government that reigns over a good economy (or to support one that leads during a crisis) might not be as self-evident. Even if one realizes that some share of economic outcomes is caused by factors beyond governments’ control, how relevant this share is, how to discount it, and even whether discounting is worth the effort, are much more complex processes than those required in the previous experiments. This is particularly true if, as Healy, Malhotra & Mo (2010) suggest, the “strong correlation between economic performance and the probability of incumbent reelection may be that voters’ general sense of well-being serves as a conduit between the state of the economy and electoral outcomes” [p.4].

Debiasing, from a cognitive stance, should therefore be harder in a setting in which respondents are expected to discount a factor that is highly influential on well-being but does not fall under

the control of the president (the price of oil or copper) when evaluating her job performance. And definitely much less likely to occur than what is implied by the mainstream economic voting literature.

In order to determine the micro foundations of Campello & Zucco Jr.’s (2015) macro results, we conducted what can be labeled as a series of “debiasing studies,” in which the goal is to eliminate misattribution in subjects’ reasoning. As Wilson & Brekke (1994) point out, debiasing studies are more likely to succeed when they help participants clear all the said hurdles. In other words, when “they make people aware of the unwanted processing, motivate people to resist it, when people are aware of the direction and magnitude of the bias and have sufficient control over their response to correct for it” (Wilson & Brekke 1994, p.133). Additionally, it is easier to correct biases that arise from decisions subject to conscious mental processes. In our case, evaluating presidents’ job performance seems to be clearly within the realm of conscious information processing, and therefore we do not worry about it.

We designed our studies to examine the several possible mechanisms described above, which can be divided into informational and cognitive. Studies #1 and #2 deal with the informational side. The first study provided respondents with information (both positive and negative) about the relative economic performance of their country (Brazil). In the second study, we dealt not with relative performance, but with information about the behavior and impact of the most relevant export commodity to their country’s economy. Here we experimented with several bits of the causal chain. We informed people about the relevance of a commodity for the country’s economy (oil in Ecuador, copper in Chile), about recent changes in the price of the commodity, about how commodity prices are not under the president’s control, and, finally, by telling respondents explicitly that their country’s economy was hurt by the falling commodity prices.

Because we believe (and have some evidence to suggest) that our subjects in Study #2 already possessed the information concerning the behavior and impact of commodities information, we also explored, in Study # 3, whether it was mental contamination, rather than insufficient information, that was driving the lack of discounting. Therefore, we experimented with manipulations that should help respondents clear the hurdles described in Wilson & Brekke (1994). In general, becoming aware of unwanted mental processing requires relatively strong and credible manipulations. We sought to make people aware of biases by telling subjects we were studying the impact of commodities, and by focusing their attention on the information that we assumed (and, in some condition checked) that they already possess, strategies previously adopted in Healy, Malhotra & Mo (2010) and Schwarz & Clore (1983). We also attempted to make them motivated to correct biases by telling respondents that most people fail to discount factors beyond presidents’ control

when assessing government performance, and that experts agree that democracy works better when these factors are discounted. This information provided people with the “direction” of the possible bias and some cues about its magnitude. These manipulations were implemented cumulatively, as psychological research suggests that all possible hurdles in the cognitive process (i.e. awareness, motivation, direction) have to be cleared in order for debiasing to be possible.

3 Study #1: Information on Relative Performance

In this study, we examined whether providing information about a country’s economic performance relative to peer countries affects the evaluation of the president. We fielded this experiment in Brazil, embedded in two waves of the 2014 Brazilian Electoral Panel Study (BEPS), carried out during the 2014 presidential electoral campaign.⁴

3.1 Design of the Study

The study consisted of three simple experimental conditions, embedded at the end of a larger survey. The outcome variable was respondents’ rating of President Dilma Rousseff’s performance in managing the economy.⁵ Prior to answering this question, participants were randomly assigned to one of the following three different conditions in which we manipulated relative information about the country’s performance. Between the manipulation item and the outcome question, respondents answered filler items about media attention.

- Control: No information about relative performance was provided
- Negative: Respondents were asked whether they knew that in the previous three years the Brazilian economy had grown less than the economies of countries “such as Argentina, Chile, and Mexico”
- Positive: Respondents were asked whether they knew that in the previous three years, the unemployment rate in Brazil had been lower than in countries “such as Argentina, Chile, and Mexico”

One obvious limitation of the study is the asymmetry between the type of information provided in the positive and in the negative manipulation, as respondents might care more about growth than unemployment (or vice-versa). However, there is no way one can provide truthful information, with

⁴The first wave of this survey was carried out with household interviews, in a nationally representative sample of 3,000 respondents. The follow-up waves interviewed subsets of this original sample over the phone.

⁵Actual wording was: “Como o(a) Sr.(a) avalia o desempenho da Presidente Dilma no que diz respeito apenas à política econômica do país? É ótimo, bom, regular, ruim ou péssimo?” (“How do you evaluate President Dilma’s management of the economy? Is it very good, good, regular, bad or very bad?”)

Table 1: Between Subjects Treatment Effects (Study #1)

Controls	Linear		Ordered Logit	
	No	Yes	No	Yes
Negative Manipulation	-0.175	-0.173	-0.303	-0.304
<i>p-value</i>	0.013	0.012	0.032	0.032
Positive Manipulation	-0.062	-0.048	-0.067	-0.040
<i>p-value</i>	0.365	0.491	0.351	0.382

Estimates are based on comparison of each treatment condition to the control condition. P-values are computed from robust standard errors.

negative and positive content, and using the same indicator and the same countries as comparison. As we prioritised holding countries fixed, and using only truthful information, we were forced to vary the particular aspect of the economy used in positive and negative primings.

3.2 Sample Descriptives

The sample consisted of 1,203 of the more than 3,000 interviewees in the BEPS. Participants were first interviewed in-person in June of 2014 as part of the initial round of the BEPS. The initial wave also included the same question used as the outcome variable in the experimental survey. In mid-July and late-August, respondents were interviewed once again, this time over the phone, and were subject to the manipulation described above. A total of 397 respondents were assigned to the control condition, 403 to the negative condition, and 402 to the positive condition. Balance statistics are reported in the Appendix. There were no observed differences between those that were re-interviewed in the July and those that were re-interviewed in August.

3.3 Results

Figure 1 reports both the between- and within-subjects results of the study. The effects of relative information are limited to the negative manipulation in the between-subjects analysis. The positive manipulation yielded no effect, and neither manipulation generated any effect in the within-subjects comparison.

Table 1 makes these graphical results clearer. Regardless of whether we simplify the analysis by transforming the outcome variable to a linear five-point scale, or whether we employ an ordered logit approach, we find that the negative manipulation did yield significant effects on the assessment of the presidents job performance, while the positive manipulation did not. These effects, however, cannot be overstated. Although statistically significant, the negative between-subjects effect corresponds to less than 0.2 standard deviations of the dependent variable.

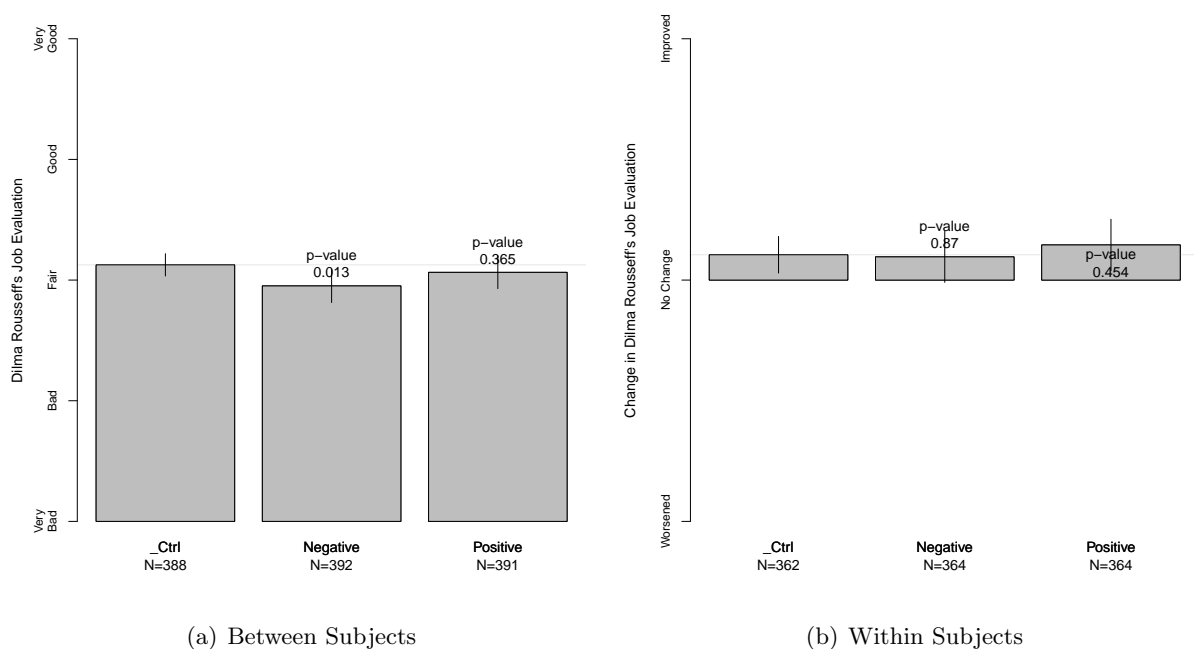


Figure 1: Treatment Effects (Study #1)

Robust standard errors are shown, and p-values refer to t-tests of each treatment condition relative to the control condition. The N is smaller in the within subjects because some respondents filed to answer the outcome question prior to treatment.

We also examined some heterogenous treatment effects, namely whether results held for partisans and non-partisans, and for respondents with different levels of political knowledge. Surprisingly, we found no effect of partisanship on the results, but found that for respondents with higher levels of political knowledge, both the positive and the negative manipulations had significant effects, on the expected direction, on voters' assessment of the president.⁶ While we would have expected that the marginal effect of information should be larger for respondents with less a priori information, higher levels of information might very well be capturing greater sophistication (and capacity to process information). In this sense, the results suggest that more sophisticated voters would be more capable of correcting misattribution when provided with information. This, in turn, implies that the prospects for such correction in the Latin American context are meager. It is important to keep in mind, however, that randomization was not performed conditional on knowledge levels, which means the study was not designed to answer this question.

⁶Political knowledge was assessed in the face-to-face wave of the BEPS, through factual questions about domestic and international politics. We produced a knowledge index based on the answers to these questions, used to various analyses embedded in the project.

4 Study #2: Information about Exogenous Shocks

In this study, we examined whether providing information about different aspects of the relationship between one country’s main export commodity and economic outcomes affects the evaluation of president’s performance. We fielded the experiment in Ecuador and Chile, two Latin American countries highly dependent on exports of a single commodity. Oil corresponds to 52% of Ecuador’s exports while copper amounts to 51% of Chilean exports.⁷ The price of both commodities has fallen over the past years, though the fall in oil prices has been more abrupt and concentrated in the last 12 months. The main difference between the two countries, however, is the fact that President Rafael Correa of Ecuador (still) has relatively high approval rates, while President of Chile Michelle Bachelet’s approval rates are lower. This might, in part, be driven by the relatively recent drop in oil prices, and also by the fact that substantial funds from China have propped up the Ecuadoran economy as oil prices have tumbled.

4.1 Design of the Study

We implemented a simple four-condition survey experiment in which we manipulated the information about the country’s main commodity that was provided to respondents. For all respondents, our very short survey started with four demographic questions, followed by an attention screener warning, and a question about their main source of political and economic information. The experimental items were then applied, and were followed by questions on party identification, education, and household items. All items in the experimental block, as well as the question immediately preceding it and the attention screener that followed were timed, to prevent respondents from searching for information on the internet prior to answering.

The experimental block, which was inserted roughly in the middle of the survey, included both our treatment manipulations and the main outcome question, which was the respondents’ evaluation of the job performance of the president. The following four manipulations were carried out:

- A (Baseline control group) Respondents were asked about their assessment of the presidents job performance without receiving any other relevant information.
- B Respondents were asked whether they had seen any news about the fact that Ecuador/Chile’s economy and government finances are hurt when oil/copper prices are low, and then asked about their evaluation of the presidents’ performance.

⁷Overall, Ecuador’s export portfolio is more concentrated than Chile’s.

- C Respondents were asked the same question as in B, and also whether they had seen any news about oil/copper prices having fallen by 50%/30% over the last year/few years.
- D Respondents were asked the same questions as in C, and also whether they has seen any news about the fact that oil/copper prices are set in international markets and that the president has no influence over them.

In conditions A, B, and C, after respondents assessed the president’s performance they were asked about their opinions about the information items they had not seen. For instance, respondents in condition A were asked, after evaluating the president, what they thought had happened to oil/copper prices, whether low prices of oil/copper hurt or benefited the country’s economy, and the extent to which presidents could influence the prices of the relevant commodity. These are not experimental items per se, but are useful to assess the baseline levels of information in the sample. Throughout the text, we refer to these as “knowledge” questions.

The attention screener item we employed included a question header that was considerably longer than any other question in the survey. The first line suggested the question was about the media. The final line asked respondents the frequency with which they read or watched the news. The answer options ranged from rarely to daily, and were therefore typical of such a question. Buried in the middle of the question header, however, there was an instruction to ignore everything else, select the option “other” and type a specific value in text box beside it. This question was also timed, so as to not stand out relative to the timed questions that preceded it. These features, combined, make this a relatively demanding attention screener, as a respondent paying even moderate attention to the survey would likely be caught. The screener is further discussed in the appendix.

4.2 Sample Descriptives

We recruited subjects by posting ads in Facebook geared towards residents of Ecuador and Chile. These ads advertised the chance to win an iPad if respondents participated in a short survey.⁸ After clicking on the ads, participants were taken to the informed consent page of our survey, hosted by Qualtrics. Those that consented were taken to the main survey. Upon completion of the survey, respondents had the option to participate in the lottery by providing us with an email address. Email addresses were kept separate from responses to the survey, and no other identifying information was collected.

We fielded the Ecuador study between 2015-08-04 and 2015-08-17. In this report, we use the dataset downloaded from Qualtrics on 2015-08-19, which contained a total of 651 responses. After

⁸The lottery and the delivery of the prize will be conducted after the final study in this series.

eliminating those that originated outside of Ecuador, under-age respondents, those that originated from repeat IP addresses, and those that dropped out prior to answering the outcome question, we were left with 542 responses. The median time to completion of these surveys was 4.2 minutes and 29.2% of respondents passed the attention screener. This share was lower than in previous experiences, probably because we set up a relatively demanding screener. The question was much longer than other questions, response to the question was timed, and the answer categories induced one to easily guess what the question was about. Although we report answers from the whole sample, we focus primarily on the sub-sample of respondents that passed the screener.

The informational study in Chile was fielded between 2015-08-13 and 2015-08-16,⁹ and data were downloaded from Qualtrics on 2015-08-16. After eliminating responses that originated outside of Chile, under-age respondents, those that originated from repeat IP addresses, and those that dropped out prior to answering the outcome question, we were left with 493 responses. Similarly to the Ecuador study, the median time to completion was 4.2 minutes and 27.8% of respondents passed the attention screener. Sample demographics and balance statistics are reported in the appendix.

4.3 Results

In both countries, we observed no treatment effects in the complete set of respondents. This is not entirely surprising as the medium we used certainly includes many “low quality” participants. We focus, therefore, solely on respondents that passed our relatively demanding attention screener.

Figure 2 describes the most basic results of the study that, as we describe in more detail in Table 2, are essentially unchanged with the inclusion of the control variables age, sex, whether the respondent identify with any party, whether the respondent has a family member living abroad, and a socioeconomic status index we constructed.¹⁰ Among high attention respondents in Ecuador there are some positive results. Those in condition B rated President Correa’s performance 0.34 higher (in a 5 point scale) than those in the control group, though this difference falls short of statistical significance (p-value=0.169 on a two-sided t-test). Respondents in C were more than half a point more favorable to the president, which amounts to non-negligible change of 0.48 standard deviations (p-value=0.022, which is lower than the target p-value at the 0.05 level adjusted for multiple comparisons). This suggest that the information added in condition B (low oil prices hurt the economy) and the information added in condition B (oil prices fell) together yielded significantly higher evaluations.

Surprisingly (for us), respondents in condition D did not rate the president higher than those in

⁹In Chile, Study #2 and #3 (below) were fielded together, but we separate them here to streamline presentation.

¹⁰More information about the control variables is provided in the balance statistics section of the Appendix.

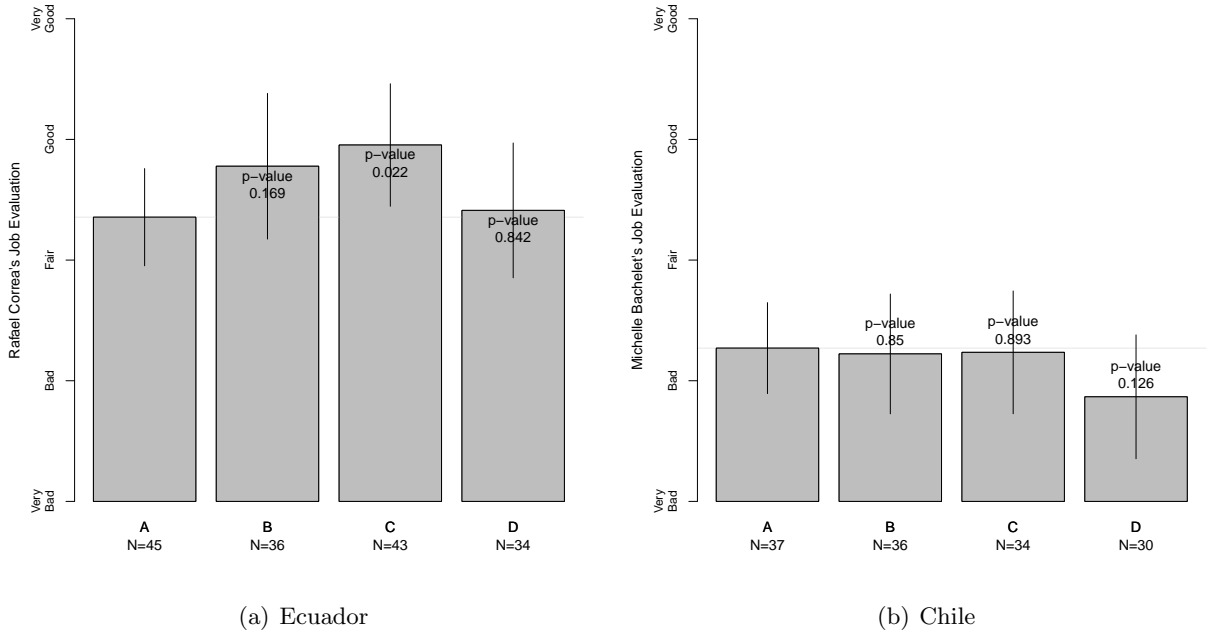


Figure 2: Government Evaluation by Manipulation (Attentive Respondents)

Figures show results for high-attention respondents only. Results for the complete sample are shown in the appendix. Robust standard errors are shown, which are algebraically equivalent to p-values of t-tests of the differences in means between treatment groups and the control group (Samii & Aronow 2011).

the control group. This is intriguing because our manipulations were cumulative, so manipulation D included the two priming pieces of information in C, plus an additional one on how the president cannot affect commodity prices. It is possible, therefore, that highlighting that presidents lack power to control commodity prices generates a negative assessment of the president that counteracts the positive effect of previous manipulations. This suggests that the information in manipulation D was not read, as we expected, as an exonerating circumstance. This possibility, in fact, is further supported by the fact that in Chile, too, manipulation D yielded lower support for the president than previous manipulations.

In Chile no condition produced any effect. One possibility is that Chileans are better informed than Ecuadoreans and, therefore, information produces less of an effect. This interpretation, however, is not supported by answers given to the knowledge questions. Table 3 shows that, in both cases, knowledge was quite high for the questions about the recent behaviour of commodity prices, and the impact of oil/copper prices on each country's economy, and even higher in Ecuador.

If any information were to have had a positive effect on government evaluation, it should have been precisely the one contained exclusively in manipulation D, as this was the only item for which about half of the respondents were mistaken. These respondents, in other words, thought that their presidents could have done something to influence commodity prices and improve the

Table 2: Treatment Effects (Study#2)

(a) Ecuador					(b) Chile				
	All Respondents		Attentive Only			All Respondents		Attentive Only	
	Mod. 1	Mod. 2	Mod. 3	Mod. 4		Mod. 1	Mod. 2	Mod. 3	Mod. 4
A (average)	2.560	2.429	2.356	2.328	A (average)	1.289	1.176	1.270	0.949
B	0.109	-0.016	0.422	0.388	B	0.090	0.077	-0.048	-0.013
<i>p-value</i>	0.510	0.927	0.167	0.223	<i>p-value</i>	0.511	0.611	0.850	0.964
C	0.121	0.015	0.598	0.492	C	-0.017	-0.035	-0.035	-0.037
<i>p-value</i>	0.454	0.925	0.021	0.054	<i>p-value</i>	0.907	0.817	0.893	0.897
D	0.087	-0.074	0.056	-0.040	D	-0.000	-0.007	-0.404	-0.336
<i>p-value</i>	0.586	0.657	0.842	0.889	<i>p-value</i>	0.998	0.963	0.126	0.242
Controls	No	Yes	No	Yes	Controls	No	Yes	No	Yes
N	542	452	158	142	N	493	444	137	130
Target p-val	0.049	0.046	0.043	0.041	Target p-val	0.034	0.037	0.042	0.041

Estimates for condition A are the average rating of the president in a 0-4 scale in the control group. All other estimates are differences between average values in each conditions and that of condition A. P-values are computed from robust standard errors. Target p-values are computed by simulation and indicate the threshold for significance at the 0.05 that accounts for the Family-Wise Error Rate due to multiple comparisons.

economy. When we told them otherwise they did not discount this, but rather, rated the president worse, probably punishing him/her for not being able to affect commodity prices.

Table 3: Correct Answers to Knowledge Questions

	% Correct	
	Ecuador	Chile
Oil/Copper Prices Fell Recently	81	70
Low Oil/Copper Prices Hurt the Economy	84	89
President's Does Not Control Prices	58	48

Overall, we judge the evidence in support of the idea that simple lack of information generates the misattribution, and that providing information can correct misattribution to be quite weak.

5 Study # 3: Manipulating the Cognitive Process

In this study we shifted away from the provision of information and attempted to correct the misattribution by employing manipulations that operate in the cognitive process and address the main sources of mental contamination. Following Wilson & Brekke (1994), we sought to raise awareness of the misattribution bias, provide motivation to correct the bias, and indicate the direction of the bias.

To do this, we implemented a simple five-condition survey experiment, whose general structure was very similar to Study #2, except that the conditions in the experimental block addressed possible sources of mental contamination. The main outcome question, as in Study #2, was respondents' evaluation of the president's job performance, though we also included a second

outcome question in some manipulations (details below) as part of the attempt to provide voters with motivation to correct the bias. Overall, we used relatively heavy handed manipulations because, based on Study #2, we suspected that there would be little or no debiasing going on.¹¹

- A (Baseline control group) Identical to the control condition in Study #2; respondents were asked about their assessment of the President’s job performance without receiving any other relevant information.
- E Respondents were told, at the start of the experimental block, that the study attempted to gauge the effects of oil (Ecuador) and copper (Chile) prices on the popularity of Ecuadoran and Chilean presidents. This was an indirect attempt to raise awareness of a potential misattribution bias, as done in Schwarz & Clore (1983) and Healy, Malhotra & Mo (2010).
- F Respondents were provided with this same statement as in Condition E, but were also informed that recent studies indicated that voters tend to evaluate presidents for things they cannot control, and then asked whether they thought they were subject to the same “error”. This was a more explicit way to raise awareness of a potential misattribution bias.
- G In addition to the awareness primes in condition F, an item was inserted stating that experts suggested that it would be better for democracy that voters only evaluated presidents by results they can control or influence. This attempted to provide a normative motivation to correct misattribution bias.
- H In addition to the manipulations in condition G, an item was inserted stating that the evaluation of presidents of commodity-rich countries tends to follow that commodity’s price even though the president cannot do anything about it. This was an attempt to make the direction of the bias clear.

In parallel to these basic manipulations that should impact the main outcome question, we also added a second outcome to conditions F, G, and H. After evaluating the president, respondents were asked to imagine a group of “experts” that knew which outcomes fell under control of the president and which did not, and asked how the respondent believed these experts would evaluate the president’s performance. This was designed as a further attempt to provide respondents with motivation to correct their bias, as it focused respondents attention on the behavior of somebody else.

¹¹We named the treatment conditions continuously relative to the Study #2 (i.e. E, F, G and H) to avoid confusion between the two.

5.1 Sample Descriptives

We employed the same recruitment and fielding strategy as in the previous study. This Ecuador study was fielded between 2015-08-11 and 2015-08-12, and the data used here were downloaded from Qualtrics on 2015-08-13. We had a total of 588 responses. After eliminating those that originated outside of Ecuador, under-age respondents, those that originated from repeat IP addresses, and those that dropped out prior to answering the outcome question, we were left with 485 responses. The median time to completion of these surveys was 4.6 minutes and 33% of respondents passed the attention screener. The cognitive process study in Chile was fielded between 2015-08-13 and 2015-08-16, and data were downloaded from Qualtrics on 2015-08-16. After eliminating those that originated outside of Chile, under-age respondents, those that originated from repeat IP addresses, and those that dropped out prior to answering the outcome question, we were left with 574 responses. The median time to completion of these surveys was 4.6 minutes and 31.5% of respondents passed the attention screener. Sample demographics and balance statistics are reported in the appendix.

5.2 Results

Manipulations of the cognitive process affected the assessment of presidents' performance mostly in the expected direction but, as Figure 3 shows (and Table 4 corroborates), the effect was only statistically significant relative to the control group for manipulation F, and again only for Ecuador. In fact, whereas one would have expected the latter treatments to yield greater effects because they dealt with more sources of mental contamination, no such trend is discernible from the data.

Analysis of the second outcome also does not provide much evidence that debiasing took place (Table 5). Recall that if respondents were unable or unmotivated to correct their own bias, but were able to spot bias in the judgement of others, they should have responded to second outcome by saying that the knowledgeable experts would have rated the president "better" than the average person. This option was the least frequent answer in the Chilean case.¹²

In fact, rather than discounting, what seems to have happened is that respondents project their own preferences onto the experts. The better an individual rated the government, the more likely she was to say that experts would have evaluated the government even better (p-value <0.001 in both Ecuador and Chile). As there were many more respondents rating Correa well as there were respondents rating Bachelet, the result was that many more respondents answered that experts would have rated the government better in Ecuador than in Chile. Therefore, evidence suggests

¹²The second outcome was available to respondents in conditions F, G, and H only, but there was no discernible difference in the responses across treatment conditions.

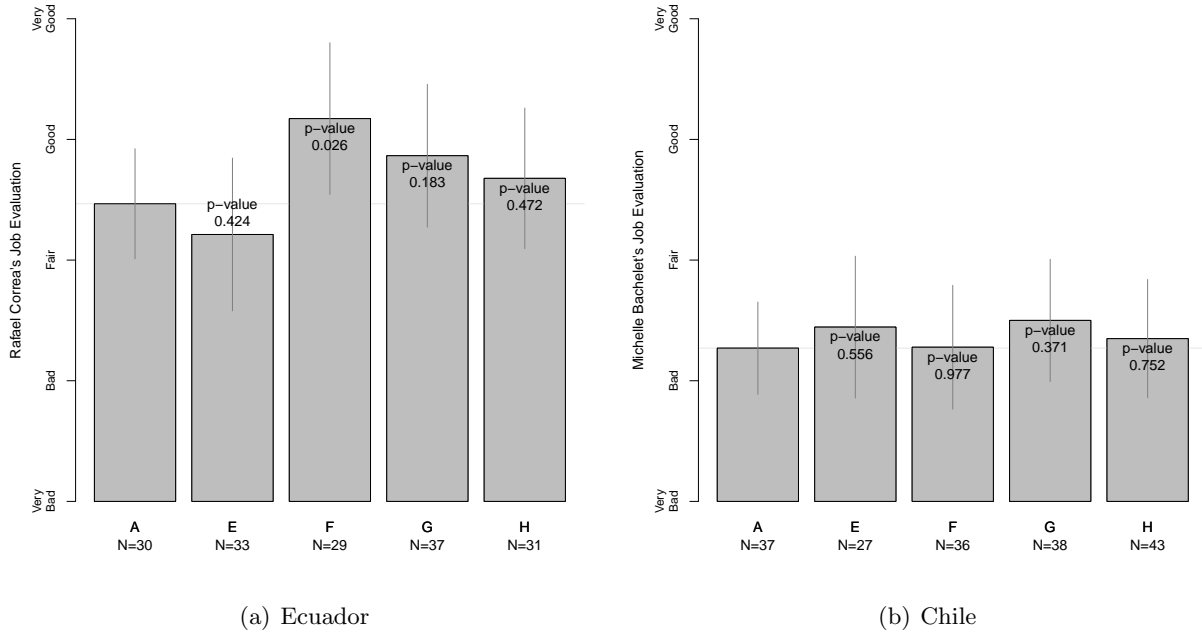


Figure 3: Government Evaluation by Cognitive Condition (Study #3)

that even though the results in Ecuador were compatible with some level of debasing, this was probably not the mechanisms that actually generated the results.

Respondents in conditions F, G and H were told, prior to the outcome question, that “studies suggest that voters often evaluate presidents based on attributes they cannot change or control” and asked whether the respondent thought she made the same “mistake.”¹³ The share of respondents that answer this question positively was 40% in Ecuador and 35% in Chile. Psychological theory suggests that lay beliefs about the mind are crucial to debiasing, and that those that know they are susceptible to mental contamination are more likely to make an effort to correct the bias (Wilson & Brekke 1994, p.188). The way in which we offered the primes in this study allows us to examine whether this is true by looking at whether there exist heterogeneous treatment effects across self-perceived susceptibility to misattribution bias.

Results displayed in Table 6 mostly corroborate this expectation, although they are also a bit noisy. We see that among those that regard themselves as susceptible to the specific type of mental contamination examined in the study, the treatment always had positive effects, and these effects were almost always larger than those observed in the sample that did not regard itself susceptible to mental contamination. For those self-described as susceptible to misattribution bias, treatment effects were statistically significant in conditions F and G in Ecuador. It should be noted, however,

¹³Actual wording was: “Estudios recientes sugieren que los ciudadanos muchas veces evalúan a los presidentes por hechos que ellos no pueden controlar ni modificar. ¿Cree que Ud. comete ese mismo error?”

Table 4: Treatment Effects (Study#3)

(a) Ecuador					(b) Chile				
	All Respondents		Attentive Only			All Respondents		Attentive Only	
	Mod. 1	Mod. 2	Mod. 3	Mod. 4		Mod. 1	Mod. 2	Mod. 3	Mod. 4
A (average)	2.713	2.423	2.467	2.074	A (average)	1.289	1.227	1.270	1.299
E	-0.217	-0.218	-0.255	-0.279	E	0.211	0.175	0.174	0.198
<i>p-value</i>	0.254	0.235	0.424	0.353	<i>p-value</i>	0.161	0.280	0.554	0.503
F	0.113	0.182	0.706	0.646	F	0.023	0.098	0.008	0.195
<i>p-value</i>	0.541	0.331	0.026	0.039	<i>p-value</i>	0.868	0.507	0.977	0.470
G	-0.030	0.065	0.398	0.397	G	0.131	0.060	0.230	0.286
<i>p-value</i>	0.870	0.719	0.183	0.140	<i>p-value</i>	0.347	0.689	0.369	0.285
H	-0.197	-0.229	0.211	0.252	H	0.099	0.062	0.079	0.048
<i>p-value</i>	0.274	0.199	0.472	0.376	<i>p-value</i>	0.474	0.678	0.751	0.857
Controls	No	Yes	No	Yes	Controls	No	Yes	No	Yes
N	485	398	160	146	N	574	500	181	171
Target p-val	0.046	0.040	0.036	0.035	Target p-val	0.047	0.041	0.035	0.037

Values for condition A are the average rating of the president in a 0-4 scale in the control group. All other estimates are differences between average values in each conditions and that of condition A. P-values are computed from robust standard errors. Target p-values are computed by simulation and indicate the threshold for significance at the 0.05 that accounts for the Family-Wise Error Rate due to multiple comparisons.

Table 5: How Would Knowledgeable Experts Rate The President?

	Ecuador	Chile
Worse than the public	14.9	30.8
The same as the public	27.4	46.4
Better than the public	57.7	22.8

that the experiment was not designed to test this specific hypothesis. Susceptibility was not manipulated or induced experimentally and we did not ask the susceptibility question to those in the control condition or in condition E¹⁴. Therefore, the comparison is not “clean”, as we compare self-perceived susceptible and non-susceptible respondents to the average respondent in the control condition. Results, in short, should be taken with a few grains of salt.

¹⁴This could be easily done by asking the exact same question *after* the outcome item in these conditions, much in the same spirit as the weather question was manipulated in Schwarz & Clore (1983).

Table 6: Treatment Effects by Recognition of Susceptibility to Biases

(a) Ecuador			(b) Chile		
	Self-Perception			Self-Perception	
	Susceptible	Not Susceptible		Susceptible	Not Susceptible
A	2.467		A	1.270	
F	0.908	0.629	F	0.330	-0.225
<i>p-value</i>	0.026	0.082	<i>p-value</i>	0.320	0.435
G	0.733	0.117	G	0.174	0.359
<i>p-value</i>	0.071	0.726	<i>p-value</i>	0.622	0.199
H	0.105	0.298	H	0.184	-0.020
<i>p-value</i>	0.759	0.396	<i>p-value</i>	0.639	0.940

Values for condition A are the average rating of the president in a 0-4 scale in the control group. We cannot disaggregate this average for respondents that perceive themselves as susceptible and not susceptible. All other estimates are differences between average values in each conditions and that of condition A. P-values are computed from robust standard errors. Susceptibility to bias was not manipulated experimentally. Treatment group E did not receive the susceptibility item, and is therefore excluded from this analysis.

6 Discussion and Next Steps

The starkest regularity to emerge from our analysis is that we had only very meager success in debiasing respondents. There was some minor debiasing in Study #1, but only for the negative condition and only for the between-subjects analysis. In Studies #2 and #3, while a few manipulations had some effect on Rafael Correa’s job evaluations, no treatment managed to affect evaluations of Michelle Bachelet.

Before considering the potential theoretical implications of these findings, which are plenty, it is necessary to engage in extra-careful scrutiny of the approach to determine whether some feature of the choice of cases, research design, or delivery medium might be preventing us from uncovering debiasing effects. Some limitations of our studies are rather obvious. First, we provide information to respondents in a fairly artificial setting. Theoretical models of signal extraction reviewed earlier assume that voters gather and observe information over time, embedded in context. It is possible, therefore, that true information acquisition and processing cannot be easily replicated in experimental settings, especially in a relative low-stake environment as the internet. This also applies to Study #1, which also failed to generate strong debiasing in a more personal form of interaction between researcher and subject. The intuition that voters’ acquire by processing facts over time might just be too fundamentally different from their response to relevant information offered at once.

Another explanation for our null results might be related to the data collection instrument used in Studies #2 and #3. It is possible that the cumulative priming design saturated respondents. It is also possible that certain manipulations had effects opposite to what we anticipated, thereby canceling out primes that did a positive effect. To correct for that, we plan to field each manip-

ulation separately (instead of cumulatively) in the informational study. Unfortunately, it is not possible to do the same in the cognitive experiment, as theory suggests that there would be no point in providing respondents with a motivation to correct a bias that they are not aware exists, or even to illustrate the direction of the bias if respondents have no motivation to correct it.

Relatively low attention rates in these surveys always raise a flag. For this reason we focused only on respondents that passed the attention screener. We are considering small changes to the questionnaire, such as removing the timer, as it is only really necessary for the few factual questions. Previous studies of other topics have succeeded in uncovering treatment effects using Facebook recruited samples and simple surveys such as the one we employed (Samuels & Zucco Jr. 2014), but given that debiasing is rather unlikely, it might require a more tightly controlled environment.

Finally, results seem to confirm the notion that debiasing is easier for respondents that accept that they are susceptible to bias. Our studies, however, were not optimally designed to test this hypothesis, so additional research is needed on this point. The fact that only a minority of respondents actually concede their susceptibility might be one of the reasons that debiasing is so hard to achieve. If confirmed, it would indicate that a majority of respondents is not “debiasable.”

One way in which the choice of cases might have influenced the null results is that data from Study #1 suggest that negative debiasing is easier than positive debiasing, a result quite established in the psychology literature; yet there are currently no countries in Latin America experiencing booming times. One alternative to deal with this constraint could be to prime respondents in Ecuador and Chile not about the recent fall in commodity prices, but rather about the sharp increase that preceded it. Particularly in the case of Ecuador, where the same president governed during the boom and the bust, we suspect that voters’ evaluations are still very influenced by previous “good times.”

If after follow-up work these null results continue to hold, they are potentially very meaningful. In line with macro level evidence that voters do not discount exogenous factors when evaluating Latin American presidents (Campello & Zucco Jr. 2015), the micro-level analysis reported here suggests that it is particularly hard to induce voters to discount such factors. Voters, in this sense, are more akin to the usually efficient but sometimes bias-prone individuals of psychology textbooks than to those that inhabit economic models, and who readily update their beliefs in light of new evidence. Put simply, debiasing, for all the conditions necessary for it to happen, is very unlikely in the real world. The weight of the fact that the economy is doing well or not on individuals daily lives is probably too intense for citizens to be tampered by complicated discounting, even for those that are aware that certain factors should be discounted.

Our next step is to separate the rationale that leads to bias from the bias itself. One way to

do it would be to conduct similar experiments in contexts in which bias has not yet formed. For example, to ask the same questions about the Chilean president to Ecuadoreans and vice-versa. If we find that voters can discount exogenous factors in a situation in which they do not have a prior and are still forming their views, this means respondents can avoid contamination, even though they cannot debias once contamination has already happened. This would point, once again, to the relevance of providing information that allows voters to discount exogenous factors — be that about relative performance or commodity price cycles — on a regular basis, so that citizens can form their views already based on that.

References

- Achen, Christopher & Larry Bartels. 2006. “Blind Retrospection: Electoral Responses to Droughts, Floods, and Shark Attacks.” Presented at the Meeting of the American Political Science Association.
- Alcañiz, Isabella & Timothy Hellwig. 2011. “Who Is to Blame? The Distribution of Responsibility in Developing Democracies.” *British Journal of Political Science* 41:389–411.
- Alesina, Alberto & Howard Rosenthal. 1995. *Partisan Politics, Divided Government, and the Economy*. Cambridge: Cambridge University Press.
- Benton, Allyson Lucinda. 2005. “Economic Hardship, Political Institutions, and Voting Behavior in Latin America.” *Comparative Political Studies* 38(4):417–442.
- Campbell, Angus, Philip E. Converse, Warren E. Miller & Donald E. Stokes. 1960. *The American Voter*. New York: John Wiley and Sons.
- Campello, Daniela & Cesar Zucco Jr. 2015. “Presidential Success and the World Economy.” Working Paper, Fundação Getúlio Vargas.
- Canes-Wrone, Brandice, Michael C. Herron & Kenneth W. Shotts. 2001. “Leadership and Pandering: A theory of executive policymaking.” *American Journal of Political Science* 43(3):532–550.
- Cochran, WG. 1968. “Effectiveness Of Adjustment By Subclassification In Removing Bias In Observational Studies.” *Biometrics* 24(2):295–&.
- Cole, Shawn, Andrew Healy & Eric Werker. 2012. “Do voters demand responsive governments? Evidence from Indian disaster relief.” *Journal of Development Economics* 97(2):167 – 181.
URL: <http://www.sciencedirect.com/science/article/pii/S0304387811000502>
- Converse, Philip E. 1969. “Of Time and Partisan Stability.” *Comparative Political Studies* 2(2):139–71.
- Duch, Raymond M. & Randolph T. Stevenson. 2008. *The Economic Vote: How Political Institutions Condition Election Result*. New York: Cambridge University Press.
- Ebeid, Michael & Jonathan Rodden. 2005. “Economic Geography and Economic Voting: Evidence from the US States.” *British Journal of Political Science* 36(3):527–547.
- Fearon, James D. 1999. Electoral accountability and the control of politicians: selecting good types versus sanctioning poor performance. In *Democracy, Accountability, and Representation*, ed. B Manin A Przeworski, SC Stokes. New York: Cambridge Univ. Press pp. 55–97.
- Ferejohn, John A. 1986. “Incumbent performance and electoral control.” *Public Choice* 50(1–3):5–25.
- Gaspar, John T. & Andrew Reeves. 2011. “Make It Rain? Retrospection and the Attentive Electorate in the Context of Natural Disasters.” *American Journal of Political Science* 55(2):340–355.
URL: <http://dx.doi.org/10.1111/j.1540-5907.2010.00503.x>
- Healy, Andrew J., Neil Malhotra & Cecilia Hyunjung Mo. 2010. “Irrelevant events affect voters’ evaluations of government performance.” *Proceedings of the National Academy of Sciences* 107(29):12804–12809.
- Healy, Andrew & Neil Malhotra. 2013. “Retrospective Voting Reconsidered.” *Annual Review of Political Science* 16:285–306.
- Hellwig, Timothy & David Samuels. 2007. “Voting in Open Economies The Electoral Consequences of Globalization.” *Comparative Political Studies* 40(3):283–306.

- Ho, Daniel E., Kosuke Imai, Gary King & Elizabeth Stuart. 2007. "Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference." *Political Analysis* 15(3):199–236.
- Huber, Gregory A., Seth J. Hill & Gabriel S. Lenz. 2012. "Sources of Bias in Retrospective Decision Making: Experimental Evidence on Voters' Limitations in Controlling Incumbents." *American Political Science Review* 106:720–741.
- Johnson, Gregg B. & Leslie A. Schwindt-Bayer. 2009. "Economic Accountability in Central America." *Journal of Politics in Latin America* 1(3):33–56.
- Kayser, Mark & Michael Peress. 2012. "Benchmarking across borders: electoral accountability and the necessity of comparison." *American Political Science Review* 106(3):661–684.
- Leigh, Andrew. 2009. "Does the World Economy Swing National Elections?" *Oxford Bulletin of Economics and Statistics* 71(2):163–181.
- Powell, G. Bingham, Jr. & Guy D. Whitten. 1993. "A Cross-National Analysis of Economic Voting: Taking Account of the Political Context." *American Journal of Political Science* 37(2):pp. 391–414.
- Samii, Cyrus & Peter M. Aronow. 2011. "On Equivalencies Between Design-Based and Regression-Based Variance Estimators for Randomized Experiments." Preprint submitted to Statistics and Probability Letters.
- Samuels, David. 2004. "Presidentialism and Accountability for the Economy in Comparative Perspective." *American Political Science Review* 98(3):425–436.
- Samuels, David & Cesar Zucco Jr. 2014. "The Power of Partisanship in Brazil: Evidence from Survey Experiments." *American Journal of Political Science* .
- Schwarz, Norbert & Gerald L Clore. 1983. "Mood, misattribution, and judgments of well-being: informative and directive functions of affective states." *Journal of personality and social psychology* 45(3):513.
- Torsten, Persson, Gerard Roland & Guido Tabellini. 1997. "Separation of powers and political accountability." *Quarterly Journal of Economics* 112:1163–202.
- Wilson, Timothy D. & Nancy Brekke. 1994. "Mental Contamination and Mental Correction: Unwanted Influences on Judgments and Evaluations." *Psychological Bulletin* 116(1):117–142.

Appendices

A Balance Across Experimental Conditions

In this appendix we report mean standardized differences on selected pre-treatment variables for each condition *relative* to the control group. These differences were computed by simply dividing the the difference in means by the pooled standard error of the two groups being compared. Although there is no clear cut test for balance, one rule of thumb often cited in the literature is to consider balanced a sample in which the largest mean standard difference is less than 1/4 of a standard deviation (Ho, Imai, King & Stuart 2007, Cochran 1968).

Randomization in the BEPS project (Study #1) achieved good balance on all socio-demographics available, excellent balance on missing responses, and excellent balance on actual answers given to vignettes (not reported), suggesting a similar level of information across conditions (Table A.7).

Table A.7: Balance on Selected Observable Attributes (Study #1)

	Negative	Positive
Positive Pre-Treatment Evaluation	-0.11	-0.10
PT ID	-0.08	0.05
Wealth Index	-0.01	0.07
Vote For Dilma	-0.00	-0.06
Gov. Responsible Econ.	0.06	-0.04
Political Knowledge	0.05	0.20

Table reports mean standard difference between each treatment condition and the control group

In Studies #2 and #3 randomization generated four groups that display fairly similar observable attributes both in Chile and in Ecuador, as reported in Tables A.8 and A.9. In these tables, female, party ID, family abroad, and attention are all binary indicators. Age is a continuous variable, and so is the socioeconomic status (SES) index that we created combining educational and wealth indicators.¹

Table A.8: Balance on Selected Observable Attributes (Study #2)

(a) Ecuador				(b) Chile			
	B	C	D		B	C	D
Age	-0.16	0.02	-0.09	Age	0.36	0.26	0.29
Female	0.19	0.15	0.10	Female	0.04	0.09	0.02
SES Index	-0.11	-0.16	-0.05	SES Index	-0.43	-0.42	-0.24
Party ID	0.16	0.18	0.28	Party ID	-0.15	-0.29	-0.02
Family Abroad	0.07	0.08	-0.01	Family Abroad	0.07	-0.05	0.09
Attention	-0.06	-0.02	-0.17	Attention	0.01	-0.05	-0.19
N	124	138	139	N	116	121	135

Table reports mean standard difference between each treatment condition and the control group

Most of our observed variables meet the 1/4 of standard deviation criteria in all conditions in both studies, with the exception of Party ID in study #2, SES Index and Age in study #2 and

¹Our SES index is the score on the first dimension from a factorial analysis performed on the answers to household item and educational level questions in the survey.

#3 in Chile only. Based on these results, we conclude that randomization was mostly successful in producing similar groups, and consequently, we can focus more directly on the unconditional average treatment effects reported in the text.

Table A.9: Balance of Selected Observable Attributes (Study #3)

	(a) Ecuador					(b) Chile			
	E	F	G	H		E	F	G	H
Age	-0.16	-0.19	-0.05	-0.16	Age	-0.01	0.22	0.19	0.20
Female	-0.12	0.05	-0.05	0.15	Female	0.11	0.27	-0.08	-0.04
SES Index	-0.15	-0.09	-0.09	-0.17	SES Index	-0.29	-0.04	-0.22	-0.24
Party ID	-0.15	0.02	-0.01	-0.06	Party ID	-0.09	-0.13	-0.06	0.13
Family Abroad	-0.00	-0.07	-0.20	-0.30	Family Abroad	0.03	0.09	-0.01	-0.04
Attention	-0.06	-0.06	0.02	-0.05	Attention	0.00	0.03	0.03	0.03
N	105	92	104	97	N	88	112	119	134

Table reports mean standard difference between each treatment condition and the control group

B The Attention Screener

Figure A.4 shows a screen shot of the attention screener used in Studies #2 and #3. The fake question was shown in bold, and the answer categories were completely compatible with it. In fact, one would be extremely tempted to guess what the question was simply from reading the answer options. For these reasons, we believe the attention screener was quite demanding, and respondents that succeeded were, in fact, paying significant attention to the study.

Algunas personas leen periódicos o ven muchos noticieros, otros no leen periódicos ni ven noticias. Esto es así en toda parte, pero nosotros sólo queremos chequear que Ud. esté leyendo estas preguntas hasta el final. Ignore la pregunta que sigue, y seleccione la opción "otra cifra", añadiendo el número cinco en el espacio proporcionado junto a esa opción. **Por favor indique ¿con qué frecuencia lee periódicos o ve noticieros?**

Nunca o muy raramente

Uno o dos días por semana

Casi todos los días

Todos los días

Varias veces por día

Otra cifra

Powered by Qualtrics

Figure A.4: The Attention Screener