Appendix A

Explicit Bayesian Analysis in Qualitative Case Research: An Empirical Example

Fairfield's (2013) article on taxing economic elites in unequal democracies includes a methodological appendix that explicitly uses process-tracing tests to elucidate causal inferences in the author's case narratives. In the following exercise, which is the first of its kind, we revise that appendix by replacing the process-tracing tests approach with explicit application of Bayesian analysis, which we view as a more appropriate methodological foundation for inference in qualitative research and scientific reasoning more broadly. The present appendix is intended both to serve as a detailed pedagogical resource and to highlight the technical challenges we encountered and the pragmatic workarounds we devised for applying Bayesian analysis to qualitative evidence.¹

We begin by introducing the Chilean tax reform case study that we will analyze (§A1). We then specify the explanatory hypotheses we will compare and assign prior probabilities (§A2). §A3 explains how we elaborate the evidence for Bayesian analysis, drawing on Fairfield's case narrative. §A4 gives a thorough explanation of the likelihoods we assign for the six pieces of evidence that form the basis for Fairfield's inference. In §A5, we apply Bayes' rule to calculate posterior probabilities on each hypothesis from our priors and our likelihoods. In §A6, as an internal consistency check, we conduct the analysis using a different ordering of the six pieces of evidence. We conclude by conducting Bayesian sensitivity analysis to assess the strength of Fairfield's (2013) causal claim (§A7).

A1. Chile's 2005 Tax Reform

In *Private Wealth and Public Revenue in Latin America: Business Power and Tax Politics,* Fairfield (2015) analyzes how and when unequal democracies can tax economic elites. Fairfield explains the scope and fate of tax policy proposals by analyzing business's instrumental (political) power and structural (investment) power. Instrumental power entails deliberate political actions like lobbying. Structural power arises from the profit-maximizing behavior of firms and investors; if policymakers anticipate that a reform will provoke disinvestment or capital flight, they may rule it out to protect growth and employment. When business actors have strong power of either type, their interests shape policy decisions. However, strategies for mobilizing public support or tempering elite opposition can facilitate incremental reforms that might not otherwise be feasible. One such strategy—a vertical equity appeal—aims to mobilize public support by emphasizing a tax increases' congruence with the widely-accepted principle that those who earn more or own more assets should bear a larger share of the tax burden.

Fairfield's (2013) article on tax reform strategies includes the case of a 2005 Chilean reform that eliminated a longstanding tax benefit for owners of eligible shares, who belonged to the richest 0.5%. During the presidential campaign, Chile's Catholic bishops forcefully denounced the country's extreme levels of inequality, thereby raising the salience of this issue. The right-wing opposition candidate Lavín responded by blaming Chile's persistent inequality on the governing center-left coalition and accusing incumbent president Lagos of failing to deliver on his promise of growth with equity. Lagos seized the opportunity to eliminate the tax subsidy by linking the reform

¹ Applications of Bayesian analysis in empirical process tracing research will be necessarily less detailed than what we present here; Sections 4 and 5 of the main article suggest possibilities for more pragmatic (if less careful) intermediate approaches between common-language narrative-based process tracing and the explicit Bayesian analysis we present here.

to the issue of inequality and thereby mobilizing public support with the following equity appeal: "The famous Article 57 bis [the tax subsidy] signifies a tremendous source of inequality. ...Instead of just talking, why don't we agree to eliminate 57 bis in less than 24 hours?"² The right-wing opposition coalition, which held a majority in the senate, accepted the challenge and voted in favor of eliminating the tax benefit, deviating from its prior position on this policy as well as the preferences of its core constituency—business owners and upper-income individuals. Before proceeding, readers may wish to review Fairfield's (2013:48-49) original case narrative.

This case was chosen for methodological attention for both substantive and practical reasons. Substantively, the 2005 reform was an emblematic case of equity-enhancing tax reform in Chile that illustrates both the importance of strategies that mobilize public support and the limitations to how much revenue they can raise in contexts of strong business power. Practically, this is a clear-cut case in which a relatively small number of key pieces of evidence establish the causal importance of the reform strategy. As the complexity of the case and the quantity of evidence the analyst draws on to make inferences increase, explicit elaboration of either process tracing tests or the explicit Bayesian reasoning we employ here may become infeasible.

A2. Hypothesis Space and Priors

Fairfield (2013) argues that President Lagos' equity appeal, which took place in an unusual context of strong electoral competition from the right-wing opposition coalition on the issue of inequality during a major electoral campaign, compelled the opposition to approve the 2005 reform in congress. The postulated causal mechanism is that in this unusual context, the equity appeal created concern within the opposition coalition that rejecting the initiative to eliminate the regressive tax benefit would damage its candidate's electoral prospects.

 H_{EA} = The equity appeal, in the context of a tight presidential race where inequality had assumed high issue-salience, drove the right-wing coalition to accept the 2005 reform in order to avoid electoral costs.

 $\sim H_{EA}$ = The right-wing coalition would have accepted the 2005 reform anyway without the equity appeal in the context of a tight presidential race where inequality had assumed high issue-salience. In other words, those factors did not have a relevant causal effect on the fate of the reform initiative.

Whereas frequentists usually consider a single null hypothesis and its negation, applying Bayes' theorem requires elaborating a complete set of mutually exclusive hypotheses. We need to explicitly state the alternatives before we can reason meaningfully about the likelihood of observing the evidence if the author's hypothesis does not hold. To that end, we decompose $\sim H_{EA}$ into three rival alternative hypotheses: an institutional hypothesis (H_I), a median-voter hypothesis (H_{MV}), and a core-constituency hypothesis (H_{CC}):

 H_I = The right-wing coalition accepted the reform because Chile's institutionalized party system and stable rules of the game motivate cross-partisan cooperation in congress and consensual politics (drawing on Flores-Macías 2010). These institutions lengthen time horizons and encourage parties to moderate their policy stances in anticipation of future rounds of negotiation on other issues.

 H_{MV} = The right-wing coalition accepted the 2005 reform in accord with a simple median voter model of redistributive politics, where electoral competition drives

² Fairfield (2013:48): Lagos, quoted in *El Mercurio*, May 10, 2005.

politicians to converge on policies that promote the median voter's material interests (e.g. Meltzer and Richard 1981). In other words, the president's equity appeal and the specific context in which it took place were irrelevant for obtaining right coalition votes in support of eliminating the tax benefit.

 H_{CC} = The right-wing coalition accepted the reform because its core constituency business and upper-income individuals—had a weaker material interest in defending the tax subsidy in 2005 compared to the 1990s, due to a decline in the assets eligible for the subsidy.

The last hypothesis is motivated by the fact (taken as background information) that a 1998 reform closed the tax subsidy to new beneficiaries while grandfathering in existing beneficiaries; the amount of qualifying stocks declined thereafter as owners gradually sold eligible shares.

It is important to note that we are assuming as part of our background information that these alternative hypotheses are mutually exclusive and exhaustive—in other words, only one of the mechanisms corresponding to the four different hypotheses may operate. Otherwise we cannot maintain that $\sim H_{EA} = H_I + H_{CC} + H_{MV}$. In the real world, one could imagine that the equity appeal might work to create consensus between the right and the left, but that logic corresponds to a more complex causal hypothesis that blends elements of both H_{EA} and H_I . Similarly, it might be the case that changing business preferences in conjunction with the equity appeal motivated the right to accept the reform, such that a combination of the causal factors in H_{EA} and H_{CC} was critical to the outcome. Allowing for causal complexity in which multiple mechanisms operate at the same time, to varying degrees or in interaction, would require elaborating additional, more complicated mutually exclusive hypotheses, which can be challenging if we wish to be precise enough to explicitly apply Bayesian analysis (Section 3.1).

We consider three different prior distributions for the four hypotheses. For the first prior, we employ the indifference principle and set equal probabilities of 25% on each hypothesis, assuming a position of maximal ignorance I_0 in which we have no reason to privilege any one of the four explanations. This approach avoids bias in favor of any of the hypotheses, at the cost of ignoring much of the background information *I* that we bring to the analysis.

The second prior distribution aims to take into consideration a large body of literature questioning the logic underlying simple median-voter models. Authors have identified numerous assumptions in these models that do not hold up against empirical evidence-not only for developed countries like the US (e.g. Hacker and Pierson 2010), but also for developing countries (e.g. Kaufman 2009). For the case of Chile, Luna (2010) analyzes in detail how right parties have successfully defended the economic interests of their core upper-income constituency while still winning broad support among low-income voters who would stand to benefit from redistribution. Accordingly, we place a low prior probability on H_{MV} of 0.001% and equal probabilities on H_{EA} , H_{I} , and H_{CC} of 33.3%. These assignments correspond to a prior log-odds ratio of 45 dB against H_{MV} relative to each of the other hypotheses. Using our sound analogy (Section 3.6), we could say that H_{MV} is "sleeping" in the background, and it would take roughly 45 dB to "wake it up" (Table 1 below). It is worth noting that from the perspective of logical Bayesianism, instead of relying on our intuition to penalize H_{MV} we should begin with indifference priors and systematically incorporate all prior evidence we possess that bears on the hypotheses. That approach is clearly infeasible in practical terms—we would need to apply Bayes' rule to every piece of qualitative and quantitative evidence that Luna (2010) provides in his extensive analysis, not to mention all other works in the large body of literature on median voter theories and redistribution.

Our third prior distribution represents the views of a highly-skeptical reader who possesses very different background information. This prior places a very low probability on H_{EA} of 0.0003%³ and sets equal probabilities on H_I , H_{CC} , and H_{MV} of 33.3%. This prior distribution penalizes H_{EA} by 50 decibels relative to any of the other three hypotheses. Continuing with our sound analogy, if the three alternative hypotheses are considered to be "in conversation," the equity appeal hypothesis corresponds to a mere "pin-drop" in the background. We will subsequently assess the strength of Fairfield's (2013) causal claim by updating this skeptical prior in light of the evidence that the case narrative brings to bear on the 2005 Chilean reform.

10	Adult hearing threshold; rustling leaves, pin-drop
20-25	Whisper
30	Quiet bedroom or library, ticking watch
45	Sufficient to wake a sleeping person
50	Moderate rainstorm
60	Typical conversation
70	Noisy restaurant, common TV level
80	Busy curbside, alarm clock
90	Passing diesel truck or motorcycle
100	Dance club, construction cite
115	Rock concert, baby screaming

Table 1: Typical sound levels (dB) (Reproduced from Section 3.4)

A3. Evidence

We will consider the six key pieces of evidence (E_1 – E_6) that Fairfield (2013) examines when analyzing the 2005 Chilean reform, summarized below (Table A.1). The evidence includes not only observations about the causal process operating within the 2005 tax reform case, but also information about previous episodes of tax reform and non-reform that bear on the hypotheses (see E_1 and E_3). Our analysis accordingly illustrates how Bayesian logic integrates both with-in case and cross-case observations—not just for mixed-method research designs that combine large-N datasets with qualitative mechanism observations (Humphreys and Jacobs 2015), but also for qualitative small-N and medium-N comparative research.

In parsing the evidence into distinct pieces, we have followed the guidelines elaborated in Section 3.2. In most cases, we have treated information from distinct types of sources as separate pieces of evidence. We have also sought a middle ground between aggregating evidence into overly course-grained pieces and disaggregating it into excessively fine-grained pieces, for the sake of facilitating reasoning about likelihoods. Notice also that we have included information about the source when articulating the evidence, in accord with our recommendations for working with testimonial evidence in Section 3.3.2.

³ Jaynes (2003:99-100) similarly employs 10^{-6} as a "very low prior probability" in his example of testing hypotheses about widget quality.

Table A.1: Summary of Evidence

E_1	The center-left discussed the reform multiple times during the 1990s but ruled it out given resistance from the right coalition, according to multiple sources on the left
E_2	Government informants suggested that Lagos' public exchange with Lavín forced the right to approve the reform
E_3	Government informants reported failed efforts to reach an agreement with business on eliminating the tax subsidy during the early years of Lagos' term (before 2005)
E_4	A right-party technical advisor reported that opposition-coalition legislators feared electoral punishment if they did not approve the reform
E_5	The right-learning national newspaper reported that Lavín's advisors viewed the reform proposal as an electoral "trap"
E_6	A right-party deputy reported that the opposition coalition reluctantly accepted reform "so as not to harm the presidential option"

It is important to note that although $E_I - E_6$ mention very specific information obtained during fieldwork, the likelihoods we assign in the following section correspond to any informationallyequivalent piece of evidence that might have arisen. For example, $P(E_6 | H_{EA} E_{prev} I)$ refers to the likelihood of a particular right-party deputy interviewed on December 23, 2005 sharing the exact comments reported, or to any other essentially equivalent story shared by a similar informant from that party, using slightly different language, on a different day, and so forth. This point may seem like a technicality, but the probability of observing the exact piece of evidence E_6 would otherwise be vanishingly small given the myriad contingencies that ultimately produced that specific conversation. Moreover, irrelevant details associated with a narrowly-defined piece of evidence would be common under each alternative hypothesis—they would simply lower the likelihood of the evidence under every hypothesis, such that their effect would cancel out of the likelihood ratios.

In general, defining the equivalence class entails optimizing a tradeoff between generality and specificity. If the equivalence class is too broad and vague, we may risk circularity by essentially asserting that "the evidence is that there was evidence in favor of the hypothesis," and there will be little basis for assessing likelihoods. If the equivalence class becomes too narrow and specific, with too many irrelevant details, the likelihoods will become vanishingly small and hence difficult to assess, since our brains are not well adapted for reasoning about small probabilities. The set of hypotheses under consideration will also guide decisions about how narrowly or broadly to define the equivalence class for the sake of effectively discriminating among the explanations.

In most cases, we implicitly define the equivalence class by the details that are omitted in stating the evidence. For example, if a quote is attributed to an informant of a particular type, a similar statement from an alternative informant of that same type would be assigned the same likelihood.

A4. Likelihoods

The key inferential task when explicitly applying Bayesian analysis in process tracing research entails assigning likelihoods, $P(E_x | H_j E_{prev} I)$, for each hypotheses, $\{H_j, j = EA, I, P, MV\}$, conditional on the evidence that we have previously incorporated into our analysis, E_{prev} , and on our background information, *I*. This section begins by discussing our strategy for assigning numerical values to these probabilities. We then present our likelihood assessments for each piece of evidence

 $E_1 - E_6$. Finally, we summarize the weights of evidence and highlight distinctions between our Bayesian analysis and Fairfield's (2013) process-tracing tests analysis.

A4.1 Assigning Numerical Values: Challenges and Strategies

Assigning numerical values to likelihoods is challenging in qualitative research. Ideally, one should reason out each probability in the problem and then calculate likelihood ratios $P(E_x|H_j E_{prev} I)/P(E_x|H_k E_{prev} I)$ to assess how much a given piece of evidence discriminates between a pair of hypotheses. In practice, however, it is very difficult to assess what absolute value a likelihood should assume when conditioning a piece of evidence on a hypothesis that simply does not fit. We know the likelihood of the evidence should be very low, but our intuition gives us little traction for discerning whether that likelihood should be lower or higher than the probability of some other piece of evidence that is extremely unlikely under another hypothesis. Our brains are not accustomed to making judgments on these scales—it is very difficult to assess differences between probabilities that are extremely small. To circumvent this problem, we opted for the following approach.

First, we set values for the likelihood of each piece of evidence E_x under the most compatible hypothesis—this task entails handling probabilities in a range for which we feel capable of making reasonable assignments. Second, we used our intuition to assess how large the weight of evidence (ten times the log of the likelihood ratio) should be for E_x relative to each rival hypothesis. This approach is natural since only the *relative* probabilities of observing the evidence under different hypotheses matter for assessing which explanation fits best, and because it is easier for our brains to perceive and interpret differences on a logarithmic scale (Section 3.4). For evidence that strongly discriminates between two hypotheses, we give the weight of evidence 30 decibels-a likelihood ratio of 10^3 . Employing the analogy of inference as a dialog with the data, 30 decibels in acoustic terms roughly corresponds to the difference between a quiet bedroom and an ordinary conversation—in other words, the data are "talking clearly." Very low probabilities are then determined by the baseline probability of E_x under the most compatible hypothesis, and the likelihood ratio relative to the rival hypotheses. We adjusted low probabilities as necessary when a clear argument could be made that a given likelihood should be higher or lower than another likelihood in the exercise. Our lowest probability assignments, for evidence that we view as exceedingly unlikely under a given hypothesis, are on the order of 10^{-5} : extremely improbable, but a healthy order of magnitude higher than being struck by lightening over a lifetime (10^{-6}) and several orders of magnitude larger than other relevant improbable perils, such as experiencing a plane crash on a major airline (10^{-7}) or winning a major lottery (10^{-8}) .

Readers may nevertheless object that our lowest probabilities are too small. In response, we emphasize that we are using a logarithmic scale because humans have evolved to deal with probabilities that vary over orders of magnitude (Section 3.4). A logarithmic scale is actually better suited to human perception than a more familiar linear scale, once we become accustomed to working in decibels. Moreover, our two-step procedure for assigning improbable likelihoods minimizes (although hardly removes) the arbitrariness of quantifying inherently qualitative data. Readers should find the likelihoods we assign under the hypothesis that fits best (ranging from 3%–60%) to be reasonable. And the analogy to sound levels helps make assessments of likelihood ratios as consistent and intuitive as possible. Together, these two factors uniquely determine likelihoods under rival hypotheses that would otherwise be extremely difficult to reliably quantify. In our experience, explicit Bayesian analysis of our case study would have been intractable and fraught with inconsistencies had we not devised the procedure outlined above. A final critical point is that the primary objective in social science should be comparing hypotheses, rather than calculating posteriors, in which case only the likelihood *ratios* matter—not the absolute value of the

probabilities. For readers whose skepticism persists, we assess how our conclusions would change if we were to compress our likelihood ratios—increasing our lowest probabilities by a factor of 50—in the last section of this appendix.

A4.2 Likelihood Assignments for E_1 – E_6

We proceed to assess likelihoods for the six pieces of evidence in turn under each of our four alternative hypotheses. Our analysis closely follows the guidelines elaborated in Section 3 of the main article. We endeavor to mentally "inhabit" the world of each hypothesis when asking how likely the evidence is if a given explanation is correct (Section 3.3.1). For testimonial evidence, we assess the likelihood that "source *S* stated *X*," paying careful attention to potential instrumental incentives and/or unmotivated biases that could affect a source's inclination to make particular statements and/or disposition to reveal or conceal information if a given hypothesis is true (Section 3.3.2). We condition likelihoods on all previously-incorporated evidence, thinking carefully about logical dependence between E_x and E_{prev} under the assumption that H_i is true (Section 3.3.3).⁴ We have sequenced the evidence as convenient to facilitate reasoning about logical dependence, by placing what we anticipated to be the most decisive evidence last (Section 3.3.4). Thinking about logical dependence among the evidence is nevertheless extremely challenging. In practice, even in this relatively simple case study, we manage to explicitly condition on only those pieces of evidence that are most clearly dependent under a given hypothesis.⁵

Finally, we identify key elements of the background information and explain how they inform our likelihood assignments, while recognizing that in practice it is impossible to fully enumerate all relevant background knowledge. Broadly speaking, our background information is informed by Fairfield's extensive fieldwork in Chile, which included 216 interviews, research in news and congressional archives, and observation of congressional proceedings, conferences, and public events relevant to tax policy. More specifically, the background information includes knowledge about effort expended to uncover relevant evidence, persistence in seeking to obtain interviews, relative ease or difficulty of reaching particular types of informants, skill at establishing rapport with and degree of trust in informants, first-hand knowledge about Chilean politics, and a wide range of contextual clues that inform interpretation of interviews and other evidence. We also take the particular set of informants interviewed as part of the background information, to avoid reasoning about the probability of reaching a specific individual or type of informant when assessing likelihoods. For example, E_2 includes a statement made by former president Lagos; we condition the likelihood of E_2 on the background information that Fairfield was able to interview Lagos on multiple occasions. Otherwise, we would have to lower the likelihood of E_2 under each hypothesis.

⁴ As explained in Section 3.3.3, this task entails asking what we learn about the likelihood of E_x from observing E_{prev} —beyond what H_j tells us, if we also know E_{prev} , are we any more or less likely to observe E_x , and by how much?

⁵ We apply our pragmatic recommendation to focus exclusively on weights of evidence in Section 4 of the main article.

E_1 = Governing-coalition informants and congressional documents indicated that the center-left coalition had discussed including a measure to definitively eliminate the tax subsidy in multiple prior tax reforms (1990, 1995, 1998, 2001) (E_{1a}).⁶ However, that measure was ultimately ruled out as infeasible on every such occasion due to resistance from the right (E_{1b}).⁷

We have endeavored to describe E_1 in terms that convey an appropriate and manageable equivalence class. What matters most for discriminating among our alternative hypotheses is the existence of multiple prior discussions about eliminating the tax benefit, not the details regarding how many times it was considered or in which years. If we were to consider these details as central to the definition of E_1 , the likelihoods would become orders of magnitude smaller, and much more difficult to assess—the chances that discussion would take place in each of these particular years under any of the hypotheses is very, very low. We include these details parenthetically and footnote several of the sources to illustrate the concrete specifics of the data uncovered.

Notice also that for convenience, we have taken E_1 to be the conjunction of two pieces of information E_{1a} and E_{1b} . We could assess E_{1a} and E_{1b} separately as distinct pieces of evidence; we could even disaggregate further so that each of the sources noted above contributes one or more pieces of evidence to be considered separately. However, there would be few analytical gains. When we are dealing with qualitative data, we need to operate at a level that facilitates reasoning, rather than trying to build up systematically from extremely specific bits of evidence. If we disaggregate too finely and if we make too many analytical steps explicit, we will become lost in minutia. Recall that the mathematics of Bayesian analysis allows us to aggregate or disaggregate data at whatever level is convenient.

$P(E_1|H_{EA} I) = 20\%$

This evidence is consistent with the hypothesis that Lagos' high-profile equity appeal in the unusual context of electoral competition from the right on the issue of inequality explains the right's acceptance of the 2005 reform, since some new factor that was not present in previous years must have changed the right's behavior.

The probability of observing E_1 in the world of H_{EA} will depend on how likely we think it is under the equity-appeal hypothesis that: (1) center-left governments would consider eliminating the tax subsidy on multiple prior occasions, (2) evidence of such discussions would be uncovered, given that they may or may not have taken place publicly, (3) the right would have resisted the initiative on all such occasions, and (4) governing-coalition informants would attribute their decision not to push forward with the initiative to resistance from the right. Regarding (1), we take as background information that center-left governments were interested in raising revenue and eliminating tax privileges for the wealthy; however, eliminating the tax subsidy may not have been discussed at all if other issues had higher priority on the reform agenda. In contrast, we view propositions (2)–(4) as highly probable. For (2), we judge the probability of discovering evidence if prior initiatives were discussed to be high, drawing on the (logically prior) background information that Fairfield obtained extensive access to finance ministry informants who shared ample information about policy deliberations that was not part of the public record. For (3), we view the probability of right resistance as high, given background information from previous research (Silva 1996, Luna 2010)

⁶ Sources: *Diario de Sesiones del Senado, Legislatura 331, Sesión 14,* July 6, 1995: 37, and *338, Sesión 13,* July 7, 1998: 64; interviews: Aninat, Montes, and all E_{1b} interviews. See Fairfield 2015.

⁷ Sources (interviews): Bitar, Executive Advisor A, Eyzaguirre, Ffrench-Davis, Finance Ministry-A, -B, -H, Jorratt, Marcel, Marfán. See Fairfield 2015.

that the right generally opposed tax increases on principle.⁸ Regarding proposition (4), given that the right held a majority in the senate during this period, we see no reason under H_{EA} that center-left informants would not identify right resistance as a major impediment to reform.

Ultimately, we (somewhat arbitrarily) assign $P(E_I|H_{EA})$ a value of 20%, in light of the possibility that any number of other progressive reform initiatives could have been prioritized in the past. In reality, this probability may be overestimated; however, recall that for the purpose of comparing hypotheses, only the relative likelihoods under the four hypotheses will matter.

$P(E_I|H_I I) = .02\%$

We set this likelihood three orders of magnitude lower than $P(E_I|H_{EA} I)$, yielding a weight of evidence of 30 dB in favor of H_{EA} . If stable institutions produced consensus on eliminating the tax subsidy in 2005, they should have produced consensus on this initiative in previous years as well, since our background information includes the fact that institutions did not change during the intervening time period. If eliminating the tax benefit had been discussed and ruled out on only one occasion, we would be less surprised in the world of H_I ; however, E_I is a conjunction of multiple instances in which institutions failed to promote right-party cooperation.

To reconcile E_1 with H_1 , we might imagine that the reform was ruled out due to internal dissent within the governing coalition, but informants blamed the right coalition for instrumental reasons. However, this scenario is highly improbable. First, as indicated in our E_1 citations, similar analysis was provided by multiple informants across different governments and different government positions (tax agency, finance ministry, congress, presidency), making the possibility of collusion on a false story less likely. Second, our background information includes a high level of confidence in the informants' knowledge and judgment, as well as the fact that the right opposed tax increases on principle and defended tax benefits as "acquired property rights" during this period (Silva 1996). Third, several of the informants cited in E_1 noted that internal dissent had been a problem for other tax issues, which suggests that had internal dissent been relevant in the case of our tax subsidy, they would have divulged that information.

$P(E_I|H_{CC}I)=10\%$

In this world, the right accepted the reform in 2005 because its core constituency no longer had sizable assets that benefitted from the tax subsidy (H_{CC}), even though the right tended to resist even modest tax increases as a matter of principle (I). To avoid a contradiction between H_{CC} and I, (in the absence of further details about the magnitude and rate of change in eligible assets) we must assume that (1) assets did not decline enough from 1998 to 2001 for the right to be willing to give up the tax subsidy, but (2) by 2005, the decline in assets had pushed the right past its threshold of resolve for resisting the reform. This timing seems plausible, especially considering that deliberation on the 2001 reform began in 2000, just two years after the 1998 reform that restricted the scope of the tax benefit. However, under H_{CC} and I, we do not have any clear prediction about when exactly we would expect the tax subsidy to be eliminated, whereas under H_{EA} and I, we have a much more specific rationale for why the right accepted that reform in 2005 as opposed to some number of years earlier or later. Under H_{CC} , the observed timeline of non-reforms and successful reform in 2005 is just one of several more or less equally plausible scenarios in which the successful

⁸ Strictly speaking, by treating this information as background, it should also inform our priors on the hypotheses. Specifically, this information would lower the prior probability on H_{MV} (and possibly H_I as well). However, we nevertheless consider priors that do not penalize H_{MV} relative to other hypotheses for the sake of being conservative, and for the sake of highlighting the impact of the six pieces of evidence from the case study.

reform occurs in a different year. These multiple feasible possibilities reduce the likelihood of the specific timeline observed. Accordingly, we judge the weight of evidence to favor H_{EA} by 3 dB, giving E_I a likelihood of 10% under H_{CC} .

An alternative scenario under H_{CC} and I could be that the decline in eligible assets after 1998 was such that the governing coalition could have eliminated the subsidy before 2005, but misjudged the opposition's threshold for resistance. However, we view this possibility as much more unlikely, given our confidence in the informants' judgment and knowledgability regarding the right's preference and resolve on tax issues.

$P(E_1|H_{MV}I) = 0.02\%$

We judge this likelihood to be roughly the same as $P(E_I|H_I I)$. In a world where the right is in the practice of catering to the median voter's material interests on redistributive issues (following a simple median voter logic where neither preferences nor issue awareness is problematized), consistent right resistance to eliminating the tax subsidy is highly surprising. As discussed under $P(E_I|H_I I)$, center-left governments might have incentives to blame the right if some other problem had precluded reform, but we view that possibility as unlikely.

In sum, while E_I does not discriminate very much between H_{EA} and H_{CC} , this evidence does cast significant doubt on the institutional hypothesis H_I and on the median voter hypothesis H_{MV} . The weight of evidence in favor of H_{EA} compared to either of these two alternatives is 30 decibels.

Discussion Point: Comparing H to rivals, instead of ~H

 E_I helps illustrate the importance of comparing specific alternative hypotheses rather than attempting to evaluate the working hypothesis directly against its full logical negation. Section 3.1 emphasized that assessing likelihoods conditional on an ill-specified $\sim H$ may be essentially impossible. For E_I , this problem is clearly apparent because the likelihoods under our three alternative hypotheses are not all equal. If we try to directly evaluate $P(E_I | \sim H_{EA} I)$, how would we mentally inhabit a world as vague as $\sim H_{EA}$, which is really an amalgamation of three very different kinds of worlds? Only after we have decomposed $\sim H_{EA}$ into H_I , H_{MV} , and H_{CC} can we reason meaningfully about likelihoods by inhabiting each alternative world in turn.

Once we have evaluated each likelihood, we can then calculate $P(E_1 | \sim H_{EA} I)$ if desired:

,

$$P(E_{I}|\sim H_{EA} I) = \frac{P(H_{I}|I) P(E_{I}|H_{I} I) + P(H_{MV}|I) P(E_{I}|H_{MV} I) + P(H_{CC}|I) P(E_{I}|H_{CC} I)}{1 - P(H_{EA}|I)}$$

which follows from Bayes' rule and the assumption that H_{EA} , H_I , H_{MV} and H_{CC} are mutually exclusive and exhaustive. Notice that the likelihood under $\sim H_{EA}$ depends on the prior distribution. Under our indifference prior and our 'skeptical reader' prior, $P(E_1 | \sim H_{EA} I)$ is the same: 4.7%. But under the median-voter irrelevance prior, this likelihood is 6%. These values are not very different; they translate into a weight of evidence for H_{EA} vs. $\sim H_{EA}$ of either 4 dB or 6 dB. However, for E_5 , which also gives likelihoods that vary depending on the rival hypothesis (see below), a similar calculation shows that the weight of evidence for H_{EA} vs. $\sim H_{EA}$ differs by a noticeable 6 dB depending on the priors. But even for the example of E_1 , where our choice of prior does not make much difference to $P(E_1 | \sim H_{EA} I)$, how would we directly intuit or justify a probability of 5-6%? In sum, comparing H_{EA} to well-specified rival hypotheses is essential.

 $E_2 = A$ finance ministry official commented to the investigator that the tax subsidy "was a pure transfer of resources to rich people; there was no way to argue differently. It was not possible for the right to oppose the reform after making that argument about inequality." Likewise, former president Lagos told the investigator that the tax subsidy "never would have been eliminated if I had not taken Lavín at his word"— i.e., if Lagos had not taken seriously Lavín's publicly-professed concern over inequality and issued an equity-appeal challenge.⁹

We treat these two statements as a single piece of evidence: regardless of which hypothesis is correct, we expect that the president and finance ministry officials have communicated extensively and share similar analyses of why the right accepted the reform. In other words, these two statements are strongly (although not completely) dependent under any hypothesis. Recall as well that we are free to aggregate or disaggregate evidence as we see fit to facilitate probability assignments, as long as we ultimately take all of the relevant evidence into account. It is also worth noting that Fairfield's research uncovered similar statements by additional Lagos administration informants. For example, a presidential advisor asserted: "the right was trapped in its discourse and had to cede," (interview, Oct. 21, 2005). This evidence further corroborates the statements in E_2 , but we consider it highly dependent on E_2 and therefore view it as adding little additional inferential weight.

Note that in E_2 we refer to a finance ministry official rather than the specific individual interviewed to denote the relevant equivalence class—similar information conveyed by another knowledgeable member of the finance ministry team would be essentially equivalent in its probative value. However, we explicitly name former president Lagos, since he is the highest relevant authority and is in a class of his own; he is in a unique position to assess the politics surrounding the 2005 reform.

$P(E_2|H_{EA} E_1 I) = 60\%$

 E_2 provides glimpses of the causal mechanism underlying H_{EA} . The first informant does not explicitly mention the equity appeal but is clearly referring to the exchange between Lavín and Lagos that culminated in Lagos' equity appeal with respect to the tax subsidy. Lagos' comment is likewise clearly a reference to the equity appeal. Because E_2 makes the Lagos administration appear savvy and effective at achieving socially-desirable goals while highlighting the right's resistance to redistribution, there should be little reason for the government to conceal this information if H_{EA} is in fact true. $P(E_2|H_{EA} E_1)$ should therefore be fairly high. While we would be surprised if we did not obtain evidence from government informants that the equity appeal mattered under H_{EA} , we choose a value of 60%, bearing in mind that we have a conjunction of two statements in E_2 .

$P(E_2|H_I E_1 I) = 6\%$

If H_I is true, government informants might nevertheless have incentives to attribute the right's support for eliminating the tax subsidy in 2005 to Lagos' equity appeal, since as elaborated above, this story portrays the government in a positive light and the right in a negative light. On the other hand, our background information gives us strong confidence in the informants' knowledgeability, analytical judgments, and sincerity. Balancing these considerations, we take $P(E_2|H_I E_I I)$ to be ten times (10 dB) lower than $P(E_2|H_{EA} E_I I)$.

$P(E_2|H_{CC} E_1 I) = P(E_2|H_{MV} E_1 I) = 6\%,$

These likelihood assignments follow a similar logic as for $P(E_2|H_I E_1 I)$.

⁹ Interviews: Oct 13, 2005 and Sept 20, 2006 respectively.

In sum, the weight of evidence for E_2 in favor of H_{EA} compared to any of the three alternatives is 10 dB, which roughly corresponds to the sound of leaves rustling in the distance, or a pin drop.

Discussion Point: Testimonial evidence and accuracy

Note that E_2 provides an example where the accuracy of the information provided by the sources depends on the hypothesis (Section 3.3.2). Under H_{EA} , the informants' statements must be taken as true, whereas under the rival hypotheses, the statements are necessarily false—the informants are either mistaken or lying.

$E_3 = A$ finance ministry informant reported that after the 2001 Anti-Evasion reform, the Lagos administration tried to reach an agreement with business to eliminate the tax subsidy on several occasions without success.¹⁰

$P(E_3|H_{EA} E_{1-2} I) = 40\%$

This evidence is consistent with H_{EA} : some new dynamic, like the equity appeal, was necessary for eliminating the tax subsidy in 2005. We see few incentives for a finance ministry informant to withhold the information in E_3 ; moreover, as part of our background information, we know that Fairfield achieved strong rapport with finance ministry informants. Nevertheless, additional evidence of unsuccessful prior efforts at eliminating the tax subsidy under H_{EA} is not very surprising in light of E_1 , so we assign a higher probability for $P(E_3|H_{EA} E_{1-2})$ compared to $P(E_1|H_{EA})$. We keep the value below 50% because it is also plausible that the government may not have bothered trying to eliminate the tax subsidy again in light of the previous difficulties.

$P(E_3|H_I E_{1-2} I) = 0.1\%$

If stable institutions alone created consensus with the right on eliminating the tax subsidy in 2005, we would not expect the Lagos administration to try to negotiate the reform directly with business a couple years earlier. Nor can we think of any reasonable instrumental incentive for a finance ministry informant to invent this episode or "misremember" something that did not happen if H_I holds, although there might be some incentive to exaggerate the number of efforts undertaken to eliminate the tax subsidy for the sake of emphasizing the government's commitment to progressive reforms. We set this probability much lower (26 dB) than $P(E_3|H_{EA}, E_{I-2})$, but higher than $P(E_I|H_I)$ for two reasons. First, E_3 does not contradict H_I as directly as E_I , since the right was not involved in the E_3 negotiations. Second, E_3 and E_I may still have some dependence under H_I , thereby making us less surprised to observe E_3 in light of E_I . If H_I is true, E_I must be viewed as a bizarre fluke (however improbable under H_I), such that if the government had approached the right about eliminating the tax subsidy once again, institutions would indeed have compelled the right to accept the reform. However, the experience of E_I may nevertheless have led the government to doubt that right politicians would behave differently, motivating the administration to approach business instead.

$P(E_3|H_{CC} E_{1-2} I) = 0.04\%$

If the right's core constituency no longer valued the tax subsidy (H_{CC}), the government should have been able to negotiate its elimination in direct talks with business. It is very unlikely that a major shift in the structure of assets occurred during the second half of the Lagos administration such that business changed its position on the tax subsidy within the timespan of just a couple years. We

¹⁰ Interview, Finance Ministry-b, Oct. 13, 2005. See Fairfield 2015.

therefore set this probability three orders of magnitude (30 dB) smaller than $P(E_3|H_{EA} E_{I-2})$. This probability ends up slightly higher than the lowest probabilities we have assigned so far—0.02% for $P(E_I|H_{I,MV})$. We view this difference as reasonable, since some other issue could conceivably have hurt government-business relations and caused negotiations to fall apart even if business did not care about the tax subsidy any more (H_{CC}), whereas it is much more difficult to rationalize repeated resistance from the right under H_I or H_{MV} . And again, E_3 and E_1 may be slightly dependent for the same reasons discussed under $P(E_3|H_I E_{I-2})$ above.

Whereas previous evidence has been reasonably consistent with the preference change hypothesis, we now have strong evidence against H_{CC} . The weight of evidence E_3 favors H_{EA} by 30 decibels.

$P(E_3|H_{MV}E_{1-2}I) = 0.1\%$

If the right caters to the median voter's material interests (H_{MV}), then it should support eliminating the tax subsidy, and there would be no reason for the government to attempt negotiating the tax reform directly with business. We set this probability equal to 0.1% for similar reasons as discussed for $P(E_3|H_I E_{I-2})$ above.

$E_4 = A$ technical advisor to the right coalition's congressional bloc told the investigator: "The government said we have to eliminate the tax subsidy, and I said that is a mistake, and they [the right-coalition legislators] said 'no, we will lose votes if we don't approve it. "¹¹

$P(E_4|H_{EA} E_{1-3} I) = 50\%$

This probability depends on how likely we think it is that a technical advisor would reveal such information if the equity appeal did in fact motivate the right to accept reform. We view this probability as fairly high—a technical advisor who holds strong views on tax policy would have few incentives to hide the role of electoral concerns in undermining his or her advice when talking to a foreign academic who did not disclose her own political or policy views. Right-wing economists in Chile as elsewhere have no shortage of technical arguments in favor of inequitable tax measures and actively promote such arguments in the public sphere. Likewise, right legislators should have little incentive to hide or misrepresent their reasons for supporting the 2005 reform in conversation with their own partisan technical advisors.

Notice that $P(E_4|H_{EA} E_{I-3} I)$, is a bit lower than $P(E_2|H_{EA} E_1 I)$ (recall that E_2 = statements from government informants on the importance of the equity appeal). We view this ranking as reasonable since it is less "instrumental" for a right informant to assert E_4 (even though there should be few incentives to hide this information) than it is for a government informant to assert E_2 , and because we view E_4 as largely independent from E_3 and E_2 .

$P(E_4|H_I E_{1-3} I) = 0.05\%$

We judge it highly unlikely that a technical advisor would report that legislators were concerned over losing votes if institutions were what mattered for the right's decision to support the 2005 reform. Accordingly, we set the likelihood of E_4 under H_1 three orders of magnitude (30 dB) lower than its likelihood under H_{EA} . One might imagine that right legislators could have some incentive to cultivate a (false) image of responsiveness to voters in any situation where the content of conversations could be leaked to the public. However, Fairfield's background information includes the fact that many members of the Chilean political elite eschew "populist" tendencies and openly

¹¹ Interview, Instituto Libertad y Desarrollo, Santiago, Chile, Nov. 25, 2005. See Fairfield 2015.

advocate pursing "technically appropriate" policies rather than catering to public opinion on economic issues.

Notice that the weight of evidence E_4 in favor of H_{EA} relative to H_I is 4 decibels higher than the corresponding weight of evidence E_3 , which is a noticeable difference. This makes sense intuitively because E_4 is a more direct statement about the mechanism underlying H_{EA} and is also less consistent with H_I .

$P(E_4|H_{CC} E_{1-3} I) = 0.05\%$

Following a similar logic, this likelihood should be roughly the same as $P(E_4 | H_I E_{1-3})$.

$P(E_4|H_{MV}E_{1-3}I) = 50\%$

The informant is simply reporting concern over votes, which should be as likely if a simple median voter logic were at work as it is if the equity appeal in the context of a major campaign where inequality had become highly salient were critical for igniting that concern.

Discussion Point: Logical dependence

Notice that in assessing these likelihoods, we have considered E_4 to be essentially independent of E_{1-3} . There may well be some dependence with E_2 —the government informants' statements about the equity appeal, but since E_2 and E_4 come from sources on opposite sides of the political spectrum and because the statements contain somewhat different information, we judge the potential dependence to be negligible for our purposes. More specifically, under H_{EA} , there need be no logical or causal dependence, because sources on both sides are simply stating the truth, and hearing one account does not appreciably increase or decrease our surprise in hearing a similar narrative from the other side. Under the other hypotheses, both the left and right informants somehow came to relate similar incorrect accounts. Either they were misinformed by the same rumors or news accounts, or both jumped to similar reasonable, if erroneous, conclusions. We judge the latter scenario most likely under any of the three rival hypotheses, so we consider any probabilistic dependence to be small.

E_5 = The newspaper El Mercurio, which is widely recognized as having strong ties to business and the right, reported that Lavín's advisors attributed Lagos's narrow victory in the 1999 presidential election to the right's rejection of a labor-rights bill that the center-left government sent to congress during the campaign. Lavín's advisors compared the 2005 bill eliminating the tax subsidy to that 1999 labor bill and commented: "The center-right is not willing to fall into the 1999 trap again."¹² Two additional articles from the same newspaper referred to similar points regarding the right's comparison of the 1999 bill and the 2005 bill.¹³

It is convenient to treat E_5 as a single piece of evidence because the existence of the second and third articles is strongly dependent on the first. The May 12 and May 13 articles in particular should be considered highly correlated, meaning that the probability of observing both does not differ much from the probability of seeing the first. Although the authors are different, a follow-up article in the same newspaper articulating similar points regarding the same policy issue is hardly surprising. The June 15 article includes some distinct sources of information and can therefore be

¹² *El Mercurio*, May 13, 2005.

¹³ El Mercurio, May 12, 2005; El Mercurio, June 15, 2005

considered less correlated with the previous articles and hence providing more independent corroboration, but regardless of which hypothesis we are evaluating, this third article is less surprising following the appearance of the prior articles. The earlier articles may well have helped to publicly disseminate a particular perspective among readers and subsequent commentators.

$P(E_5|H_{EA} E_{1-4} I) = 3\%$

This evidence, which stresses the timing of the reform, the difficult position it created for the right, and anticipated electoral costs—is consistent with the hypothesized mechanism underlying H_{EA} , although it does not explicitly refer to Lagos' equity appeal. However, we judge the likelihood to be low because it is surprising for sources on the right to openly admit that the government's strategy put them in a tight place— E_5 strikes us as quite embarrassing.

We set $P(E_5|H_{EA} E_{I-4} I)$ a bit lower than $P(E_2|H_{I, CC, MV} E_1 I)$, the likelihood of hearing government informants instrumentally or mistakenly emphasize the importance of the equity appeal under any one of the rival hypotheses where the equity appeal did not matter—we are more surprised to find the right admitting E_5 in the press than we would be to learn that the government informants were being instrumental or mistaken in their analysis.

$P(E_5|H_I E_{1-4} I) = 0.003\%$

 E_5 is much less likely in the world of H_I compared to the H_{EA} world. If institutions motivated the right's decision on the tax subsidy in 2005, there would be no instrumental reason for the right to state that it felt trapped and anticipated electoral costs to rejecting the reform. We set this likelihood three orders of magnitude (30 dB) less than $P(E_5|H_{EA} E_{I-4})$, and slightly lower than $P(E_4|H_I E_{I-3})$ since E_5 seems much more embarrassing than E_4 .

$P(E_5|H_{CC} E_{1-4} I) = 0.003\%$

 E_5 is also highly implausible if what really mattered for the right's decision on the tax subsidy in 2005 were changing preferences among its core constituency (H_{CC}). If H_{CC} holds, we would expect a right informant to state that the tax subsidy was simply not an important tax benefit.

$P(E_5|H_{MV}E_{1-4}I) = 0.03\%$

 E_5 implies that the right had voted against public opinion in the past, associated that decision with electoral punishment, and sought to learn from its mistake—a scenario consistent with H_{MV} . However, we would not expect the right to publicly announce that it had fallen into a trap in 1999. In addition, the emphasis on timing in E_5 suggests that outside of presidential campaigns, the right would not have feared electoral punishment for deviating from the median voter's material interest, which would be surprising under our basic version of a median-voter theory. Given these considerations, we judge the likelihood of E_5 under H_{MV} to be two orders of magnitude lower than under H_{EA} . This assignment yields a weight of evidence of 20 decibels in favor of the equity appeal hypothesis, roughly equivalent to the difference between a normal conversation and an alarm clock.

Discussion Point: Logical dependence

Note that under any of our hypotheses, E_5 and E_4 could have some dependence, since the newspaper stories could conceivably have influenced right informants' perceptions or memories. However, E_5 seems much more surprising than E_4 under any hypothesis, so we view any potential dependency between E_4 and E_5 as having little meaningful upward effect on the likelihoods of E_5 . Any conditioning on E_4 will only have an effect at the margins. E_6 = When asked about the 2005 reform, a right-party deputy with long-term experience on the congressional finance committee told the investigator: "Our candidate made a commitment, and it was also a difficult moment for him. Therefore the political decision was made to support what the candidate said; we had to take maximum safeguards so that it would not be a disaster. The opposition demonstrated that this time it would accept things that usually it was not disposed to accept so as not to harm the presidential option—in this case it would do something popular."¹⁴

$P(E_6|H_{EA} E_{1-5} I) = 5\%$

 E_6 captures the causal mechanism underlying H_{EA} more completely than any of the pieces of evidence previously analyzed. According to the informant, the right was concerned that its presidential candidate would lose votes if their legislators defended the tax subsidy, and that concern drove the right to deviate from the decision it otherwise would have made on the reform. Given the reasonable assumption that average citizens would not have been familiar with, or at least would not have been thinking about the tax subsidy—an obscure benefit for wealthy stockowners—prior to the exchange between Lavín and Lagos, we can infer that Lagos' equity appeal drove the right's manifest concern over public opinion, even though the informant does not explicitly refer to that appeal.

The likelihood of E_6 again depends on whether we think such a right informant would admit concern over votes in this manner if H_{EA} holds. On the one hand, we have already observed other right sources providing similar evidence (E_5) , and rapport with informants is part of our background information. Nevertheless, we judge the probability of observing E_6 to be low, because E_6 is surprisingly candid and much more explicit than E_5 , in a manner that runs against the expected direction of instrumental bias. It seems strategically disadvantageous and embarrassing for a right party deputy to state that the government succeeded in driving his party to do something it otherwise would not have done, and to acknowledge that the party did not share Lavín's purported enthusiasm for eliminating the tax benefit to promote equity. An instrumental response would have instead entailed no comment, or a denial that the government's strategy mattered, or agreement with the government's rationale for reform, in line with Lavín's public statement following Lagos' equity challenge (see E_7 below). Overall, these considerations lead us to assign a low value of 5% for $P(E_6|H_{EA} E_{1-5})$ —slightly higher than $P(E_5|H_{EA} E_{1-4})$ because in light of E_5 , we are a bit less surprised to hear another right informant admitting similarly embarrassing points.

$P(E_6|H_I E_{1-5} I) = 0.005\%$

We judge it far less likely that a right informant would spell out the mechanism underlying H_{EA} if alternative hypothesis H_I holds instead. The informant did not simply say that the right agreed to eliminate the tax subsidy because public opinion supported the reform, a plausible instrumental, socially desirable response that could make the right appear democratic and responsive to the majority interest. Instead, the informant indicated that the right was in a tough spot and felt pressured by public opinion against its will to support a reform it did not like. If institutions and consensual politics (H_I) were what really motivated the right to accept reform, we would not expect an informant to tell a potentially embarrassing story about feeling forced to do something it did not want to do for the sake of protecting its candidate's electoral prospects. We assign $P(E_6|H_I E_{I-5})$ a

¹⁴ Interview: UDI, Dec. 23, 2005. See Fairfield 2015. Note that this evidence is similar in structure to the "smoking gun" evidence that Bennett (2015: 279) highlights in his discussion of Tannenwald's (2007) research: a decision-maker who disagreed with the policy decision (in Tannenwald's case, non-use of nuclear weapons; in Fairfield's case, elimination of the tax subsidy) essentially articulates the author's hypothesized explanation for why that decision was made (Tannenwald: normative constraints; Fairfield: well-timed equity appeal).

value of 0.005%, three orders of magnitude (30 dB) smaller than $P(E_6|H_{EA} E_{1-5})$, but a bit higher than $P(E_5|H_I E_{1-4})$ since there is some potential dependence between E_6 and E_5 (E_6 is less surprising in light of E_5). We do not boost the probability very much, however, because a scenario where H_I is true but an incorrect story about why the right accepted the 2005 reform emerged in the press (E_5) and then diffused among the right remains highly unlikely.

$P(E_6|H_{CC} E_{1-5} I) = 0.005\%$

Our rationale follows that described above for $P(E_6|H_{CC} E_{1-5})$. We view E_6 to be highly implausible if what really mattered for the right's 2005 decision on the tax subsidy were changing preferences among its core constituency (H_{CC}). If H_{CC} holds, we would expect a right informant to state that the tax subsidy was simply not an important tax benefit.

$P(E_6|H_{MV}E_{1-5}I) = 0.006\%$

While the electoral logic in the informant's statement does not contradict a simple median voter model (H_{MV}), the clear implication that the right would have rejected the reform if its presidential candidate had not been in a tight place indicates that responsiveness to public opinion on this issue was a deviation from the right's usual behavior on taxation and redistribution. However, this evidence is certainly more consistent with a median voter hypothesis compared to H_I or H_{CC} , which act through non-electoral mechanisms. We therefore set $P(E_6|H_{MV}E_{I-5}I)$ slightly higher than the likelihood under H_I and H_{CC} .

In sum, we can view E_6 as a smoking gun for H_{EA} —the likelihood is fairly low under this hypothesis, but the weight of evidence in favor of H_{EA} compared to each of the three alternatives is high—29 decibels relative to H_{MV} and 30 decibels relative to H_{CC} and H_I .

We consider one final additional observation for purely pedagogical purposes:

$E_7 = A$ right party deputy responded when asked about the 2005 reform that Lavín agreed with Lagos' proposal and the right therefore supported the initiative in congress.¹⁵

Intuitively, this evidence is not informative—the likelihood of E_7 should be very similar under each of the four hypotheses, perhaps around 50%. This statement is what we would expect to hear from the right—not admitting any internal discontent with Lavín's declaration. If H_{EA} holds, we would not be very surprised to hear E_7 because it is instrumentally preferable for the right not to acknowledge that the government's strategy forced the opposition to do something it preferred not to do. If the institutional hypothesis H_I holds, E_7 is not very surprising because the right is portraying the reform as consensual and non-controversial. E_7 is consistent with H_{CC} and H_{MV} as well, since either changing preferences among business or a simple median-voter logic could explain this informant's assertion that the right was willing to go along with Lavín's support for eliminating the tax subsidy. Of course, we would need to condition the likelihood of observing E_7 on each of the prior pieces of evidence E_1-E_6 ; however, we could obviate this complication by incorporating this piece of evidence first (as E_0 instead of E_7). This statement therefore does little to help discriminate between the four hypotheses and is not relevant for the causal analysis. Fairfield (2013) accordingly does not discuss this piece of evidence.

¹⁵ Interview: Dittborn, 2005. See Fairfield 2015.

A4.3 Weights of Evidence and Comparison to Process-Tracing Tests

Figure A.1 summarizes our likelihood assignments for each piece of evidence $E_1 - E_6$. The figure displays the weight of evidence in favor of the equity appeal hypothesis relative to each alternative hypothesis. The larger the weight of evidence, the more probative value that piece of evidence provides against the alternative hypothesis in question.



Figure A.1: Weight of Evidence (dB), E₁-E₆

Summary of Evidence

E_1	The center-left discussed the reform multiple times during the 1990s but ruled it out given resistance from the right coalition, according to multiple sources on the left
E_2	Government informants suggested that Lagos' public exchange with Lavín forced the right to approve the reform
E_3	Government informants reported failed efforts to reach an agreement with business on eliminating the tax subsidy during the early years of Lagos' term (before 2005)
E_4	A right-party technical advisor reported that opposition-coalition legislators feared electoral punishment if they did not approve the reform
E_5	The right-learning national newspaper reported that Lavín's advisors viewed the reform proposal as an electoral "trap"
E_6	A right-party deputy reported that the opposition coalition reluctantly accepted reform "so as not to harm the presidential option"

A few differences are worth highlighting between our Bayesian analysis and the original process-tracing tests appendix (Fairfield 2013) regarding the probative value of discreet pieces of evidence. Figure A.1 indicates that E_2 (government informants on the equity appeal) should be

considered overall the least probative piece of evidence, whereas the process-tracing tests appendix took that information to be strongly supportive of H_{EA} , although E_2 was correctly identified as less decisive than similar statements from the right ($E_4 - E_6$). And whereas the process-tracing tests appendix viewed E_1 and E_3 as only weakly supporting H_{EA} , our Bayesian analysis assigns a strong weight of evidence to these pieces of information relative to the alternative hypotheses we consider here: 25–30 dB (with the exception that E_1 does not discriminate much between H_{CC} and H_{EA}).

The lesson is that explicitly elaborating alternative hypotheses, rather than attempting to assess a hypothesis (the equity appeal had an effect) against its negation (it had no effect), can help us better assess the probative value of our evidence. This is one illustration of why Bayesian analysis is preferable to the process-tracing tests approach. Stated more strongly, the weight of evidence depends *by definition* on which hypotheses we compare; we cannot judge how decisive the evidence is with respect to our working hypothesis alone, without considering concrete alternatives.

A5. Inference via Bayes' Rule

We can now apply Bayes' theorem to calculate posterior probabilities for the hypotheses in light of the evidence:

$$P(H_k|E|I) = \frac{P(H_k|I)P(E|H_k|I)}{P(E|I)} = \frac{P(H_k|I)P(E|H_k|I)}{\sum P(H_k|I)P(E|H_n|I)},$$
(A.1)

E represents the conjunction of all six pieces of evidence E_{I-6} , and the sum in the denominator runs over all four hypotheses. Recall that we are treating the four hypotheses as mutually exclusive and exhaustive; this assumption is taken as part of the background information *I*. Expanding the denominator and suppressing the background information *I* to save space, we have:

$$P(H_k|E) = \frac{P(H_k) P(E|H_k)}{P(H_{EA}) P(E|H_{EA}) + P(H_l) P(E|H_l) + P(H_P) P(E|H_P) + P(H_{MV}) P(E|H_{MV})} , \qquad (A.2)$$

where

$$P(E|H) = P(E_6|H E_{1-5}) P(E_5|H E_{1-4}) P(E_4|H E_{1-3}) P(E_3|H E_{1-2}) P(E_2|H E_1) P(E_1|H) , \qquad (A.3)$$

because we can always break down the joint probability of some composite evidence $E = E_{ab}$ as follows: $P(E_a E_b | H) = P(E_b | H E_a) P(E_a | H)$, in other words, the likelihood of all the evidence is the probability of one piece of evidence E_a conditional on the hypothesis and on the rest of the evidence, E_b .

The charts below illustrate how the probabilities for the hypotheses change after each piece of evidence is considered, for each of the three scenarios corresponding to different priors on the hypotheses (Figure A.2). In each scenario, the posterior probability on H_{EA} reaches near certainty, while the probability on the closest competing hypothesis falls to at most 10^{-7} (Table A.2). Starting from the most unfavorable prior on Fairfield's explanation—0.0003% in Scenario 3 corresponding to a highly-skeptical reader—our confidence in H_{EA} increases to 97% after incorporating only the first four pieces of evidence. The log-scale charts illustrate how subsequent pieces of evidence cast more and more doubt on the alternative explanations, reducing their posterior probabilities by additional orders of magnitude.

Table A.2

a) Prior and Posterior Probabilities on the Hypotheses

	Scenario 1:		Scenario 2:		Scenario 3:	
	Indifference		Median-Voter Irrelevance		Skeptical Reader	
	Prior	Posterior	Prior	Posterior	Prior	Posterior
H _{EA}	25%	1.0	33.3%	1.0	0.0003%	1.0
H _I	25%	2.5 E-16	33.3%	2.5 E-16	33.3%	2.8 E-11
H _{CC}	25%	5.0 E-14	33.3%	5.0 E-14	33.3%	5.6 E-9
H _{MV}	25%	3.0 E-12	0.001%	9.0 E-17	33.3%	3.3 E-7

b) Prior and Posterior Odds Ratios for H_{EA} Relative to Rivals (in decibels)

	Scenario 1:		Scenario 2:		Scenario 3:	
	Indifference		Median-Voter Irrelevance		Skeptical Reader	
	Prior	Posterior	Prior	Posterior	Prior	Posterior
H_{EA} : H_{I}	0	156	0	156	-50	106
H_{EA} : H_{CC}	0	133	0	133	-50	83
H_{EA} : H_{MV}	0	115	45	160	-50	65



Figure A.2: Probabilities of Hypotheses After Incorporating Evidence (E₁-E₆)



A6. Reordering the Evidence

We now carry out our analysis using a different ordering of the six pieces of evidence. This exercise provides an important internal consistency check and further illustrates the challenges of explicitly applying Bayesian analysis to qualitative research.

The rules of conditional probability demand that the order in which evidence is incorporated in Bayesian analysis must not affect the final posterior probabilities on the hypotheses (Section 3.3.4). When attempting to quantify inherently qualitative data, however, we cannot expect to exactly reproduce our results—there is too much arbitrariness inherent in assigning numerical values to likelihoods. Accordingly, we would view agreement to within a factor of two or so (about 3 dB) as acceptable. On our first pass, however, we ended up with discrepancies of several orders of magnitude between $P(E'_{1-6}|H_i I)$ for the reordered evidence and $P(E_{1-6}|H_i I)$ from the original exercise for some hypotheses. To redress this problem, we then iteratively adjusted our numerical values for both orderings by carefully comparing the probability assigned to each likelihood in the new ordering with the likelihood of that respective piece of evidence in the original ordering, and thinking about how conditioning on a different body of previously-incorporated evidence should affect the relative numerical assignments.

This reordering exercise also provides an opportunity to assess how the sequencing of evidence affects the difficulty of conditioning on previously-incorporated evidence. The most noteworthy difference in the new ordering scheme (Table A.3) is that we place the right-party deputy's elaboration of the mechanism underlying H_{EA} —the most decisive single piece of evidence against the three alternative hypotheses—first instead of last, and we move the similar but less discriminating statements from government informants to the end. We anticipated that it would be easier to assess likelihoods when the most decisive pieces of evidence come last (as in the original ordering), because those likelihoods would be large regardless of the evidence we analyzed previously. Nevertheless, we found conditioning on previous-incorporated evidence to be challenging for both sequencings—especially in cases where previous pieces of evidence would have to be considered a fluke under the given hypothesis.

Sequence 2	Sequence 1	Evidence
<i>E</i> ₁ '	E_6	A right-party deputy reported that the opposition coalition reluctantly accepted reform "so as not to harm the presidential option"
<i>E</i> ₂ ′	E_4	A right-party technical advisor reported that opposition-coalition legislators feared electoral punishment if they did not approve the reform
E_{3}'	E_1	The center-left discussed the reform multiple times during the 1990s but ruled it out given resistance from the right coalition, according to multiple sources on the left
<i>E</i> 4'	E_3	Government informants reported failed efforts to reach an agreement with business on eliminating the tax subsidy during the early years of Lagos' term (before 2005)
<i>E</i> 5'	E_5	The right-learning national newspaper reported that Lavín's advisors viewed the reform proposal as an electoral "trap"
E ₆ '	E_2	Government informants suggested that Lagos' public exchange with Lavín forced the right to approve the reform

Table A.3: Reordering the Evidence

We explain below the rationale for our likelihood assignments in the new ordering scheme. In practice, we have kept most likelihood ratios roughly the same for each piece of evidence across the two sequencing schemes; when conditioning on a different body of prior evidence, we generally shift the likelihoods under each hypothesis by a constant factor compared to their values in the original ordering. While there is no reason to expect that likelihood ratios should remain the same in general, this approach simplifies the exercise, and we found no compelling reason in our case to alter any of the likelihood ratios. Readers who do not wish to delve into the details may skip to the final section of this appendix.

 $E_1' = E_6 =$ When asked about the 2005 reform, a right-party deputy with long-term experience on the congressional finance committee and intimate knowledge of the party's internal decisionmaking processes told the investigator: "Our candidate made a commitment, and it was also a difficult moment for him. Therefore the political decision was made to support what the candidate said; we had to take maximum safeguards so that it would not be a disaster. The opposition demonstrated that this time it would accept things that usually it was not disposed to accept so as not to harm the presidential option—in this case it would do something popular."

$P(E_I'|H_{EA}I) = 1\%$

 E_{I}' is the first surprising evidence from the right that we incorporate in this new ordering, whereas the similar information reported by right sources in the press (E_{5}) was incorporated before this piece of evidence in the original exercise. We set $P(E_{I}'|H_{EA} I)$ a factor of three lower than $P(E_{5}|H_{EA} E_{I-4} I)$ I) because E_{I}' is the more candid, more detailed, and hence more surprising evidence. We set $P(E_{I}'|H_{EA} I)$ a factor of five lower than $P(E_{6}|H_{EA} E_{I-5} I)$ because in the original exercise, we had to condition on E_{5} , which made E_{6} less surprising that it would otherwise be.

$P(E_1'|H_I I) = 0.001\%$

Following our discussion of E_6 under H_I in the original exercise, we assign $P(E_I'|H_I I)$ a value three orders of magnitude (30 dB) smaller than $P(E_I'|H_{EA} I)$ to convey the much lower probability of observing this evidence under the institutional hypothesis. We set $P(E_I'|H_I I)$ a factor of five lower than $P(E_6|H_{EA} E_{I-5} I)$ because under the new ordering, this is the first piece of evidence we incorporate, whereas in the original ordering we conditioned on previous evidence which included similar information reported by right sources in the press (E_5).

$P(E_1'|H_{CC}I) = 0.001\%$

$P(E_1'|H_{MV}I) = 0.0012\%$

Following similar logic to that described for $P(E_1'|H_I I)$ above, we set $P(E_1'|H_{CC} I)$ and $P(E_1'|H_{MV} I)$ five times lower than $P(E_6|H_{CC} E_{1-5} I)$ and $P(E_6|H_{MV} E_{1-5} I)$ respectively.

$E_2' = E_4 = A$ technical advisor to the right coalition's congressional bloc told the investigator: "The government said we have to eliminate the tax subsidy, and I said that is a mistake, and they [the right-coalition legislators] said 'no, we will lose votes if we don't approve it."

$P(E_2'|H_{EA} E_1' I) = 60\%$

After observing E_1' , we are not as surprised to find another source on the right corroborating the electoral motivation. Since E_2' has some dependence on E_1' , we set $P(E_2'|H_{EA}, E_1', I)$ slightly higher than $P(E_4|H_{EA}, E_{1-3}I)$ in the original exercise, where we had not yet taken into account the UDI deputy's comments.

$P(E_2'|H_I E_1' I) = 0.06\%$

On its own, E_2' would be as unlikely as E_1' under H_I , but E_2' and E_1' have some dependence because they contain similar information from informants on the right. We assign $P(E_2'|H_I E_1'I)$ a value that is three orders of magnitude (30 dB) smaller than $P(E_2'|H_{EA} E_1'I)$ but larger than $P(E_1'|H_I I)$. Note that the dependence between E_2' and E_1' also raises $P(E_2'|H_I E_1'I)$ above $P(E_4|H_I E_{1-3} I)$ in the original exercise.

$P(E_2'|H_{CC} E_1' I) = 0.06\%$

Following a similar logic, this probability should be basically the same as $P(E_2'|H_1E_1'I)$.

$P(E_2'|H_{MV}E_1'|I) = 60\%$

Following the logic discussed in the original sequencing regarding $P(E_4 | H_{MV} E_{1-3} I)$, $P(E_2' | H_{MV} E_1' I)$ should be essentially equal to $P(E_2' | H_{EA} E_1' I)$. Given some dependence between E_2' and E_1' , $P(E_2' | H_{MV} E_1' I)$ is slightly higher than $P(E_4 | H_{MV} E_{1-3} I)$. Note that under H_{MV} , we must view those elements of E_1' that go beyond a strict median voter logic as a fluke, where the informant was either mistaken or lying. However, the elements of E_1' that simply express concern over votes are consistent with H_{MV} , and those elements do have some degree of dependence with E_2' —we are now hearing another informant on the right indicate concern over votes.

 $E_3' = E_1$ = Governing-coalition informants and congressional documents indicated that the center-left coalition had discussed including a measure to definitively eliminate the tax subsidy in multiple prior tax reforms (1990, 1995, 1998, 2001). However, that measure was ultimately ruled out as infeasible on every such occasion due to resistance from the right.

 $P(E_3'|H_{EA} E'_{1-2} I) = 20\%$

 $P(E_3'|H_I E'_{1-2} I) = 0.02\%$

 $P(E_3'|H_{CC} E'_{1-2} I) = 10\%$

 $P(E_3'|H_{MV}E'_{1-2}I) = 0.02\%$

We set these probabilities the same as the respective $P(E_1|H_i I)$'s in the original exercise since we view E_3' as more or less independent from E_1' and E_2' under all the hypotheses. Note that in practice, it would be extremely difficult to condition the likelihood of E_3' on E_1' and E_2' under any

of the alternative hypotheses (H_I , H_{CC} , H_{MV}), since E_1' is an extremely rare event under all of these hypotheses and E_2' is also extremely rare under H_I and H_{CC} . The question is whether this prior evidence makes E_3' any more or less consistent with the alternative hypotheses, and it is very hard to evaluate given that we are in highly improbable situations that make little sense—we would have to be in a world of bizarre coincidences or massive misunderstandings. It is difficult to even assess whether the prior information would lead us to increase or decrease the likelihood of observing E_3' .

 $E_4' = E_3 = A$ finance ministry informant reported that after the 2001 Anti-Evasion reform, the Lagos administration tried to reach an agreement with business to eliminate the tax subsidy on several occasions without success.

 $P(E_4'|H_{EA} E'_{1-3} I) = 40\%$

 $P(E_4'|H_I E'_{1-3} I) = 0.1\%$

 $P(E_4'|H_{CC} E'_{1-3} I) = 0.04\%$

 $P(E_4'|H_{MV}E'_{1-3}I) = 0.1\%$

We again set these probabilities equal to the respective $P(E_3 | H_i E_{1-2} I)$'s in the original exercise because the new ordering of the evidence does not introduce any clearly distinct dependencies upon which we must condition (E_3 'and E_4 ' have some dependence, but $E_3' = E_1$ came before $E_4' = E_3$ in the original exercise as well).

 $E_5' = E_5 =$ The newspaper El Mercurio, which is widely recognized as having strong ties to business and the right, reported that Lavín's advisors attributed Lagos's narrow victory in the 1999 presidential election to the right's rejection of a labor-rights bill that the center-left government sent to congress during the campaign. Lavín's advisors compared the 2005 bill eliminating the tax subsidy to that 1999 labor bill and commented: "The center-right is not willing to fall into the 1999 trap again." Two additional articles from the same newspaper referred to similar points regarding the right's comparison of the 1999 bill and the 2005 bill.

$P(E_5'|H_{EA} E'_{1-4} I) = 15\%$

 E_5' is consistent with the hypothesized mechanism underlying H_{EA} , similar to E_1' . As discussed for $P(E_1'|H_{EA} I)$, it is unlikely that sources on the right would openly admit that the government's equity appeal put them in a tight place. However, given that an UDI deputy already outlined a similar rationale (E_1') , it becomes more likely that another source on the right would also admit this logic—stressing the timing of the reform, the difficult position it created for the right, and anticipated electoral costs. Accordingly, we assign a probability of 15%, (roughly) one order of magnitude higher than $P(E_1'|H_{EA} I)$.

$P(E_5'|H_I E'_{1-4} I) = 0.015\%$

As with E_I' , observing E_5' is unlikely if institutions motivated the right's decision on the tax subsidy, so before conditioning on previously-analyzed evidence, $P(E_5'|H_I I)$ should be about as low

as $P(E_I'|H_I I)$, which we had assigned 0.001%. However, E_5' should have some dependence on E_I' . Under H_I , these stories are not true, but if one of these stories were to circulate, it is less surprising to hear a similar story from a different source within the right. We therefore set $P(E_5'|H_I E'_{I-4} I)$ equal to 0.015%, three orders of magnitude (30 dB) lower than $P(E_5'|H_{EA} E'_{I-4} I)$ and roughly one order of magnitude higher than $P(E_I'|H_I I)$.

 $P(E_5'|H_{CC} E'_{1-4} I) = 0.015\%$ following a similar logic as for $P(E_5'|H_I E'_{1-4} I)$.

$P(E_5'|H_{MV}E'_{1-4}I) = 0.15\%$

As with E_5 , E_5' should be much less likely under H_{MV} than under H_{EA} , but more plausible than under H_I and H_{CC} . As with the conditional probabilities under H_I and H_{CC} , we assign a value that preserves the likelihood ratio relative to H_{EA} from the original exercise, since we judge this evidence equally probative under the new ordering.¹⁶

Note that each $P(E_5'|H_i E'_{I-4} I)$ is higher (by a factor of 5) than the corresponding $P(E_5|H_i E_{I-4} I)$ in the original exercise because of the dependence between E_5' and E_1' in this new ordering. Under H_{EA} , this dependency arises because we update our expectations regarding how likely right informants are to acknowledge (aspects of) the potentially embarrassing causal mechanism, whereas under the alternative hypotheses, the dependency arises because a story may have circulated even if it is incorrect. In general, there is no reason to expect that the factor by which we increase the likelihoods under H_{EA} should be the same as the factor by which we increase the likelihoods under the alternative hypotheses in light of the dependencies. However, we see no way to reliably quantify these relative effects and therefore opt for a common factor.

 $E_6' = E_2 = A$ finance ministry official commented to the investigator that the tax subsidy "was a pure transfer of resources to rich people; there was no way to argue differently. It was not possible for the right to oppose the reform after making that argument about inequality." Likewise, former president Lagos told the investigator that the tax subsidy "never would have been eliminated if I had not taken Lavín at his word"— i.e., if Lagos had not taken seriously Lavín's publicly-professed concern over inequality and issued an equity-appeal challenge.

$P(E_6'|H_{EA} E'_{1-5} I) = 70\%$

The likelihood of this evidence conditional on H_{EA} alone, $P(E_6'|H_{EA} I)$, should be much greater than the probability of hearing sources on the right confess a similar story, for example, $P(E_1'|H_{EA} I)$ which we set to 1%. As noted in the original exercise, the rationale is that E_6' makes the government appear savvy and effective at achieving socially-desirable goals while highlighting the right's resistance to redistribution. Moreover, E_6' is not very surprising in light of our similar prior evidence from right sources (E_1' and E_5'). We therefore set $P(E_6'|H_{EA} E'_{1-5} I)$ equal to 70%, slightly higher than $P(E_2|H_{EA} E_1 I)$, because the new ordering entails conditioning on different prior evidence which has some dependence.

¹⁶As should be the case, this likelihood is higher than $P(E_1 | H_{MV}, I)$ since E_5' does not make the right look as bad as E_1' , but lower than $P(E_2 | H_{MV}, E_1', I)$ since E_1' was a more median-voter compatible statement from an informant on the right.

$P(E_6'|H_I E'_{1-5}I) = P(E_6'|H_{CC} E'_{1-5}I) = P(E_6'|H_{MV} E'_{1-5}I) = 7\%$

As in the original exercise, we set each of these three conditional probabilities ten times lower than $P(E_6'|H_{EA} E'_{1-5} I)$. Note that these probabilities are also slightly larger than the corresponding $P(E_2 | H_i E_1 I)$'s because E_6' is somewhat dependent on the prior evidence E_1' and E_5' . As explained above, private communications among the political elite and news articles could result in a shared analysis (however incorrect under these alternative hypotheses) regarding why the right accepted the reform.

Before continuing, it is instructive to conduct a consistency check across the two orderings on the likelihoods involving the right-candidate campaign advisors' analysis $(E_5' = E_5)$ and the right-party deputy's statement $(E_1' = E_6)$, the two pieces of evidence that are most strongly dependent. Since the joint probability of two propositions *A* and *B* can be broken down either as P(A B|H) = P(A|H) P(B|HA) or as P(A B|H) = P(B|H) P(A|HB), we have:

$$\frac{P(B|HA)}{P(A|HB)} = \frac{P(B|H)}{P(A|H)} , \qquad (A.4)$$

Applying this relationship to our two pieces of evidence using the first ordering scheme, we have:

$$\frac{P(E_6 \mid H E_5 E_{1-4})}{P(E_5 \mid H E_6 E_{1-4})} = \frac{P(E_6 \mid H E_{1-4})}{P(E_5 \mid H E_{1-4})} .$$
(A.5)

If we consider E_6 to be independent of E_{1-4} , an assumption that we made in practice when assigning likelihoods, then we can replace the numerator on the right hand side of (A.5) with $P(E_1'|H)$. If we also take the right campaign advisors evidence (E_5) as independent from the government informants' statements about the equity appeal (E_2)—another assumption that we made when assigning likelihoods in the first ordering—then the denominator on the left-hand side of (A.5) becomes: $P(E_5 | H E_6 E_{1-4}) = P(E_5 | H E_6 E_1 E_{3-4}) = P(E_5' | H E'_{1-4})$, where we have relabeled the pieces of evidence according to the second (primed) ordering scheme. Equation (A.5) can then be rewritten as:

$$\frac{P(E_6 \mid H \mid E_{1-5})}{P(E_5' \mid H \mid E'_{1-4})} = \frac{P(E_1' \mid H)}{P(E_5 \mid H \mid E_{1-4})} .$$
(A.6)

Both the left-hand side and the right-hand side of (A.6) can be calculated directly from our likelihood assignments. For each hypothesis, the equation is satisfied. For example, for H_{EA} we have 0.05/0.15 = 0.33 = 0.001/0.003.

Of course, the assumptions we have made about independence probably do not hold exactly. E_5 and E_2 , for example, may well have some dependence. If the equity appeal hypothesis is false, one could imagine mechanisms through which the political elite might nevertheless converge on a common perception of the equity appeal's importance. If H_{EA} is correct, it could be the case that sources in E_5 learned about the equity appeal's effect in part from the sources in E_2 , or vice versa. Any dependence is probably small, however. Under H_{EA} , there are many ways that informants could learn about the equity appeal's importance, while under $\sim H_{EA}$, informants might still independently jump to the reasonable albeit incorrect conclusions expressed in E_2 and E_5 (see similar discussion regarding E_2 and E_4 in Section A4.2). Again, it can be very difficult to assess

logical and causal dependencies in qualitative data given the multiple and complex ways in which such dependencies could arise.

We proceed to calculate posterior probabilities on the four hypotheses by applying the Bayes' rule (equation A.2) as before. The new sequencing produces posteriors that are essentially identical to those calculated with the original ordering, although as discussed previously, this consistency was achieved only after extensive deliberation and iterative adjustments of the likelihoods across the two sequencings. The charts below show how our degree of belief in each hypothesis changes after incorporating each piece of evidence following the new sequencing; for comparison we reproduce the charts corresponding to the first ordering of evidence as well (Figure A.3). In the first and second scenarios (equal priors and median-voter irrelevance priors), the first piece of evidence alone—the right party deputy's candid statement—boosts our confidence in H_{EA} above 99%. In the skeptical reader scenario, this second ordering of the evidence establishes H_{EA} as the leading explanation more quickly than the first sequencing—we reach 84% confidence in the equity appeal hypothesis after incorporating just the first three pieces of evidence. It is interesting to note that the probability on H_{MV} increases as the first two pieces of evidence are taken into account, reaching 99%. This result arises because H_{EA} starts out with such a low prior, H_{MV} fits best with E_1 and E_2 among the three initially much more likely hypotheses, and we have assumed that one of the four hypotheses is correct. However, the very low likelihood of E_3 (right party resistance to eliminating the tax subsidy in the past) under H_{MV} relative to H_{EA} establishes the equity appeal hypothesis as the leading explanation.



Figure A.3: Probabilities of Hypotheses After Incorporating Evidence: E'₁-E'₆ (left panels) vs. E₁-E₋₆ (right panels)

3) Skeptical Reader Priors

A7. Bayesian Sensitivity Analysis

We now assess the strength of Fairfield's (2013) causal claim. The original case narrative included E_1 , E_2 , E_5 and E_6 . Starting from the skeptical reader priors, we reach 99.4% confidence in H_{EA} after incorporating these four pieces of evidence;¹⁷ the leading alternative hypothesis in this scenario, H_{CC} , is 23 dB less likely than H_{EA} . For a skeptical reader to sustain that the case study does not include sufficient evidence to substantiate the equity-appeal explanation, the relative prior odds against H_{EA} would have to be even lower than we have assumed for this exercise.¹⁸ A prior probability of 10^{-8} on H_{EA} , corresponding to relative log-odds against the hypothesis of 72 dB, would leave the equity appeal hypothesis as plausible as the preferences hypothesis (but still 28 dB more likely than the institutional hypothesis and 17 dB more likely than the median voter hypothesis) in light of the four pieces of evidence. Once we bring in E_3 and E_4 , which Fairfield (2013) included in the process tracing appendix to further substantiate the equity appeal argument, an extremely small prior probability of 10^{-12} on H_{EA} would be needed to leave the posterior probability on H_{EA} similar to the posterior probability on the leading alternative (now H_{MV}).¹⁹ The initial relative log-odds against H_{EA} would be 115 dB, which is extremely large—roughly corresponding to the noise level of a live rock concert. In other words, a skeptical reader would have to feel that his/her background information is "screaming" against H_{EA} .

Of course, a skeptic might contest the likelihood ratios we have assigned for the evidence. However, there are six orders of magnitude to contend with before H_{EA} can be called into question in favor of a rival hypothesis. Table A.4 shows the relative prior odds against H_{EA} that would leave the equity appeal hypothesis as likely as the leading rival hypothesis for three scenarios in which we compress the likelihood ratios of our evidence. In scenario (a) we arbitrarily reduce the likelihood ratios for E_1 and E_3 through E_6 by a factor of ten.²⁰ For E_2 , we set the likelihood under each alternative hypothesis to half the likelihood under H_{EA} to represent a lower degree of confidence in the government informants' judgment and sincerity and hence a higher probability of hearing them declare that the equity appeal mattered if H_{EA} is not correct. The changes introduced in this scenario decrease the relative prior odds against H_{EA} needed for parity with H_{MV} (the leading rival) in light of the evidence from 115 dB to 68 dB. In scenario (b) we compress the likelihood ratios for E_1 and E_3 through E_6 by another factor of five. This scenario reduces the relative prior odds against H_{EA} required for equivalence with H_{MV} to slightly over 40 dB. However, in scenario (c) where we simultaneously lower the prior on H_{MV} to 0.1% while keeping the priors on H_I and H_{CC} equal, the posterior probability on H_{EA} remains higher than the rivals until the relative prior odds against H_{EA} compared to H_{CC} increase to 58 dB.

In sum, to maintain that the case study does not include adequate evidence to substantiate H_{EA} —operationalized as at least 10 dB in favor of H_{EA} relative to the leading alternative—a reader must have an extremely high prior bias against the equity appeal hypothesis and substantial confidence in the median voter hypothesis and/or maintain that the evidence is far less discriminating (in terms of likelihood ratios) than we have argued.

¹⁷Since we judge E_5 and E_6 to be essentially independent of E_3 and E_4 , we can proceed using the likelihoods previously assigned.

¹⁸Assuming the likelihood ratios remain unaltered.

¹⁹A prior-independent way to assess the leverage gained by including E_3 and E_4 is to examine the added weight of evidence in favor of H_{EA} : 56 dB relative to H_I , 60 dB relative to H_{CC} , and 26 dB relative to H_{MV} . These numbers can be obtained by adding the weights of evidence displayed for E_3 and E_4 in Figure A.1.

²⁰In this and the following scenarios we leave the small likelihood ratio for E_1 under H_{EA} vs. H_{CC} unchanged.

	Likelihood ratios ($E_1 - E_5$) reduced by:			
In order to achieve:	a) Factor of 10*	b) Factor of 50**		
	Equal priors on H_I , P	H_{CC} , and H_{MV}	c) Equal priors on H_I , H_{CC} ; H_{MV} prior =0.1%	
Posterior parity with leading rival hypothesis	–68 (noisy restaurant)	-40.5	-58	
Relative posterior odds of 10 dB in favor of H_{EA}	–58 (typical conversation)	–30.5 (watch ticking)	–48 (rainstorm)	

Table A.4: Required Prior Odds against *H*_{EA} (dB) Relative to Most Likely Alternative

*Reduces weight of evidence by 10 dB

**Reduces weight of evidence by 17 dB

References

- Bennett, Andrew. 2015. "Disciplining Our Conjectures: Systematizing Process Tracing with Bayesian Analysis." In Andrew Bennett and Jeffrey Checkel, eds, *Process Tracing in the Social Sciences: From Metaphor to Analytic Tool.* Cambridge University Press, 276–98.
- Fairfield, Tasha. 2013. "Going Where the Money Is: Strategies for Taxing Economic Elites in Unequal Democracies." *World Development* 47 (2013): 42–57.
- Fairfield, Tasha. 2015. Private Wealth and Public Revenue in Latin America: Business Power and Tax Politics. Cambridge University Press.
- Flores-Macías, Gustavo. 2010. "Statist vs. Pro-market: Explaining Leftist Governments' Economic Policies in Latin America." *Comparative Politics*, 42(4), 413–433.
- Hacker, Jacob, and Paul Pierson. 2010. "Winner-Take-All-Politics: Public Policy, Political Organization, and the Precipitous Rise of Top Incomes in the United States." *Politics and Society* 38 (2): 152-204.
- Humphreys, Macartan, and Alan Jacobs. 2015. "Mixing Methods: A Bayesian Approach." *American Political Science Review* 109(4):653-73.
- Jaynes, E.T. 2003. Probability Theory: The Logic of Science. Cambridge University Press.
- Kaufman, Robert. 2009. "The Political Effects of Inequality in Latin America: Some Inconvenient Facts." *Comparative Politics* 41 (3): 359-79.
- Luna, Juan Pablo. 2014. Segmented Representation: Political Party Strategies in Unequal Democracies. Oxford University Press.
- Meltzer, Allan, and Scott Richard. 1981. "A Rational Theory of the Size of Government." *Journal* of *Political Economy* 89 (5): 914–27.
- Tannenwald, Nina. 2007. The Nuclear Taboo: The United States and the Non-Use of Nuclear Weapons since 1945. Cambridge University Press.

Appendix B. Testimonial Evidence and Source Accuracy

In our discussion of "testimonial" evidence consisting of reports, responses, or recollections from human sources, we stress that we should consider the information learned to be "source S stated X" (Section 3.3.2). In other words, logical Bayesianism accounts for the possibility that the claims people make may be inaccurate by treating the testimony—not the alleged facts—as our evidence. This best practice is not always followed in the Bayesian process tracing literature. Some scholars instead attempt to articulate a notion of accuracy that assesses the reliability of the source independently of the hypotheses under consideration (Beach and Pedersen 2013:126-28).²¹ This appendix provides a detailed mathematical exposition of why attempting to make a global assessment of the accuracy of information X provided by source S is generally ill-advised.

To illustrate, let *E* represent the evidence that informant *S* made a statement *X* (in some specific context, e.g. an interview with the author). Evaluating each likelihood $P(E|H_i I)$ requires assessing the informant's potential motives to assert *X* under the given hypothesis H_i , as well as using our background information (which is the same under all hypotheses) to assess the informant's overall sincerity, knowledgeability, integrity, and judgment. In general, the accuracy of the statement *X*—which might be understood as the probability P(X|E I) that *X* is true given that source *S* asserted it—depends on the hypothesis under consideration. The motives we attribute to an informant and hence the probability that s/he is speaking the truth may vary across hypotheses. Furthermore, it may be the case that a given hypothesis directly implies that *X* is true, and hence the statement is accurate, regardless of whether we trust the informant more generally or whether we believe s/he is in a position to have reliable information. Under a different hypothesis, we may instead ascertain that *X* cannot be true, in which case the statement is necessarily inaccurate, and *E* must have occurred because the informant was either mistaken or lying, however reliable we expect that person to be in general.

If we expand the overall "probability of accuracy" P(X|E|I) using the assumption (included in our background information) that we are considering mutually exclusive and exhaustive hypotheses $H_{I}-H_{N}$, we find:

$$P(X|E I) = P(X (H_1 + H_2 + \dots + H_N)|E I) = P(X H_1|E I) + \dots + P(X H_N|E I)$$

= $P(X|E H_1 I) P(H_1|E I) + \dots + P(X|H_N E I) P(H_N|E I)$, (B.1)

where $P(X|E|H_i|I)$ can be regarded as the "hypothesis-dependent accuracy"—the conditional probability that X is true given both that the informant asserted the statement (E) and that a particular hypothesis H_i holds. But note that we now have posterior probability factors $P(H_i|E|I)$ which need to be calculated using Bayes' rule: $P(H_i|E|I) = P(E|H_i|I) P(H_i|I) / P(E|I)$, and we are back to the task of assessing $P(E|H_i|I)$, which we can—and in practice must—evaluate directly, without recourse to P(X|E|I).

We can only move from assessing $P(E | H_i I)$ (the likelihood of hearing the source make a claim) to considering $P(X|H_i I)$ (the likelihood of the asserted facts) in special cases where (i) we judge the accuracy of a statement to be very high across all hypotheses under consideration, and (ii) we judge incentives for the informant to reveal X if it is in fact the truth to be nearly the same across all of the hypotheses. Suppose we wish to calculate likelihood ratios $P(E | H_i I) / P(E | H_j I)$. Because X and $\sim X$ are always mutually exclusive and exhaustive (X is either true or false),

²¹ See also Bennett and Checkel (2015: 24-25).

$$\frac{P(E|H_i I)}{P(E|H_j I)} = \frac{P(E(X + \sim X)|H_i I)}{P(E(X + \sim X|H_j I))} = \frac{P(EX|H_i I) + P(E \sim X|H_i I)}{P(EX|H_j I) + P(E \sim X|H_j I)},$$
(B.2)

If the hypothesis-dependent accuracy of *X* is invariably very high, such that $P(\sim X | E H_n, I)$ is negligibly small for every H_n , then the joint probability $P(E \sim X | H_n, I) = P(\sim X | E H_n, I) P(E | H_n I)$ is also negligibly small because $0 \le P(E | H_n I) \le 1$. Equation (*B*.2) then becomes:

$$\frac{P(E|H_i I)}{P(E|H_j I)} \approx \frac{P(E X|H_i I)}{P(E X|H_j I)} = \frac{P(X|H_i I) P(E|X H_i I)}{P(X|H_j I) P(E|X H_j I)},$$
(B.3)

If it is also the case that incentives for the informant to state *X* when *X* is true do not vary appreciably across the hypotheses, then the second factor in the numerator and denominator above are almost equal and approximately cancel out, leaving us with:

$$\frac{P(E|H_i I)}{P(E|H_j I)} \approx \frac{P(X|H_i I)}{P(X|H_j I)} , \qquad (B.4)$$

where we can now replace E with X in our likelihood ratio assessments. It is important to note that if X flatly contradicts a hypothesis H_n , then we cannot proceed in this manner, because conditioning on the conjunction of X and H_n as in equation (B.3) above would be nonsensical. In such cases, we must have $P(X|H_n I) = 0$, which implies that $P(E X|H_n I) = 0$, and therefore, regardless of how strongly our background information inclines us to trust our informant, we must infer that the informant was either mistaken or prevaricating: $P(E |H_n I) = P(E \sim X|H_n I)$. However, if we have an otherwise high level of trust in the informant, we would assign a low probability to the likelihood $P(E|H_n I)$. To the extent that X and H_n are not flatly contradictory but are jointly highly improbable, we can also expect $P(E|H_n I)$ to be very small if we generally do trust this informant.

With the exception of the specific circumstances discussed above, we will tend to get trapped in logical circles if we try to evaluate $P(E|H_n I)$ by thinking about the accuracy P(X|E I) or even the conditional accuracies $P(X|E H_n I)$. On the other hand, there may be cases for which it is helpful to decompose the probability $P(E|H_n I)$ as:

$$P(E|H_n I) = P(E|X H_n I) P(X|H_n I) + P(E|\sim X H_n I)[1 - P(X|H_n I)]$$
(B.5)

where we first think about how likely it is for X to hold under a hypothesis, and then consider how likely it would be for S to inform us of X if X is true.

For readers familiar with Beach and Pedersen's (2013:127) approach, it may be worth a closer look at how and where their attempt to articulate a hypothesis-independent "accuracy of the evidence" goes wrong. The authors base their discussion on Howson and Urbach's (2006:107-113) treatment of an issue related to the Quine-Duhem Thesis; namely, how evidence can differentially affect our credence in a primary hypothesis and an auxiliary hypothesis. Howson and Urbach derive the following special form of the likelihood ratio in their illustration of how evidence E that might seem to cast doubt on a hypothesis H can in some cases increase its posterior probability while lowering our degree of belief in an auxiliary assumption a:

$$\frac{P(E|\sim H)}{P(E|H)} = \frac{P(E|\sim H a) P(a) + P(E|\sim H \sim a) P(\sim a)}{P(E|H \sim a) P(\sim a)}.$$
(B.6)

(Note that for simplicity, we follow Howson and Urbach's original notation, which contrary to best practice does not explicitly include the background information.) In Howson and Urbach's atomic physics example, the auxiliary assumption a happens to concern the accuracy of the measurement process that produced the evidence.²²

However, when citing equation (*B*.6), Beach and Pedersen omit to note that it rests on two highly restrictive assumptions (Howson and Urbach 2006:110,108):

(i) the probability P(a) of the auxiliary assumption *a* is independent of the primary hypothesis *H* under consideration, such that $P(a|H) = P(a) = P(a|\sim H)$,

(ii) the evidence *E* in question is totally incompatible with the conjunction of the primary hypothesis and the auxiliary assumption about the accuracy of the measurement, such that P(E|Ha) = 0.

As explained in detail above, assumption (i) generally cannot hold when humans serve as our "measuring devices" and we take their testimony as our observed evidence; to reiterate, our assumptions about a source's instrumental incentives and/or unmotivated biases may vary by necessity depending on the hypothesis at hand. Meanwhile, assumption (ii) suggests that the proposition *a* cannot embody any useful claim about accuracy in general. Using the product rule, we can see that assumption (ii) also implies that:

$$P(E | H) P(a | E H) = P(a E | H) = P(a | H) P(E | a H) = 0$$

and we must have either P(E|H) = 0, or P(a | E H) = 0. That is, either the observed evidence flatly contradicts the hypothesis, thereby ruling it out, or the proposition *a* is known to be false whatever the observed evidence *E*. In either case, the proposition about accuracy can do no inferential work. When Howson and Urbach's two assumptions do *not* hold, the likelihood ratio instead takes the more general form:

$$\frac{P(E|\sim H)}{P(E|H)} = \frac{P(E|\sim H a) P(a|\sim H) + P(E|\sim H \sim a) P(\sim a|\sim H)}{P(E|H a) P(a|H) + P(E|H \sim a) P(\sim a|H)},$$
(B.7)

and we see that no simple and general statements can be made about how the probability P(a) affects the value of the posterior odds ratio.

In sum, equation (B.6) is not relevant to testimony from human sources, at least in the way that Beach and Pedersen intend. Even if equation (B.6) were applicable, it is worth noting that Beach and Pedersen's (2013:127-8) conclusions nevertheless would be overly strong. They assert that if the probability $P(\sim a)$ is high, then "E does little to update our confidence in the veracity of the hypothesis." Yet if $P(\sim a)$ is high (close to unity), then the likelihood ratio in equation (B.6) approximately reduces to $P(E|\sim H \sim a)/P(E|H \sim a)$, which does not in itself allow us to infer anything definitive about how our relative belief in H versus $\sim H$ should change, without knowing more about what E, H, and $\sim a$ represent in the problem at hand, and, for social science applications, thinking about the motivations and circumstances that may affect the source's claims in the world of hypothesis H vs. the world of hypothesis $\sim H$.

²² Specifically, *H* is the hypothesis that all atomic weights are integral multiples of hydrogen's atomic weight; $\sim H$ is the hypothesis that atomic weights are randomly distributed; *a* represents a set of assumptions about the absolute accuracy of the measurement technique, and *E* is a measured value of the ratio of the atomic weight of chlorine to hydrogen, 35.83, that was judged incompatible with the conjunction of *H* and *a* (Howson and Urbach's (2007:108).